Measuring Competition with a Flexible Supply Model

Ying Fan*
University of Michigan, CEPR and NBER

Christopher Sullivan†
University of Wisconsin - Madison

Abstract

This paper provides a theoretically founded empirical model to simultaneously investigate firm competition and estimate markups. The model nests the standard oligopoly model, but also allows for firm collusion. Different from conduct parameter models, our model is consistent with a series of theoretical models. We show that a nonparametric marginal cost function can be identified, which gives an estimate of markups. We then apply our model to measure competition in the airline industry.

Key words: markup, firm competition, collusion

1 Introduction

A central question of the empirical Industrial Organization literature is to understand the nature of firm competition in a market and to measure firm markups. They have important implications for consumer welfare, firm profits, and market efficiency.

One strand of the literature estimates markups based on a standard oligopoly model. The key idea is that once demand is estimated from the data, markups can be inferred based on an oligopoly model of supply. This approach dates backed to Rosse (1967), and is widely used in the Industrial Organization literature. However, this approach relies on an oligopoly model assumed by researchers, typically, a classic oligopoly model with Nash equilibrium. In other words, this approach makes an assumption on the nature of firm competition and estimates firm markups.

Another strand of the literature proposes a model deviating from a standard oligopoly model by including so-called conduct parameters and estimates these conduct parameters together with markups. Examples date back to Bresnahan (1982), Lau (1982) and Porter (1983), and include papers as recent as Ciliberto and Williams (2014) and Miller and Weinberg (2017). In such a model, when the conduct parameters take a certain specific value, the model becomes a specific standard form of firm competition (e.g., a classic differentiated Bertrand model or perfect collusion). By nesting the standard oligopoly models, it is therefore more flexible than the above approach.

*Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; yingfan@umich.edu.
†Department of Economics, University of Wisconsin, Madison, WI 53706; cjsullivan@wisc.edu.

However, unfortunately, when the conduct parameters are not of these specific values, it is unclear what kind of firm competition such a model describes.[1] It is, therefore, also unclear whether the estimated markups reflect the true underlying markups. This is mainly because such an approach lacks a theoretical foundation.

An alternative is, therefore, to write down an infinitely repeated game, so that collusion is possible, and estimate the model directly. This approach faces several challenges, too. First, it requires that researchers make assumptions on firm behavior and the nature of the market environment's evolution. For example, we need to assume the strategies taken by firms. Are they using the grim trigger strategies or carrot-and-stick strategies? If the latter, how long is the punishment period? We also need to make an assumption on how the state variables evolve. A Markov process of one order or higher orders? Second, estimating such a dynamic game is typically challenging computationally and imposes a high requirement on data.[2]

In this paper, we propose a theoretically founded empirical model that is easy to estimate and can be used to simultaneously investigate firm competition and estimate markups. We have three goals. First, we derive a model that nests the classic oligopoly models and also allows for the possibility of collusion. Second, we show the identification of nonparametric cost functions. Third, we compare the performance of our approach to that of the two aforementioned strands of literature.

To achieve these goals, we set up an infinitely repeated game and derive its optimality conditions. We show that these optimality conditions are the same as those of a static model where a firm's objective function depends on not only its own profit, but also its competitors' profits and its competitors' deviating profits, which is the highest profit that a firm can achieve holding this firm's opponents' decisions fixed. Considering rivals' deviating profit, which is the difference between our model and a conduct parameter model, is important because a firm needs to take into account how its action affects its opponents' incentive compatibility. We also try to reconcile the difference between our model and the conduct parameter model. We provide an example in which the two models are equivalent. However, this is the only example we can find. It requires homogeneous products, symmetric firms and a linear demand. We show that when any of these assumptions is violated, the two models are not equivalent.

Furthermore, we demonstrate that our model is consistent with a series of supergames, implying that it is robust to a set of model specifications. The key parameters to be estimated include those on each opponent's profit and deviating profit, as well as a marginal cost function. It is immediately obvious from our derivation of the model that the weights vary over time with changes in the market

[1] See Corts (1999) and Reiss and Wolak (2007) for discussions on this point.

[2] See Fershtman and Pakes (2000) for a discussion on how to compute such models. Igami and Sugaya (2017) examine the vitamin cartels in the 90s through the lens of an infinitely repeated game with a grim trigger strategy. This industry is characterized by homogenous goods and relatively good measures of marginal costs. Igami and Sugaya (2017) directly rely on marginal cost data, and avoid the computational challenge by focusing on a Cournot model.

environment. Intuitively, these weights are closely related to the degree of collusion. To what extent firms can collude depends on the market environment. Therefore, these parameters vary by firm pairs and time. However, we show that under the assumption that firms pool incentives across markets and interact as in Bernheim and Whinston (1990)'s seminal paper on multi-market contact, these weight parameters are invariant across markets. Cross-sectional variation, therefore, can be used for identifying these parameters together with the marginal cost function. Based on this intuition, we formally show that the weight parameters and a nonparametric marginal cost function can be identified.

We conduct Monte Carlo simulations to evaluate the performance of our model. Specifically, we set up a supergame and simulate data by solving for its equilibrium. Using these simulated data, we estimate the marginal cost parameters based on our model, the standard oligopoly model and the conduct parameter model. We compare the performance of these models by comparing the estimated marginal cost parameters based on each model to the true marginal cost parameters. We also compare markups and welfare measures.

This paper is related to the literature of estimating markups, documenting their changes and their relationship with market environment. As mentioned, a large number of papers estimate demand and infer markups based on an oligopoly model of supply. An incomplete list of examples include Berry et al. (1995), Goldberg (1995), Nevo (2001), Villas-Boas (2007), Berry and Jia (2010), Fan (2013) and Eizenberg (2014).[3] Compared to this group of papers, our model is more flexible in firm conduct. Another group of papers infer markups based on the production function estimation (instead of demand function estimation) and the assumption of cost minimization (instead of equilibrium conditions implied by a specific competition model). Examples include De Loecker and Warzynski (2012), De Loecker et al. (2016) and Loecker and Eeckhoutz (2017).[4] The latter approach does not impose a specific form of firm conduct. It, however, requires that researchers have a precise measure of inputs for estimating the production function and data on the expenditure share of one variable input. When such data are not available, our paper provides an alternative approach that is also flexible with respect to the assumption of firm conduct, though it does rule out certain supply models such as Green and Porter (1984).

By allowing the possibility of collusion in our model, this paper is also related to a literature studying collusion. Examples include Porter (1983), Slade (1987), Sudhir (2001), Ciliberto and Williams (2014), Miller and Weinberg (2017) and Khwaja and Shim (2017). Many of these papers are based on a conduct parameter model, which lacks a theoretical foundation. Other papers estimate a set of models assuming different forms of firm conduct and use non-nested hypotheses test to select among the competing models, for example, Bresnahan (1987), Gasmi et al. (1992) and Nevo (1998). Different from these papers, we estimate one model that nests different forms of

[3]See also Berry and Haile (2014) for the identification of nonparametric cost functions using this approach when demand is described by the widely used discrete-choice model.

[4]See also Loecker and Scott (2016) for a comparison of the demand approach and the production approach.

firm conduct. Moreover, our framework allows us to go beyond testing firm conduct to estimate markups and welfare.

We proceed with the rest of the paper as follows. We briefly discuss the standard oligopoly model and the conduct parameter model in Section 2. In Section 3, we set up an infinitely repeat game and derive its static representation. We also compare and highlight the similarity and the difference between our model and the conduct parameter model. We discuss the identification in Section 4 and present the simulation results in Section 5. Finally, we conclude in Section 7.

2 The Standard Oligopoly Model and the Conduct Parameter Model

For expositional simplicity, consider a market with only two firms. Each firm i ($i = 1, 2$) chooses its action x_{imt} in market m in period t. The profit functions are given by $\pi_1(x_{1mt}, x_{2mt}, z_{mt})$ and $\pi_2(x_{1mt}, x_{2mt}, z_{mt})$, where z_{mt} are profit shiftors, for example, demand shiftors that shift or rotate demand. In a standard static oligopoly model, the first-order conditions are

$$\frac{\partial \pi_i(x_{1mt}, x_{2mt}, z_{mt})}{\partial x_{imt}} = 0. \tag{1}$$

The conduct parameter models "extend" the standard oligopoly model by allowing the first-order condition to deviate from (1) as follows:

$$\frac{\partial \pi_i(x_{1mt}, x_{2mt}, z_{mt})}{\partial x_{imt}} + \lambda_{ij}\frac{\partial \pi_j(x_{1mt}, x_{2mt}, z_{mt})}{\partial x_{imt}} = 0, \tag{2}$$

where λ_{ij} is the conduct parameter. When $\lambda_{ij} = 0$, the model becomes a standard oligopoly model. When $\lambda_{ij} = 1$, the model implies perfect collusion. In general, these conduct parameters are interpreted as how much one firm internalizes the effect of its action on its competitors. In other words, the model is equivalent to a non-cooperative oligopoly model where the objective function of firm i is

$$\tilde{\pi}_i(x_{1mt}, x_{2mt}, z_{mt}) = \pi_i(x_{1mt}, x_{2mt}, z_{mt}) + \lambda_{ij}\pi_j(x_{1mt}, x_{2mt}, z_{mt}), \tag{3}$$

where λ_{ij} is the weight that firm i assigns to firm j's profit. These parameters are typically assumed to be constant over time and across markets, and are identified exploiting cross-sectional or temporal variations in the profit shiftors z_{mt}.

While these profit weight parameters are intuitively related to firm conduct, such models do not have a theoretical foundation. Note that firms in the model are still assumed to be independent rather than formally integrated. It is unclear why they include their competitors' profits into their own objective functions. Because the conduct parameters are the profit weight parameters, we refer to this type of conduct parameter model as a profit-weight conduct parameter model.

Another type of conduct parameter model is motivated by the troubled conjecture variations literature. Since the problems with such models are well recognized, we render the discussion of such models and the comparison of them to the profit-weight conduct parameter model in (3) to Supplemental Appendix SA.

Because there is no theoretical foundation, for a value of λ_{ij} between 0 and 1, it is unclear how to interpret these parameters.[5] For the same reason, when $\lambda_{ij} \in (0, 1)$, it is unclear whether the profit-weight conduct parameter model is consistent with any form of firm competition. Such models, therefore, cannot be used to estimate the underlying marginal costs. In the next section, we provide a model that can be used to simultaneously test whether firms collude and estimate marginal cost functions.

3 Our Model

Our proposed model, in a two-firm economy, is

$$\tilde{\pi}_i\left(x_{1mt}, x_{2mt}, z_{mt}\right) = \pi_i\left(x_{1mt}, x_{2mt}, z_{mt}\right) + \lambda_{ijt}\pi_j\left(x_{1mt}, x_{2mt}, z_{mt}\right) - \rho_{jt}\lambda_{ijt}\pi_j^d\left(x_{imt}, z_{mt}\right) \qquad (4)$$

where π_j^d is firm j's profit evaluated at its best-response to its opponent's action x_{imt}, which is known as the deviating profit in infinitely repeated game models. This model differs from the profit-weight conduct parameter model in the last term. The necessity of including this term will be clear as we derive the model below. Intuitively, firm i needs to take into account how its action affects its opponent j's incentive compatibility, which depends on the comparison of j's on equilibrium payoff and its deviating payoff.

A supergame consistent with the model

We first present a supergame that is consisten with the model in (4). We later discuss what other supergames are consistent with the model. Consider an infinitely repeated game with two firms and grim trigger strategies. Let $\pi_{im}\left(x_{1mt}, x_{2mt}, z_{mt}\right)$ be the stage profit for firm i in market m, and δ be the discount factor. Suppose z_{mt} follows a stationary first-order Markov process. Consider a Pareto optimal supergame equilibrium where any deviation in any market is punished by reversing to one-shot Nash equilibrium strategies forever in all markets, which is the harshest punishment. A supergame equilibrium is Pareto optimal if the payoff of one firm cannot be increased without decreasing the payoff of the other firm. Therefore, a Pareto optimal supergame equilibrium

[5] Focusing on the conjecture variation type of conduct models, Corts (1999) points out that the conduct parameter is not a good measure for the degree of collusion. In a simulation, Corts (1999) sets up a supergame and allow the discount factor to vary from 0 to 1 (i.e., the outcome varies from Nash equilibrium with no collusion to perfect collusion). He finds that estimated conduct parameter do not move monotonically, making it difficult to interpret the conduct parameter as a measure for the degree of collusion. A similar point is also mentioned in Reiss and Wolak (2007): "Currently there is no satisfactory economic interpretation of this parameter as a measure of firm behavior."

$\left(x_1^*\left(z_{mt}\right), x_2^*\left(z_{mt}\right)\right)$ satisfies

$$
\begin{aligned}
& \left(x_1^*\left(z_{mt}\right), x_2^*\left(z_{mt}\right)\right) \\
= & \underset{\left\{\left(x_{1mt}, x_{2mt}\right), m=1, \ldots, M\right\}}{\arg\max} \sum_m \left[\omega_1 \pi_1\left(x_{1mt}, x_{2mt}, z_{mt}\right)+\omega_2 \pi_2\left(x_{1mt}, x_{2mt}, z_{mt}\right)\right] \\
& s.t. \quad \textstyle\sum_m \left[\pi_i\left(x_{1mt}, x_{2mt}, z_{mt}\right)+\sum_{\tau=1}^{\infty} \delta^\tau E\left\{\pi_i\left(x_1^*\left(z_{mt+\tau}\right), x_2^*\left(z_{mt+\tau}\right), z_{mt+\tau}\right) | z_{mt}\right\}\right] \\
& \qquad \geq \textstyle\sum_m \left[\pi_i^d\left(x_{-imt}, z_{mt}\right)+\sum_{\tau=1}^{\infty} \delta^\tau E\left\{\pi_i\left(x_1^{NE}\left(z_{mt+\tau}\right), x_2^{NE}\left(z_{mt+\tau}\right), z_{mt+\tau}\right) | z_{mt}\right\}\right] \\
& \qquad \text{for } i=1,2,
\end{aligned} \tag{5}
$$

where ω_1 is normalized to be 1, $\omega_2>0$, and $\pi_i^d\left(x_{-it}, z_t\right)$ and $x_i^{NE}\left(z_{t+\tau}\right)$ represent, respectively, the highest deviating profit and the Nash equilibrium strategy. The constraints in (5) are the incentive compatibility constraints for each firm. They define the set of outcomes that can be sustained as a collusive outcome. In this supergame, firms pool incentives across markets. These incentive compatibility constraints are the same as those in Bernheim and Whinston (1990)'s multi-market contact paper. Denote these incentive compatibility constraints as $IC_i\left(\boldsymbol{x}_{1t}, \boldsymbol{x}_{2t}, \boldsymbol{z}_t\right) \geq 0$.

The optimality conditions of (5) are

$$
\begin{aligned}
\omega_1 \frac{\partial \pi_{1mt}}{\partial x_{1mt}}+\omega_2 \frac{\partial \pi_{2mt}}{\partial x_{1mt}}+\gamma_{1t} \frac{\partial \pi_{1mt}}{\partial x_{1mt}}+\gamma_{2t}\left(\frac{\partial \pi_{2mt}}{\partial x_{1mt}}-\frac{\partial \pi_{2mt}^d}{\partial x_{1mt}}\right) & = 0 \\
\omega_1 \frac{\partial \pi_{1mt}}{\partial x_{2mt}}+\omega_2 \frac{\partial \pi_{2mt}}{\partial x_{2mt}}+\gamma_{1t}\left(\frac{\partial \pi_{1mt}}{\partial x_{2mt}}-\frac{\partial \pi_{1mt}^d}{\partial x_{2mt}}\right)+\gamma_{2t} \frac{\partial \pi_{2mt}}{\partial x_{2mt}} & = 0,
\end{aligned} \tag{6}
$$

where γ_{1t} and γ_{2t} are the Karush–Kuhn–Tucker multipliers (KKT multipliers, henceforth) corresponding to the two incentive compatibility constraints in (5). They satisfy $\gamma_{it} \geq 0$ and $\gamma_{it} IC_i\left(\boldsymbol{x}_{1t}, \boldsymbol{x}_{2t}, \boldsymbol{z}_t\right)=0$ for $i=1,2$. These KKT multipliers are indexed by t because the constraints $IC_i\left(\boldsymbol{x}_{1t}, \boldsymbol{x}_{2t}, \boldsymbol{z}_t\right)$ vary over time. Since the incentives are pooled across markets (similar to Bernheim and Whinston (1990)), they do not, however, vary across markets. Rearranging (6) yields

$$
\begin{aligned}
\frac{\partial \pi_{1mt}}{\partial x_{1mt}}+\frac{\omega_2+\gamma_{2t}}{\omega_1+\gamma_{1t}} \frac{\partial \pi_{2mt}}{\partial x_{1mt}}-\frac{\gamma_{2t}}{\omega_1+\gamma_{1t}} \frac{\partial \pi_{2mt}^d}{\partial x_{1mt}} & = 0 \\
\frac{\partial \pi_{2mt}}{\partial x_{2mt}}+\frac{\omega_1+\gamma_{1t}}{\omega_2+\gamma_{2t}} \frac{\partial \pi_{1mt}}{\partial x_{2mt}}-\frac{\gamma_{1t}}{\omega_2+\gamma_{2t}} \frac{\partial \pi_{1mt}^d}{\partial x_{2mt}} & = 0.
\end{aligned} \tag{7}
$$

Let $\lambda_{ijt}=\frac{\omega_j+\gamma_{jt}}{\omega_i+\gamma_{it}}$ and $\rho_{jt}=\frac{\gamma_{jt}}{\omega_j+\gamma_{jt}}$. Then, (7) can be further simplified to

$$
\frac{\partial \pi_{imt}}{\partial x_{imt}}+\lambda_{ijt} \frac{\partial \pi_{jmt}}{\partial x_{imt}}-\rho_{jt} \lambda_{ijt} \frac{\partial \pi_{jmt}^d}{\partial x_{imt}}=0. \tag{8}
$$

The optimality condition in (8) is equivalent to that of a static non-cooperative model where firm i's objective function is given in (4). In other words, our model in (4) can be considered the static representation of the supergame. Since the optimality conditions (8) will be used for estimation, with a slight abuse of terminology, we also refer to the optimality conditions in (8), instead of model (4), as our model.

As an internal consistency check, we check whether the model (8) becomes $\frac{\partial \pi_{imt}}{\partial x_{imt}} = 0$ (i.e., a one-shot Nash equilibrium condition) when $\delta \to 0$, and $\omega_i \frac{\partial \pi_{imt}}{\partial x_{imt}} + \omega_j \frac{\partial \pi_{jmt}}{\partial x_{imt}} = 0$ (i.e., the perfect collusion condition) when $\delta \to 1$. Note that when $\delta \to 0$, the KKT multipliers $\gamma_{it} \to \infty$.[6] Therefore, $\rho_{jt} = \frac{\gamma_{jt}}{\omega_j + \gamma_{jt}} \to 1$. At the same time, $\frac{\partial \pi_{jmt}}{\partial x_{imt}} - \frac{\partial \pi^d_{jmt}}{\partial x_{imt}} \to 0$. Therefore, $\frac{\partial \pi_{imt}}{\partial x_{imt}} = 0$. When $\delta \to 1$, however, the incentive compatibility constraints become non-binding, implying that $\gamma_{it} \to 0$. As a result, $\lambda_{ijt} \to \frac{\omega_j}{\omega_i}$, $\rho_{jt} \to 0$, and (8) becomes $\frac{\partial \pi_{imt}}{\partial x_{imt}} + \frac{\omega_j}{\omega_i} \frac{\partial \pi_{jmt}}{\partial x_{imt}} = 0$.

The model can be easily extended to N firms. The corresponding optimality condition is

$$\frac{\partial \pi_{imt}}{\partial x_{imt}} + \sum_{j \neq i, j \in \mathcal{J}_{mt}} \left(\lambda_{ijt} \frac{\partial \pi_{jmt}}{\partial x_{imt}} - \rho_{jt} \lambda_{ijt} \frac{\partial \pi^d_{jmt}}{\partial x_{imt}} \right) = 0, \tag{9}$$

where $\mathcal{J}_{mt}$ is the set of firms in market m at time t.

Other supergames consistent with the model

The model (8) is also consistent with a supergame where the Nash reversion punishment of a finite T periods is used (instead of infinite periods as in the grim trigger strategy). In this case, the incentive compatibility constraints become

$$\begin{aligned} & \sum_m \left[\pi_i \left(x_{1mt}, x_{2mt}, z_{mt} \right) + \sum_{\tau=1}^{T} \delta^\tau E \left\{ \pi_i \left(x^{NE}_{1m} \left(z_{mt+\tau} \right), x^{NE}_{2m} \left(z_{mt+\tau} \right), z_{mt+\tau} \right) | z_{mt} \right\} \right] \\ \leq & \sum_m \left[\pi^d_i \left(x_{-imt}, z_{mt} \right) + \sum_{\tau=1}^{T} \delta^\tau E \left\{ \pi_i \left(x^{NE}_{1m} \left(z_{mt+\tau} \right), x^{NE}_{2m} \left(z_{mt+\tau} \right), z_{mt+\tau} \right) | z_{mt} \right\} \right]. \end{aligned} \tag{10}$$

The optimality conditions remain the same as (8).[7]

Moreover, the process of z_{mt} can be different in different markets and can be a higher-order Markov process. In such a case, the expectation operator $E\{\cdot | z_{mt}\}$ in (5) becomes a market/time specific operator $E_m \{\cdot | z_{mt}, z_{mt-1}, ...\}$. But the optimality conditions again remain the same.

The model does not impose $\omega_i = 1$ for all i. It is thus silent about whether there are transfers among firms. When transfers are allowed, the outcome should be such that the joint profit is maximized and thus $\omega_i = 1$ for all i. The model allows this possibility, but does not impose it. We can also relax the assumption of Pareto optimality, permitting firms to collude at equilibria that lie in the interior of the set of feasible actions defined by the incentive constraints.[8]

[6] When $\delta = 0$, $\frac{\partial \pi_{jmt}}{\partial x_{imt}} = \frac{\partial \pi^d_{jmt}}{\partial x_{imt}}$. Therefore, (6) implies that $\gamma_{1t} = -\left(\omega_1 \frac{\partial \pi_{1mt}}{\partial x_{1mt}} + \omega_2 \frac{\partial \pi_{2mt}}{\partial x_{1mt}} \right) / \frac{\partial \pi_{1mt}}{\partial x_{1mt}}$. Given that $\frac{\partial \pi_{1mt}}{\partial x_{1mt}} = 0$ at Nash equilibrium, $\gamma_{1t} = \infty$. Similarly, $\gamma_{2t} = \infty$, too.

[7] This is not to say that even the parameters such as λ_{ijt} remain the same. In fact, as the incentive compatibility constrains change, these parameters most likely change too.

[8] We will formally show this in future versions of the paper.

The fact that the model is consistent with a series of supergames is an advantage: researchers do not need to specify certain details of the underlying supergame. In a similar vein, one does not need to impose or estimate the discount factor δ. As a result, the model is more robust to possible model mis-specifications.

This model, however, does rule out some supergames. Or, put it differently, some underlying supergames are inconsistent with this static representation. For example, the process of z_{mt} cannot be influenced by firms' actions x_{imt}. Otherwise, the optimality condition would have to include how x_{imt} affects the continuation value. Similarly, the model requires perfect monitoring. Otherwise, the optimality condition would have to include how x_{imt} affects the probability of regime shifting, from cooperation to punishment and vice versa (see, for example, Green and Porter (1984)). It also requires that the process z_{mt} is stationary. Otherwise, we cannot define a Pareto optimal supergame equilibrium as a fixed point to (5). Finally, it requires the equilibrium strategies to be such that firms revert to the one-shot Nash equilibrium during the punishment phase. This ensures that firms will not deviate from punishment, allowing us to drop the incentive compatibility constraints during the punishment period from (5).

Just as being consistent with a set of supergames is an advantage of the model, being inconsistent with another set is a disadvantage. However, by nesting a standard oligopoly model, our model is more flexible than the latter. It allows us to estimate markups in industries and during time periods where we are not sure if a standard oligopoly model captures firm competition well. One can also say that our model is more flexible than a conduct parameter model because it nests the latter by setting ρ_j to be zero. More importantly, our model is derived from a well-specified theoretical model.

Comparison to the profit-weight conduct parameter model

Our model differs from the profit-weight conduct parameter model in the last term: in our model, firm i considers not only the effect of its action on its competitor's profit $\left(\text{i.e., } \frac{\partial \pi_{jmt}}{\partial x_{imt}}\right)$, but also how its action affects its opponent's deviating profit (i.e., $\frac{\partial \pi^d_{jmt}}{\partial x_{imt}}$). Intuitively, both effects impact the constraint set and thus the outcome.

One could further rewrite our model (8) as $\frac{\partial \pi_{imt}}{\partial x_{imt}} + \left(\lambda_{ijt} - \rho_{jt}\lambda_{ijt}\frac{\partial \pi^d_{jmt}}{\partial x_{imt}}/\frac{\partial \pi_{jmt}}{\partial x_{imt}}\right)\frac{\partial \pi_{jmt}}{\partial x_{imt}} = 0$ so that it resembles the profit-weight conduct parameter model (2). However, $\frac{\partial \pi^d_{jmt}}{\partial x_{imt}}/\frac{\partial \pi_{jmt}}{\partial x_{imt}}$ is i, j, m, t-specific, which means the term $\left(\lambda_{ijt} - \rho_{jt}\lambda_{ijt}\frac{\partial \pi^d_{jmt}}{\partial x_{imt}}/\frac{\partial \pi_{jmt}}{\partial x_{imt}}\right)$ varies at the level of the data variation, and thus cannot be considered a parameter. Therefore, the two equations (2) and (8) are equivalent if and only if $\frac{\partial \pi^d_{jmt}}{\partial x_{imt}}/\frac{\partial \pi_{jmt}}{\partial x_{imt}}$ is invariant in m. This is in general not true. In Appendix A, we show that when the two firms produce homogenous products, face a linear market demand and have the same marginal cost, the ratio $\frac{\partial \pi^d_{jmt}}{\partial x_{imt}}/\frac{\partial \pi_{jmt}}{\partial x_{imt}}$ is indeed invariant across markets. This is the only example we can find. In a Bertrand setting, when demand is not linear, or when the two firms are asymmetric, the ratio becomes market specific.

Interpretation of the parameters λ_{ijt} and ρ_{jt}

In a profit-weight conduct parameter model $\frac{\partial \pi_{imt}}{\partial x_{imt}} + \lambda_{ijt} \frac{\partial \pi_{jmt}}{\partial x_{imt}} = 0$, the conduct parameter λ_{ijt} is interpreted as how much firm i cares about firm j when making a decision, with a larger value of λ_{ijt} indicating a higher degree of collusion. Our model is $\frac{\partial \pi_{imt}}{\partial x_{imt}} + \lambda_{ijt} \frac{\partial \pi_{jmt}}{\partial x_{imt}} - \rho_{jt} \lambda_{ijt} \frac{\partial \pi^d_{jmt}}{\partial x_{imt}} = 0$. As we will argue below, the parameter λ_{ijt} actually does not capture the degree of collusion.

Note that $\lambda_{ijt} = \frac{\omega_j + \gamma_{jt}}{\omega_i + \gamma_{it}}$, implying that $\lambda_{ijt} = 1/\lambda_{jit}$. Therefore, λ_{ijt} and λ_{jit} do not move in the same direction. They cannot both capture the degree of collusion. But it is also conceptually impossible that λ_{ijt} captures the degree of collusion while λ_{jit} is negatively correlated with it. What they really capture is some asymmetry between the two firms. These parameters depend on (ω_i, ω_j) and $(\gamma_{it}, \gamma_{jt})$. One could consider the profit weights ω_i and ω_j in the objective function of (5) to be determined by the bargaining power between the two firms. Therefore, λ_{ijt} captures the asymmetry in the bargaining power. Similarly, since the KKT multipliers γ_{it} and γ_{jt} are related to the slackness of each firm's incentive compatibility constraint, λ_{ijt} also captures the asymmetry in the slackness. In sum, $\lambda_{ijt} = 1/\lambda_{jit}$ captures the asymmetry between the two firms in terms of the slackness of their incentive compatibility constraints and their bargaining power. It has nothing to do with the degree of collusion, or firm conduct.

What is closely related to the degree of collusion is ρ_{jt}. As explained, $\rho_{jt} \to 1$ when $\delta \to 0$. At the other extreme, when the discount factor $\delta = 1$, the constraints in (5) are non-binding, implying $\gamma_{it} = 0$ and $\rho_{jt} = \frac{\gamma_{jt}}{\omega_j + \gamma_{jt}} = 0$. In sum, $\rho_{jt} = 1$ when the outcome approaches a Nash equilibrium without collusion (when $\delta = 0$) and $\rho_{jt} = 0$ when the outcome is perfect collusion (when $\delta = 1$). What about $\delta \in (0, 1)$? Is ρ_{jt} decreasing in δ? Appendix shows that this is the case. Note that as δ increases, the solution to (5) yields a higher joint profit for the firms $(\omega_1 \pi_1 + \omega_2 \pi_2)$, indicating a higher degree of collusion. Therefore, a monotonic relationship between ρ_{jt} and δ implies a monotonic relationship between ρ_{jt} and the degree of collusion increases, at least as the discount factor varies.

That said, we do not think one should label ρ_{jt} as a measure of firm conduct, or use the comparison of its the value across industries to suggest one industry is more competitive than the other. This is because the parameter ρ_{jt} is intrinsically industry-specific, depending on the market environment in an industry, such as demand, the number of firms, firms' efficiency levels, etc. A comparison of this parameter across industries does not necessarily inform us about the underlying form of firm competition. As mentioned, instead of providing a measure of the degree of collusion, the focus of this paper is to provide a model for estimating marginal costs and hence markups allowing the possibility of collusion. Given that focus, we now turn to the identification and the estimation.

4 Identification and Estimation

The objective is to estimate marginal costs. We now make marginal costs explicit in the profit function. Specifically, let $c_{imt} = f(\boldsymbol{w}_{imt}) + \omega_{imt}$ be the marginal cost function. With a slight abuse of notation, let the profit function $\pi_i(\boldsymbol{x}_{mt}, \boldsymbol{z}_{mt})$ in the previous section be $\pi_i(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}, c_{imt}) = \pi_i(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}, f(\boldsymbol{w}_{imt}) + \omega_{imt})$, where $\boldsymbol{x}_{mt}$ is the vector of all firms' actions, $\boldsymbol{y}_{mt}$ is the demand shiftor, $\boldsymbol{w}_{imt}$ is the observable cost shiftor and ω_{imt} is the unobservable cost shiftor . We assume that the function π_i is known based on the demand estimation. In the demand estimation, there might be unobservable demand shocks. These shocks become known after the demand estimation. We therefore consider them part of the demand shiftors y_{mt}. For example, in a Cournot setting, the profit function is $\pi_i(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}, c_{imt}) = [p(\sum_i x_{imt}, \boldsymbol{y}_{mt}) - c_{imt}]\, x_{imt}$. In a Bertrand setting, $\pi_i(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}, c_{imt}) = (x_{imt} - c_{imt})\, M_{mt} s_i(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt})$, where M_{mt} is the market size and s_i is the market share function.

With this notation, our model is therefore

$$\begin{aligned}
&\frac{\partial \pi_i}{\partial x_{imt}}(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}, c_{imt}) \\
&+ \sum_{j \neq i, j \in \mathcal{J}_{mt}} \left(\lambda_{ijt} \frac{\partial \pi_{jm}}{\partial x_{imt}}(\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}, c_{jmt}) - \rho_{jt}\lambda_{ijt} \frac{\partial \pi_{jm}^d}{\partial x_{imt}}(\boldsymbol{x}_{mt} \backslash x_{jmt}, \boldsymbol{y}_{mt}, c_{jmt}) \right) = 0.
\end{aligned} \tag{11}$$

where $\lambda_{ijt} = \frac{\omega_j + \gamma_{jt}}{\omega_i + \gamma_{it}}$ and $\rho_{jt} = \frac{\gamma_{jt}}{\omega_j + \gamma_{jt}}$. Let $\tau_{jt} = \omega_j + \gamma_{jt}$. Then, $\lambda_{ijt} = \frac{\tau_{jt}}{\tau_{it}}$. Since $\omega_i > 0$ and $\gamma_{it} \geq 0$, $\tau_{it} > 0$ and $0 \leq \rho_{it} \leq 1$. Moreover, since ω_1 is normalized to be 1, $\tau_1 = \frac{1}{1-\rho_1}$. Therefore, what we need to identify is $\tau_{2t}, ..., \tau_{Nt}, \rho_{1t}, ..., \rho_{Nt}$, and the marginal cost function $f(\cdot)$, where $\tau_{it} > 0$ and $0 \leq \rho_{it} \leq 1$.

In a Cournot setting with homogenous products, the action x_{imt} is the quantity produced by a firm, and (11) becomes

$$\begin{aligned}
&(p_{mt} - c_{imt}) + p_m^{'}(X_{mt}, \boldsymbol{y}_{mt})\, x_{imt} \\
&+ \sum_{j \neq i, j \in \mathcal{J}_{mt}} \frac{\tau_{jt}}{\tau_{it}} p_m^{'}(X_{mt}, \boldsymbol{y}_{mt})\, x_{jmt} \\
&- \sum_{j \neq i, j \in \mathcal{J}_{mt}} \rho_{jt} \frac{\tau_{jt}}{\tau_{it}} p_m^{'}\left(X_{-j,mt} + x^{BR}(X_{-j,mt}, c_{jmt}, \boldsymbol{y}_{mt})\right) x^{BR}(X_{-j,mt}, c_{jmt}, \boldsymbol{y}_{mt}) = 0
\end{aligned} \tag{12}$$

where $X_{mt} = \sum_{i \in \mathcal{J}_{mt}} x_{imt}$ is the total quantity, $X_{-j,mt}$ is firm j's opponents' total quantity and $x^{BR}(X_{-j,mt}, c_{jmt}, \boldsymbol{y}_{mt})$ is firm j's best-response. The identification question is, therefore, whether we can identify the parameters $\tau_{2t}, ..., \tau_{Nt}, \rho_{1t}, ..., \rho_{Nt}$, and the function $f(\cdot)$ with data $(\boldsymbol{x}, p, \boldsymbol{y}, \boldsymbol{w})$.

In a Bertrand setting with differentiated products and single-product firms, the action x_{imt} is the price charged by a firm, and (11) becomes

$$
\begin{aligned}
&(x_{imt} - c_{imt}) \frac{\partial s_i}{\partial x_{imt}} (\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}) + s_{imt} \\
&+ \sum_{j \neq i, j \in \mathcal{J}_{mt}} \frac{\tau_{jt}}{\tau_{it}} (x_{jmt} - c_{jmt}) \frac{\partial s_j}{\partial x_{imt}} (\boldsymbol{x}_{mt}, \boldsymbol{y}_{mt}) \\
&- \sum_{j \neq i, j \in \mathcal{J}_{mt}} \rho_{jt} \frac{\tau_{jt}}{\tau_{it}} \left(x_j^{BR} (\boldsymbol{x}_{-j,mt}, \boldsymbol{y}_{mt}, c_{jmt}) - c_{jmt} \right) \frac{\partial s_j}{\partial x_{imt}} \left(\boldsymbol{x}_{-j,mt}, x_j^{BR} (\boldsymbol{x}_{-j,mt}, \boldsymbol{y}_{mt}, c_{jmt}), \boldsymbol{y}_{mt} \right) = 0,
\end{aligned} \tag{13}
$$

where $\boldsymbol{x}_{-j,mt} = \boldsymbol{x}_{mt} \backslash x_{jmt}$ is the vector of all firms' prices except firm j, $x_j^{BR} (\boldsymbol{x}_{-j,mt}, \boldsymbol{y}_{mt}, c_{jmt})$ is firm j's best response. The identification question in this setting is whether the parameters $\tau_{2t}, ..., \tau_{Nt}, \rho_{1t}, ..., \rho_{Nt}$, and the function $f(\cdot)$ with data $(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{y}, \boldsymbol{w})$.

First, we show that for any given $\tau_{2t}, ..., \tau_{Nt}, \rho_{1t}, ..., \rho_{Nt}$ such that $\tau_{it} > 0$ and $0 \leq \rho_{it} \leq 1$, and for any given dataset, there is a unique vector of marginal costs $(c_{1t}, ..., c_{Nt})$ that satisfies (12) or (13). Different from a standard Cournot or a Bertrand model, where we only have the first line in (12) or (13), marginal costs enters the effect of firm i's action on firm j's deviating profit nonlinearly in the third line in (12) or (13). Therefore, the invertibility is not immediately obvious. Appendix B gives the proof of the invertibility for any Cournot model and a Logit Bertrand model. While we have not formally extended our proof to more complex demand systems, we have conducted simulations which suggest that the model is invertible for nested logit and discrete consumer type logit demand systems.

The invertibility result gives us the following estimation equation

$$c_{imt} (\boldsymbol{x}_{mt}, \boldsymbol{v}_{mt}, \boldsymbol{y}_{mt}; \boldsymbol{\tau}_t, \boldsymbol{\rho}_t) = f(\boldsymbol{w}_{imt}) + \omega_{imt},$$

where the variable $\boldsymbol{v}_{mt}$ is the market price p_{mt} in the Cournot setting and the vector of market shares $\boldsymbol{s}_{mt}$ in the Bertrand setting. Both $\boldsymbol{x}_{mt}$ (the actions) and $\boldsymbol{v}_{mt}$ (some outcomes) are endogenous, while the demand shiftors $\boldsymbol{y}_{mt}$ and the cost shiftors $\boldsymbol{w}_{imt}$ are assumed to be exogenous.

Since the parameters $\boldsymbol{\tau}_t$ and $\boldsymbol{\rho}_t$ vary over time, we cannot rely on cross-time variations to identify them. However, they are stable across markets. This is because when firms pool incentives across markets as in Bernheim and Whinston (1990), there is only one incentive compatibility constraint per firm per period. As a result, cross-market variations can be used to identify these parameters.

5 Simulations

In this section, we perform Monte Carlo simulations to evaluate the performance of our model relative to two benchmarks: the standard Nash Bertrand oligopoly model and the profit weight conduct parameter model. To do so, we set up a repeated game and simulate data by solving for the Pareto optimal equilibrium. Specifically, we consider two single-product firms (indexed by j)

which compete in 100 independent geographic markets (indexed by m) each period t. In order to simplify the computation of the equilibrium, we assume that the demand and cost shocks faced by the firms in each market are perfectly persistent over time. Because shocks are perfectly persistent, we drop the subscript t in what follows. We define two primitives, consumer utility and marginal cost. We impose logit demand, such that the utility that consumer i receives from purchasing the product produced by firm j in geographic market m is given as:

$$U_{ijm} = \beta_x x_{jm} + \alpha p_{jm} + \beta_m + \xi_{jm} + \epsilon_{ijm} \tag{14}$$

where x_{jm} is a scalar observable product characteristic, β_m is a market fixed effect, ξ_{jm} is an unobserved market specific taste for product j, and ϵ_{ijm} is a idiosyncratic demand shock. We assign observable characteristics for firm 1 and 2 in market m according to $(x_{1m}, x_{2m}) \in K \otimes K$, where $K = \{4, 5, 7, 8, 10\}$. This gives us 25 distinct combinations of product characteristics for firms 1 and 2. As there are 100 geographic markets, we assume that each combination of product characteristics appears in 4 geographic markets. For each geographic market, we draw a fixed effect β_m from $\mathcal{N}(-1, 1)$, and for each firm market pair we draw the unobservable characteristic ξ_{jm} from $\mathcal{N}(0, 1)$. As is standard, ϵ_{ijm} is distributed type 1 extreme value. Finally, we impose that $\beta_x = 1.5$ and that $\alpha = -0.5$.

Next, we define the constant marginal cost to firm j in market m as:

$$c_{jm} = \mu_0 + \mu_1 w_{jm} + \eta_{jm} \tag{15}$$

where the scalar observable cost covariate, w_{jm} is drawn from $\mathcal{N}(3, 1)$ and the unobserved cost shock, η_{jm}, is drawn from $\mathcal{N}(0, 1)$. We impose $\mu_0 = 1.5$ and $\mu_1 = 1$. To complete the set up, we draw the market size from a lognormal (0, 1) distribution. With demand and cost specified, we solve for prices p_{jm} as the Pareto optimal stationary equilibrium to this game for a set of discount factors δ. In particular, we solve the game 10 times, for each $\delta \in \{0, 0.05, ..., 0.45\}$. We do not consider higher values of δ because at $\delta = 0.45$, the firms charge the monopoly prices.

Once we have solved the model for a given value of δ, we have data on prices and market shares. Using this data, we estimate marginal costs c_{jm} in each of the three models of interest. While cost can be inverted analytically from the Nash Bertrand first order condition, the profit weight model and our model require us to estimate parameters of the first order condition. As has been discussed in Bresnahan (1982) and Berry and Haile (2014), in order to distinguish cost from conduct, instruments are needed. In our setting, the market size, market fixed effect, unobserved characteristic ξ_{jm}, and the rival's observable cost shifter are all valid instruments which we use in estimating both the profit weight model and our model.

Let c_{jm} be true marginal cost of firm j in market m and let $\hat{c}_{jm}$ be the estimated marginal cost. Table 1 presents summary statistics for the squared error in the marginal cost estimates for the Bertrand model, profit weight model, and our model. The results reported in Table 1 for the Bertrand model are not surprising. We are able to back out the true marginal cost using the Bertrand first order conditions when the firms are not colluding and $\delta = 0$. As we would suspect, the squared error increases monotonically with the discount rate as the prices increase from the Bertrand prices to the monopoly prices.

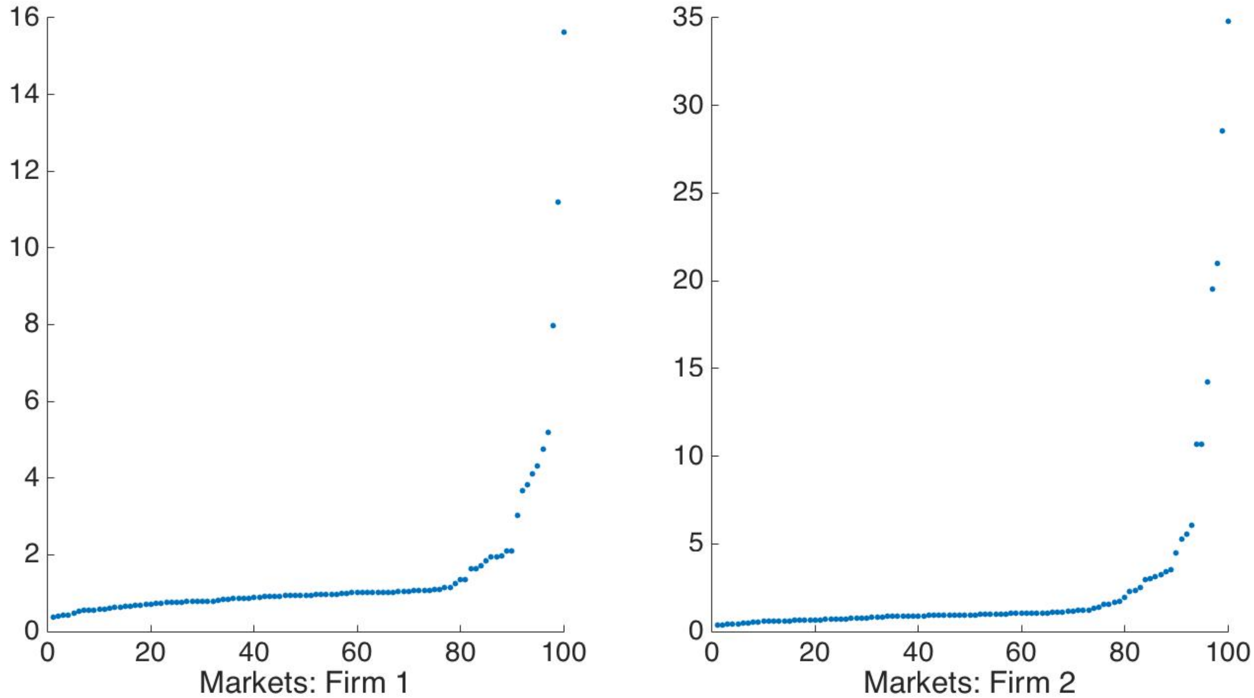
Table 1: Summary statistics of $(\hat{c}_{jm} - c_{jm})^2$

Model		$\delta = 0$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
Bertrand	mean	1.0e-24	0.86	3.62	8.98	18.45	34.47	55.83	91.79	101.81	103.26
	sd	2.3e-24	0.86	3.46	8.47	18.48	39.26	63.88	112.11	126.00	128.02
Profit Weight	mean	1.0e-24	1.62	1.84	3.16	8.64	24.20	28.39	6.63	0.30	0.04
	sd	2.3e-24	3.72	4.58	6.83	17.68	51.06	52.65	11.48	0.48	0.06
Our Model	mean	-	0.65	0.86	1.53	2.01	0.20	0.16	0.17	0.19	0.02
	sd	-	0.62	1.06	1.84	2.93	0.38	0.25	0.22	0.34	0.03

We also see in Table 1 that the profit weight approach does well at obtaining accurate estimates of the marginal cost when firms are either not colluding or perfectly colluding. In both cases, the first order condition of the repeated game given by (8) is identical to the profit weight model with weights of 0 in the case of no collusion and 1 in the case of perfect collusion. However, our model outperforms the profit weight model for $\delta \in (0, 0.45)$. Ex ante, we would expect our model to perform at least as well as the profit weight approach given that the first order conditions of our model are identical to those derived from the Pareto equilibrium.

As was discussed in Section 3, our model and the profit weight model are equivalent only when $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ is invariant across markets. In Figure 1, we plot $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ when $\delta = 0.3$ across the m markets for each firm. Each point in the plots is a firm-market pair. The left panel presents the observations for firm 1 and the right panel for firm 2. For visual purposes, we have arranged the graphs such that, in each, $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ is monotonically increasing. One can see that $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ is far from constant across markets for intermediate values of δ, and thus it is unsurprising that the squared error of the cost estimates is large for the Profit Weight model in this case.

Figure 1: Figure 1:Plot of $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ Across Firm-Market Pairs, $\delta = 0.3$



Note: This figure displays a scatter plot of $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ for each of the 200 firm-market pairs. The left panel $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ plots the observations for firm 1 and the right for firm 2. For visual purposes, the firm-market pairs have been arranged such that, in each panel, $\frac{\partial \pi^d_{jmt}/\partial x_{imt}}{\partial \pi_{jmt}/\partial x_{imt}}$ is monotonically increasing.

6 Empirical Section

[[To be added]]

7 Conclusion

[[To be added]]

References

Bernheim, B. Douglas and Michael D. Whinston (1990), "Multimarket Contact and Collusive Behavior." *RAND Journal of Economics*, 21, 1–26.

Berry, Steven and Philip Haile (2014), "Identification in differentiated products markets using market level data." *Econometrica*, 82, 1749–1797.

Berry, Steven and Panle Jia (2010), "Tracing the Woes: An Empirical Analysis of the Airline Industry." *American Economic Journal: Microeconomics*, 2, 1–43.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63, 841–90.

Bresnahan, Timothy (1982), "The oligopoly solution concept is identified." *Economics Letters*, 10, 87–92.

Bresnahan, Timothy (1987), "Competition and collusion in the american automobile industry: The 1955 price war." *Journal of Industrial Economics*, 35, 457–482.

Ciliberto, Federico and Jonathan W. Williams (2014), "Does multimarket contact facilitate tacit collusion? inference on conduct parameters in the airline industry." *RAND Journal of Economics*, 45, 764–791.

Corts, Kenneth (1999), "Conduct parameters and the measurement of market power." *Journal of Econometrics*, 88, 227–250.

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal, and Nina Pavcnik (2016), "Prices, markups, and trade reform." *Econometrica*, 84, 445–510.

De Loecker, Jan and Frederic Warzynski (2012), "Markups and firm-level export status." *American Economic Review*, 102, 2437–71.

Eizenberg, Alon (2014), "Upstream Innovation and Product Variety in the U.S. Home PC Market." *Review of Economic Studies*, 81, 1003–1045.

Fan, Ying (2013), "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market." *American Economic Review*, 103, 1598–1628.

Fershtman, Chaim and Ariel Pakes (2000), "A dynamic oligopoly with collusion and price wars." *RAND Journal of Economics*, 31, 207–236.

Gasmi, Farid, Jean-Jacques Laffont, and Quang Vuong (1992), "Econometric analysisof collusive behaviorin a soft-drink market." *Journal of Economics & Management Strategy*, 1, 277–311.

Goldberg, Pinelopi (1995), "Product differentiation and oligopoly in international markets: The case of the u.s. automobile industry." *Econometrica*, 63, 891–951.

Green, Edward and Robert Porter (1984), "Noncooperative collusion under imperfect information." *Econometrica*, 52, 87–100.

Igami, Mitsuru and Takuo Sugaya (2017), "Measuring the incentive to collude: The vitamin cartels, 1990–1999." working paper, Yale University.

Khwaja, Ahmed and Beomjoon Shim (2017), "The collusive effect of multimarket contact on prices: Evidence from retail lumber markets." working paper, Yale University.

Lau, Lawrence J. (1982), "On identifying the degree of competitiveness from industry price and output data." *Economics Letters*, 10, 93–99.

Loecker, Jan De and Jan Eeckhoutz (2017), "The rise of market power and the macroeconomic implications." working paper, Princeton University.

Loecker, Jan De and Paul T. Scott (2016), "Estimating market power evidence from the us brewing industry." Working Paper 22957, National Bureau of Economic Research.

Miller, Nathan and Matthew Weinberg (2017), "Understanding the price effects of the millercoors joint venture." *Econometrica*, 85, 17631791.

Nevo, Aviv (1998), "Identity of the oligopoly solution concept in a differentiated products industry." *Economics Letters*, 59, 391–395.

Nevo, Aviv (2001), "Measuring market power in the ready-to-eat cereal industry." *Econometrica*, 69, 307–342.

Porter, Robert (1983), "A study of cartel stability: The joint executive committee, 1880-1886." *Bell Journal of Economics*, 14, 301–314.

Reiss, Peter C. and Frank A. Wolak (2007), "Structural Econometric Modeling: Rationales and Examples from Industrial Organization." In *Handbook of Econometrics* (J.J. Heckman and E.E. Leamer, eds.), volume 6 of *Handbook of Econometrics*, chapter 64, Elsevier.

Rosse, James (1967), "Daily newspapers, monopolistic competition, and economies of scale." *American Economic Review*, 57, 522533.

Slade, Margaret (1987), "Interfirm rivalry in a repeated game: An empirical test of tacit collusion." *Journal of Industrial Economics*, 35, 499–516.

Sudhir, K. (2001), "Competitive pricing behavior in the auto market: A structural analysis." *Marketing Science*, 20, 42–60.

Villas-Boas, Sofia Berto (2007), "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data." *Review of Economic Studies*, 74, 625–652.

Appendices

A When Our Model and the Profit-weight Conduct Parameter Model are Equivalent

In this appendix section, we provide one example where our model and the profit-weight conduct parameter are equivalent. We show, however, that this equivalence result does not hold with any deviation from this example. For expositional simplicity, we suppress the subscript t in this section.

Consider a Cournot model with two firms, linear demand $p_m = a_m - b_m (x_{1m} + x_{2m})$, and zero marginal cost. Our model in the Cournot setting (12) becomes

$$\begin{aligned} a_m - b_m (x_{1m} + x_{2m}) - b_m x_{1m} - b_m \lambda_{12} x_{2m} + \rho_2 \lambda_{12} b_m \frac{a_m - b_m x_{1m}}{2b_m} &= 0 \\ a_m - b_m (x_{1m} + x_{2m}) - b_m x_{2m} - b_m \lambda_{21} x_{1m} + \rho_1 \lambda_{21} b_m \frac{a_m - b_m x_{2m}}{2b_m} &= 0, \end{aligned} \tag{A.1}$$

which can be simplified to be

$$\begin{aligned} \frac{a_m}{b_m} - (x_{1m} + x_{2m}) - x_{1m} - \lambda_{12} x_{2m} + \rho_2 \lambda_{12} \frac{\frac{a_m}{b_m} - x_{1m}}{2} &= 0 \\ \frac{a_m}{b_m} - (x_{1m} + x_{2m}) - x_{2m} - \lambda_{21} x_{2m} + \rho_2 \lambda_{21} \frac{\frac{a_m}{b_m} - x_{2m}}{2} &= 0. \end{aligned} \tag{A.2}$$

It turns out that the solution is of the following format

$$x_{im} = \frac{a_m}{b_m} g_i (\lambda_{12}, \lambda_{21}, \rho_1, \rho_2), \tag{A.3}$$

where $g_i(\cdot)$ is a function independent of a_m and b_m. In other words, the solution x_{im} is proportional to $\frac{a_m}{b_m}$.

Therefore, the ratio of interest

$$\frac{\partial \pi^d_{jmt}}{\partial x_{imt}} / \frac{\partial \pi_{jmt}}{\partial x_{imt}} = \frac{-b_m \frac{a_m - b_m x_{im}}{2b_m}}{-b_m x_{jm}} = \frac{1 - g_i (\lambda_{12}, \lambda_{21}, \rho_1, \rho_2)}{2 g_j (\lambda_{12}, \lambda_{21}, \rho_1, \rho_2)}, \tag{A.4}$$

which is independent of (a_m, b_m) and thus invariant across markets. In this case, we can define $\tilde{\lambda}_{ijt} = \lambda_{ijt} - \rho_{jt} \lambda_{ijt} \frac{\partial \pi^d_{jmt}}{\partial x_{imt}} / \frac{\partial \pi_{jmt}}{\partial x_{imt}}$ and rewrite our model (8) $\frac{\partial \pi_{imt}}{\partial x_{imt}} + \lambda_{ijt} \frac{\partial \pi_{jmt}}{\partial x_{imt}} - \rho_{jt} \lambda_{ijt} \frac{\partial \pi^d_{jmt}}{\partial x_{imt}} = 0$ as $\frac{\partial \pi_{imt}}{\partial x_{imt}} + \tilde{\lambda}_{ijt} \frac{\partial \pi_{jmt}}{\partial x_{imt}} = 0$, which is the profit-weight conduct parameter model.

However, the result that $\frac{\partial \pi^d_{jmt}}{\partial x_{imt}} / \frac{\partial \pi_{jmt}}{\partial x_{imt}}$ is invariant across markets is specific to the assumptions listed above. For example, the result does not hold when the demand function is nonlinear or when firms are asymmetric in marginal costs. It does not hold when we switch to a Bertrand setting with differentiated products either. In sum, our model and the profit-weight conduct parameter model are not equivalent in general.

B Proof of Invertibility

B.1 Cournot Setting

Throughout this section, we suppress the subscripts m and t and the demand shiftor $\boldsymbol{y}_{mt}$. In the Cournot setting, $\pi_i(\boldsymbol{x}, c_i) = [p(\sum_i x_i) - c_i]\, x_i$. Our model becomes

$$\begin{aligned}(p - c_i) + p^{'}(X)\, x_i + \sum_{j \neq i} \frac{\tau_j}{\tau_i} p^{'}(X)\, x_j \\ - \sum_{j \neq i} \rho_j \frac{\tau_j}{\tau_i} p^{'}\left(X_{-j} + x^{BR}(X_{-j}, c_j)\right) x^{BR}(X_{-j}, c_j) = 0.\end{aligned} \tag{B.5}$$

For any given data $(\boldsymbol{x}, p)$ and parameters $(\boldsymbol{\tau}, \boldsymbol{\rho})$, define a function $F : [0, p]^N \to R^N$ as

$$\begin{aligned} F(\boldsymbol{c})_i \;&=\; -(p - c_i) - p^{'}(X)\, x_i - \sum_{j \neq i} \frac{\tau_j}{\tau_i} p^{'}(X)\, x_j \\ &\quad + \sum_{j \neq i} \rho_j \frac{\tau_j}{\tau_i} p^{'}\left(X_{-j} + x^{BR}(X_{-j}, c_j)\right) x^{BR}(X_{-j}, c_j)\,. \end{aligned} \tag{B.6}$$

We want to show that $F(\boldsymbol{c}) = 0$ has a unique solution. By the Gale-Nikaido-Inada Theorem, it is sufficient to show that the Jacobian matrix of F is a p-matrix.

Let $J(\boldsymbol{c})$ denote the Jacobian matrix of F, where $J(\boldsymbol{c})_{ii} = 1$ and $J(\boldsymbol{c})_{ji} = -\rho_j \frac{\tau_j}{\tau_i} \frac{\partial x_j^{BR}}{\partial x_i}$. Note that the best-response x_j^{BR} satisfies the following first-order condition

$$p^{'}\left(X_{-j} + x^{BR}(X_{-j}, c_j)\right) x^{BR}(X_{-j}, c_j) + p\left(X_{-j} + x^{BR}(X_{-j}, c_j)\right) - c_j = 0. \tag{B.7}$$

Taking the total differentiation with respect to x_i gives

$$p'' x_j^{BR} \left(1 + \frac{\partial x_j^{BR}}{\partial x_i}\right) + p' \frac{\partial x_j^{BR}}{\partial x_i} + p' \left(1 + \frac{\partial x_j^{BR}}{\partial x_i}\right) = 0. \tag{B.8}$$

Therefore,

$$\frac{\partial x_j^{BR}}{\partial x_i} = -\frac{p'' x_j^{BR} + p'}{p'' x_j^{BR} + 2p'}. \tag{B.9}$$

The second-order condition implies that $p'' x_j^{BR} + p' \leq 0$, which in turn implies that $p'' x_j^{BR} + 2p' < 0$. Under the assumption that $p'' \geq 0$, $\frac{\partial x_j^{BR}}{\partial x_i} \in \left[-\frac{1}{2}, 0\right)$. From (B.9), we can see that $\frac{\partial x_j^{BR}}{\partial x_i}$ is independent of i. This is because x_i affects x_j^{BR} only through shifting $X_{-j} = \sum_{i \neq j} x_i$. We can therefore denote $-\rho_j \frac{\partial x_j^{BR}}{\partial x_i}$ by h_j. Since $0 \leq \rho_j \leq 1, h_j \in \left[0, \frac{1}{2}\right]$.

Therefore, a k^{th}order principal submatrix of $J(\boldsymbol{c})$ is of the following format:

$$\begin{bmatrix} 1 & \frac{\tau_{i_2}}{\tau_{i_1}} h_{i_2} & \cdots & \frac{\tau_{i_k}}{\tau_{i_1}} h_{i_k} \\ \frac{\tau_{i_1}}{\tau_{i_2}} h_{i_1} & 1 & & \frac{\tau_{i_k}}{\tau_{i_2}} h_{i_k} \\ \vdots & & \ddots & \\ \frac{\tau_{i_1}}{\tau_{i_k}} h_{i_1} & \frac{\tau_{i_2}}{\tau_{i_k}} h_{i_2} & & 1 \end{bmatrix}, \tag{B.10}$$

where $(i_1, ..., i_k)$ is a subset of $(1, ..., N)$. Its determinant equals

$$\begin{vmatrix} 1 & h_{i_2} & \cdots & h_{i_k} \\ h_{i_1} & 1 & & h_{i_k} \\ \vdots & & \ddots & \\ h_{i_1} & h_{i_2} & & 1 \end{vmatrix} = (h_{i_1} h_{i2} \cdots h_{ik}) \begin{vmatrix} \frac{1}{h_{i1}} & 1 & \cdots & 1 \\ 1 & \frac{1}{h_{i_2}} & & 1 \\ \vdots & & \ddots & 1 \\ 1 & 1 & & \frac{1}{h_{i_k}} \end{vmatrix} = (h_{i_1} h_{i2} \cdots h_{ik}) \det A. \tag{B.11}$$

Since $h_j \geq 0$, we only need to show that $\det A$ is non-negative. In what follows, we show that A is a positive definite.

Let $a = (a_1, ..., a_k)^T \neq 0$ be any nonzero vector. Then,

$$\begin{aligned} a^T A a &= \sum_l \frac{a_l^2}{h_{il}} + 2 \sum\sum_{l \neq l'} a_l a_{l'} \qquad \text{(B.12)} \\ &\geq 2 \sum_l a_l^2 + 2 \sum\sum_{l \neq l'} a_l a_{l'} \text{ because } 0 \leq h_l \leq 1/2 \\ &> \sum_l a_l^2 + 2 \sum\sum_{l \neq l'} a_l a_{l'} \\ &= \left(\sum_l a_l \right)^2 > 0. \end{aligned}$$

B.2 Bertrand Setting

We start with showing the invertibility of a Logit model. In a Bertrand setting, $\pi_i(\boldsymbol{x}, c_i) = (x_i - c_i) M s_i(\boldsymbol{x})$. Our model becomes

$$\begin{aligned} &(x_i - c_i) \frac{\partial s_i}{\partial x_i}(\boldsymbol{x}) + s_i + \sum_{j \neq i} \frac{\tau_j}{\tau_i} (x_j - c_j) \frac{\partial s_j}{\partial x_i}(\boldsymbol{x}) \qquad \text{(B.13)} \\ &- \sum_{j \neq i} \rho_j \frac{\tau_j}{\tau_i} \left(x_j^{BR}(\boldsymbol{x}_{-j,}, c_j) - c_j \right) \frac{\partial s_j}{\partial x_i} \left(\boldsymbol{x}_{-j}, x_j^{BR}(\boldsymbol{x}_{-j,}, c_j) \right) = 0. \end{aligned}$$

For any given data $(\boldsymbol{x}, s)$ and parameters $(\boldsymbol{\tau}, \boldsymbol{\rho})$, define a function $F : [0, \bar{c}_1] \times \cdots [0, \bar{c}_N] \to R^N$ as

$$F(\boldsymbol{c})_i = (x_i - c_i) \frac{\partial s_i}{\partial x_i}(\boldsymbol{x}) + s_i + \sum_{j \neq i} \frac{\tau_j}{\tau_i}(x_j - c_j) \frac{\partial s_j}{\partial x_i}(\boldsymbol{x}) \tag{B.14}$$
$$- \sum_{j \neq i} \rho_j \frac{\tau_j}{\tau_i} \left(x_j^{BR}(\boldsymbol{x}_{-j,}, c_j) - c_j\right) \frac{\partial s_j}{\partial x_i}\left(\boldsymbol{x}_{-j}, x_j^{BR}(\boldsymbol{x}_{-j,}, c_j)\right),$$

where $\bar{c}_i$ is the marginal cost of firm i such that at its best-response to $\boldsymbol{x}_{-i}$ in the data, its market share equals to the market share in data. In other words, $\bar{c}_i$ is the solution to $s_i\left(\boldsymbol{x}_{-i}, x_i^{BR}(\boldsymbol{x}_{-i,}, c_i)\right) = s_i$, where s_i on the right-hand side is the market share in data. In Supplemental Appendix, we that at the solution of the supergame (5) in a Bertrand setting, $x_i^{BR} \leq x_i$. In other words, when one firm deviates unilaterally, it must decrease its price. This is quite intuitive as we expect firms in a collusion to raise prices to the extent that each firm has an incentive to undercut. This implies that $s_i^{BR} \geq s_i$. Note that s_i^{BR} is strictly decreasing in c_i. Therefore, c_i must be smaller than $\bar{c}_i$.

We again denote the Jacobian matrix by $J(\boldsymbol{c})$, where the diagonal element $J(\boldsymbol{c})_{ii} = -\frac{\partial s_i}{\partial x_i}(\boldsymbol{x})$. In a Logit model, it is $\alpha s_i (1 - s_i)$, where α is the absolute value of the price coefficient. The off-diagonal element is

$$J(\boldsymbol{c})_{ij} = -\frac{\tau_j}{\tau_i}\alpha \frac{\partial s_j}{\partial x_i} + \rho_j \frac{\tau_j}{\tau_i} \frac{\partial s_j^{BR_j}}{\partial x_i} = -\frac{\tau_j}{\tau_i} s_i s_j + \rho_j \frac{\tau_j}{\tau_i}\left(\alpha s_i^{BR_j} s_j^{BR_j} - \alpha s_j^{BR_j}\left(1 - s_j^{BR_j}\right) \frac{\partial x_j^{BR_j}}{\partial x_i}\right), \tag{B.15}$$

where we use the subscript BR_j to indicate that a function is evaluated at firm j's best response. For example, $s_i^{BR_j}$ and $s_i^{BR_i}$ are different. The best-response $x_j^{BR_j}$ satisfies the following first-order condition

$$-\alpha\left(x_j^{BR_j} - c_j\right) s_j^{BR_j}\left(1 - s_j^{BR_j}\right) + s_j^{BR_j} = 0, \tag{B.16}$$

which implies $-\alpha\left(x_j^{BR} - c_j\right)\left(1 - s_j^{BR}\right) + 1 = 0$. Taking the total differentiation with respect to x_i yields

$$\frac{\partial x_j^{BR_j}}{\partial x_i}\left(1 - s_j^{BR_j}\right) + \alpha\left(x_j^{BR_j} - c_j\right) s_j^{BR_j}\left[\left(1 - s_j^{BR_j}\right)\frac{\partial x_j^{BR_j}}{\partial x_i} - s_i^{BR_j}\right] = 0. \tag{B.17}$$

Combing (B.16) and (B.17), we have $\frac{\partial x_j^{BR_j}}{\partial x_i} = \frac{s_i^{BR_j} s_j^{BR_j}}{1 - s_j^{BR}}$. Moreover, since $s_i^{BR_j}/s_i = \left(1 - s_j^{BR_j}\right)/(1 - s_j)$, we have

$$J(\boldsymbol{c})_{ij} = -\frac{\tau_j}{\tau_i}\alpha s_i s_j + \alpha \rho_j \frac{\tau_j}{\tau_i}\frac{1}{1 - s_j} s_i s_j^{BR_j}\left(1 - s_j^{BR_j}\right)^2. \tag{B.18}$$

A k^{th} order principal submatrix of $J(\boldsymbol{c})$ is of the following format

$$\begin{bmatrix} J_{i_1 i_1} & J_{i_1 i_2} & \cdots & J_{i_1 i_k} \\ J_{i_2 i_1} & J_{i_2 i_2} & & J_{i_2 i_k} \\ \vdots & & \ddots & \vdots \\ J_{i_k i_1} & J_{i_k i_2} & \cdots & J_{i_k i_k} \end{bmatrix},$$

where $(i_1, ..., i_k)$ is a subset of $(1, ..., N)$. Since the diagonal elements are all positive, the sign of its determinant is the same as the sign of the determinant of the following matrix

$$A_k = \begin{bmatrix} 1 & \frac{J_{i_1 i_2}}{J_{i_1 i_1}} & \cdots & \frac{J_{i_1 i_k}}{J_{i_1 i_1}} \\ \frac{J_{i_2 i_1}}{J_{i_2 i_2}} & 1 & & \frac{J_{i_2 i_k}}{J_{i_2 i_2}} \\ \vdots & & \ddots & \vdots \\ \frac{J_{i_k i_1}}{J_{i_k i_k}} & \frac{J_{i_k i_2}}{J_{i_k i_k}} & \cdots & 1 \end{bmatrix}.$$

Note that

$$\begin{aligned} \frac{J_{ij}}{J_{ii}} &= \frac{-\frac{\tau_j}{\tau_i}\alpha s_i s_j + \alpha \rho_j \frac{\tau_j}{\tau_i}\frac{1}{1-s_j} s_i s_j^{BR_j}\left(1 - s_j^{BR_j}\right)^2}{\alpha s_i (1-s_i)} \\ &= \frac{\tau_j}{\tau_i}\frac{-s_j + \rho_j \frac{1}{1-s_j} s_j^{BR_j}\left(1 - s_j^{BR_j}\right)^2}{1-s_i} \\ &\triangleq \frac{\tau_j}{\tau_i}\frac{h_j}{g_i}. \end{aligned} \tag{B.19}$$

Therefore,

$$\det A_k = \begin{vmatrix} 1 & \frac{h_{i_2}}{g_{i1}} & \cdots & \frac{h_{i_k}}{g_{i_1}} \\ \frac{h_{i_1}}{g_{i_2}} & 1 & & \frac{h_{i_k}}{g_{i_2}} \\ \vdots & & \ddots & \vdots \\ \frac{h_{i_1}}{g_{i_k}} & \frac{h_{i_2}}{g_{i_k}} & \cdots & 1 \end{vmatrix} = \prod_{l=1,\ldots,k} \frac{h_{i_l}}{g_{i_l}} \cdot \begin{vmatrix} \frac{g_{i_1}}{h_{i_1}} & 1 & \cdots & 1 \\ 1 & \frac{g_{i_2}}{h_{i_2}} & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & \frac{g_{i_k}}{h_{i_k}} \end{vmatrix}$$

Define $f_i = \frac{h_i}{g_i} = \frac{-s_i}{1-s_i} + \frac{\rho_i s_i^{BR_i}\left(1-s_i^{BR_i}\right)^2}{(1-s_i)^2}$. Since $s_i^{BR_i} \geq s_i$, $-\frac{s_i}{1-s_i} \leq f_i \leq 1$. We can rewrite

$\det A_k$ as

$$
\begin{aligned}
\det A_k &= \prod_{l=1,\ldots,k} f_{i_l} \cdot \begin{vmatrix} \frac{1}{f_{i_1}} & 1 & \cdots & 1 \\ 1 & \frac{1}{f_{i_2}} & & 1 \\ \vdots & & \ddots & \vdots \\ 1 & 1 & \cdots & \frac{1}{f_{i_k}} \end{vmatrix} \qquad \text{(B.20)} \\
&= \prod_{l=1,\ldots,k} f_{i_l} \cdot \begin{vmatrix} \frac{1}{f_{i_1}} & 1 & \cdots & 1 & 1 \\ 1 & \frac{1}{f_{i_2}} & & 1 & 1 \\ \vdots & & \ddots & \vdots & \vdots \\ 1 & 1 & & \frac{1}{f_{i_{k-1}}} & 1 \\ 0 & 0 & \cdots & 1-\frac{1}{f_{i_{k-1}}} & \frac{1}{f_{i_k}}-1 \end{vmatrix} \\
&= \prod_{l=1,\ldots,k} f_{i_l} \left[-\left(1-\frac{1}{f_{i_{k-1}}}\right) \begin{vmatrix} \frac{1}{f_{i_1}} & 1 & \cdots & 1 & 1 \\ 1 & \frac{1}{f_{i_2}} & & 1 & 1 \\ \vdots & & \ddots & \vdots & \vdots \\ 1 & 1 & & \frac{1}{f_{i_{k-2}}} & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{vmatrix} + \left(\frac{1}{f_{i_k}}-1\right) \frac{\det A_{k-1}}{\prod_{l=1,\ldots,k-1} f_{i_l}} \right] \\
&= \prod_{l=1,\ldots,k} f_{i_l} \left[-\left(1-\frac{1}{f_{i_{k-1}}}\right) \prod_{l=1,\ldots,k-2} \left(\frac{1}{f_{i_l}}-1\right) + \left(\frac{1}{f_{i_k}}-1\right) \frac{\det A_{k-1}}{\prod_{l=1,\ldots,k-1} f_{i_l}} \right] \\
&= (1-f_{i_1}) \cdots \left(1-f_{i_{k-1}}\right) f_{i_k} + (1-f_{i_k}) \det A_{k-1}.
\end{aligned}
$$

We now use induction to show that $\det A_k > 0$. First of all, $\det A_2 = 1 - f_{i_1} f_{i_2}$. When f_{i_1} and f_{i_2} are of different signs, $\det A_2 > 0$. When both are positive, since both are smaller than 1, $\det A_2 > 0$. When both are negative, $\det A_2$ is increasing in both f_{i_1} and f_{i_2}. Given that $f_i \geq -\frac{s_i}{1-s_i}$, $\det A_2 \geq 1 - \frac{s_{i_1} s_{i_2}}{(1-s_{i_1})(1-s_{i_2})} = \frac{1-s_{i_1}-s_{i_2}}{(1-s_{i_1})(1-s_{i_2})} > 0$ because $0 < s_{i_1}, s_{i_2} < 1$ and $s_{i_1} + s_{i_2} < 1$. Therefore, $\det A_2 > 0$.

Then, we show that $\det A_{k-1} > 0$ for any $(k-1)$-element subset of $(1, ..., N)$ implies $\det A_k > 0$ for any k-element subset. Since $\det A_k = (1-f_{i_1}) \cdots \left(1-f_{i_{k-1}}\right) f_{i_k} + (1-f_{i_k}) \det A_{k-1}$ and $f_{i_l} < 1$, $\det A_k > 0$ if $f_{i_k} > 0$. Note that $(f_{i_1}, \cdots, f_{i_k})$ are exchangeable. Therefore, if there is any $f_{i_l} > 0$, $\det A_k > 0$. We now show that when all $(f_{i_1}, \cdots, f_{i_k})$ are negative, $\det A_k$ is also positive. This proof is carried out in three steps.

In Step 1, we again use induction to show that $\det A_k < (1-f_{i_1}) \cdots (1-f_{i_k})$ when $(f_{i_1}, \cdots, f_{i_k})$ are all negative. When $k=2$, $\det A_2 = 1 - f_{i_1} f_{i_2} < 1 < (1-f_{i_1})(1-f_{i_2})$. Suppose $\det A_{k-1} < (1-f_{i_1}) \cdots \left(1-f_{i_{k-1}}\right)$. Then, $\det A_k < (1-f_{i_k}) \det A_{k-1} < (1-f_{i_1}) \cdots \left(1-f_{i_{k-1}}\right)(1-f_{i_k})$.

In Step 2, we show that $\det A_k$ is increasing in f_{i_l} when $(f_{i_1}, \cdots, f_{i_k})$ are all negative. Note that

$\frac{\partial \det A_k}{\partial f_{i_k}} = (1 - f_{i_1}) \cdots \left(1 - f_{i_{k-1}}\right) - \det A_{k-1} > 0$ based on the result in Step 1. By exchangeability, this means that $\frac{\partial \det A_k}{\partial f_{i_l}} > 0$ for any $l = 1, ..., k$. For example, when $l = 1$, by exchangeability, $\det A_k = (1 - f_{i_2}) \cdots (1 - f_{i_k}) f_{i_1} + (1 - f_{i_1}) \det A_{k-1}(f_{i_2}, ..., f_{i_k})$ where $\det A_{k-1}$ is now a function of $(f_{i_2}, ..., f_{i_k})$. The result in Step 1 still holds, i.e., $\det A_{k-1}(f_{i_2}, ..., f_{i_k}) < (1 - f_{i_2}) \cdots (1 - f_{i_k})$. Therefore, $\frac{\partial \det A_k}{\partial f_{i_1}} > 0$. From Step 2, we know that when $(f_{i_1}, \cdots, f_{i_k})$ are all negative, the lower bound of $\det A_k$ is when $f_{i_l} = -\frac{s_{i_l}}{1-s_{i_l}}$.

In Step 3, we use induction to show that this lower bound of $\det A_k$ is $\frac{1-\sum_{l=1}^{k} s_{i_l}}{\prod_{l=1}^{k}(1-s_{i_l})}$. We have shown in Step 1 that when $f_{i_1} = -\frac{s_{i_1}}{1-s_{i_1}}$ and $f_{i_2} = -\frac{s_{i_2}}{1-s_{i_2}}$, $\det A_2 = \frac{1-s_{i_1}-s_{i_2}}{(1-s_{i_1})(1-s_{i_2})}$. Suppose when $\left(f_{i_1}, ..., f_{i_{k-1}}\right)$ are all at their respective lower bound, $\det A_{k-1} = \frac{1-\sum_{l=1}^{k-1} s_{i_l}}{\prod_{l=1}^{k-1}(1-s_{i_l})}$. Then, the lower bound of $\det A_k = -\left(\prod_{l=1}^{k-1} \frac{1}{1-s_{i_{k-1}}}\right) \frac{s_{i_k}}{1-s_{i_k}} + \frac{1}{1-s_{i_k}} \frac{1-\sum_{l=1}^{k-1} s_{i_l}}{\prod_{l=1}^{k}(1-s_{i_l})} = \frac{1-\sum_{l=1}^{k} s_{i_l}}{\prod_{l=1}^{k}(1-s_{i_l})}$. Since $0 < s_{i_l} < 1$ and $\sum_{l=1}^{k} s_{i_l} < 1$, $\det A_k > 0$.

SA Two Types of Conduct Parameter Models

The main body of the paper focuses on the comparison of our model to the profit-weight conduct parameter models (as well as the classic oligopoly models). As mentioned, there is another type of conduct parameter models, i.e., the conjecture variation models. In this appendix section, we present both types of conduct parameter models and explain how they are related. For expositional simplicity, we suppress the subscript m in this appendix.

In a standard static oligopoly model, the first-order conditions are

$$\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}} = 0. \tag{SA.1}$$

Typically, a firm's action has a direct effect and an indirect effect on its profit. For example, consider a Cournot model where x_{it} represents the quantity of firm i, the inverse market demand is $p(x_{1t} + x_{2t}, z_t)$ and the marginal cost is constant c_{it}. In such a model, $\pi_i (x_{1t}, x_{2t}, z_t) = p(x_{1t} + x_{2t}, z_t) x_{it} - c_{it} x_{it}$. The marginal profit of x_i is $\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_i} = [p(x_{1t} + x_{2t}, z_t) - c_{it}] + p'(x_{1t} + x_{2t}, z_t) x_{it}$, where the first term is the direct effect of producing more quantity and the second term the indirect effect on profit through driving down the market price. Similarly, in a Bertrand model where x_{it} represents the price of firm i and the demand for the firm is given by $q_i (x_{1t}, x_{2t}, z_t)$. The profit function is $\pi_i (x_{1t}, x_{2t}, z_t) = (x_{it} - c_{it}) q_i (x_{1t}, x_{2t}, z_t)$; and the corresponding marginal profit is $\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}} = q_{it} + (x_{it} - c_{it}) \frac{\partial q_{it}}{\partial x_{it}}$, which is again the sum of the direct effect of increasing the price on profit (i.e., q_i) and the indirect effect on profit by decreasing the quantity (i.e., $(x_i - c_i) \frac{\partial q_i}{\partial x_i}$). Let $\left(\frac{\partial \pi_i (x_1, x_2)}{\partial x_i}\right)_D$ and $\left(\frac{\partial \pi_i (x_1, x_2)}{\partial x_i}\right)_{ID}$ represent the direct and the indirect effects, respectively. Therefore, in a standard oligopoly model, the first-order condition (SA.1) can be rewritten as

$$\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}} = \left(\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}}\right)_D + \left(\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}}\right)_{ID} = 0. \tag{SA.2}$$

A typical conduct parameter model assumes that

$$\left(\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}}\right)_D + \lambda_{ii} \left(\frac{\partial \pi_i (x_{1t}, x_{2t}, z_t)}{\partial x_{it}}\right)_{ID} + \lambda_{ij} \frac{\partial \pi_j (x_{1t}, x_{2t}, z_t)}{\partial x_{it}} = 0, \tag{SA.3}$$

where λ_{ii} and λ_{ij} are the conduct parameters.

For example, consider a market with homogeneous products where firms choose quantities. Setting $\lambda_{ij} = 0$, equation (SA.3) becomes

$$x_{it} = c_{it} - \lambda_{ii} p' (x_{1t} + x_{2t}, z_t) x_{it}, \tag{SA.4}$$

the same as in Bresnahan (1982). In a market with heterogeneous products where single-product firms choose prices, equation (SA.3) becomes

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} c_{1t} \\ c_{2t} \end{pmatrix} + \begin{pmatrix} \lambda_{11}\frac{\partial q_{1t}}{\partial x_{1t}} & \lambda_{21}\frac{\partial q_{1t}}{\partial x_{2t}} \\ \lambda_{12}\frac{\partial q_{2t}}{\partial x_{1t}} & \lambda_{22}\frac{\partial q_{2t}}{\partial x_{2t}} \end{pmatrix}^{-1} \begin{pmatrix} q_{1t} \\ q_{2t} \end{pmatrix}, \tag{SA.5}$$

the same as the framework in Nevo (1998).

The first set of conduct parameters λ_{ii} are motivated by the troubled conjecture variations literature. For example, in the Cournot setting, $\lambda_{ii}-1$ is interpreted as a firm's expectations about how its rivals react to an increase in its quantity. Such a model with conduct parameters λ_{ii} is a conjecture variations conduct parameter model. However, Daughety (1985) and Lindh (1992) show that the Cournot conjectures where $\lambda_{ii}=1$ are the only consistent equilibrium conjectures.

Papers in the recent literature (e.g., Sudhir (2001), Black et al. (2004), Ciliberto and Williams (2014), Miller and Weinberg (2017) and Khwaja and Shim (2017)) do not rely on the concept of conjecture variations, and set $\lambda_{ii}=1$ as it should be, and focus instead on λ_{ij}. The model is therefore

$$\frac{\partial \pi_i\left(x_{1t}, x_{2t}, z_t\right)}{\partial x_{it}} + \lambda_{ij}\frac{\partial \pi_j\left(x_{1t}, x_{2t}, z_t\right)}{\partial x_{it}} = 0, \tag{SA.6}$$

where λ_{ij} are the conduct parameters to be estimated. This is the profit-weight conduct parameter model where these conduct parameters λ_{ij} capture how much firm i internalizes the effect of its action on its competitor j in its decision making. However, as mentioned in Section 2, both the conjecture variations type of conduct parameter models in (SA.4) and the profit-weight type of conduct parameter models in (SA.6) lack of theoretical foundation.

Supplemental Appendix References

Black, Michael, Gregory Crawford, Shihua Lu, and Halbert White (2004), "A virtual stakes approach to measuring competition in product markets." Working paper.

Bresnahan, Timothy (1982), "The oligopoly solution concept is identified." *Economics Letters*, 10, 87–92.

Ciliberto, Federico and Jonathan W. Williams (2014), "Does multimarket contact facilitate tacit collusion? inference on conduct parameters in the airline industry." *RAND Journal of Economics*, 45, 764–791.

Daughety, Andrew F. (1985), "Reconsidering Cournot: The Cournot Equilibrium is Consistent." *RAND Journal of Economics*, 16, 368–379.

Khwaja, Ahmed and Beomjoon Shim (2017), "The collusive effect of multimarket contact on prices: Evidence from retail lumber markets." working paper, Yale University.

Lindh, Thomas (1992), "The inconsistency of consistent conjectures: Coming back to Cournot." *Journal of Economic Behavior & Organization*, 18, 69–90.

Miller, Nathan and Matthew Weinberg (2017), "Understanding the price effects of the millercoors joint venture." *Econometrica*, 85, 17631791.

Nevo, Aviv (1998), "Identity of the oligopoly solution concept in a differentiated products industry." *Economics Letters*, 59, 391–395.

Sudhir, K. (2001), "Competitive pricing behavior in the auto market: A structural analysis." *Marketing Science*, 20, 42–60.