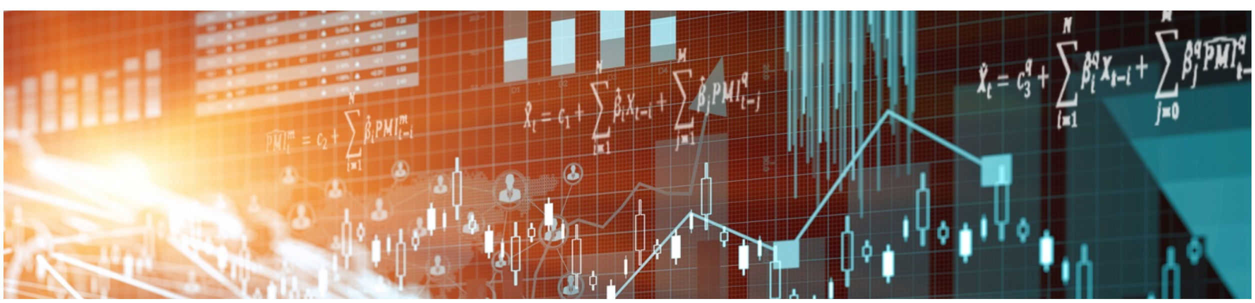
Staff Working Paper/Document de travail du personnel 2017-51

Identifying the Degree of Collusion Under Proportional Reduction



by Oleksandr Shcherbakov and Naoki Wakamori

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2017-51

November 2017

Identifying the Degree of Collusion Under Proportional Reduction

by

Oleksandr Shcherbakov[1] and Naoki Wakamori[2]

[1] Currency Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
oshcherbakov@bankofcanada.ca

[2] University of Tokyo
Tokyo, Japan 113-0033
nwakamo@e.u-tokyo.ac.jp

ISSN 1701-9397

Acknowledgements

We are grateful to Frank Neher, Volker Nocke, Kathleen Nosal, Martin Peitz, Lily Samkharadze, and Philipp Schmidt-Dengler. We also wish to thank the participants at the 4th Mannheim Centre of Competition and Innovation (MaCCI) I.O. Day, the 19th Conference of SFB-TR 15, the 2015 MaCCI Annual Conference, the 2015 International Industrial Organization Conference (IIOC), the 2015 European Economic Association Meeting, and the 2017 European Association Research in Industrial Economics (EARIE). This project was mostly completed when the authors were at University of Mannheim and financial support from the German Science Foundation (SFB-TR 15) is gratefully acknowledged.

Abstract

Proportional reduction is a common cartel practice in which cartel members reduce their output proportionately. We develop a method to quantify this reduction relative to a benchmark market equilibrium scenario and relate the reduction to the traditional conduct parameter. Our measure is continuous, allowing us to have an intuitive interpretation as the "degree of collusion" and nesting the earlier models in the existing literature. Furthermore, our methodology addresses Corts' (1999) critique by estimating time-varying degree of collusion from a short panel of firm-level observations, exploiting firms' ex post heterogeneity. We illustrate the method using the Joint Executive Committee railroad cartel data.

Bank topics: Econometric and statistical methods; Market structure and pricing
JEL codes: D22, L41, C36

Résumé

La réduction proportionnelle de la production est une pratique courante des cartels. Nous élaborons une méthode de calcul qui quantifie ces réductions en fonction d'un point d'équilibre de référence établi à partir de scénarios et faisons un lien avec un paramètre traditionnel de conduite. La mesure que nous obtenons est continue et peut être intuitivement interprétée comme une mesure du degré de collusion. Elle s'insère dans les modèles proposés par d'autres auteurs. En outre, avec cette approche, et à l'aide de données de panel court provenant d'entreprises, nous répondons à la critique de Cort (1999) en estimant le degré de collusion variable dans le temps et en exploitant l'hétérogénéité a posteriori des firmes. Pour illustrer le fonctionnement de notre méthode, nous utilisons les données du comité conjoint de direction du cartel du chemin de fer.

Sujets : Méthodes économétriques et statistiques ; Structure de marché et fixation des prix
Codes JEL : D22, L41, C36

Non-Technical Summary

Measuring market power and assessing industry conduct remain among the major challenges in empirical industrial organization. These questions have important implications for welfare analysis and antitrust regulation. The earlier literature on collusion, which is often referred to as *conjectural variation* literature, typically assumes that cartels maximize joint industry profit. Under this assumption it is possible to derive an estimable equation with a parameter that takes three distinct values corresponding to three different modes of competition: perfect competition (the most competitive), Cournot competition, and monopoly (the least competitive). The estimated parameter then indicates which mode of competition is likely to be supported with the data. As pointed out by Corts (1999), however, this methodology has some fundamental issues: (1) the estimated value might be different from all three values, which results in internal inconsistency, and (2) the method must assume that the degree of collusion is the same over time, which is very unlikely in practice.

To address these limitations of the traditional methodology in the literature, we develop a new method to quantify the degree of collusion. Instead of assuming that cartel members maximize joint profit, we assume that they proportionally reduce their output from benchmark outcome when engaged in collusion. This assumption is called "proportional reduction," which is formally introduced by Schmalensee (1987). Even though the assumption seems to be restrictive, there are many cartels that use proportional reduction in practice. Section 2 of the paper supports our assumption and documents many examples that use proportional reduction or some variation in practice, including a famous worldwide Vitamin cartel and some cartels found in Austria, Finland, Germany, Japan and Norway.

We then formally introduce our model in Section 3 and show that our method allows us to have a continuous and time-varying measure of degree of collusion, which overcomes the aforementioned critique by Corts (1999). The main reason why we can overcome the issues in the earlier literature is explicitly exploiting firm-level data, while the existing literature does not fully utilize this variation. The paper provides a thorough discussion of identification and extendibility of our method. Monte Carlo simulations in Section 4 demonstrate validity of our method.

To illustrate our method, we use well-documented data on the Joint Executive Committee – a legal railroad cartel in the U.S. in 1880–1886. Since our approach is based on a structural model, we provide identification conditions within the proposed parametric structure. We begin our analysis by showing that the firms observed in this data were indeed engaged in Cournot competition by conducting a non-parametric test developed by Cavajal et al. (2013). The test results show that at least 90 percent of data is consistent with Cournot competition, which enables us to apply our method. Our final estimation results show that, even not knowing when a cartel occurred, our methodology can detect most of the periods when firms colluded and reduced their output by about 30 percent compared with the static Cournot equilibrium output.

1 Introduction

Measuring market power and assessing industry conduct remain among the major challenges in empirical industrial organization. These questions have important implications for welfare analysis and antitrust regulation. A wide variety of empirical models have been developed to measure the degree of competition in markets where reliable cost data are not available. The problem frequently boils down to estimating a "conduct parameter" that summarizes the level of competition in an industry. Typically, an econometrician specifies a supply relation where the conduct parameter takes on distinct values nesting perfect competition (Bertrand), Cournot, and perfect cartel (Monopoly) models. Estimated parameter values are then interpreted as the degree of collusiveness. In reality, however, the estimated parameter values are often significantly different from the values describing either of the conduct regimes, making it harder to interpret.[1]

A problem that is more serious than the internal inconsistency between a theoretical model and its empirical implications is raised by Corts (1999), who shows that the estimated parameter values may fail to measure market power because of dynamic considerations of the firms. When firms are efficiently colluding, changes in the economic environment may affect the degree of collusion (e.g., cartel sustainability as described in Rotemberg and Saloner, 1986), suggesting that the conduct parameter would change over time and would be an endogenous variable. Thus, across-time variation in the demand and supply conditions may fail to identify the industry conduct.

In this study, we propose an alternative way to evaluate the industry conduct that overcomes many of the aforementioned problems in the literature. The key to our method is an assumption on the way collusion is implemented. Instead of assuming that the objective function of a cartel is known, e.g., joint profit maximization, we assume that firms employ proportional reduction (PR) collusive technology (as discussed in Schmalensee, 1987). Under the PR assumption, cartel members reduce their outputs proportionately relative to a benchmark market equilibrium output. We provide empirical evidence in support of our assumption in the next section.

Our method has several advantages over the traditional conduct parameter approach. First, our parameter takes values on a continuous interval, having a simple interpretation as the percentage reduction in the output relative to a well-defined benchmark competitive equilibrium outcome. Second, we show that firms' heterogeneity provides useful variation, which can be used to estimate time-varying degree of collusion from a relatively short panel of firm-level observations. This source of identification is present even when firms are symmetric ex ante, i.e., before realizations of iid innovations to their costs. The ability to estimate a time-varying degree of collusion is important to address Corts' critique. Finally, while illustration of the method in this paper is provided using a very simple static framework, the method is extendable to more complex settings with dynamically optimizing agents and more flexible forms of the demand and cost functions. Therefore while strong in itself, our assumption about collusive technology can help to accommodate a wide variety of complex strategic interactions and can be used when a researcher prefers to stay agnostic about the objective function of a cartel.

[1] In such cases, the industry competitiveness is evaluated in terms of the number of firms playing a particular equilibrium. This interpretation of the conjectural variation parameter is sometimes referenced as the "*as-if*" interpretation. For example, an industry with N firms is as competitive as if it were Cournot equilibrium with K players.

To illustrate our method, we use the data from the Joint Executive Committee railroad cartel, which is thoroughly and extensively studied by Porter (1983) and Ellison (1994). This cartel is well documented, as the cartel was formed in 1879, which is prior to the passage of the Sherman Act in 1880. Thus, the data contain monthly-level information on whether or not the firms successfully colluded, as well as standard firm-level information on quantity – shipment volumes for grain and flour in this context – and prices. Assuming that the firms used PR collusive technology when they colluded and the proportion was constant over the sample periods, we find that, in collusive periods, the firms reduced their output by about 30% compared with the static Cournot equilibrium output. We further estimate the parameters assuming that we do not know whether the firms successfully colluded or not. Our estimated parameters detect the non-collusive periods quite well, which validates our methodology.

The rest of the paper is organized as follows. First, we provide theoretical and empirical evidence in support of our key assumption as well as discuss related literature. Section 3 begins with an illustrative example and then describes a general framework with a linear demand and constant marginal cost functions. Then we discuss alternative cost specifications as well as testable implications of the model. Discussion of caveats and possible extensions of our framework are presented in Section 4. Section 5 evaluates finite-sample properties of our model with a known data-generating process using Monte Carlo simulations. We illustrate an application of our method using the well-known data of the Joint Executive Committee railroad cartel in 1880–1886 in Section 6. Section 7 concludes.

2 Related Literature

The main focus of this paper is PR collusive technology and thus we begin by discussing the plausibility of PR from both theoretical and empirical points of view. We then relate our methodology to the traditional conjectural variation literature.

2.1 Proportional Reduction

To the best of our knowledge, Schmalensee (1987) is the first paper that explicitly defines PR in the literature, listing four distinct collusive technologies: (1) full collusion with side payments, (2) market division, (3) market sharing, and (4) proportional reduction. The most profitable one is *full collusion with side payments*, where only the most efficient firms produce. This type of collusion is unrealistic from the legal aspect and the remaining three ways of colluding do not require side payments.[2] Collusion implemented through *market division* occurs when each firm is assigned to a part of the market and charges its optimal monopoly price in this segment. The possibility of arbitrage makes such a technology difficult to implement in practice. *Market sharing* collusive technology involves assigning production quotas. For example, the quotas may be chosen to equate the critical discount factor among the cartel members, which would maximize sustainability of the cartel. Such arrangements generally would require solving a non-trivial bargaining problem,

[2]Without side payments, collusion would imply positive production levels even for the least efficient firms, making the maximization of joint industry profit unfeasible.

particularly when the firms are imperfectly informed about their rivals' costs. The last type is *proportional reduction*, which can be viewed a special case of market sharing. The simplicity of PR's implementation, compared to that of market sharing, makes it attractive for firms. Another advantage of PR is that frequently used concentration measures (e.g., HHI or C_n) would be observationally equivalent to a competitive outcome, as the distribution of market shares does not change between competitive and collusive regimes. Therefore, from the perspective of firms considering forming a cartel, PR has this very appealing feature.

From a microeconomic theoretical point of view, PR is a very natural assumption when assuming symmetric firms, which is common in the theoretical literature. It is hard to see why symmetric players that maximize joint industry profit would reduce their output *non-proportionally*.[3] In the case of asymmetric firms, the assumption becomes binding, as PR may not maximize total industry profit.

In the empirical literature, the classification of cartels is slightly different from Schmalensee (1987). For example, a recent paper by Hyytinen et al. (2017) categorizes cartels into (1) price-based cartels, and (2) allocation-based cartels, the latter of which includes proportional reduction. The latter type is more prevalent in practice, and Marshall and Marx (2008) study 20 major industrial-product cartel decisions of the European Commission between 2000 and 2005 in an attempt to follow up on a remarkable but underdeveloped theme in Stigler (1964), that is, the prevalence of market share allocations among explicit cartels. Marshall and Marx (2008) argue that "a market share allocation may be a superior choice as long as compliance with the agreement can be monitored" (p.1). Moreover, they claim that with a limited information exchange and a market share allocation, the firms could be able to eliminate all punishment periods associated with tacit collusion.[4] The authors also provide empirical evidence on the use of reversion to pre-collusive play as a punishment for deviations from collusion. This has been explicitly mentioned in Congressional testimony involving dyestuffs manufacturing (p.12).

Furthermore, Hyytinen et al. (2017) and Fink et al. (2015) study Finnish and Austrian cartels, respectively, and they find that allocation-based cartels are by far more popular in manufacturing sectors. In particular, Hyytinen et al. (2017) study contracts of about 900 legal Finnish cartels and find that market allocation-based cartels are more popular in manufacturing (73%), while price-based cartels are more popular in non-manufacturing (78%). A similar conclusion follows from Fink et al. (2015), who study legal cartels in Austria. The authors find that quota cartels were very common in the manufacturing sector of Austria. Given these findings, our method appears to be more applicable to manufacturing rather than to a service sector.

Establishing the fact that the market share allocation is indeed one of the most common cartel formats, a natural question one may ask next is the following: Among these market share allocations, do we indeed observe PR in practice? Our answer is yes. Before listing several examples, it is important to discuss implementation of PR, as it

[3] Symmetry and joint profit maximization are standard assumptions in the literature, in which case existing methods of identifying the parameter of interest (e.g., Bresnahan, 1982, Lau, 1982) are directly applicable within our framework.

[4] Note that unlike customer and geographic allocations, a market share allocation need not be invalidated by changes in the economic environment.

requires setting up a benchmark equilibrium with respect to which firms proportionally reduce their outputs. In practice, there are two possibilities: the firms reduce output proportionately based on (1) production capacity, or (2) the average market share from the prior periods. We believe that in either case, our PR assumption would be justified. Consider the former case. Kreps and Scheinkman (1983) show that a two-stage game, with the quantity pre-commitment in the first stage and price competition in the second stage, under some assumptions has a unique NE with the Cournot outcome. Hence, if firms invest in their capacity assuming competition in the subsequent stages, collusion with quotas proportional to the production capacity would be equivalent to our PR assumption. Similarly, in a homogeneous product environment, quotas based on the average of market shares from the prior periods should be proportional to the Cournot market shares.

In what follows, we summarize some empirical evidence in support of our assumption. Harrington et al. (2014) attempt to better understand why and how firms that initially agreed to participate in a cartel would subsequently choose to cheat. The authors find that the most important source of disagreement lies in the market allocation, where the most commonly documented method to allocate quotas is to use historical market shares. For example, the organic peroxides cartel used sales from 1969–70 to set collusive quotas for 1971. Cartels in the vitamins A and E markets in the early 1990s set market shares at 1988 levels and firms agreed to maintain these shares in response to market growth. For the folic acid cartel, cartel member Roche negotiated with the Japanese cartel members as a group and ultimately settled on market shares based on 1990 sales, which gave Roche a share equal to 42%. The Japanese producers then allocated their 58% share amongst themselves according to their 1990 market shares. The citric acid and zinc phosphate cartels used the average of firms' sales over the previous three years. The sorbates cartel set the allocation for 1978 between Hoechst and the four Japanese producers based on sales volumes in 1977 for each region of the world, and the Japanese producers allocated their aggregate share according to 1973–77 sales. A similar situation occurred in the German cement cartel, which established quotas based on historical market shares (p.11). The authors conclude that "cartel formation often implies freezing the relative positions of firms in that collusive market shares are set equal to the competitive market shares at that time" (p.2). If growth in market demand is insufficient to utilize new capacity expansion, then a market sharing agreement based on the historical market shares may make the cartel unstable.

Roller and Steen (2006) study the Norwegian cement cartel in 1955–1968, which had to decide on the total amount of cement sold domestically and on a sharing rule. The latter in turn determines how the rent is allocated among the cartel members. The cartel decided to reward domestic market shares based on the members' share of total capacity. As a result, cartel members raced to increase their sales quotas by building more capacity. By refuting alternative explanations for the over-investment in production capacity, such as (1) unrealistically high anticipation of increased future consumption and (2) import deterrence, the authors conclude that the incentives created by the cartel's sharing rule is the most plausible explanation for the large capacity investments in Norway. The authors mention similar examples of the market sharing arrangements in Germany in the 1920s and 1930s and in Japan in the 1950s and 1960s. Another prominent example of a cartel that divides the market share according to production capacity is the so-called Lysine cartel in 1992–1995. The most recent examples of production-sharing rules are found in the agricultural cooperatives that purchase whatever their members have been able to produce and then decide how much to sell at home (e.g., Bergman, 1997, as cited in Roller and Steen, 2006).

Matsui (1989) provides empirical evidence on the Japanese "recession cartels." When the performance of an industry became worse off because of business fluctuations, the government not only allowed firms to form a recession cartel but also suggested that they should do so (p.451). Similar to the Norwegian cement cartel, the market sharing rule allocated collusive shares in proportion to the production capacity. Observed excess capacity is thus consistent with the incentives provided by the sharing mechanism. Indeed, if firms have more than one control but agree to collude in only one of them, there is a clear strategic reason to affect the market sharing arrangements by the choice of another control. It is worth noting that the phenomena was described by several authors and is sometimes called a "disadvantaged semi-collusion" (e.g., Fershtman and Gandal, 1994). Of course, the over-investment makes sense if the firms anticipate that collusion will re-occur regularly in the future, which is usually the case with the government-supported cartels.[5]

Of course, there are some cartels that use market share allocation, which might be different from PR. We thoroughly discuss this possibility in Section 4.1 via numerical simulations. Moreover, in this paper we do not differentiate between legal and tacit cartels.[6] However, there is a nice insight into the mechanics of implementation of PR-like collusive technologies in Cave and Salant (1995). The authors discuss the institutional arrangements under which governments may act as a collusion-facilitating device. They study the choice of quotas by legal volume-restricting cartels and find that the same voting institution has appeared in different countries at different points in time. Such an institution appears helpful in resolving the fierce internal conflict over choosing the level of collusion. The mechanics of the voting process involve assigning a scalar q_i (e.g., capacity, historical shares) to each firm and then choosing a common scale factor F, such that each firm is allowed to produce in the interval $[0; Fq_i]$. The authors specify conditions such that there exists a unique voting equilibrium in which the ideal point (in terms of F) of the median indexed firm is weakly preferred to any other point by a majority of voters. As a consequence, the total industry profit is not maximized as would be the case if (and only if) the ideal point of the largest firm is chosen. This setup is consistent with our PR assumption if the scalars chosen are equal to the Cournot equilibrium output levels. Hence, we believe Cave and Salant (1995) provide additional theoretical motivation for the underlying assumption in our method.

2.2 Alternative Approaches to Identifying Market Power

Our work is closely related to the early studies in conjectural variation literature by Bresnahan (1982), Lau (1982), Porter (1983) and Ellison (1994). Ellison (1994) provides a comprehensive empirical comparison of competing theories of collusion by Green and Porter (1984) versus Rotemberg and Saloner (1986). In these (and many other) articles, in order to derive an empirical specification for estimation a researcher has to assume that the objective function of a cartel is known, e.g., joint profit maximization. In reality, the objective function of a cartel is rarely known. It may be quite complex and depend

[5]Interestingly, a similar story applies to voluntary export restraints (VERs). Yano (1989) shows that if a quota is expected to be imposed on foreign exporting firms, it intensifies the exporting firms' competition in the pre-quota period.

[6]This is because we are agnostic about objective functions of the cartels.

not only on the current and future states of the demand and supply conditions, but also on the probability of disclosure (which, in turn, may be a function of the level of collusion itself) and expected punishment by the antitrust authorities. Given feasible (potentially implementable in the real world) equilibrium supporting strategies, not all levels of collusion can be sustained, as described by Rotemberg and Saloner (1986). Analysis of more complex settings, when the collusion occurs only along one of the dimensions, e.g., price fixing with competition in quality or capacity, is provided in Fershtman and Gandal (1994). Availability of reliable cost data facilitates estimation of industry conduct considerably. For example, Genesove and Mullin (1998) conduct a comprehensive comparison of various ways to estimate industry conduct and marginal costs in the sugar industry. Wolfram (1999) also considers a model with time-varying conduct parameters when direct measures of marginal costs are available. However, her identification still relies on the time-series variation in the data because in a duopolistic market the variation across firms is limited. A more structural way to address the Corts' critique can be found in Puller (2009).

Assumption of PR in collusive periods facilitates identification of the degree of monopolization within a given model structure, including functional form assumptions on the demand and cost functions. It is worth noting that, differently from the earlier conjectural variation literature, we do not require assumptions about non-separability of demand in some observable demand "rotators," as in Bresnahan (1982) and Lau (1982). In this sense, we can be flexible when choosing demand and cost specifications. We provide a traditional identification argument based on the familiar rank conditions for linear demand and constant (or linear in quantity) cost functions. We do not discuss non-parametric identification in this study. A formal non-parametric identification argument based on variation in exogenous demand and/or cost shifters and market size for similar models can be found in Berry and Haile (2014). The closest paper on non-parametric identification in homogeneous product markets is Cherchye et al. (2013). The authors provide characterization of the Cournot, Perfect Competition, Perfect Collusion, and Conjectural Variation models. The characterization is based on three conditions for consistency with a corresponding model. Cherchye et al. (2013) discuss characterization of a conjectural variations model in terms of the firms' conjectures about the likely responses of their rivals. Their Theorem 3.1. provides the necessary and sufficient conditions for the reduced form price and individual quantity functions to be *conjectural variation consistent*. Our framework is different in an important way because our monopolization parameter θ measures the degree of output reduction actually implemented and not the hypothetical beliefs about other firms' responses, as in traditional conjectural variation models.[7] This difference is crucial, as discussed in Corts (1999) and Reiss and Wolak (2007), and is one of the major flaws of the traditional literature on collusion.

There is a large and growing literature on identifying collusion in differentiated products industries. Nevo (1998) discusses two typical approaches to identification of collusion. The first, "menu approach," is directly derived from theory and suggests estimating a finite set of alternative models of firms' conduct. Each of the models corresponds to a matrix (sometimes called an "ownership matrix") of zeros and ones, where zeros (ones) switch off (on) corresponding elements in the matrix of cross-price derivatives. Each of the matrices is consistent with an assumption about joint profit maximization for a given subset of products. Discrimination between these alternatives is based on the model fit.

[7]We provide comparison of values for both measures in Table 1.

The second approach has a weaker theoretical justification and replaces the ownership matrix with a set of parameters (supposedly between zero and one) to estimate. The set of conjectural variation parameters allows firms to internalize the effect of their control variables on the market shares of their rivals in a continuous way. For example, instead of making a binary decision of whether to internalize it or not, the firms decide to what extent the cross-price elasticity in the first-order conditions matters for various competing products. Identification of parameter within the ownership matrix may be hard because it requires finding a large number of exogenous demand shifters uncorrelated with the shocks in the firms' first order conditions. For example, Ciliberto and Williams (2014) in their study of collusion and multi-market contact in the airline industry use exogenous variation in the number of airport gates leased to different firms as exogenous instruments.[8] Our approach can be extended to differentiated product markets, as we briefly discuss in Section 4 and also in Appendix D. We intentionally abstract from such a complexity in the main text in order to keep our presentation transparent.[9] In practice, competition authorities have time and resource constraints when making decisions, and we believe that an easily implementable but reliable methodology that has microfoundation will be helpful for them.

3 The Model

In this section, we outline our framework by presenting a model with linear demand and constant marginal cost functions. Potential extensions of the model are discussed in Section 4. We begin with an example to illustrate the idea of our methodology.

3.1 An Illustrative Example

Consider a homogeneous product market where two firms, A and B, compete in quantities over two periods, 1 and 2. The demand and cost functions are given by

$$\begin{aligned} P_t(Q) &= a - b(q_A + q_B) + d_t, \\ C_i(q) &= c_i q, \end{aligned}$$

where d_t denotes an observed demand shifter at period t. Then, the Cournot outcome for period t is given by

$$(q^c_{A,t}, q^c_{B,t}) = \left(\frac{1}{3b}(a + d_t - 2c_A + c_B), \frac{1}{3b}(a + d_t + c_A - 2c_B) \right).$$

Now consider a situation where two firms collude using PR technique, i.e., the firms agree to reduce their outputs by the same percentage relative to the Cournot outcome. Define θ_t as a reciprocal number of the production amount under PR relative to the Cournot outcome. For example, if the firms reduce their output by 20% in period 1 and 0% in period 2, meaning that they produce 80% and 100% of the Cournot outcomes in period 1 and 2, respectively, then θ_t must be $\theta_1 = 1/(1 - 0.2) = 1/0.8 = 1.25$ and

[8] The authors also permit time-varying conduct parameter.

[9] Application to a differentiated product setup would require additional assumptions on the demand and cost functions.

$\theta_2 = 1/(1-0) = 1/1 = 1$. Econometricians can observe $(q^{obs}_{A,1}, q^{obs}_{B,1})$ and $(q^{obs}_{A,2}, q^{obs}_{B,2})$ in the data, though they cannot observe θ_t or cost parameters, (c_A, c_B). In this case, can econometricians recover $(\theta_1, \theta_2, c_A, c_B)$ from what they observe? The answer crucially depends on whether the demand shifters vary across period, i.e., $d_1 \neq d_2$ or $d_1 = d_2$. We discuss these two cases below, assuming that the demand parameters are known in order to keep our discussion transparent.

Case 1: The demand varies over time Suppose that the parameters take the following values: $a = 15$, $b = 1$, $d_1 = 0$, $d_2 = 9$, $c_A = 1$, and $c_B = 2$. Then, the Cournot outcomes are given by

$$(q^c_{A,1}, q^c_{B,1}) = (5, 4), \quad \text{and} \quad (q^c_{A,2}, q^c_{B,2}) = (8, 7).$$

First, consider a case where the degree of collusion changes over time, say $\theta_1 = 1.25$ and $\theta_2 = 1$, i.e., two firms reduce their outputs by 20% relative to the Cournot quantity in period 1, but these firms fail to collude in period 2, producing the Cournot quantity. In this case, what econometricians can observe in the data are

$$(q^{obs}_{A,1}, q^{obs}_{B,1}) = (4, 3.2), \quad \text{and} \quad (q^{obs}_{A,2}, q^{obs}_{B,2}) = (8, 7).$$

The data, then, allow us to have four first-order conditions with four unknowns:

$$\begin{aligned} 15 - c_A - 11.2\theta_1 &= 0, & 15 - c_B - 10.4\theta_1 &= 0, \\ 24 - c_A - 23\theta_2 &= 0, & 24 - c_B - 22\theta_2 &= 0, \end{aligned}$$

using the relationship $q^c_{j,t} = \theta_t q^{obs}_{j,t}$. Solving this system of equations enables us to obtain $(c_A, c_B, \theta_1, \theta_2) = (1, 2, 1.25, 1)$, implying that the parameters can be recovered. Naturally, the model allows us to identify time-invariant degree of collusion. Suppose that two firms reduce their outputs by 20% relative to the Cournot outcome in both periods, i.e., $\theta_1 = \theta_2 = 1.25$. In this case, the data must look like

$$(q^{obs}_{A,1}, q^{obs}_{B,1}) = (4, 3.2), \quad \text{and} \quad (q^{obs}_{A,2}, q^{obs}_{B,2}) = (6.4, 5.6).$$

Again, our model gives four first-order conditions with four unknowns, which enables us to recover the parameters of interest, $(c_A, c_B, \theta_1, \theta_2) = (1, 2, 1.25, 1.25)$.

Case 2: The demand does not vary over time In order to see the importance of the demand shifters, d_t, assume that $d_t = 0$ for both periods, yielding exactly the same Cournot outcomes:

$$(q^c_{A,1}, q^c_{B,1}) = (5, 4), \quad \text{and} \quad (q^c_{A,2}, q^c_{B,2}) = (5, 4).$$

First, consider a case where the degree of collusion varies over time, $\theta_1 = 1.25$ and $\theta_2 = 1.0$. In this case the data must be

$$(q^{obs}_{A,1}, q^{obs}_{B,1}) = (4, 3.2), \quad \text{and} \quad (q^{obs}_{A,2}, q^{obs}_{B,2}) = (5, 4),$$

which allows us to have the following four first-order conditions:

$$\begin{aligned} 15 - c_A - 11.2\theta_1 &= 0, & 15 - c_B - 10.4\theta_1 &= 0, \\ 15 - c_A - 14\theta_2 &= 0, & 15 - c_B - 13\theta_2 &= 0. \end{aligned}$$

This system of equations does not have a solution, which does not allow us to identify the degree of collusion. The same conclusion holds for a case where the degree of collusion does not change over time, i.e., $\theta_1 = \theta_2$. In this case, there are four parameters of interest, $(c_A, c_B, \theta_1, \theta_2)$, though there are essentially only two relevant first-order conditions, which prevents us from recovering four parameters. Therefore, when the demand does not vary over time, the model does not allow us to identify the degree of collusion, even it is time-invariant.

There are two important observations to make about this numerical example: (1) the importance of the demand shifters, and (2) the role played by heterogeneity in cost functions. For the first point, these two cases demonstrate that the demand shifter is one of the important sources for identification. Moreover, when demand shifters are available, our methodology allows us to identify even the time-varying degree of collusion, which can address the critique by Corts (1999). Then, these observations raise the question why the model can identify time-varying θ_t when the demand shifters are available. Intuitively, in a Cournot model with linear cost functions, when the demand changes, which corresponds to the change in d in our model, the market share must be changed. In our example, the market share for firm A is given by

$$s_{A,t} = \frac{a + d_t - 2c_A + c_B}{2a + 2d_t - c_A - c_B},$$

regardless of the degree of collusion. Assuming that the cost is known, only the change in d leads to the change in market shares. This change in market share enables us to identify θ; when observing the change in the market share in the data, deviating from the model prediction, then it must be through the change in the degree of collusion. Case 2 confirms this intuition; with the absence of the demand shifters, the market share for firm A is always constant. In case of $\theta_1 = \theta_2$, the market share for firm A is always $5/(5+4)$. Similarly, in case of $\theta_1 \neq \theta_2$, the market share for firm A in period 1 is $4/(4+3.2)$ and the market share in period 2 is $5/(5+4)$, which is identical to $4/(4+3.2)$. This feature prevents us from identifying the degree of collusion.

Second, the heterogeneity in cost functions plays a central role in this model. If two firms have the same technology and produce the same output level, we do not have a sufficient number of relevant first-order conditions. Thus, heterogeneous costs, resulting in heterogeneous market share, are important. Note that although non-linearity in cost functions seems to complicate the model and one might worry that such models do not allow us to identify the degree of collusion, this is not the case. More complicated functional forms create more heterogeneous responses to the change in the demand, which enables us to identify the degree of collusion. This issue is also thoroughly discussed in the next section.

3.2 Our Main Model

Consider a homogeneous product market with N firms competing in quantity over time, $t = 1, 2, \ldots, \infty$. Suppose each firm is characterized by a cost function denoted by $C_i(q_{it}, z_{it})$, where q_{it} is output and z_{it} is a vector of cost shifters. Let the inverse demand function be given by $P_t = P(Q_t, Y_t)$, where $Q_t = \sum_{i=1}^{N} q_{it}$ denotes total industry output and Y_t is a vector of demand shifters. The per-period reward function is given by

$$\pi_{it} = P(Q_t, Y_t) q_{it} - C_i(q_{it}, z_{it}). \tag{1}$$

The firms in the industry interact repeatedly and can be engaged in tacit collusion agreements. Instead of making an assumption on the objective function of the cartel, which is typically unknown to econometricians, we make the following assumption on the way the collusion is implemented.

Assumption 1: *In any collusive period, firms reduce their individual output proportionally to the baseline Cournot quantities, i.e.,*

$$q_{it}^{C} = \theta_t q_{it}^{PR}, \ \forall i, t,$$

where q_{it}^{C} and q_{it}^{PR} denote one-period Cournot and collusive output levels under PR respectively, and $\theta_t \geq 1$ is the inverse of the percentage reduction in output.

Assumption 1 implies that knowing θ_t allows us to compute the counterfactual Cournot quantity by "inflating" observed output q_{it}^{PR} by a factor of θ_t. For example, suppose that in the collusive period each firm reduces its output by 10% relative to the Cournot quantity. Then, $\theta_t = 1/(1-0.1) = 1.11$. Under Assumption 1, the degree of collusion can be summarized by the parameter θ_t. Hence, our ultimate objective is to estimate θ_t from the observed data.

Before proceeding with how to recover θ_t, it is worth noting that we intentionally abstain from developing a particular structural model of collusion, i.e., our model avoids specifying the objective function of the cartel or the bargaining process, which we cannot learn from the data. However, one can think of simple collusion supported by grim trigger strategies with Cournot-Nash as the punishment phase. Lemma 1 in Appendix B shows that PR technology is profitable for all firms in the neighborhood of the Cournot equilibrium quantity. Therefore, it is straightforward to prove that there exists a common discount factor $\beta = \min\{\beta_1, \ldots, \beta_N\}$, $\beta_i \in (0,1)\ \forall i$, such that the collusion is sustainable.

We now assume a linear inverse demand function $P(Q_t, Y_t) = \alpha_0 + \alpha_1 Q_t + \alpha_2 Y_t + \nu_t^d$ and suppose we observe $(q_{it}^{PR}, z_{it}, P_t^{PR}, Y_t),\ i = 1, \ldots, N;\ t = 1, \ldots, T$, in the data. Under Assumption 1, the following relationship must hold:

$$\begin{aligned} P(Q_t^C, Y_t) &= P(Q_t^C, Y_t) - P(Q_t^{PR}, Y_t) + P(Q_t^{PR}, Y_t) \\ &= \alpha_1 (\theta_t - 1) Q_t^{PR} + P_t^{PR}, \end{aligned}$$

where $Q_t^{PR} = \sum q_{it}^{PR}$ and $P_t^{PR} = P(Q_t^{PR}, Y_t)$ are collusive total output and equilibrium price respectively, and $Q_t^C = \sum q_{it}^C$ and $P(Q_t^C, Y_t)$ are unobserved (counterfactual) Cournot total output and equilibrium price, respectively.

On the supply side, we assume a constant marginal cost function, i.e., $\partial C_i(q_{it}, z_{it})/\partial q_{it} = \beta_{0i} + z_{it}\beta + \nu_{it}^s$, where z_{it} is a vector of observed cost shifters in the data and ν_{it}^s is unobserved cost component. This assumption is for transparency and we can easily relax this cost function to a more general class, which is discussed in Section 3.3. In a Cournot NE, first-order conditions for firm i are given by

$$\alpha_1 q_{it}^C + P_t^C - \beta_{0i} - z_{it}\beta - \nu_{it}^s = 0. \tag{2}$$

Note that equation (2) would not hold with equality when evaluated at $\left(q_{it}^{PR}, P_t^{PR}\right)$, as there would be incentives to deviate from the collusive quantity by expanding the output. However, we know the relationship between the collusive and competitive regimes and, therefore, can "restore" individual first-order conditions in terms of collusive values and the parameter θ_t as follows:

$$\alpha_1 \theta_t q_{it}^{PR} + \alpha_1 (\theta_t - 1) Q_t^{PR} + P_t^{PR} - \beta_{0i} - z_{it}\beta - \nu_{it}^s = 0. \tag{3}$$

Even though one may attempt to identify both θ_t and α_1 using just the supply relation (3), we focus on identification of the conduct and cost function parameters and assume, throughout the rest of the paper, that α_1 can be consistently estimated using conventional instrumental variable techniques.[10]

From equation (3) one can already see that variation in q_{it}^{PR} across firms, holding Q_t^{PR} and P_t^{PR} fixed within a cross-section, provides additional identification power. More formally, identification of our parameter of interest relies on the availability of firm-level exogenous cost shifters, which is summarized in the following Assumption 2.

Assumption 2: *Data contains information on exogenous demand and firm-level cost shifters, (Y_t and $z_{1t}, \ldots, z_{Nt}$ respectively), such that demand-side and cost-side innovations satisfy*

$$\mathrm{E}\left[\nu_{it}^s | z_{it}, z_{-it}, Y_t\right] = \mathrm{E}\left[\nu_t^d | Z_t, Y_t\right] = 0,$$

where $Z_t = \sum_i z_{it}$ and z_{-it} are cost shifters for firms other than the firm i.

For example, if $\beta_{0i} = \beta_0, \forall i$ and we observe just one cost shifter satisfying Assumption 2 in the data, we can identify all parameters in the model given that $T > 1$.[11] To see this, rewrite equation (3) as

$$\nu_{it}^s = P_t^{PR} - \beta_0 - z_{it}\beta - \alpha_1 Q_t^{PR} + \theta_t \alpha_1 (q_{it}^{PR} + Q_t^{PR}),$$

and define $X = (\mathbf{1}_{NT}, \mathbf{z}, \mathbf{q}, \mathbf{I}_T \otimes \mathbf{x}_t)$ and $Z = (\mathbf{1}_{NT}, \mathbf{z}, \mathbf{I}_T \otimes \mathbf{z}_{-it})$ where

- $\mathbf{1}_{NT}$ and $\mathbf{1}_T$ are vectors of ones of sizes $N \times T$ and T respectively;
- $\mathbf{z} = (z_{11}, \cdots, z_{N1}, \cdots, z_{1,T}, \cdots, z_{NT})'$;
- $\mathbf{q} = \left(\alpha_1 Q_1^{PR}, \cdots, \alpha_1 Q_1^{PR}, \cdots, \alpha_1 Q_T^{PR}, \cdots, \alpha_1 Q_T^{PR}\right)'$;
- $\mathbf{I}_T$ is identity matrix of size T;
- $\mathbf{x}_t = \left(\alpha_1 \left(q_{1t}^{PR} + Q_t^{PR}\right), \cdots, \alpha_1 \left(q_{Nt}^{PR} + Q_t^{PR}\right)\right)'$;
- $\mathbf{z}_{-it} = \left(\sum_{j \neq 1} z_{jt}, \cdots, \sum_{j \neq N} z_{jt}\right)'$;
- $\otimes$ denotes Kronecker product.

The standard identification conditions for IV methods requires $Z'X$ to have a full column rank. With just one period of data, it is clear that X does not have a full column rank, given the first and third columns. Hence, separate identification of the constant term in the marginal cost specification and the time-varying parameter θ_t requires data for at least two time periods.[12]

More generally, there are two sources of variation that help to identify the degree of output reduction. The first one is variation across asymmetric firms. The second one is variation over time in the demand and supply conditions. Under the PR assumption, asymptotics is in terms of $N \times T^*$, where T^* is the number of time periods with constant

[10] It is rather the degree of collusion that may change in response to a changing economic environment.
[11] The system would be over-identified if in addition we observe demand shifters.
[12] Of course, to estimate demand parameter α_1, one would need longer time series.

conduct parameter θ_t. Of course, to fully utilize this property a researcher should specify observable cost shifters at the firm level. Regarding the non-parametric identification of our parameter, we reference Cherchye et al. (2013), who provide necessary and sufficient conditions for identification of alternative models of firms' conduct, including traditional conjectural variation models. While their approach can be applied in our settings, we believe that our PR assumption is already sufficiently strong to make a non-parametric identification argument unnecessary.

3.3 Functional Form Assumptions and Testable Implications

To make our exposition transparent, the previous subsection assumes the constant marginal cost specification. As mentioned earlier, however, this assumption can be relaxed. In this section we demonstrate a cost function specification that includes quadratic term and show that our identification argument does not rely on the constant marginal cost assumption. Furthermore, under two cost specifications, we further derive the testable implications of our model regarding whether or not our PR assumption is supported in the data.

In order to generalize our model, we begin by considering the firm i's maximization problem in a competitive regime where the per-period profit function is given by equation (1). Assuming away any dynamic effects of the quantity choice, which is discussed in Section 4 below, the first-order conditions for quantity choice are given by

$$\text{FOC}[q_{it}] : \frac{\partial P(Q_t, Y_t)}{\partial Q_t} q_{it} + P_t - \frac{\partial C_i(q_{it}, z_{it})}{\partial q_{it}} = 0. \tag{4}$$

When firms are in a collusive regime under PR, equation (4) can be written in terms of observables $(P_t^{PR}, Q_t^{PR}, q_{it}^{PR})$ and the parameter θ_t as follows:

$$P_t^{PR} + \Big[P(\theta_t Q_t^{PR}, Y_t) - P(Q_t^{PR}, Y_t)\Big] + \frac{\partial P(\theta_t Q_t, Y_t)}{\partial Q_t} \theta_t q_{it} - \frac{\partial C_i(\theta_t q_{it}, z_{it})}{\partial q_{it}} = 0, \tag{5}$$

where $\Big[P(\theta_t Q_t^{PR}, Y_t) - P(Q_t^{PR}, Y_t)\Big]$ represents a "collusive markup" over the Cournot price level, i.e., the difference between the observed outcomes and hypothetical competitive outcomes. This term measures price differences in the case of movement along the demand curve from the observed output levels to competitive output levels.

3.3.1 Relaxing the constant marginal cost assumption

Consider the particular case of when the marginal cost function contains a linear-in-q_{it} term, i.e., when

$$\frac{\partial C_i(q_{it}, z_{it})}{\partial q_{it}} = \beta_{0i} + \beta_q q_{it} + z_{it}\beta + \nu_{it}^d.$$

In the earlier literature it is well known that this case substantially complicates estimation of the conduct parameter (the problem description and potential solutions are discussed in Bresnahan, 1982, Lau, 1982). One of the frequently employed solutions would be to find exogenous variables affecting elasticity of the demand, i.e., in addition to the demand shifters, one would need to find some demand "rotators." Interestingly, when assumption 1 holds, we can estimate parameters of the model without the demand rotators.

To see this point, consider equation (3), which now becomes

$$(\alpha_1 - \beta_q)\theta_t q_{it}^{PR} + \alpha_1 (\theta_t - 1) Q_t^{PR} + P_t^{PR} - \beta_0 - z_{it}\beta - \varepsilon_{it} = 0. \tag{6}$$

As before, we assume that the slope of the demand function α_1 is estimated using the demand relationship (or that appropriate moment conditions are included into the GMM criterion function). Therefore, we can identify θ_t from the coefficient on Q_t^{PR}, while β_q is identified by the coefficient on q_{it}^{PR}. Note that now in order to disentangle β_q and θ, we need to have both estimates of $\widehat{(\alpha_1 - \beta_q)}\theta_t$ and $\alpha_1 \widehat{(\theta_t - 1)}$, where the latter requires variation in Q_t. Moreover, as both $(\theta_t - 1)$ and Q_t^{PR} in the second term are time t specific, our methodology must assume that θ_t must contain more than one time period. However, we believe that this restriction is not binding in many cases. For example, while we have weekly-level data, it is hard to imagine that θ_t changes every week, and thus we can estimate θ_t for each month. In summary, the traditional conjectural variation literature in presence of a linear-in-q term would require demand rotators, whereas our methodology does not require such demand rotators but needs to assume that θ_t is fixed for, at least, several periods.

3.3.2 Testable implications

Our approach relies on a strong behavioral assumption that firms proportionally reduce their outputs from the static Cournot quantity. It is natural to ask whether there are any testable implications of this assumption. Two recent studies address non-parametric identification of Cournot and conjectural variation models. The first study is by Cherchye et al. (2013), who require virtually no functional form restrictions except for twice continuous differentiability of the reduced form price and individual quantity functions. The authors provide a set of conditions that the reduced form functions must satisfy in order to be consistent with a particular behavioral model (Cournot, conjectural variation, etc.). The second study is by Carvajal et al. (2013), who use a revealed preference approach to characterize the Cournot model. Interestingly, the characterization allows them to test for constant level of collusion. The key assumption is monotone cost functions, i.e., there cannot be any shifters of the costs over time with all variation coming from changes in the demand conditions. The test is applied in the context of our application.

Even though these two recent non-parametric tests provide a nice opportunity to test the data for consistency with PR (provided their underlying assumptions are satisfied), they cannot be used to measure the degree of collusion. Besides, in practice, econometricians would typically assume some particular demand and cost functions. Thus, in what follows we focus on deriving testable implications within a given parametric structure, which can be more easily implemented than the aforementioned non-parametric tests.

Consider the general form of the first-order conditions evaluated at the observed realizations of variables given by equation (5). In this equation there are three elements that are affected by the choice of functional forms for demand and costs:

(i) Difference between the observed (collusive) price and the Cournot equilibrium price,

$$P(\theta_t Q_t, Y_t) - P(Q_t, Y_t).$$

(ii) Slope of the inverse demand at the Cournot equilibrium quantity,

$$\frac{\partial P(\theta_t Q_t, Y_t)}{\partial Q_t}.$$

(iii) Each firm's marginal costs evaluated at hypothetical Cournot quantities,

$$\frac{\partial C_i(\theta_t q_{it}, z_{it})}{\partial q_{it}}.$$

In the paper we assume that inverse demand function can be estimated separately, i.e., we do not rely only on the firms' first-order conditions to recover the inverse demand function and its slope.[13] Therefore, elements (i) and (ii) in the list are not very likely to confound identification of θ. At the same time, identification of the marginal cost function and the parameter of interest do rely on the same set of first-order conditions. However, since θ enters several additively separable functions, we potentially can test restrictions on the coefficients. For example, if we are willing to assume PR and that the degree of collusion is constant over time in the context of our duopoly model with linear demand, we have two equations,

$$\begin{aligned} P_t + \alpha_1(\theta - 1)q_{2t} + \alpha_1(2\theta - 1 + \beta_q)q_{1t} - \ldots = 0, \\ P_t + \alpha_1(2\theta - 1 + \beta_q)q_{2t} + \alpha_1(\theta - 1)q_{1t} - \ldots = 0, \end{aligned}$$

or

$$\begin{aligned} P_t + \gamma_1 q_{2t} + \gamma_2 q_{1t} - \ldots = 0, \\ P_t + \gamma_2 q_{2t} + \gamma_1 q_{1t} - \ldots = 0. \end{aligned}$$

If marginal cost is not linear in q, i.e., $\beta_q = 0$,

$$\gamma_2 = \alpha_1(2\theta - 1) \implies \theta = \frac{1}{2}\left(\frac{\gamma_2}{\alpha_1} + 1\right).$$

Also,

$$\gamma_1 = \alpha_1(\theta - 1) \implies \theta = \frac{\gamma_1}{\alpha_1} + 1,$$

which implies

$$\frac{1}{2}\left(\frac{\gamma_2}{\alpha_1} + 1\right) = \frac{\gamma_1}{\alpha_1} + 1 \implies \gamma_2 = 2\gamma_1 + \alpha_1.$$

Therefore, we can test for presence of confounding terms in the unknown marginal cost function by testing coefficient restriction $\gamma_2 = 2\gamma_1 + \alpha_1$. Of course, such a test requires long time series with constant level of collusion.

[13] An identification argument for derivatives, finite differences and function extensions can be found in Matzkin (2012).

What if firms reduce output non-proportionally as a result of some explicit or implicit collusion, i.e., such that $\theta_i \neq \theta_j$? For example, in a duopoly with a linear demand system we have two equations:

$$P_t + \alpha_1(\theta_2 - 1)q_{2t} + \alpha_1(2\theta_1 - 1)q_{1t} - \ldots = 0,$$
$$P_t + \alpha_1(2\theta_2 - 1)q_{2t} + \alpha_1(\theta_1 - 1)q_{1t} - \ldots = 0.$$

If we have enough time-series observations, we can estimate

$$P_t + \gamma_1 q_{2t} + \gamma_3 q_{1t} - \ldots = 0,$$
$$P_t + \gamma_2 q_{2t} + \gamma_4 q_{1t} - \ldots = 0,$$

and test the following restrictions whether $\gamma_1 = \gamma_4$ and $\gamma_2 = \gamma_3$. Interestingly, even in cases where the marginal costs are linear in q_i we still can test for PR because the system of equations would be given by

$$P_t + \alpha_1(\theta_2 - 1)q_{2t} + \alpha_1(2\theta_1 - 1 + \beta_q)q_{1t} - \ldots = 0,$$
$$P_t + \alpha_1(2\theta_2 - 1 + \beta_q)q_{2t} + \alpha_1(\theta_1 - 1)q_{1t} - \ldots = 0.$$

Of course, the test would require strong assumptions that the level of collusion is constant for long enough time periods because we no longer can rely on the degrees of freedom coming from across-firm variation in the first-order conditions (we lose them to learn θ_i).

4 Discussions and Extensions

This section provides additional discussion of our framework. The first subsection discusses the relationship between a general class of market sharing rules and PR. Although PR is one particular type of market sharing, other types of market share allocation (e.g., the ones maximizing cartel sustainability) might be employed by the firms. We explore the relationship between alternative market sharing allocations using numerical examples. Then we relate our PR parameter to the conduct parameter in the traditional conjectural variation literature. Finally we discuss some extensions of our model: (1) incorporating product differentiation; (2) modeling forward-looking decisions of the firms; and (3) proposing a test for the changes in regimes.

4.1 Market Sharing and Proportional Reduction

As we discuss in the Introduction, Marshall and Marx (2008) provide a number of arguments in favor of cartel organization based on market shares allocation. Unfortunately, it is hard to find direct evidence of PR collusive technology. This is because alternative collusive techniques may be difficult to distinguish if they result in output quotas allocations. For example, it is relatively easy to discriminate between (1) profit maximization with side payments, where only the most efficient firms produce; (2) market division, where each seller sells to a geographically separate market; or (3) customer allocation, where each seller sells to a given subset of buyers. PR and market sharing (MS) collusive technologies may have different quota allocation rules. For example, firms may choose to maximize total industry profit subject to maximum cartel sustainability, i.e., when all firms have common critical discount factor. This rule does not guarantee that the market shares of individual firms remain constant relative to the Nash equilibrium.

In order to explore the relationship between output quotas implied by MS and PR technologies, we simulate collusive and competitive scenarios with 3, 5, and 9 firms. In calculating equilibrium under the MS collusive technology we assume that firms maximize joint industry profit subject to maximum cartel sustainability. The latter, in turn, requires critical discount factors to be equal across firms. We assume that firms face a linear demand function, $P_t = 100 - Q_t,\ Q_t = \sum_{i=1}^{N} q_{it}$, and constant marginal costs, $c_i > 0$, such that $\frac{1}{N}\sum_{i=1}^{N} c_i = 5.0$. Then we obtain Nash, Collusive, and Deviation profits for the various number of firms and their costs asymmetries. In particular, to get optimal collusive profits we solve

$$\max_{q_1,\ldots,q_N} \left\{ \sum_{i=1}^{N} \pi_i(q_1,\ldots,q_N) \right\},\ s.t.\quad \frac{\pi_i^d - \pi_i^c}{\pi_i^d - \pi_i^n} = \frac{\pi_j^d - \pi_j^c}{\pi_j^d - \pi_j^n} \quad \forall i,j,$$

where π_i^d, π_i^n, and π_i^c are deviation, Nash, and collusive profit for firm i, respectively. This exercise is similar in spirit to the contract curves in Schmalensee (1987) (figures 1 and 2 on p.355). Our results are summarized in Table 1 below.

Table 1: A Comparison between Market Sharing and Proportional Reduction

	Identical Firms				Small Differences				Large Differences			
N	c_i	q_i^C	q_i^M	q_i^C/q_i^M	c_i	q_i^C	q_i^M	q_i^C/q_i^M	c_i	q_i^C	q_i^M	q_i^C/q_i^M
	5	23.75	13.67	0.67	4	24.75	14.36	0.67	1	27.75	16.49	0.68
3	5	23.75	13.67	0.67	5	23.75	13.68	0.67	5	23.75	13.69	0.67
	5	23.75	13.67	0.67	6	22.75	13.00	0.66	9	19.75	11.06	0.66
	5	15.83	8.68	0.60	3	17.83	9.77	0.60	1	19.83	10.93	0.60
	5	15.83	8.68	0.60	4	16.83	9.22	0.60	3	17.83	9.78	0.60
5	5	15.83	8.68	0.60	5	15.83	8.68	0.60	5	15.83	8.68	0.60
	5	15.83	8.68	0.60	6	14.83	8.16	0.60	7	13.83	7.65	0.61
	5	15.83	8.68	0.60	7	13.83	7.66	0.61	9	11.83	6.69	0.62
	5	9.50	5.64	0.56	3	11.50	6.51	0.53	1	13.50	7.46	0.51
	5	9.50	5.64	0.56	3	11.50	6.51	0.53	2	12.50	6.97	0.52
	5	9.50	5.64	0.56	4	10.50	6.06	0.54	3	11.50	6.50	0.53
	5	9.50	5.64	0.56	4	10.50	6.06	0.54	4	10.50	6.05	0.54
9	5	9.50	5.64	0.56	5	9.50	5.64	0.55	5	9.50	5.62	0.55
	5	9.50	5.64	0.56	6	8.50	5.24	0.57	6	8.50	5.22	0.57
	5	9.50	5.64	0.56	6	8.50	5.24	0.57	7	7.50	4.84	0.60
	5	9.50	5.64	0.56	7	7.50	4.86	0.60	8	6.50	4.49	0.64
	5	9.50	5.64	0.56	7	7.50	4.86	0.60	9	5.50	4.17	0.70

Note: Collusive quantities are computed under assumption that firms maximize total industry profit subject to constraint that all firms have common critical discount factor. We assume linear demand function $P = 100 - Q$, and constant marginal costs, s.t. $\frac{1}{N}\sum_{i=1}^{N} c_i = 5.0$.

Each block corresponds to one simulation, and nine simulation results are demonstrated in the table. For instance, the top-left block shows the simulation results for a case of three firms, having the same marginal costs of five. The symbol q_i^n denotes the quantities produced in a Cournot-Nash equilibrium, whereas q_i^c denotes the quantities produced under the optimal MS. When all firms have the same marginal costs of five, the ratio of q_i^n/q_i^c are the same across firms, 0.67, implying that MS yields the equivalent outcome to that of PR. For each number of firms, we simulate three sets of costs: homogeneous, heterogeneous but relatively similar, and very heterogeneous. As in the top-left, top-middle and top-right blocks, if the number of firms is three, regardless of the degree of heterogeneity, the optimal MS must be very similar to the outcome of PR.

The main message from the simulation exercise is that for relatively similar firms (with respect to their cost functions) PR may be a good approximation to MS collusive technology. This is not surprising because if the firms are identical, all collusive technologies would result in PR of output.

Another important observation is that similar firms have equal market shares under both competitive and collusive regimes. Indeed, if several firms have very similar cost functions it is conceivable that their competitive and collusive output levels are symmetric as well. This provides an interesting extension to our framework to the cases where firms can be divided into relatively homogeneous (with respect to cost functions) groups based on their observed market shares. Then our monopolization parameter can be estimated for each group of firms. By doing so one has to assume that firms having identical market shares also face identical demand and cost functions. For example, in Table 1 firms 1&2, 3&4, 6&7, and 8&9 have the same market shares under competitive and (an alternative to PR) MS collusive technology. Of course, estimating parameters for each group of the firms puts additional requirements on the data or would require additional assumptions about the time invariance of the parameters to exploit variation over time in observables.

One of the benefits of our method is that one does not have to assume that all firms are colluding. The framework is easily applicable to an industry with a few dominant players and a competitive fringe. As long as we are willing to make assumptions regarding the identities of colluding and free-riding firms, the method can be directly applied (e.g., cheating firms would choose their output levels with $\theta_{it} = 1$ if the baseline NE is Cournot).

4.2 Relation to the Conduct Parameter

It must be useful to compare our measure of market power to the conduct parameter from the earlier literature (e.g., Bresnahan, 1982). Typically, the existing literature identifies market power as a conduct parameter, λ, nesting three types of first-order conditions within one equation,

$$q_{it}\frac{\partial P}{\partial Q}\lambda + P_t - mc_i = 0, \tag{7}$$

where λ can take three distinct values, depending on the underlying scenario of industry conduct. Moreover, assuming symmetric firms, i.e., $q_{it} = q_{jt}\ \forall j$, we have $Q_t = Nq_{it}$ and, under three different modes of competition, equation (7) can be rewritten as

$$\begin{aligned}
&\text{Bertrand:}\quad \lambda = 0 && P_t - mc_i = 0,\\
&\text{Cournot:}\quad \lambda = 1 && q_{it}\frac{\partial P}{\partial Q} + P_t - mc_i = 0,\\
&\text{Perfect collusion:}\quad \lambda = \frac{Q_t}{q_{it}} = N && Q_t\frac{\partial P}{\partial Q} + P_t - mc_i = 0.
\end{aligned}$$

In practice, estimation can be done using aggregate quantities. Summing equations (7) over the firms and dividing by N yields the following equation that nests alternative collusive regimes:

$$\frac{\lambda}{N}Q_t\frac{\partial P}{\partial Q} + P_t - \frac{1}{N}\sum_i mc_i = 0. \tag{8}$$

On the other hand, a similar operation applied to our "restored" first-order condition under PR assumption, which is given by equation (3), gives us

$$\left[\frac{\theta + N\theta - N}{N}\right] Q_t^{PR} \frac{\partial P}{\partial Q} + P_t^{PR} - \frac{1}{N}\sum_i mc_i = 0. \tag{9}$$

A comparison of equations (8) and (9), which are in an aggregate form, immediately leads to two important observations. First, if we do not exploit firms' heterogeneity and use the variables aggregated at market-level, our methodology would face the same problem as the traditional conjectural variation literature does. Because firms are symmetric and individual first-order conditions do not provide additional information, our method would need to deal with a known identification problem when marginal cost function contains terms that are linear in Q_t. In other words, estimation would require the availability of demand rotators. Interestingly, non-parametric identification results from Cherchye et al. (2013) would apply to this formulation.[14]

Second, a comparison of two equations enables us to find a relationship between our model and the traditional conjectural variation model. As there is a one-to-one relationship between our parameter and a traditional measure of industry conduct, Table 2 summarizes the values of our parameter θ for each theoretically admissible value of λ as a function of the number of firms, N; as mentioned earlier, in the traditional conjectural variation model, Bertrand, Cournot and Monopoly models correspond to $\lambda = 0, 1$, and N, respectively. Under the symmetric assumption, our model can also nest these three models, which correspond to $\theta = N/(N+1), 1$, and $2N/(N+1)$. When we consider asymmetric firms, however, one can see the advantage of our measure; In the conjectural variation literature, when estimated parameter is different from $0, 1$, or N, the estimation results cannot be straightforwardly interpreted. On the other hand, our θ is defined on (the positive part of) a real line and has a straightforward interpretation as a reduction percentage from the Cournot quantity. Therefore, our model can be seen as a natural extension of the existing conjectural variation models.

Table 2: Traditional Measure of Industry Conduct vs θ

Mode of Competition	Conduct Parameter λ	Proportional Reduction θ
Bertrand	0	N/(N+1)
Cournot	1	1
Monopoly	N	2N/(N+1)

4.3 Some Extensions

Differentiated products So far we have considered homogeneous product markets. Potentially, the method can be applied to differentiated product markets. At the same time, it is not a straightforward extension. Under homogeneous products all firms face the same price and proportional quantity reduction appears natural. In a differentiated

[14] Since we are using individual firm first-order conditions, we do not elaborate on traditional identification strategies here.

product industry both quantities and prices are different. So, should the firms reduce output proportionally or should they instead raise prices proportionally?

If firms are similar with respect to both cost and demand parameters, then prices and quantities would change proportionally. With asymmetric firms in a differentiated product market, the difference in market shares between collusive and competitive regimes depends not only on the firms' cost functions but also on the degree of substitution or complementarity between the products.

In Appendix D we provide an example of a differentiated product market where firms maximize joint industry profits in a collusive regime. It is worth noting that application of our method to a differentiated product industry requires additional assumptions on the elasticity of the demand functions of individual firms and on the relationship between the level of demand and costs.

Dynamics Another potential extension would be to use a more structural approach and to model firms' maximization problem as a dynamic game. For example, Assumption 1 can be used within the framework of Fershtman and Pakes (2000). However, this would require explicit assumptions on the objective function of the colluding firms as well as specifying punishment strategies, which is exactly what we want to avoid in this study. As long as the baseline scenario (relative to which firms reduce their outputs) is given by a static NE, it remains an NE of a dynamic game. Hence, our parameter estimates can still be interpreted relative to a well-defined alternative conduct regime. An example of a more structural approach, addressing the critique by Corts (1999) when the firms are engaged in efficient collusion, is given in Puller (2009).

Testing for change in regimes If firms are colluding, with $\theta_t = \theta$ fixed within all periods in the observed data, the distribution of market shares would be identical to the one under the baseline scenario (Cournot in our case). However, if the regime of collusion changes at some point in time (e.g., as a result of price war), one can search for such evidence by inspecting variation in the aggregate or individual outputs and the distribution of market shares. One possible example would be to check if a concentration measure (say, HHI or C_N) is statistically different in periods before and after the time of the potential change in the conduct regime with a similar test (e.g., difference in means) used for the aggregate or individual levels of output. If the test rejects that the distribution of market shares are different while the difference in the output levels is significant, this would be consistent with a change in the degree of collusion. Alternatively, one can use non-parametric tests developed by Cherchye et al. (2013) and Carvajal et al. (2013) to test data for consistency with the conjectural variation model.[15]

Can Cournot dynamics be observationally equivalent to PR? Another important issue to discuss is identification of cost structure and the monopolization parameter θ_i. Since estimation of the cost function parameters requires assumptions about the firms' conduct, it is not clear if exogenous variation in the demand and supply conditions can replicate PR even in a competitive regime. For this to be true, in a Cournot NE, firms

[15]In the latter case the test is whether the degree of collusion is constant across firms, which would be consistent with our assumption.

must respond to exogenous variation in a proportional way. For example, consider a change in the demand conditions Y_t. Proportional change in output of each firm implies that

$$\frac{\partial q_{it}/\partial Y_t}{q_{it}} = \frac{\partial q_{jt}/\partial Y_t}{q_{jt}} \Rightarrow \frac{\partial q_{it}/\partial Y_t}{\partial q_{jt}/\partial Y_t} = \frac{q_{it}}{q_{jt}}, \forall i, j.$$

Through the implicit function theorem and Cournot first-order conditions, it is easy to show that to replicate PR as a result of aggregate demand shocks one must impose very strong restrictions on the underlying demand and cost functions. In particular, it is not possible in the case of linear demand and constant marginal costs (unless the firms are truly identical), while if the marginal costs are linear in quantity, the following would be required,

$$\frac{2\alpha_1 - \beta_q^{(j)}}{2\alpha_1 - \beta_q^{(i)}} = \frac{q_{it}}{q_{jt}}, \forall i, j,$$

where $\beta^{(i)}$ and $\beta^{(j)}$ are cost parameters for firms i and j. Similar conclusions can be made regarding the cost shocks. While it might be possible to reverse-engineer a model (and/or competitive equilibrium concept) where firms do respond to some exogenous variation by replicating PR collusive technology, we believe that for a very wide class of parametric empirical specifications used for estimation this is not true.

5 Monte Carlo Simulations

In order to demonstrate the performance of our method and evaluate the properties of our estimator, we conducted Monte Carlo simulations. The details of the simulation design are as follows. Inverse demand and marginal cost functions are given by

$$P_t = \alpha_0 + \alpha_1 Q_t + \alpha_2 Y_t + \nu_t^d,$$
$$mc_i(q_{it}, z_{it}) = \beta_0 + \beta_1 z_{it} + \nu_{it}^s.$$

To make our simulations realistic, we have chosen the following parameter values: demand-side parameters are given by $\alpha_0 = 500$, $\alpha_1 = -1.0$, and $\alpha_2 = 1.0$, and supply side parameters are given by $\beta_{0i} = 10.0\ \forall i$ and $\beta_1 = 1.0$. The observable demand shifter, Y_t, the unobservable demand innovation, ν_t^d, the observed cost shifter, z_{it}, and the unobserved cost shock ν_{it}^s are randomly drawn from normal distributions, $Y_t \overset{iid}{\sim} N(0, 100)$, $\nu_t^d \overset{iid}{\sim} N(0, 1)$, $z_{it} \overset{iid}{\sim} N(1, 4)$ and $\nu_{it}^s \overset{iid}{\sim} N(0, 0.04)$, respectively. In every period, firms operated in one of three randomly chosen regimes with $\theta_t \in \{1.0, 1.2, 1.4\}$, where $\theta = 1.0$ implies Cournot NE. To see the effects of the number of firms, N, and time periods, T, a set of pairs of (N, T) is chosen from $\{10, 20, 30\} \times \{10, 20, 30\}$.

We simulate a data set 10,000 times and each time estimate parameters of the model using 2-step optimal GMM. The GMM criterion function is constructed using two sets of moment restrictions implied by Assumption 2. In particular, demand-side moment conditions are constructed by interacting ν_t^d with (i) a constant, (ii) demand shifters, and (iii) a sum of firm-level cost shifters. Supply-side moment conditions are obtained using products of ν_{it}^s with (z_{it}, z_{-it}, Y_t) and dummy variables for each regime. The weighting matrix is assumed to have a block-diagonal structure.

Table 3: Monte Carlo Simulation for $N = 10, 30$ and $T = 10, 30$

Regime 1: True Parameter Value = 1.000						
	T=10			T=30		
	$\bar{\theta}$	Std. Dev.	ASE	$\bar{\theta}$	Std. Dev.	ASE
N = 10	1.000	0.005	0.004	1.000	0.002	0.002
N = 30	1.000	0.002	0.002	1.000	0.001	0.001
Regime 2: True Parameter Value = 1.200						
	T=10			T=30		
	$\bar{\theta}$	Std. Dev.	ASE	$\bar{\theta}$	Std. Dev.	ASE
N = 10	1.200	0.012	0.010	1.200	0.006	0.006
N = 30	1.200	0.008	0.007	1.200	0.005	0.004
Regime 3: True Parameter Value = 1.400						
	T=10			T=30		
	$\bar{\theta}$	Std. Dev.	ASE	$\bar{\theta}$	Std. Dev.	ASE
N = 10	1.400	0.019	0.016	1.400	0.010	0.010
N = 30	1.400	0.015	0.012	1.400	0.008	0.008

Note: $\bar{\theta} = \sum_{s=1}^{ns} \hat{\theta}_s$ and "Std. Dev." is defined as $\sqrt{\frac{1}{ns-1}\sum_{s=1}^{ns}(\hat{\theta}_s - \bar{\theta})}$. ASE is the average of standard errors for each simulation.

As our interest lies only in the estimates of the conduct parameter, Table 3 conveniently summarizes average estimates of θ_t, denoted by $\bar{\theta}$, standard deviation and average values of the estimated standard errors, denoted by Std. Dev. and ASE, respectively, for $(10, 30) \times (10, 30)$ sample sizes. The full set of estimation results can be found in Appendix C. In all cases, parameter estimates are precise and the standard deviations of the estimated coefficients are consistent with the mean values of the standard errors. As expected, the estimates become more accurate as the number of firms and/or the number of time-series observations increases. Monte Carlo simulations suggest that a longer panel (larger T) provides more precise parameter estimates than a wider panel (larger N). We believe that this is because an increased number of time periods contributes to both the demand- and supply-side moment conditions, whereas an increased number of firms affects only the supply-side set of moment conditions.

6 Application: The Joint Executive Committee

In order to illustrate how our method works with real data, we apply our methodology to the Joint Executive Committee (JEC) railroad cartel data from Porter (1983) and Ellison (1994). The JEC was a legal cartel that controlled freight shipments from Chicago to the Atlantic seaboard in the 1880s. The cartel was created in 1879 – that is prior to the Sherman Act of 1880. The data contains firm-level information on prices and shipment volumes for grain and flour. Moreover, information about the availability of alternative transportation routes through the Great Lakes is observed in the data, which serves as one of the demand shifters. In each year, the shipping industry could operate only several months during summer and it substituted railroad transportation. As a result, the demand for railroad transportation was substantially affected and thus it can serve as a demand shifter. A detailed description of the data can be found in Porter (1983) and Ellison (1994). It is worth noting that we provide this application primarily for illustrative purposes and the estimation results could be improved, if more detailed information were available, in particular on the individual firms' cost shifters.

Before discussing our empirical specification, it might be interesting to test whether the JEC data is consistent with the PR assumption. As mentioned in Subsection 3.3, two recent papers, Cherchye et al. (2013) and Carvajal et al. (2013), develop such a test. Unfortunately Cherchye et al. (2013) cannot be applied to our data because we do not have a continuous demand shifter to use in reduced form functions. However, we can apply the test developed by Carvajal et al. (2013) as it imposes virtually no functional form restrictions except for monotone cost functions. While we clearly understand that the monotone cost assumption is unlikely to be satisfied in our settings (this should be apparent because we employ cost shifters in our estimation), we can apply the test to two adjacent time periods where change in the cost function is unlikely to be dramatic. With this caveat in mind we briefly summarize results of the test.

The test identifies the set of marginal costs that rationalizes the data. In practice, the algorithm constructs an upper and lower bounds on the marginal costs. No shifts or twists of the demand (or cost) function are required. The test relies on the *Common Ratio property*

$$\frac{P_t - mc_{1t}}{q_{1t}} = \frac{P_t - mc_{2t}}{q_{2t}} = \cdots = \frac{P_t - mc_{Nt}}{q_{Nt}},$$

and *Co-monotone property* stated as

$$(mc_{it'} - mc_{it})(q_{it'} - q_{it}) \geq 0.$$

For conjectural variation models, the authors suggest an extension of the Common Ratio property to *Generalized Common Ratio property*,

$$\frac{P_t - mc_{1t}}{\theta_1 q_{1t}} = \frac{P_t - mc_{2t}}{\theta_2 q_{2t}} = \cdots = \frac{P_t - mc_{Nt}}{\theta_N q_{Nt}}.$$

Note that the test cannot be used to recover the level of collusion because if the property is satisfied for some θ, it would be satisfied for $\lambda\theta$ for any $\lambda > 0$.

We implemented the test by using data from $t \in [1, 327]$ to $t + 1$ for five major firms – continuous establishments that stayed in the cartel for the entire sample, which covers 328 weeks. Our results suggest that the same level of collusion cannot be rejected in 90% of cases (295 out of 327). Hence, 90% of observations in our data are consistent with PR.[16]

6.1 Empirical Specification and Estimation

Let θ, α and β denote vectors of PR parameters, demand and cost function parameters, respectively. In reality, the JEC railroad cartel used market share allocation, which

[16]It is worth noting that the test requires substantial variation in prices because larger variation will produce tighter bounds on the rationalizable levels of monotone marginal costs. In the case of JEC data, prices may stay constant over several periods thus adversely affecting the power of the test.

includes PR as a special case and can generate the similar output-level to PR, as discussed in Section 4.1. Moreover, each firm in the JEC cartel held at least 10% of the market share, which makes us believe that PR is a sensible assumption. Therefore, we assume that the cartel members of the JEC railroad use PR collusive technology here with parameter θ_t. Assume that the per-period profits of the firms within the JEC are given by

$$\pi(q_{it}, z_{it}, \nu_t^d, \nu_{it}^s; \theta, \alpha, \beta) = P(\theta_t Q_t, Y_t, \nu_t^d; \alpha)\theta_t q_{it} - C_i(\theta_t q_{it}, z_{it}, \nu_{it}^s; \beta),$$

where $Q_t = \sum_{i=1}^{N_t} q_{it}$, N_t is the number of firms in period t, Y_t is a vector of observed demand shifters, z_{it} is a vector of individual cost shifters, and (ν_t^d, ν_{it}^s) is a pair of demand and supply-side shocks, respectively. We assume the following functional forms:

$$\begin{aligned} P(Q_t, Y_t, \nu_t^d; \alpha) &= \alpha_0 + \alpha_1 Q_t + \alpha_2 Y_t + \nu_t^d, \\ C_i(q_{it}, z_{it}, \nu_{it}^s; \beta) &= F_i + (\beta_{0i} + \beta_1 z_{it} + \nu_{it}^s)\, q_{it}. \end{aligned}$$

When reporting estimation results, the case where $\beta_{0i} \neq \beta_{0j}$ is referenced as "fixed effect" (FE), and the restriction of $\beta_{0i} = \beta_0$, $\forall i$ is denoted as "levels" (LE). In the data, we observe shipment volumes for both grain and flour. Because of the potential (dis-)economies of scope, we define flour shipments to be an observable cost shifter z_{it} when evaluating collusion in the market for grain.[17]

Under our assumption of PR collusive technology, static Cournot first-order conditions are given by (3). In order to estimate parameters of the model, we estimate the demand and supply relations jointly. In particular, for any given vector of parameters, we isolate demand and supply shocks using the following system of equations:

$$\begin{cases} \nu_t^d = P_t - \alpha_0 - \alpha_1 Q_t - \alpha_2 Y_t, \\ \nu_{it}^s = \beta_{0i} + \beta_1 z_{it} - (\alpha_1 \theta_t q_{it}^{PR} + \alpha_1 (\theta_t - 1)\, Q_t^{PR} + P_t^{PR}). \end{cases}$$

Our estimation is based on the orthogonality restrictions following from the conditional independence assumptions,

$$\mathrm{E}[\nu_t^d | Y_t, Z_t] = \mathrm{E}[\nu_{it}^s | Y_t, z_{it}, z_{-it}] = 0,$$

where $Z_t = \sum_{i=1}^{N_t} z_{it}$ and $z_{-it} = \sum_{j \neq i} z_{jt}$. In practice, we interact z_{-it} with a set of dummy variables, one for each of the collusive regimes. We construct sample analogs of the population moment conditions, $G_d^N(Y_t, Z_t; \alpha)$ and $G_s^N(Y_t, z_{it}, z_{-it}; \alpha, \beta, \theta)$:

$$G^N(Y_t, Z_t, z_{it}, z_{-it}; \alpha, \beta, \theta) = \begin{bmatrix} G_d^N(Y_t, Z_t; \alpha) \\ G_s^N(Y_t, z_{it}, z_{-it}; \alpha, \beta, \theta) \end{bmatrix},$$

and estimate parameters using the following GMM criterion function:

$$(\alpha^*, \beta^*, \theta^*) = \underset{(\alpha,\beta,\theta)}{\arg\min} \left\{ G^N (Y_t, Z_t, z_{it}, z_{-it}; \alpha, \beta, \theta)' \cdot W \cdot G^N (Y_t, Z_t, z_{it}, z_{-it}; \alpha, \beta, \theta) \right\},$$

with a block-diagonal weighting matrix W.[18]

[17]We admit potential caveats related to the assumption of exogenous flour shipment volumes; however, available data do not provide us with better instrumental variables.

[18]In the first stage, the weighting matrix is obtained as the inner product of the instrumental variables matrices, which would be optimal for linear model. In the second (and consecutive) stage(s), we compute the optimal weighting matrix using empirical variance of the moment conditions.

6.2 Estimation Results

In 6.2.1, for the sake of increasing accuracy of the estimates for the degree of collusion, we first use all available information, including the *cartel indicator*. As mentioned earlier, the cartel was formed prior to the passage of the Sherman Act and the cartel was legal. Thus, the industry publication documented whether they successfully colluded or failed to do so for each month. This information is very helpful to divide the periods into the non-collusive and collusive periods, which should generate different values of θ_t. On the other hand, the next subsection 6.2.2 assumes that econometricians do not observe whether or not the cartel members engaged in collusion and estimate the degree of collusion to examine whether our methodology can detect the cartel.

6.2.1 Overall results

As our main focus is again on the degree of collusion, Table 4 lists the inverse of the estimated degree of collusion ($1/\hat{\theta}$) in the FE specification. The full set of estimation results are documented in Appendix A. Parameter estimates obtained from the LE specification are similar and can be also found in Tables 7, 9, 11, and 13 in Appendix A.

Table 4: Summary of the Monopolization Parameter Estimates using FE Specifications

	Model (i) Table 8	Model (ii) Table 10		Model (iii) Table 12		Model (iv) Table 14			
		C = 0	C = 1	C=0	C=1	C=0	C=0	C=1	C=1
N						L=0	L=1	L=0	L=1
5				–	0.40	–	–	0.44	0.54
6	1.31	1.51	0.71	1.46	0.55	0.91	–	0.65	0.66
7				0.68	0.64	–	0.93	0.75	0.82
8				1.36	0.81	1.42	–	0.92	1.17

Note: N, C, and L denote the number of firms, the cartel indicator, and the Great Lake operation dummy, respectively. Parameter estimates for the cases (N=6,C=0,L=1) and (N=8,C=0,L=1) are not statistically significant at any reasonable significance levels and therefore are not reported.

Model (i) assumes that θ is constant for the entire sample period, regardless of the number of firms or other observables (see the results in the first column in Table 4). The detailed estimation results for this specification are documented in Table 8. According to the results, firms produced on average 31% more output than they would produce under the Cournot scenario.[19] Similarly, Model (ii) (in the second and third columns) assumes that the cartel is maintained at the same level of θ_1 during all collusive periods and that the firms produce ($1/\theta_0$-1)% more in competitive periods than they would do in Cournot. The estimates imply that, in the collusive period, the output was reduced to about 71% of hypothetical Cournot quantity. During price wars, on the other hand, firms produced 51% more than they would do in Cournot.

[19]Interestingly, this result is consistent with the findings in Porter (1983), where for constant θ the firms' behavior in collusive periods was roughly consistent with Cournot equilibrium (pp. 309-310).

Estimation results become more plausible when the conduct regimes are defined as unique combinations of the number of firms and the indicator of collusion, because it is possible that these firms would target a different level of reductions, depending on the number of member firms. The results from Model (iii) under the fourth and fifth columns in Table 8 indicate that whenever the cartel indicator is equal to one, these firms produced 40% to 81% of the Cournot quantity. In the meantime, when the cartel broke down and the firms were involved in price wars, firms produce more than they would do in the Cournot equilibrium, except for the case of seven firms. Interestingly, the estimated degree of monopolization monotonically declines in the number of firms, which is consistent with the presumption that larger cartels are less sustainable.

It is natural to believe that the firms collude on different levels depending on the existence of a competitor to the cartel, the Great Lakes, and thus we further use finer categorization in Model (iv). This specification assumes that the degree of monopolization depends on the number of firms, collusive indicator and the state of demand, i.e., whether the Great Lakes were open for navigation. The estimation results for this case are summarized in the last four columns of Table 4. Our estimates suggest that the degree of monopolization declines in the number of firms and is generally lower at lower states of demand. The latter speaks against the counter-cyclical cartel pricing patterns as in the model by Rotemberg and Saloner (1986). According to their predictions, a cartel would reduce the degree of monopolization at high states of demand to reduce incentives for cheating. Instead, we find that when facing competition from the Great Lakes transportation routes, JEC members reduce their level of collusion.[20]

Since estimated parameter values imply a relatively high degree of collusion compared with a hypothetical Cournot equilibrium, we conducted the following experiment. Given our estimates of the cost function parameters, we calculated optimal monopoly and perfectly competitive quantity levels for each firm. The smallest optimal monopoly output among the colluding firms defines a lower bound on the total quantity of the cartel, while the largest (Bertrand) competitive quantity among the participating firms would impose an upper bound consistent with rational behavior. Figure 1 summarizes the results for the firm fixed-effect specification. Figure 3 in Appendix A presents same statistics for the specification in levels. As is apparent from the top panel of the figure, in most cases observed quantities stay in-between the upper and lower bounds. In particular, for the FE specification, in 202 out of 328 weeks (62%) JEC produces more than the standalone monopoly quantity for the least efficient firm in a given week, and for the specification in levels this occurs 205 out of 328 times. The same observation can be made when output levels are averaged for each of the potential collusive regimes (bottom panel).

To further confirm our estimation results, the own price elasticity of the demand is calculated and presented in Table 5. As expected, the degree of monopolization is positively related to the absolute value of price elasticity, i.e., the higher the degree of monopolization the larger price elasticity of demand with a correlation coefficient of 0.77. On average, during collusive regimes the price elasticity of demand is -5.11, which is almost twice the elasticity during non-collusive regimes of -2.75.

[20] Again, this finding is similar to the one in Ellison (1994), where no evidence of the countercyclical pricing was found.

Figure 1: Upper and Low Bounds on the Total Output by Week (top panel) and by Regime (bottom panel) for Estimates with Firm FE's (Table 14)

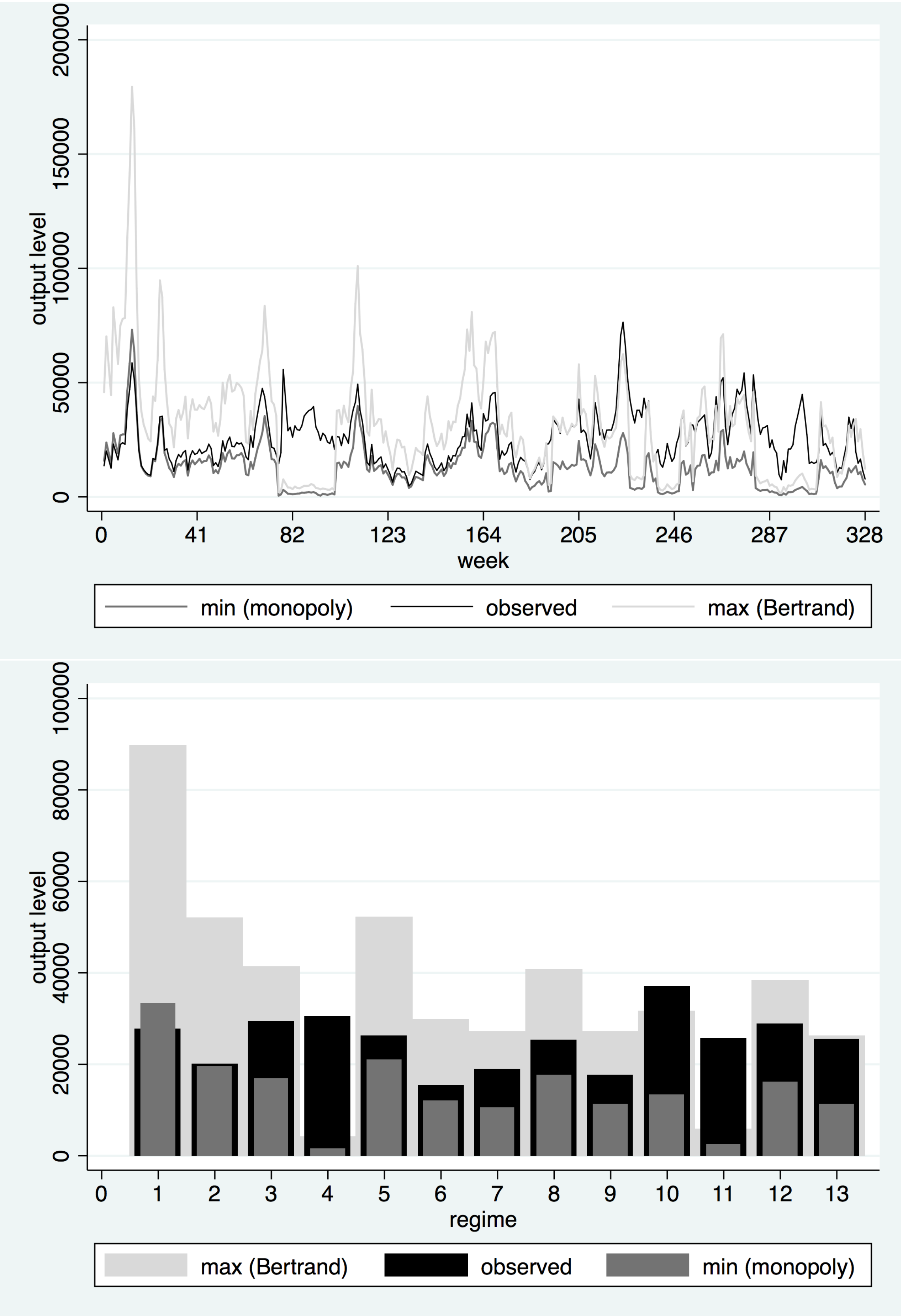


Table 5: Estimated Parameters versus Price Elasticity across Regimes

Regime	θ	% of Cournot	$p-$Elasticity
N=5, C=1, L=0	2.29	0.44	-6.00
N=5, C=1, L=1	1.85	0.54	-6.34
N=6, C=0, L=0	1.10	0.91	-3.17
N=6, C=1, L=0	1.54	0.65	-4.31
N=6, C=1, L=1	1.51	0.66	-6.75
N=7, C=0, L=1	1.07	0.93	-4.52
N=7, C=1, L=0	1.33	0.75	-4.98
N=7, C=1, L=1	1.21	0.82	-5.23
N=8, C=0, L=0	0.70	1.42	-2.72
N=8, C=1, L=0	1.09	0.92	-3.28
N=8, C=1, L=1	0.85	1.17	-3.88
Avg.	1.32	0.84	-4.65

Lastly, we conducted several robustness checks of our specifications. First, we excluded observations with six and eight firms when the cartel indicator is zero and the Great Lakes are open for navigation (see the note in Table 4). Estimation results do not change qualitatively, as can be seen from Tables 15 and 16 in Appendix A. Second, we estimated the model using two alternative specifications for the cost function. Namely, in Table 17 we report estimation results where the marginal cost function is given by either

$$mc_i(q_{it}, z_{it}; \beta) = \beta_{0i} + \beta_1 z_{it} + \beta_2 q_{it} + \nu^s_{it}$$

or

$$mc_i(q_{it}, z_{it}; \beta) = \beta_{0i} + \beta_1 z_{it} + (\beta_2 + 1) q_{it}^{\beta_2} + \nu^s_{it}.$$

Columns 2 and 4 of Table 17 summarize the results. It turns out that including a linear or non-linear term in quantity does not effect our estimates of the conduct parameter substantially. Besides, the coefficients on the own quantity variable in the cost functions are statistically not different from zero at any reasonable significance level. Unfortunately, we do not have other instrumental variables to explore much richer specifications.

6.2.2 Absence of the cartel indicator

So far we have used the cartel indicator, reported in the data, to tabulate regimes with a constant level of collusion. In practice, however, econometricians or competition authorities do not know whether or not firms collude. Thus, we must be able to define regimes relying only on observed variation in the output levels and market shares, not the cartel indicator. Therefore, without using the cartel indicator, we conduct two final empirical exercises: (i) we create our own index describing potential regimes of JEC operations and estimate the model with the new index, and (ii) we estimate the model at the monthly-level assuming that θ_t is constant within a month.

Our new indicator In order to create our own index of collusion, we inspect the data for candidate collusive periods. Our criteria require a stable distribution of market shares and reduction in output relative to the adjacent time intervals. To test the stability of market shares, we use a t-test for difference in means, which accounts for serial correlation. In particular, the test compares sub-intervals within a given interval.[21] We find nine such intervals with 662 observations in total. Table 6 reports parameter estimates for each of the collusive regimes with full estimation results listed in Table 18 in Appendix A.

Table 6: Estimation Results for 9 Selected Periods satisfying PR Assumption, 662 obs.

Regimes.	LE Specification			FE Specification			Cartel Index
	1^{st} Est.	2^{nd} Est.	% Reduction	1^{st} Est.	2^{nd} Est.	% Reduction	
θ_1 (N=6, 68-75)	1.548 (0.150)	1.531 (0.144)	0.65	1.534 (0.148)	1.522 (0.145)	0.66	0.71
θ_2 (N=6, 116-131)	1.447 (0.138)	1.418 (0.131)	0.71	1.423 (0.138)	1.394 (0.132)	0.72	1.00
θ_3 (N=6, 131-166)	1.650 (0.165)	1.616 (0.156)	0.62	1.630 (0.167)	1.600 (0.160)	0.63	0.97
θ_4 (N=7, 171-181, 324)	1.583 (0.168)	1.545 (0.159)	0.65	1.572 (0.171)	1.536 (0.163)	0.65	0.83
θ_5 (N=8, 184-189)	1.694 (0.185)	1.651 (0.174)	0.61	1.672 (0.187)	1.632 (0.178)	0.61	1.00
θ_6 (N=8, 191-196)	1.025 (0.058)	1.013 (0.056)	0.99	1.016 (0.058)	1.004 (0.056)	1.00	0.67
θ_7 (N=8, 254-259)	1.200 (0.070)	1.184 (0.067)	0.84	1.186 (0.072)	1.170 (0.069)	0.85	0.83
θ_8 (N=8, 258-263)	1.109 (0.063)	1.074 (0.058)	0.93	1.091 (0.065)	1.051 (0.061)	0.95	0.83
θ_9 (N=8, 313-318)	1.212 (0.127)	1.231 (0.128)	0.81	1.192 (0.128)	1.220 (0.130)	0.82	0.67

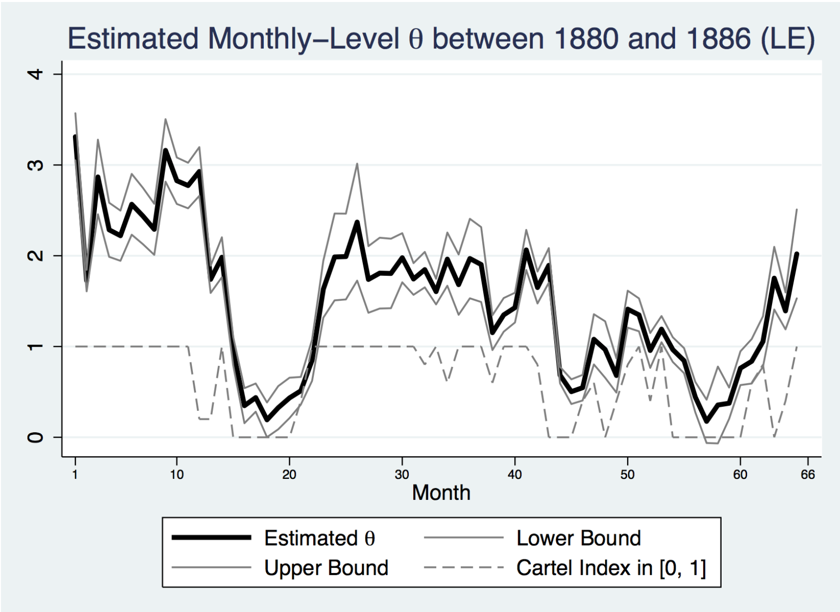
Note: 1^{st} and 2^{nd} Est. report 1st and 2nd stage GMM esimates. % Reduction demonstrates the how much firms reduce their output compared to the Cournot outcomes. Cartel Index is calculated the average value of the cartel indicator during the sample periods.

For all regimes our estimates suggest at least some degree of collusion with the output levels below static Cournot NE, as the percent reduction is almost always below one for both LE and FE specifications. To check the validity of our method, we create a Cartel Index, the average value of the reported cartel indicator during each period, expecting that the percent reduction and the Cartel Index are negatively correlated. If the Cartel Index is zero, for example, we must expect that the firms compete severely, yielding close to the Cournot output. The correlation coefficient between the percent reduction and the Cartel Index for the FE specification is -0.56, which indicates that the estimates are likely to be able to detect the existence of the cartel.

[21] We assumed AR(1) process for serial correlation and computed equivalent sample size using approximation $\hat{n}^e = n\,\frac{1-\hat{\rho}_1}{1+\hat{\rho}_1}$. We then computed the statistic for $H_0 : \mathrm{E}[HHI_1] = \mathrm{E}[HHI_2]$ using $t = \frac{\mu_1 - \mu_2}{\left(\sigma_1^2/n_1^e + \sigma_2^2/n_2^e\right)^2}$ with significance level 0.05.

Monthly-level As a last step, we estimate monthly-level θ_t to examine whether our methodology can detect the cartel for each month.[22] Figure 2 plots the estimated monthly-level θ_t. The black solid line shows the estimated value, whereas the gray solid lines indicate the confidence interval. To examine the performance of our methodology, the gray dashed line records the Cartel Index, which is an average value of the cartel indicator within a month. Whenever our estimated θ's go below one, the firms indeed failed to collude, indicated by the Cartel Index falling below one. Therefore, this observation validates our methodology.

Figure 2: Monthly Level θ



7 Conclusions

In this paper we develop a method to estimate the time-varying degree of industry monopolization. The methodology does not impose any restrictions on the objective function of colluding firms. Instead, we impose an assumption on how collusion is implemented. We believe that our method has several advantages over the traditional empirical literature on collusion. First of all, proportional reduction would be a natural way to implement collusion with symmetric firms. Therefore, most of the earlier literature on estimating conduct parameters can be viewed as a special case of our model. Asymmetricity in the firms' cost functions provides useful variation that can be utilized to identify the degree of industry monopolization conditional on observing firm-level cost shifters. Second, the parameter measuring the degree of industry monopolization is a continuous measure relating observed levels of output to the hypothetical stage game Nash equilibrium. As

[22]Although our methodology allows us to estimate θ_t for weekly-level in principle, the JEC had a small number of firms, between five and eight, depending on the time period. Estimating one parameter (weekly-level θ_t), relying on only five to eight observations, might not yield statistically significant results. Therefore, we estimate the model with monthly-level θ_t for stacking at least 20 observations for estimating θ_t for each period.

a result, it has a simple interpretation as the percentage of output reduction relative to a well-defined competitive equilibrium. Third, the fact that we do not require explicit assumptions about the objective function of the cartel allows us to accommodate a wide range of fairly complex models of collusion as long as the proportional reduction assumption is satisfied. The latter fact can be empirically tested. Fourth, we show that the variation in output levels across asymmetric firms allows time-varying estimates of the degree of monopolization. This way one can address the critique by Corts (1999) of the conjectural variation literature when the industry conduct is endogenous to the changing demand and supply conditions. Finally, we believe that the simplicity of the method is appealing to industry practitioners because estimation can be done using standard statistical software. Perhaps the best application of our framework would be at the stage of pre-screening procedures conducted by antitrust authorities when they are deciding whether to thoroughly investigate a case or dismiss it.

Monte Carlo simulations illustrate finite-sample properties of the parameter estimates and show that our method performs well even with medium sample sizes consisting of 100 to 300 data points. Thus, the parameter of interest can be estimated from relatively short panels of firm-level observations. To further investigate the practicality of our method, we use the Joint Executive Committee railroad cartel from the 19^{th} century. Our analysis using the available cartel indicator demonstrates that it strongly correlates with the estimated degree of collusion. Finally, we estimate the time-varying degree of monopolization at a monthly level. Estimation results imply substantial variability in the degree of collusion over time, with the output levels during price wars sometimes exceeding quantities predicted by the Cournot equilibrium.

References

Bajari, P. and Ye, L. (2003). Deciding between competition and collusion. *The Review of Economics and Statistics*, 85(4):971–989.

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

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

Carvajal, A., Deb, R., Fenske, J., and Quah, J. (2013). Revealed preference tests of the cournot model. *Econometrica*, 81(6):2351–2379.

Cave, J. and Salant, S. W. (1995). Cartel quotas under majority rules. *The American Economic Review*, 85(1):82–102.

Cherchye, L., Demuynck, T., and Rock, B. D. (2013). The empirical content of Cournot competition. *Journal of Economic Theory*, 148(4):1552–1581.

Ciliberto, F. and Williams, J. W. (2014). Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry. *The RAND Journal of Economics*, 45(4):764–791.

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

Ellison, G. (1994). Theories of cartel stability and the joint executive committee. *RAND Journal of Economics*, 25:37–57.

Fershtman, C. and Gandal, N. (1994). Disadvantageous semicollusion. *International Journal of Industrial Organization*, 12:141–154.

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

Fink, N., Schmidt-Dengler, P., Stahl, K., and Zulehner, C. (2015). Registered cartels in Austria - an overview. ZEW discussion paper 14058.

Genesove, D. and Mullin, W. (1998). Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914. *The RAND Journal of Economics*, 29(2):355–377.

Green, E. and Porter, R. (1984). Noncooperative collusion under imperfect price information. *Econometrica*, 52:87–100.

Harrington, J., Huschelrath, K., Laitenberger, U., and Smuda, F. (2014). The discontent cartel member and cartel collapse: the case of the German cement cartel. ZEW Discussion Paper No. 14-084.

Hyytinen, A., Steen, F., and Toivanen, O. (2017). Anatomy of cartel contracts. Working paper.

Kreps, D. and Scheinkman, J. (1983). Quantity precommitment and Bertrand competition yield Cournot outcomes. *The Bell Journal of Economics*, 14(2):326–337.

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

Marshall, R. and Marx, L. (2008). Explicit collusion and market share allocation. Unpublished manuscript, Duke University.

Matsui, A. (1989). Consumer-benefited cartels under strategic capital investment competition. *International Journal of Industrial Organization*, 7(4):451–470.

Matzkin, R. (2012). Nonparametric identification. Department of Economics, UCLA.

Nevo, A. (1998). Identification of the oligopoly solution concept in a differentiated-products industry. *Economics Letters*, 59:391–395.

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

Puller, S. (2009). Estimation of competitive conduct when firms are efficiently colluding: addressing the Corts critique. *Applied Economics Letters*, 16:1497–1500.

Reiss, P. and Wolak, F. (2007). Structural econometric modeling: Rationales and examples from industrial organization. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 6A, pages 4277–4415. Elsevier B.V. Part A.

Roller, L.-H. and Steen, F. (2006). On the workings of a cartel: evidence from the Norwegian cement industry. *The American Economic Review*, 96(1):321–338.

Rotemberg, J. and Saloner, G. (1986). A supergame-theoretic model of price wars during booms. *The American Economic Review*, 76:390–407.

Schmalensee, R. (1987). Competitive advantage and collusive optima. *International Journal of Industrial Organization*, 5:351–367.

Stigler, G. (1964). A theory of oligopoly. *The Journal of Political Economy*, 72(1):44–61.

Wolfram, C. (1999). Measuring duopoly power in the British electricity spot market. *The American Economic Review*, 89(4):805–826.

Yano, M. (1989). Voluntary export restraints and expectations: an analysis of export quotas in oligopolistic markets. *International Economic Review*, 30(4):707–723.

Appendix A Estimation Results

Table 7: Constant Conduct Parameter, $mc_i = \beta_0 + \beta_1 z_{it} + \nu_{it}$ (LE)

Param.	1^{st} Stage	%Cournot	2^{nd} Stage	%Cournot	Cont.-Update	%Cournot
α_0	35847.846		36107.822		36100.903	
	(1846.092)		(1851.162)		(1851.018)	
α_1	-0.294		-0.302		-0.302	
	(0.052)		(0.052)		(0.052)	
α_2	-6510.624		-6604.251		-6601.853	
	(893.286)		(895.631)		(895.564)	
θ	0.677	1.48	0.763	1.31	0.761	1.31
	(0.078)		(0.069)		(0.069)	
β_0	25472.798		24388.318		24333.122	
	(452.696)		(461.001)		(461.278)	
β_1	0.381		0.371		0.367	
	(0.160)		(0.162)		(0.162)	
$f-val$	2794.6981		277.2121		268.8319	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The column labeled % Cournot denotes how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 8: Constant Conduct Parameter, $mc_i = \beta_{0i} + \beta_1 z_{it} + \nu_{it}$ (FE)

Param.	1^{st} Stage	%Cournot	2^{nd} Stage	%Cournot	Cont.-Update	%Cournot
α_0	35856.839		36429.472		36413.757	
	(1846.005)		(1857.807)		(1857.448)	
α_1	-0.294		-0.311		-0.311	
	(0.052)		(.052)		(0.052)	
α_2	-6512.883		-6719.050		-6713.762	
	(893.302)		(898.722)		(898.556)	
θ	0.666	1.50	0.749	1.34	0.722	1.39
	(0.083)		(0.074)		(0.077)	
β_1	0.496		0.587		0.815	
	(0.216)		(0.224)		(0.238)	
$f-val$	2735.5657		281.4734		271.1445	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The column labeled % Cournot denotes how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 9: Regimes Defined by the Cartel Indicator only, $mc_i = \beta_0 + \beta_1 z_{it} + \nu_{it}$ (LE)

Param.	1^{st} Stage	%Cournot	2^{nd} Stage	%Cournot	Cont.-Update	%Cournot
α_0	35852.778		36396.104		36374.390	
	(1845.908)		(1857.047)		(1856.558)	
α_1	-0.294		-0.310		-0.310	
	(0.052)		(0.052)		(0.052)	
α_2	-6511.863		-6707.138		-6699.672	
	(893.259)		(898.378)		(898.154)	
θ_0 (C=0)	0.637	1.57	0.663	1.51	0.659	1.52
	(0.076)		(0.070)		(0.071)	
θ_1 (C=1)	1.398	0.72	1.409	0.71	1.424	0.70
	(0.099)		(0.096)		(0.098)	
β_0	22669.676		21900.779		21762.925	
	(413.988)		(420.671)		(423.670)	
β_1	0.051		-0.016		-0.054	
	(0.146)		(0.147)		(0.147)	
$f - val$	1870.773		238.657		227.135	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The column labeled % Cournot denotes how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 10: Regimes Defined by the Cartel Indicator only, $mc_i = \beta_{0i} + \beta_1 z_{it} + \nu_{it}$ (FE)

Param.	1^{st} Stage	%Cournot	2^{nd} Stage	%Cournot	Cont.-Update	%Cournot
α_0	35855.274		36602.543		36597.780	
	(1845.726)		(1861.554)		(1861.425)	
α_1	-0.294		-0.316		-0.316	
	(0.052)		(0.052)		(0.052)	
α_2	-6512.490		-6780.817		-6779.655	
	(893.222)		(900.470)		(900.409)	
θ_0 (C=0)	0.625	1.60	0.670	1.49	0.659	1.52
	(0.078)		(0.069)		(0.070)	
θ_1 (C=1)	1.371	0.73	1.372	0.73	1.363	0.73
	(0.108)		(0.101)		(0.100)	
β_1	0.156		0.196		0.281	
	(0.193)		(0.197)		(0.202)	
$f - val$	1844.615		270.430		264.910	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The column labeled % Cournot denotes how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 11: Regimes Defined by N and Cartel Indicator, $mc_i = \beta_0 + \beta_1 z_{it} + \nu_{it}$ (LE)

Param.	1^{st} Stage	%	2^{nd} Stage	%	Cont.-Update	%
α_0	35834.452		35546.398		35545.746	
	(1845.755)		(1840.689)		(1840.681)	
α_1	-0.294		-0.285		-0.285	
	(0.052)		(0.052)		(0.052)	
α_2	-6507.259		-6403.890		-6403.764	
	(893.140)		(890.764)		(890.759)	
θ_1 (N=5, C=1)	2.550	0.39	2.346	0.43	2.304	0.43
	(0.322)		(0.291)		(0.283)	
θ_2 (N=6, C=0)	0.907	1.10	0.629	1.59	0.401	2.49
	(0.086)		(0.106)		(0.135)	
θ_3 (N=6, C=1)	2.023	0.49	1.840	0.54	1.817	0.55
	(0.208)		(0.183)		(0.180)	
θ_4 (N=7, C=0)	1.648	0.61	1.565	0.64	1.516	0.66
	(0.148)		(0.138)		(0.131)	
θ_5 (N=7, C=1)	1.686	0.59	1.651	0.61	1.619	0.62
	(0.145		(0.142)		(0.137)	
θ_6 (N=8, C=0)	0.876	1.14	0.772	1.30	0.743	1.35
	(0.048)		(0.057)		(0.061)	
θ_7 (N=8, C=1)	1.364	0.73	1.297	0.77	1.262	0.79
	(0.092)		(0.085)		(0.080)	
β_0	20294.021		20954.043		21236.366	
	(412.154)		(405.708)		(411.161)	
β_1	-0.321		-0.297		-0.263	
	(0.152)		(0.149)		(0.151)	
$f - val$	1014.420		163.301		164.366	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 12: Regimes Defined by N and Cartel Indicator, $mc_i = \beta_{0i} + \beta_1 z_{it} + \nu_{it}$ (FE)

Param.	1^{st} Stage	%	2^{nd} Stage	%	Cont.-Update	%
α_0	35837.177		35657.267		35645.881	
	(1845.907)		(1842.674)		(1842.474)	
α_1	-0.294		-0.288		-0.288	
	(0.052)		(0.052)		(0.052)	
α_2	-6507.943		-6443.455		-6439.437	
	(893.192)		(891.680)		(891.586)	
θ_1 (N=5, C=1)	2.471	0.40	2.488	0.40	2.511	0.40
	(0.311)		(0.319)		(0.324)	
θ_2 (N=6, C=0)	0.817	1.22	0.687	1.46	0.589	1.70
	(0.088)		(0.096)		(0.105)	
θ_3 (N=6, C=1)	1.904	0.53	1.802	0.55	1.778	0.56
	(0.201)		(0.187)		(0.184)	
θ_4 (N=7, C=0)	1.546	0.65	1.467	0.68	1.431	0.70
	(0.144)		(0.134)		(0.130)	
θ_5 (N=7, C=1)	1.595	0.63	1.572	0.64	1.551	0.64
	(0.139)		(0.138)		(0.134)	
θ_6 (N=8, C=0)	0.799	1.25	0.738	1.36	0.723	1.38
	(0.057)		(0.062)		(0.063)	
θ_7 (N=8, C=1)	1.274	0.78	1.240	0.81	1.221	0.82
	(0.089)		(0.086)		(0.084)	
β_1	0.011		-0.003		0.006	
	(0.189)		(0.186)		(0.186)	
$f - val$	1003.630		203.199		216.512	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 13: Regimes Defined by N, Cartel Indicator and State of Demand, $mc_i = \beta_0 + \beta_1 z_{it} + \nu_{it}$ (LE)

Param.	1^{st} Stage	%	2^{nd} Stage	%	Cont.-Update	%
α_0	35839.113		35739.791		35741.825	
	(1845.186)		(1843.389)		(1843.426)	
α_1	-0.294		-0.291		-0.291	
	(0.052)		(0.052)		(0.052)	
α_2	-6508.430		-6472.906		-6473.643	
	(893.011)		(892.168)		(892.186)	
θ_1 (N=5, C=1, L=0)	2.257	0.44	2.348	0.43	2.354	0.42
	(0.279)		(0.298)		(0.299)	
θ_2 (N=5, C=1, L=1)	1.859	0.54	1.923	0.52	1.924	0.52
	(0.185)		(0.198)		(0.198)	
θ_3 (N=6, C=0, L=0)	1.114	0.90	1.155	0.87	1.155	0.87
	(0.114)		(0.117)		(0.117)	
θ_4 (N=6, C=0, L=1)	0.137	7.30	0.165	6.06	0.166	6.02
	(0.152)		(0.149)		(0.149)	
θ_5 (N=6, C=1, L=0)	1.574	0.64	1.616	0.62	1.617	0.62
	(0.132)		(0.140)		(0.140)	
θ_6 (N=6, C=1, L=1)	1.535	0.65	1.604	0.62	1.606	0.62
	(0.136)		(0.147)		(0.147)	
θ_7 (N=7, C=0, L=1)	1.091	0.92	1.143	0.87	1.145	0.87
	(0.077)		(0.082)		(0.082)	
θ_8 (N=7, C=1, L=0)	1.356	0.74	1.393	0.72	1.394	0.72
	(0.088)		(0.094)		(0.094)	
θ_9 (N=7, C=1, L=1)	1.231	0.81	1.284	0.78	1.286	0.78
	(0.100)		(0.106)		(0.106)	
θ_{10} (N=8, C=0, L=0)	0.736	1.36	0.752	1.33	0.752	1.33
	(0.057)		(0.056)		(0.056)	
θ_{11} (N=8, C=0, L=1)	0.211	4.74	0.240	4.17	0.242	4.13
	(0.141)		(0.138)		(0.137)	
θ_{12} (N=8, C=1, L=0)	1.125	0.89	1.154	0.87	1.155	0.87
	(0.062)		(0.065)		(0.065)	
θ_{13} (N=8, C=1, L=1)	0.869	1.15	0.904	1.11	0.905	1.10
	(0.057)		(0.056)		(0.056)	
β_0	23810.582		23529.588		23519.281	
	(409.084)		(410.955)		(411.084)	
β_1	-0.294		-0.338		-0.339	
	(0.130)		(0.132)		(0.132)	
$f-val$	8.251		3.376		3.311	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 14: Regimes Defined by N, Cartel Indicator and State of Demand, $mc_i = \beta_{i0} + \beta_1 z_{it} + \nu_{it}$ (FE)

Param.	1^{st} Stage	%	2^{nd} Stage	%	Cont.-Update	%
α_0	35839.791		35831.157		35831.293	
	(1845.995)		(1845.835)		(1845.838)	
α_1	-0.294		-0.294		-0.294	
	(0.052)		(0.052)		(0.052)	
α_2	-6508.600		-6505.512		-6505.561	
	(893.227)		(893.152)		(893.153)	
θ_1 (N=5, C=1, L=0)	2.212	0.45	2.287	0.44	2.294	0.44
	(0.274)		(0.288)		(0.290)	
θ_2 (N=5, C=1, L=1)	1.800	0.56	1.847	0.54	1.849	0.54
	(0.184)		(0.192)		(0.193)	
θ_3 (N=6, C=0, L=0)	1.056	0.95	1.095	0.91	1.094	0.91
	(0.112)		(0.114)		(0.114)	
θ_4 (N=6, C=0, L=1)	0.089	11.24	0.118	8.47	0.119	8.40
	(0.155)		(0.151)		(0.150)	
θ_5 (N=6, C=1, L=0)	1.508	0.66	1.540	0.65	1.541	0.65
	(0.135)		(0.140)		(0.140)	
θ_6 (N=6, C=1, L=1)	1.456	0.69	1.509	0.66	1.511	0.66
	(0.140)		(0.147)		(0.147)	
θ_7 (N=7, C=0, L=1)	1.033	0.97	1.074	0.93	1.075	0.93
	(0.082)		(0.085)		(0.085)	
θ_8 (N=7, C=1, L=0)	1.306	0.77	1.332	0.75	1.333	0.75
	(0.090)		(0.094)		(0.094)	
θ_9 (N=7, C=1, L=1)	1.173	0.85	1.213	0.82	1.215	0.82
	(0.103)		(0.107)		(0.107)	
θ_{10} (N=8, C=0, L=0)	0.690	1.45	0.704	1.42	0.705	1.42
	(0.062)		(0.061)		(0.061)	
θ_{11} (N=8, C=0, L=1)	0.166	6.02	0.195	5.13	0.197	5.08
	(0.143)		(0.139)		(0.139)	
θ_{12} (N=8, C=1, L=0)	1.068	0.94	1.090	0.92	1.091	0.92
	(0.067)		(0.069)		(0.069)	
θ_{13} (N=8, C=1, L=1)	0.823	1.22	0.852	1.17	0.853	1.17
	(0.060)		(0.060)		(0.060)	
c_1	-0.106		-0.143		-0.143	
	(0.163)		(0.164)		(0.164)	
$f - val$	6.659		2.842		2.798	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 15: Reduced Sample, $mc_i = \beta_0 + \beta_1 z_{it} + \nu_{it}$ (LE)

Param.	1^{st} Stage	%	2^{nd} Stage	%	Cont.-Update	%
α_0	33652.344		33480.469		33485.943	
	(2085.344)		(2084.071)		(2084.114)	
α_1	-0.220		-0.215		-0.215	
	(0.061)		(0.061)		(0.061)	
α_2	-3198.509		-3125.633		-3128.026	
	(1050.390)		(1048.918)		(1048.963)	
θ_1 (N=5, C=1, L=0)	2.678	0.37	2.846	0.35	2.853	0.35
	(0.525)		(0.585)		(0.586)	
θ_2 (N=5, C=1, L=1)	2.122	0.47	2.265	0.44	2.265	0.44
	(0.349)		(0.398)		(0.397)	
θ_3 (N=6, C=0, L=0)	1.139	0.88	1.221	0.82	1.221	0.82
	(0.160)		(0.174)		(0.174)	
θ_4 (N=6, C=1, L=0)	1.747	0.57	1.842	0.54	1.843	0.54
	(0.243)		(0.274)		(0.274)	
θ_5 (N=6, C=1, L=1)	1.661	0.60	1.801	0.56	1.803	0.55
	(0.235)		(0.274)		(0.274)	
θ_6 (N=7, C=0, L=1)	1.080	0.93	1.191	0.84	1.193	0.84
	(0.110)		(0.127)		(0.127)	
θ_7 (N=7, C=1, L=0)	1.463	0.68	1.542	0.65	1.543	0.65
	(0.159)		(0.184)		(0.184)	
θ_8 (N=7, C=1, L=1)	1.265	0.79	1.378	0.73	1.379	0.73
	(0.151)		(0.176)		(0.176)	
θ_9 (N=8, C=0, L=0)	0.639	1.56	0.676	1.48	0.677	1.48
	(0.106)		(0.100)		(0.100)	
θ_{10} (N=8, C=1, L=0)	1.145	0.87	1.212	0.83	1.213	0.82
	(0.093)		(0.108)		(0.108)	
θ_{11} (N=8, C=1, L=1)	0.804	1.24	0.874	1.14	0.875	1.14
	(0.091)		(0.087)		(0.086)	
β_0	24175.641		23713.288		23705.834	
	(455.280)		(461.587)		(461.803)	
β_1	-0.249		-0.291		-0.291	
	(0.139)		(0.142)		(0.142)	
$f - val$	9.710		3.550		3.439	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 16: Reduced Sample, $mc_i = \beta_{i0} + \beta_1 z_{it} + \nu_{it}$ (FE)

Param.	1^{st} Stage	%	2^{nd} Stage	%	Cont.-Update	%
α_0	33653.329		33664.271		33663.994	
	(2087.356)		(2087.450)		(2087.447)	
α_1	-0.220		-0.220		-0.220	
	(0.061)		(0.061)		(0.061)	
α_2	-3198.860		-3203.499		-3203.382	
	(1051.114)		(1051.213)		(1051.210)	
θ_1 (N=5, C=1, L=0)	2.619	0.38	2.734	0.37	2.744	0.36
	(0.517)		(0.549)		(0.551)	
θ_2 (N=5, C=1, L=1)	2.044	0.49	2.129	0.47	2.130	0.47
	(0.345)		(0.368)		(0.368)	
θ_3 (N=6, C=0, L=0)	1.061	0.94	1.134	0.88	1.131	0.88
	(0.155)		(0.163)		(0.163)	
θ_4 (N=6, C=1, L=0)	1.658	0.60	1.716	0.58	1.717	0.58
	(0.241)		(0.255)		(0.256)	
θ_5 (N=6, C=1, L=1)	1.559	0.64	1.645	0.61	1.647	0.61
	(0.236)		(0.255)		(0.256)	
θ_6 (N=7, C=0, L=1)	0.995	1.01	1.071	0.93	1.072	0.93
	(0.119)		(0.126)		(0.127)	
θ_7 (N=7, C=1, L=0)	1.391	0.72	1.435	0.70	1.436	0.70
	(0.158)		(0.169)		(0.169)	
θ_8 (N=7, C=1, L=1)	1.180	0.85	1.253	0.80	1.255	0.80
	(0.155)		(0.166)		(0.167)	
θ_9 (N=8, C=0, L=0)	0.573	1.75	0.602	1.66	0.602	1.66
	(0.117)		(0.111)		(0.111)	
θ_{10} (N=8, C=1, L=0)	1.062	0.94	1.103	0.91	1.104	0.91
	(0.099)		(0.105)		(0.105)	
θ_{11} (N=8, C=1, L=1)	0.735	1.36	0.786	1.27	0.787	1.27
	(0.100)		(0.095)		(0.095)	
β_1	-0.056		-0.074		-0.075	
	(0.175)		(0.177)		(0.177)	
$f - val$	6.764		2.630		2.564	

Note: The second, fourth and sixth columns report the GMM estimates for 1st stage, 2nd stage and continuously updated method. The columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row reports the values of the objective function evaluated at the estimates.

Table 17: Alternative Specifications for Marginal Cost Function, GMM 2nd Stage

Param.	$mc_{it} = \beta_{i0} + \beta_1 z_{it} + \beta_2 q_{it}$		$mc_{it} = \beta_{i0} + \beta_1 z_{it} + (\beta_2 + 1) q_{it}^{\beta_2}$	
	Coef/S.E.	%	Coef/S.E.	%
α_0	35805.207		35419.501	
	(1845.260)		(1838.694)	
α_1	-0.293		-0.281	
	(0.052)		(0.052)	
α_2	-6509.235		-6370.224	
	(892.857)		(889.753)	
θ_1 (N=5, C=1, L=0)	2.609	0.38	2.117	0.47
	(1.430)		(1.118)	
θ_2 (N=5, C=1, L=1)	2.085	0.48	1.686	0.59
	(1.063)		(0.874)	
θ_3 (N=6, C=0, L=0)	1.193	0.84	0.981	1.02
	(0.470)		(0.427)	
θ_4 (N=6, C=0, L=1)	0.101	9.90	0.074	13.51
	(0.175)		(0.148)	
θ_5 (N=6, C=1, L=0)	1.689	0.59	1.411	0.71
	(0.689)		(0.620)	
θ_6 (N=6, C=1, L=1)	1.649	0.61	1.372	0.73
	(0.642)		(0.601)	
θ_7 (N=7, C=0, L=1)	1.152	0.87	0.975	1.03
	(0.362)		(0.366)	
θ_8 (N=7, C=1, L=0)	1.443	0.69	1.228	0.81
	(0.501)		(0.478)	
θ_9 (N=7, C=1, L=1)	1.307	0.77	1.107	0.90
	(0.429)		(0.427)	
θ_{10} (N=8, C=0, L=0)	0.741	1.35	0.632	1.58
	(0.194)		(0.204)	
θ_{11} (N=8, C=0, L=1)	0.187	5.35	0.148	6.76
	(0.153)		(0.138)	
θ_{12} (N=8, C=1, L=0)	1.158	0.86	1.002	1.00
	(0.322)		(0.333)	
θ_{13} (N=8, C=1, L=1)	0.906	1.10	0.772	1.30
	(0.253)		(0.263)	
β_1	-0.014		-0.115	
	(0.371)		(0.343)	
β_2	-0.230		0.773	
	(0.880)		(0.507)	
$f-val$	3.735		2.339	

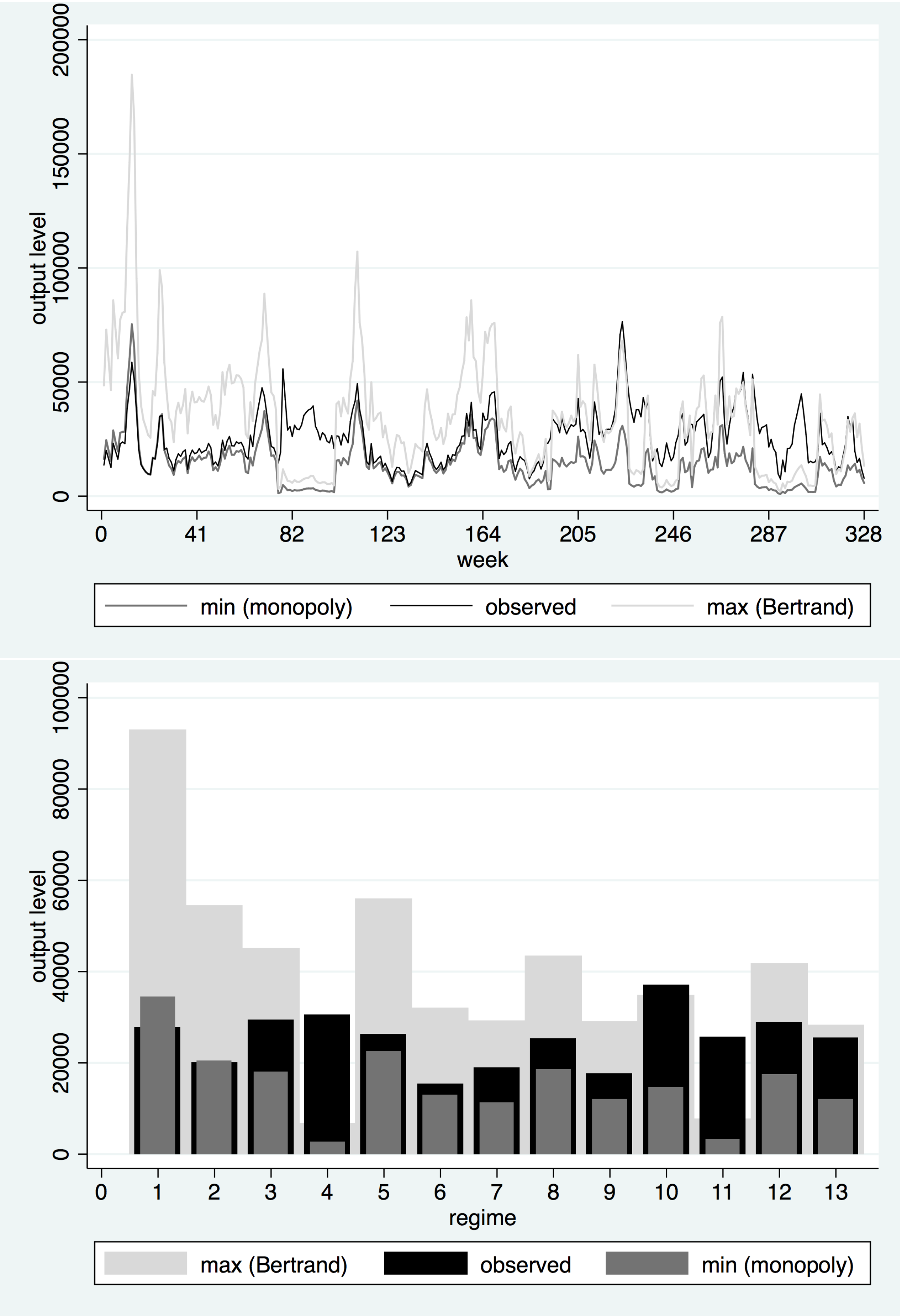
Note: The second and fourth columns report GMM estimates under the assumption of $mc_{it} = \beta_{i0} + \beta_1 z_{it} + \beta_2 q_{it}$ and $mc_{it} = \beta_{i0} + \beta_1 z_{it} + (\beta_2 + 1) q_{it}^{\beta_2}$, respectively. The numbers in the brackets are standard errors. The third and fifth columns report how much firms reduce their outputs compared with the Cournot outcomes. The last row reports the values of the GMM objective function, evaluated at the estimates.

Table 18: Estimation Results for 9 Selected Periods Satisfying PR Assumption, 662 Obs.

Param.	Levels (LE)			Firm Fixed-effects (FE)		
	1^{st}	2^{nd}	%	1^{st}	2^{nd}	%
α_0	35329.948	35522.669		35326.973	35427.255	
	(2123.370)	(2130.429)		(2125.596)	(2129.231)	
α_1	-0.281	-0.286		-0.281	-0.284	
	(0.058)	(0.058)		(0.058)	(0.058)	
α_2	-6356.179	-6427.317		-6355.350	-6392.351	
	(1018.042)	(1020.952)		(1018.649)	(1020.149)	
θ_1 (N=6, 68-75)	1.548	1.531	0.65	1.534	1.522	0.66
	(0.150)	(0.144)		(0.148)	(0.145)	
θ_2 (N=6, 116-131)	1.447	1.418	0.71	1.423	1.394	0.72
	(0.138)	(0.131)		(0.138)	(0.132)	
θ_3 (N=6, 131-166)	1.650	1.616	0.62	1.630	1.600	0.63
	(0.165)	(0.156)		(0.167)	(0.160)	
θ_4 (N=7, 171-181, 324)	1.583	1.545	0.65	1.572	1.536	0.65
	(0.168)	(0.159)		(0.171)	(0.163)	
θ_5 (N=8, 184-189)	1.694	1.651	0.61	1.672	1.632	0.61
	(0.185)	(0.174)		(0.187)	(0.178)	
θ_6 (N=8, 191-196)	1.025	1.013	0.99	1.016	1.004	1.00
	(0.058)	(0.056)		(0.058)	(0.056)	
θ_7 (N=8, 254-259)	1.200	1.184	0.84	1.186	1.170	0.85
	(0.070)	(0.067)		(0.072)	(0.069)	
θ_8 (N=8, 258-263)	1.109	1.074	0.93	1.091	1.051	0.95
	(0.063)	(0.058)		(0.065)	(0.061)	
θ_9 (N=8, 313-318)	1.212	1.231	0.81	1.192	1.220	0.82
	(0.127)	(0.128)		(0.128)	(0.130)	
β_0	22220.813	22305.381		-		
	(319.017)	(317.656)				
β_1	-0.229	-0.214		-0.170	-0.150	
	(0.104)	(0.104)		(0.128)	(0.127)	
$f-val$	37.892	3.158		46.627	4.106	

Note: The second, third, fifth and sixth columns report the estimates for the 1st and 2nd stage GMM. The fourth and seventh columns, labeled as %, report how much firms reduce their output compared with Cournot outcome. The last row shows the values of the objective function, evaluated at the estimates.

Figure 3: Upper and Lower Bounds on the Total Output by Week (Top Panel) and by Regime (Bottom Panel) for Estimates in Levels (Table 13)



Appendix B Profitability of PR Collusive Technology

Lemma 1 *Proportional reduction collusive technology is profitable for all firms in the neighborhood of Cournot equilibrium.*

Proof A Cournot competitor first-order conditions are given by

$$P'(Q_t)q_{it} + P(Q_t) - C_i'(q_{it}) = 0.$$

Consider a cartel, which sets overall industry output to $\bar{Q}_t = Q_t^{Cournot}$ and assigns market shares such that $\bar{Q}_t s_{it} = q_{it}^{Cournot}, \forall i = 1, \ldots, n$, where s_{it} is the market share of firm i in period t. Then, the profit of a cartel member is given by $\pi^m(s_{it}, \bar{Q}_t) = P(\bar{Q}_t)\bar{Q}_t s_{it} - C_i(\bar{Q}_t s_{it})$ and, by construction, is identical to the non-cooperative Cournot outcome.

Consider a derivative of this profit function with respect to $\bar{Q}_t$,

$$\begin{aligned}
\frac{\partial \pi^m(s_{it}, \bar{Q}_t)}{\partial \bar{Q}_t} &= P'(\bar{Q}_t)\bar{Q}_t s_{it} + P(\bar{Q}_t)s_{it} - C_i'(\bar{Q}_t s_{it})s_{it} \\
&= C_i'(q_{it}) - P(\bar{Q}_t) + P(\bar{Q}_t)s_{it} - C_i'(\bar{Q}_t s_{it})s_{it} \\
&= (1 - s_{it})\left(C_i'(q_{it}) - P(\bar{Q}_t)\right) < 0,
\end{aligned}$$

where the second equality is obtained by replacing $P'(\bar{Q}_t)\bar{Q}_t s_{it}$ with $C_i'(q_{it}) - P(\bar{Q}_t)$ and the inequality follows from the fact that $C_i'(q_{it}) - P(\bar{Q}_t) = P'(Q_t)q_{it} < 0$. ∎

Appendix C MC Simulations

The data generating process for our Monte Carlo simulations is as follows. We assume the following inverse demand and cost functions:

$$P_t = \alpha_0 + \alpha_1 Q_t + \alpha_2 Y_t + \nu_t^d,$$
$$mc_i(q_{it}, z_{it}) = \beta_0 + \beta_1 z_{it} + \nu_{it}^s.$$

Table 19 summarizes parameter values and the distribution of the variables. We simulated data 10,000 times for each of the following combinations of (N, T): $(10, 10)$, $(10, 20)$, $(10, 30)$, $(20, 10)$, $(20, 20)$, $(20, 30)$, $(30, 10)$, $(30, 20)$, and $(30, 30)$. Each time, parameters were estimated using 2-step optimal GMM.

Table 19: Summary of Parameter Values for Data-generating Process in MC-Simulations

Parameter / Variable	Value / Distribution
α_0	500
α_1	-1.0
α_2	1.0
β_0	10.0
β_1	1.0
Y_t	N(0,100)
ν_t^d	N(0,1)
z_{it}	N(1,4)
ν_{it}^s	N(0,0.04)
θ	$\{1.0, 1.2, 1.4\}$

We present the summary statistics for a typical data set generated for $N = 30, T = 30$ in Table 20.

Table 20: Summary Statistics for Simulated Data, N=30, T=30

Variable	Mean	P50	Min.	Max.	S.D.
q_{it}	13.397	13.306	7.688	21.319	2.369
Q_t	401.918	396.420	331.500	488.790	51.116
P_t	100.003	106.385	26.574	166.800	49.782
Y	1.944	1.789	-16.740	26.532	9.000
z_{it}	1.102	1.107	-4.765	7.932	1.969
z_{-it}	31.967	29.714	12.246	58.177	10.174
Regime 1	0.267	0.000	0.000	1.000	0.442
Regime 2	0.467	0.000	0.000	1.000	0.499
Regime 3	0.267	0.000	0.000	1.000	0.442
Y_t	1.944	1.789	-16.740	26.532	9.000

Appendix D Collusion in Differentiated Product Markets

Assume a differentiated product duopoly with marginal costs c_1 and c_2 and the following demand system:

$$q_1(p_1, p_2) = \alpha_1 - \beta_1 p_1 + p_2$$
$$q_2(p_1, p_2) = \alpha_2 - \beta_2 p_2 + p_1.$$

- Nash solution:

$$p_1^n = \frac{2\beta_1\beta_2 c_1 + 2\alpha_1\beta_2 + \beta_2 c_2 + \alpha_2}{4\beta_1\beta_2 - 1}, \ p_2^n = \frac{2\beta_1\beta_2 c_2 + 2\alpha_2\beta_1 + \beta_1 c_1 + \alpha_1}{4\beta_1\beta_2 - 1}$$
$$q_1^n = \frac{\beta_1(-2\beta_1\beta_2 c_1 + 2\alpha_1\beta_2 + \beta_2 c_2 + \alpha_2 + c_1)}{4\beta_1\beta_2 - 1}, \ q_2^n = \frac{\beta_2(-2\beta_2\beta_1 c_2 + 2\alpha_2\beta_1 + \beta_1 c_1 + \alpha_1 + c_2)}{4\beta_1\beta_2 - 1}$$

- Perfect collusion solution:

$$p_1^c = \frac{1}{2}\frac{\beta_1\beta_2 c_1 + \alpha_1\beta_2 + \alpha_2 - c_1}{\beta_1\beta_2 - 1}, \ p_2^c = \frac{1}{2}\frac{\beta_1\beta_2 c_2 + \alpha_2\beta_1 + \alpha_1 - c_2}{\beta_1\beta_2 - 1}$$
$$q_1^c = \frac{1}{2}(-\beta_1 c_1 + \alpha_1 + c_2), \ q_2^c = \frac{1}{2}(-\beta_2 c_2 + \alpha_2 + c_1).$$

If we are willing to assume that own and cross-price elasticities are the same for both firms, i.e., $\beta_1 = \beta_2 = \beta$, we can conclude that if

$$c_2 - c_1 = \frac{\alpha_2 - \alpha_1}{\beta + 1} \implies q_1^n = q_2^n \text{ and } q_1^c = q_2^c, \tag{10}$$

then the firms would reduce their output proportionally relative to the outputs under the differentiated Bertrand solution.[23]

Table 21: Example: Collusion in Differentiated Product Markets

Parameter Values						Nash Equilibrium				Collusive Equilibrium				q_1^c/q_1^n	q_2^c/q_2^n
α_1	β_1	α_2	β_2	c_1	c_2	p_1^n	p_2^n	q_1^n	q_2^n	p_1^c	p_2^c	q_1^c	q_2^c		
30	3	30	3	5	5	9.00	9.00	12.00	12.00	10.00	10.00	10.00	10.00	0.83	0.83
30	3	30	3	10	10	12.00	12.00	6.00	6.00	12.50	12.50	5.00	5.00	0.83	0.83
50	3	30	3	10	5	15.00	10.00	15.00	15.00	16.25	11.25	12.50	12.50	0.83	0.83
50	2	50	2	5	5	20.00	20.00	30.00	30.00	27.50	27.50	22.50	22.50	0.75	0.75
50	4	50	4	5	5	10.00	10.00	20.00	20.00	10.83	10.83	17.50	17.50	0.88	0.88
50	3	50	3	5	3	12.83	11.97	23.49	26.91	15.00	14.00	19.00	23.00	0.81	0.85
50	3	50	3	5	4	12.91	12.49	23.74	25.46	15.00	14.50	19.50	21.50	0.82	0.84
50	3	50	3	5	5	13.00	13.00	24.00	24.00	15.00	15.00	20.00	20.00	0.83	0.83
50	3	50	3	5	6	13.09	13.51	24.26	22.54	15.00	15.50	20.50	18.50	0.85	0.82
50	3	50	3	5	7	13.17	14.03	24.51	21.09	15.00	16.00	21.00	17.00	0.86	0.81

The first five rows in Table 21 report demand and cost parameters that make collusive and competitive output market shares equal. The last five rows fix demand-side parameters (at symmetric values) and illustrate total industry maximizing output for various levels of marginal cost parameters.

Note that differentiated product markets require additional restrictions for proportional reduction collusive technology to make sense. At the same time, at least some of these restrictions

[23] In this case firms would have equal market shares under both regimes.

are empirically testable. For example, demand parameters can be estimated without specifying the equilibrium concept on the supply side. This way one may address the question about firms' symmetry on the demand side. Estimation of the cost functions, in turn, depends on the assumptions about firms' conduct. However, by specifying the null hypothesis about the firms' behavior (collusion or competition), one can estimate marginal cost functions and then test whether demand- and supply-side parameters satisfy the required restrictions (see condition 10). When there is no information on which firms may be colluding, one can run tests in the spirit of Bajari and Ye (2003) by estimating the parameter of interest for various subsets of the firms.