

Paha, Johannes

Working Paper

Simulation and prosecution of a cartel with endogenous cartel formation

MAGKS Joint Discussion Paper Series in Economics, No. 07-2010

Provided in Cooperation with:

Faculty of Business Administration and Economics, University of Marburg

Suggested Citation: Paha, Johannes (2010) : Simulation and prosecution of a cartel with endogenous cartel formation, MAGKS Joint Discussion Paper Series in Economics, No. 07-2010, Philipps-University Marburg, Faculty of Business Administration and Economics, Marburg

This Version is available at:

<https://hdl.handle.net/10419/56581>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

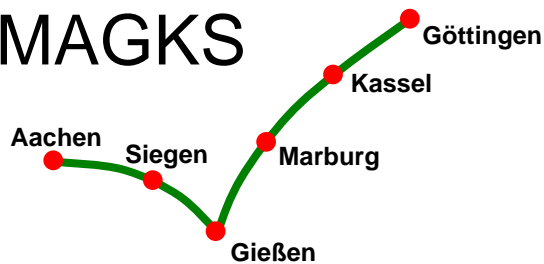
Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

MAGKS



**Joint Discussion Paper
Series in Economics**

by the Universities of
Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 07-2010

Johannes Paha

**Simulation and Prosecution of a Cartel
with Endogenous Cartel Formation**

Johannes Paha
Justus-Liebig-University Gießen
Licher Straße 62, D-35394 Gießen
Email: johannes.paha@wirtschaft.uni-giessen.de

This paper can be downloaded from
http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29

Coordination: Bernd Hayo • Philipps University Marburg • Faculty of Business Administration &
Economics • Universitätsstr. 24, D-35032 Marburg
Tel: ++49-6421-2823091, Fax: ++49-6421-2823088 e-mail: hayo@wiwi.uni-marburg.de

SIMULATION AND PROSECUTION OF A CARTEL WITH ENDOGENOUS CARTEL FORMATION

Johannes Paha*

ABSTRACT

In many cases, collusive agreements are formed by asymmetric firms and include only a subset of the firms active in the cartelized industry. This paper endogenizes the process of cartel formation in a numeric simulation model where firms differ in marginal costs and production technologies. The paper models the incentive to collude in a differentiated products Bertrand-oligopoly. Cartels are the outcomes of a dynamic formation game in mixed strategies. I find that the Nash-equilibrium of this complex game can be obtained efficiently by a Differential Evolution stochastic optimization algorithm. It turns out that large firms have a higher probability to collude than small firms. Since firms' characteristics evolve over time, the simulation is used to generate data of costs, prices, output-quantities, and profits. This data forms the basis for an evaluation of empirical methods used in the detection of cartels.

Keywords: Collusion, Cartel Detection, Cartel Formation, Differential Evolution, Heuristic Optimization, Industry Simulation

JEL Codes: C51, C69, C72, D43, L12, L13, L40

* Johannes Paha is a research associate at the

Chair for Industrial Organization, Regulation and Antitrust (VWL 1)
Justus-Liebig-University Giessen
Licher Straße 62
D-35394 Giessen

email johannes.paha@wirtschaft.uni-giessen.de
phone +49 – 641 – 99 22052
fax +49 – 641 – 99 22059
web http://wiwi.uni-giessen.de/ma/dat/goetz/Johannes_Paha%2C%20M.A./

Earlier versions of this paper were presented at Earie 09, Ljubljana, the 3rd RNIC conference, Vienna, and several workshops in Giessen and Marburg. I would like to thank all participants of these conferences and workshops for their valuable comments. A special thanks goes to Georg Götz, my colleagues in Giessen, and numerous researchers – among them Joe Harrington – who read the paper and gave me advise on it.

1 INTRODUCTION

Which firms are most likely to form a cartel? Based on a simulation model for collusive industries, I find that large firms are more inclined to form a cartel than small firms. The reason for this is in line with the findings of recent research (Bos and Harrington 2010). Small firms contribute little to raising the level of market prices but gain much from undercutting large cartelists' prices. Analogously, large firms contribute much to raising the level of market prices but gain little from not participating in a cartel since this prohibits prices from being raised sizeably.

These results are derived from an industry simulation model. Therefore, I contribute to the literature on the endogenous formation of cartels in an infinitely repeated game framework (see Bos and Harrington 2010, Vasconcelos 2005, and Compte et al. 2002). The paper adds to this literature by assuming heterogeneous firms *and* heterogeneous products. In addition, it tackles games' complexity by numerically determining the Nash-equilibrium of the cartel-formation game using a Differential Evolution (Storn and Price 1997) stochastic optimization algorithm. As such, it contributes to a relatively new strand of literature which proposes to apply optimization methods from the field of operations research to solving game theoretic problems (see e.g. Beck et al. 2007).

The paper closest to my research is Bos and Harrington (2010). They analyze firms' incentive to collude in a framework where “ n firms compete in an infinitely repeated capacity-constrained price game with homogeneous products” (p. 95). They “assume that demand is allocated within the cartel proportional to a firm's capacity” (p. 97). For stabilizing the cartel, they use Friedman's (1971) standard grim trigger strategy.

In my model, firms are heterogeneous in marginal costs and offer differentiated products. The model allows for i.i.d. cost shocks such that cartels are stabilized by a trigger strategy à la Rotemberg and Saloner (1986). This strategy enables cartelists to stabilize the collusive agreement by lowering the price in cases where cost-shocks would induce some firm to deviate. Since firms are assumed to revert to the one-shot Nash-equilibrium price, this trigger strategy generates time series of firms' prices that resemble those of a Green and Porter (1984)-type strategy. Cartel-formation is not deterministic but allows for a random element by modeling its outcome as a Nash-equilibrium in mixed strategies (Prokop 1999). As a consequence, even a small firm could be a cartel-member, or a large firm could be a fringe-firm. A crucial feature/outcome of both models is that cartels are rarely complete. Instead, cartelists have to deal with a competitive fringe. This describes many real collusive agreements better than the assumption of an all-inclusive cartel (Bos and Harrington 2010: 92). Additionally, in my model prices and quantities are the outcome of firms' decision to maximize profits. This increases the parsimony of the model since an explicit allocation

rule for demand is not needed.

Despite the differences in their modeling structure, both models give quite similar predictions. This raises the confidence that one may put in the robustness of these results. Both models find that large firms benefit more from being a cartel-member than small firms. This holds true despite a larger cartel benefits all firms irrespective of their size or their affiliation to the set of cartelists or the set of fringe-firms.

The numerical nature of the model adds a further convenient feature. The model allows for generating data on firms' prices, output quantities, and profits. Such data is used to evaluate and advance empirical methods used in the detection and prosecution of cartels. In particular, it is shown how Bajari and Ye's (2003) test for the detection of cartelists can be made applicable for standard goods markets although the test has originally been proposed for auction markets. Such assessments of empirical methods can be done more easily with simulated rather than real data. Like in a windtunnel, industry simulations allow for multifold, controlled variations in the generated data. This facilitates the examination of empirical methods' robustness in different environmental conditions. Additionally, the estimated parameters need not merely be tested against *plausible* values but can be tested against the *true* underlying parameters. As a further advantage, evaluations based on simulated data bridge a gap between assessments of methods' technical properties (biasedness and efficiency) and assessments of methods' actual properties (robustness of results under varying industry conditions).

The paper is structured as follows. Section 2.1 presents the structure of the stage game with respect to demand and supply conditions. Section 2.2 elaborates on the stability of the cartel and outlines the relevant trigger strategy with a price war-like pricing-pattern on the equilibrium path. Section 2.3 shows that a cartel can be modeled as a Nash-equilibrium solution of a formation-game in mixed strategies. In section 3 I show how this equilibrium can be attained at reasonable computational costs as the minimum of some objective function. This optimum is found by a Differential Evolution stochastic search algorithm (Storn and Price 1997). This is the main *methodological* contribution of this work. In section 3.4 I econometrically examine the determinants of cartel participation in 200 simulated industries. It is found that (*ceteris paribus*) large firms have a higher probability of joining the cartel than small firms. In section 4, data generated by this oligopoly simulation model is used for evaluating and enhancing Bajari and Ye's (2003) cartel-screening methods, i.e. their test of exchangeability and their test of conditional independence. Section 5 concludes.

2 SIMULATING COLLUSION

In this section the simulation model is presented that is used for generating the industry data. The basic model is outlined in section 2.1. The demand-side of the model is based on Shubik and Levitan's (1980) well-known utility function for differentiated products. Firms' production technology is characterized by marginal costs that are constant in output and vary across firms and over time. In this market environment firms must decide whether to maximize their profits independently or jointly. If they decide to form a cartel, firms prevent deviations from the collusive agreement by engaging in price wars on the equilibrium path as is shown in section 2.2. Section 2.3 describes the process of cartel-formation. Based on the ideas of Prokop (1999), it is shown that a cartel can arise as a Nash-equilibrium outcome when firms play a mixed strategy.

2.1 The Basic Model

Stigler (1964: 45) proposes that homogenous product markets are rather the exception than the rule. Therefore, my simulation model is based on the following representative agent's utility function for differentiated products (Shubik and Levitan 1980), where vectors and matrices are denoted in bold.¹

$$U(\mathbf{q}) = v \mathbf{q}' \boldsymbol{\iota} - \frac{n}{2 \cdot (1 + \mu)} \left[\mathbf{q}' \mathbf{q} + \frac{\mu}{n} (\mathbf{q}' \boldsymbol{\iota})^2 \right] \quad (1)$$

In this function \mathbf{q} is a $(n \times 1)$ -vector whose elements are the quantities q_i of n products. Each product is produced by exactly one firm so that there are n firms in the industry. The number of firms is modeled as fixed. This may be motivated by sunk costs being sufficiently high so that there are no firms outside the industry for whom it would be profitable to enter. $\boldsymbol{\iota}$ is a $(n \times 1)$ -vector where each element takes a value of 1. v is a positive parameter and $\mu \in [0, \infty)$ represents the degree of substitutability² of the n products. According to this utility function the representative agent consumes some quantity of each good. Maximizing utility with regard to quantity gives the set of inverse demand functions:

$$\frac{dU}{d\mathbf{q}'} = \mathbf{p} = v \boldsymbol{\iota} - \frac{1}{1 + \mu} (n \mathbf{q} + \mu (\mathbf{q}' \boldsymbol{\iota}) \boldsymbol{\iota}) \quad (2)$$

¹ In appendix A.1, the most important equations of this sections are presented in non-vector notation.

² For $\mu = \infty$ goods are perfect substitutes. For $\mu = 0$ goods are independent.

The system of corresponding demand functions can be written as:

$$q = \frac{1}{n} \left[v \iota - p(1 + \mu) + \frac{\mu}{n} (p' \iota) \iota \right] \tag{3}$$

Each product is produced by a one-product firm at marginal cost c_i . Firms may be asymmetric in their cost-structure. At the output market firms compete à la Bertrand in prices. This gives the below vector of competitive equilibrium-prices:

$$p = (I(2n + 2\mu n - \mu) - \mu \iota \iota')^{-1} \cdot (\iota n v + (n + n\mu - \mu) \cdot c) \tag{4}$$

Firms 1 to m (with $m \leq n$) may reduce competition by forming a cartel and maximizing profits jointly (Stigler 1964: 44, 45). In this case, prices will rise and equation (4) becomes

$$p = [I(2n + 2\mu n - \mu) - \mu(\iota \iota' + A)]^{-1} \cdot [\iota n v + ((n + n\mu - \mu) - \mu A) \cdot c] \tag{5}$$

where A is a $(n \times n)$ -matrix of the form shown in (6). The non-zero elements in the upper left part are of dimension $(m \times m)$ and stand for the effect of the joint profit maximization in the cartel.

$$A = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & & & & & \\ 1 & 1 & 0 & & & & & \\ \vdots & & & \ddots & & & & \vdots \\ 1 & & & & 0 & & & \\ 0 & & & & \dots & & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & & & \dots & & & 0 \end{pmatrix} \tag{6}$$

Please note that equation (4) is a special case of equation (5) with $m = 0$, i.e. no collusion occurs.

In this model, marginal costs have three features. (i) Marginal costs are firm-specific which makes firms asymmetric. This allows for analyses of differences in firms' participation-probabilities in a cartel. In section 3.4 it is shown that a firm is less likely to join a cartel if its marginal costs exceed the industry-mean. (ii) Marginal costs are modeled to be more similar if firms produce more homogenous goods. This assumption is reasonable since more homogenous goods supposedly are produced by similar production technologies with similar production costs. (iii) Cost-shocks are assumed to occur in every period such that marginal costs follow a random walk (Harrington 2008: 241). This introduces dynamics to the simulation model which is a prerequisite for using the data as a basis for evaluating empirical methods for the detection of cartels. With time-invariant marginal costs one would only observe one market outcome for collusion and one for competition. Observing only two outcomes prohibits any econometric analysis. In the remainder of this section, I describe the implementation of these features.

In the first period, marginal costs of firm i , i.e. $c_{i,t=1}$, are generated according to equation (7) in conjunction with conditions (8) and (9).

$$c_{i,t=1} = a_1 \cdot v + a_{2i,t} \cdot s_t \quad (7)$$

$$\left. \begin{aligned} a_1 &\in]0; 1[\\ a_{2i,t} &\sim No\left(\frac{a_3+1}{2}, \mu^{-2/a_4}\right) \in [a_3; 1] \\ a_4 &\geq 1 \\ a_5 &\in]0; 1[\end{aligned} \right\} \quad (8)$$

The base level of marginal costs is determined as the percentage a_1 of the variable v , which is closely related to goods' reservation price. Cost-asymmetry among firms is modeled by adding a firm-specific term $a_{2i,t}s_t$ to the base level of marginal costs. The asymmetry-term has the following features. (i) The firm-specific technology-parameter $a_{2i,t}$ is drawn randomly from a normal distribution in the interval $[a_3; 1]$ with its expected value $E(a_{2i,t})$ being the mean of the interval $[a_3; 1]$. (ii) The variance of the technology-parameter $a_{2i,t}$, determines how strongly marginal costs differ across firms. Since similar products are produced by similar technologies the variance of $a_{2i,t}$ decreases in the degree of product homogeneity μ . (iii) a_4 is an additional measure for the dispersion of cost shocks. The higher a_4 the greater is the variance of $a_{2i,t}$ and, thus, the more dispersed are cost shocks. a_4 is needed to parameterize and fine-tune the model. Cost shocks s_t are drawn randomly from a uniform distribution in the interval given by equation (9). $a_5 \in [0; 1]$ determines the amplitude of cost-shocks. Setting $a_5 = 0$ gives marginal costs that are symmetric across firms and constant over time. This interval ensures that marginal costs cannot become negative.

$$s_t \in [-a_5 \cdot (a_1 \cdot v); a_5 \cdot (a_1 \cdot v)] \quad (9)$$

In subsequent periods, marginal costs are generated according to equation (10) in conjunction with conditions (11) and (12).

$$c_{i,t} = c_{i,t-1} + a_{2i,t} \cdot s_t \quad (10)$$

$$a_{2i,t} = \begin{cases} 1 & \text{for } s_t < 0 \\ a_{2i,t} \sim No\left(\frac{a_3+1}{2}, \mu^{-2/a_4}\right) \in [a_3; 1] & \text{for } s_t > 0 \end{cases} \quad (11)$$

$$s_t \in \left[-a_5 \cdot \min_i(c_{i,t-1}); a_5 \cdot \min_i(c_{i,t-1})\right] \quad (12)$$

As in Harrington (2008: 241) marginal costs are assumed to follow a random walk. Thus, in every period t a random, scaled shock term s_t is added to last period's marginal costs of each firm i . The marginal cost-shock s_t is the same for all firms and can be considered a fluctuation in input prices. It

is drawn randomly from a uniform distribution in the interval $[-a_5 \cdot \min_i(c_{i,t-1}); a_5 \cdot \min_i(c_{i,t-1})]$, where $\min_i(c_{i,t-1})$ is the minimum (over all firms i) of last period's marginal costs. Again, this ensures that marginal costs cannot become negative. Equation (11) shows that firms' reaction to cost-shocks is asymmetric. If a cost shock is negative ($s_t < 0$) firms fully exploit the reduction in marginal costs (i.e. $a_{2i,t} = 1 \forall i$). If a cost shock is positive ($s_t > 0$) firms mitigate the cost shock by e.g. restructuring production. In the latter case, the value of the technology-parameter $a_{2i,t}$ is drawn randomly from a normal distribution with the above properties.

2.2 Cartel Stability

Most cartels are inherently instable and, therefore, require a stabilizing mechanism. In this section, I argue that the cartel is (in most cases) stabilized by employing a grim trigger strategy (Friedman 1971). With sufficiently impatient firms a grim trigger strategy may not be sufficient to mitigate firms' present incentive to deviate. Therefore, cartelists additionally engage in price wars for preventing deviations (Rotemberg and Saloner 1986). I assume that a cartel does not necessarily have to consist of all n firms in the industry. I rather deal with the more general case where some firms join the cartel while others remain independent. As is standard in the literature on cartels, I assume that no more than one cartel exists at a given time (Kuipers and Olaizola 2008: 407).

A cartel is called *stable* if it satisfies two stability criteria (D'Aspremont et al. 1983: 21). First, a cartel is called *internally stable* if it is unprofitable for any cartel-member to deviate from the cartel agreement. Second, a cartel is called *externally stable* if it is unprofitable for firms outside the cartel to join it. In the context of this paper, industries are large and products differentiated. This makes cartels externally stable but internally instable. Under some conditions (e.g. a small enough industry with symmetric firms and homogenous products) cartels can be internally stable even in a one-shot game. Contributions to this interesting strand in the oligopoly-literature have e.g. been made by Selten (1973), D'Aspremont et al. (1983), Diamantoudi (2005), and Kuipers and Olaizola (2008). It does hold for the types of industries studied here, which then are characterized by internal instability.

Cartels are externally stable since cartelists provide a positive externality to firms outside the cartel (see Deneckere and Davidson 1985 and Stigler 1950). A firm i can ceteris paribus make a higher profit π_{di} by acting as a competitive fringe-firm than by acting as a cartelist (π_i) (Deneckere and Davidson 1985: 477). This is because fringe firms may increase prices under the cartel's price umbrella and *expand* quantity, while cartelists must *reduce* output in order to maintain a high price-level. If the firm decides to collude, the cartel is of size m . If it remains a fringe firm, the cartel is of size $m-1$. These facts are expressed in condition (13).

$$\pi_i(m) < \pi_{di}(m-1) \quad (13)$$

Fringe firms enjoy the benefits of deviators from the cartel agreement without being punished. Hence, no firm outside an existing cartel has an incentive to join it. Stigler³ (1950: 25) provides a similar argument for mergers. Namely, that “the promoter of a merger is likely to receive much encouragement from each firm – almost every encouragement, in fact, except participation.”

Similarly, existing cartels are internally instable because cartelists have a desire to increase their profits by secretly undercutting cartel prices (Stigler 1964: 46). By deviating, i.e. supplying a larger quantity at reduced prices, a cartelist makes a higher profit π_{di} than by acting collusively (π_i) (D'Aspremont et al. 1983: 19). Therefore, cartels are instable⁴ in a one-shot game. It has been shown by Friedman (1971) that cartels can be stabilized by appropriate (here: grim) trigger strategies in a dynamic game. Hence, after an observed deviation from the cartel, all cartelists will revert to a competitive equilibrium forever. As a consequence, no firm will deviate from the collusive agreement unless it is very impatient. In such cases the cartel can be stabilized by engaging in price wars for making deviations unprofitable (see e.g. Green and Porter (1984), Rotemberg and Saloner (1986), and Haltiwanger and Harrington (1991)). Price wars are generally defined as periods where cartelists set prices in the range between the competitive price and below the jointly profit maximizing price. In particular, I assume that firms revert to the one-shot competitive Nash-equilibrium. The specific modeling assumptions are explicated below.

Assume the industries of interest are infinitely lived. Assume for now that firms already have established some cartel. In sections 2.3 and 3 it is shown which firms are most likely to form the cartel. The firms observe marginal costs at the beginning of the period. Hence, they can accurately anticipate their current profits in the existing cartel π_i , the profits in competition π_{ci} , and the profits they make when deviating from the cartel π_{di} . Since cost shocks are distributed i.i.d. firms' best guess about costs – and thus profits – in any future period is their respective current values, i.e. the conditions $E(\pi_i) = \pi_i$, $E(\pi_{ci}) = \pi_{ci}$, and $E(\pi_{di}) = \pi_{di}$ hold. The cartel may be discovered (at the beginning of a period) by an antitrust authority with probability P (Hinloopen 2006). After detection by the antitrust authority, competition in the industry is restored forever with firms making expected profits π_{ci} . Fines imposed on cartelists by the competition authority and payment of damages are normalized to zero. Let V_{ci} denote the expected present value of competitive profits in this case.

3 Please note that the effects of a merger are similar to those of cartel-formation. In the latter case, firms simply maintain legal independence.

4 As noted above, under some conditions cartels can also be stable in a static setting. However, such industries are beyond the scope of this paper.

$$V_{ci} = \pi_{ci} / (1 - \delta) \quad (14)$$

If the firm stays in the cartel, it makes profit π_i in the current period. For the following periods, it expects to make the competitive present value of profits V_{ci} after a detection by the antitrust authority (with probability P) and the collusive present value of profits V_i otherwise. These present values are discounted by the discount factor $\delta = 1/(1+r)$ with discount rate r .

$$V_i = \pi_i + \delta [P \cdot V_{ci} + (1-P) \cdot V_i] \quad (15)$$

Being an input parameter to the model, the discount rate r is determined exogenously in the interval $[0; \min_i((p_{ci} - c_i)/c_i)]$. Using this upper bound ensures that even the least profitable firm makes non-negative profits in competition. Otherwise it would exit the industry, which is beyond the scope of this model.

If a cartelist i deviates from the collusive agreement, it enjoys deviation profits for one period only. In the following periods, the other cartelists play a grim trigger strategy (Friedman 1971) and revert to the competitive equilibrium forever. Hence, the present value of deviation profits V_{di} can be expressed as follows.

$$V_{di} = \pi_{di} + \delta \cdot V_{ci} \quad (16)$$

For a cartel to be stable, the present value of collusive profits must be at least as great as the present value of deviation profits. Therefore, the following standard condition must hold.

$$\frac{\pi_{di} - \pi_i}{\pi_{di} - \pi_{ci}} \leq \delta \cdot (1 - P) \quad (17)$$

If firms are sufficiently patient (δ high) and the antitrust authority is sufficiently ineffective (P low), some cartels can be sustained at jointly profit maximizing prices. The remaining collusive agreements can be stabilized by lowering the collusive output price below the joint profit-maximizing price until the above condition is satisfied for all firms in the cartel (Rotemberg and Saloner 1986, Haltiwanger and Harrington 1991). In large industries with asymmetric firms it is challenging even for an omniscient researcher to determine the set of prices that equalize the left-hand side and the right-hand side of equation (17). In this case, the cartel would *just* be stable. Calculating this set of prices is disparately more difficult for the (in reality imperfectly informed) firms themselves. Cartelists are assumed to overcome this problem by explicitly coordinating on the one stable solution that is known to every firm, i.e. setting competitive prices. Therefore, firms' strategy space is restricted to setting either jointly profit maximizing prices (making profit π_{jpi}) or engaging in a price war setting competitive prices (making profit π_{ci}).

$$\pi_i = \begin{cases} \pi_{jpi} & \text{if } \forall j \in [1, m] \frac{\pi_{dj} - \pi_{jpi}}{\pi_{dj} - \pi_{cj}} \leq \delta \cdot (1 - P) \\ \pi_{ci} & \text{if } \exists j \in [1, m] \frac{\pi_{dj} - \pi_{jpi}}{\pi_{dj} - \pi_{cj}} > \delta \cdot (1 - P) \end{cases} \quad (18)$$

Rotemberg and Saloner (1986: 395) note that this pricing strategy is similar to the one proposed by Green and Porter (1984). The decision of engaging in a price war is made in each period. In the following period, when firms learn about their new marginal costs, a new decision is made.

Deviation profits π_{di} must be defined more closely. This is because, theoretically, $d \in [1; m]$

cartel-firms can jointly deviate from a cartel with m members, giving rise to potentially $\sum_{d=1}^m \binom{m}{d}$ new equilibria. Since there is no uncontested method to select one of these multiple equilibria, I go with the more widespread solution (which is also chosen by Bos and Harrington (2010: 106)) and define deviation profits according to the concept of *stability* rather than *strong stability* (Hart and Kurz 1983: 1048). A cartel is called strongly stable if “for all possible coalitions of firms in the cartel there is no incentive to jointly defect from the cartel [(strong internal stability) and if ...] there is no incentive to jointly add to the cartel” (strong external stability) (Olaizola 2007: 224). A cartel is called stable if for any single firm there is no individual incentive to join or deviate from the cartel. Therefore, each firm is assumed to build expectations about deviation profits π_{di} based on the reasoning that only itself but no other firm deviates from the collusive agreement.

The above strategy gives a self-enforcing, internally and externally stable cartel. At a first glance, one might see no need for cartelists to communicate in order to sustain the cartel. Nonetheless, the collusive agreement should be considered an illegal cartel. First, stabilizing the cartel requires active communication in agreeing on the strategy that is played in a price-war (i.e. reverting to the competitive equilibrium). Second and as is shown in section 2.3, the cartel is illegal since communication and coordination is needed for the establishment of the cartel. In this context, Spagnolo (2008: 260) points out that the property of self-enforcement is important even for illegal collusive agreements since “individual opportunism cannot be limited by explicit contracts enforced by the legal system”. Moreover, by reviewing experimental evidence Kühn (2008: 126) questions whether *real* tacitly colluding firms – without communication – would attain optimal equilibria and punishment strategies at all.

2.3 Cartel Formation

It was argued above that cartelists provide a positive externality to firms outside the cartel (see e.g. Deneckere and Davidson 1985 and Stigler 1950). Therefore and everything else equal, a firm makes

a higher profit when acting as a competitive fringe-firm than by acting as a cartel member (see equation (17)). As a consequence, any firm would like all others to form the cartel but remain independent itself (Stigler 1950: 25). Firms must be aware that a cartel will not be established if all firms behave according to that rationale.

Therefore, I model the cartel formation game as proposed by Prokop (1999). He assumes that firms play a simultaneous-moves game in two stages. “In the first stage, the firms simultaneously decide whether to join the cartel or to remain in the competitive fringe” (Prokop 1999: 248). In the second stage, the firms interact in the product market. Assume, each firm i has two pure strategies g , where $g = 0$ means staying in the competitive fringe and $g = 1$ joining the cartel. I denote the strategy space $\{0, 1\}$ of firm i by G_i . Firms have to choose the strategy g_i that maximizes their payoff function $h_i(g)$.

$$h_i(g) = h_i(g_1, \dots, g_n) = \begin{cases} E(V_i) & \text{if } g_i = 1 \\ E(V_{di}) & \text{if } g_i = 0 \end{cases} \quad (19)$$

Let P_{pw} denote the probability that colluding firms enter in a price war in some period. Then the firm i 's expected cartel profit can be expressed as the weighted sum of the jointly maximized profit π_{jpi} and the competitive profit π_{ci} .

$$\pi_i = (1 - P_{pw}) \cdot \pi_{jpi} + P_{pw} \cdot \pi_{ci} \quad (20)$$

Plugging this into equation (15) the expected present value of cartel profits can be expressed as

$$E(V_i) = \frac{(1 - P_{pw}) \cdot \pi_{jpi} + P_{pw} \cdot \pi_{ci} + \delta \cdot P \cdot \pi_{ci} / (1 - \delta)}{1 - (1 - P) \cdot \delta} \quad (21)$$

Similarly, the expected present value of profits in the fringe is

$$E(V_{di}) = \frac{(1 - P_{pw}) \cdot \pi_{di} + P_{pw} \cdot \pi_{ci} + \delta \cdot P \cdot \pi_{ci} / (1 - \delta)}{1 - (1 - P) \cdot \delta} \quad (22)$$

I.e. a fringe (cartel) firm enjoys profits π_{di} (π_{jpi}) as long as the cartel is not discovered by the competition authority which occurs with probability P and as long as no price war occurs with probability P_{pw} . Given $\pi_{jpi} < \pi_{di}$, it is straightforward to show that $E(V_i) < E(V_{di})$. From this, one may derive two results. First, an existing cartel must not be expected to grow any further. Second, by employing a pure strategy based on an assessment of expected present values of profits, no firm has an incentive to join the cartel. In this case no cartel is established and the industry remains competitive with firm i expecting the present value of profits $E(V_{ci})$. From firms' point of view this outcome is suboptimal since $E(V_{ci}) < E(V_i)$. As a result, a cartel cannot arise as an equilibrium outcome in pure strategies.

Firms may solve the above formation-problem by playing a mixed strategy. Denote the set of mixed strategies of firm i by $J_i = \{j_i: 0 \leq j_i \leq 1\}$, i.e. firm i announces to join the cartel with probability j_i and remain independent with probability $(1 - j_i)$. On the one hand, this builds up confidence among firms that a cartel will be established. On the other hand, each firm is left with the chance to remain in the fringe even when negotiating about its participation in the cartel. In such an environment, where a cartel is a possible outcome without severe limitations of firms strategies the collusive outcome can indeed arise. The challenge for each firm lies in announcing a participation-probability j_i such that its expected present value of profits $H_i(\mathbf{j})$ is maximized.

$$H_i(\mathbf{j}) = j_i \cdot E(V_i) + (1 - j_i) \cdot E(V_{di}) \quad (23)$$

As in Prokop (1999: 249) the Nash equilibrium concept is applied to this game. A strategy n -tuple $\hat{\mathbf{j}} = (\hat{j}_1, \dots, \hat{j}_n)$ is called a Nash equilibrium if for all firms i and for all $j_i \in J_i$ the condition $H_i(\hat{\mathbf{j}}) \geq H_i(\hat{\mathbf{j}}_{-j}, j_i)$ holds where $\hat{\mathbf{j}}_{-j} = (\hat{j}_1, \dots, \hat{j}_{j-1}, \hat{j}_{j+1}, \dots, \hat{j}_n)$. Since an equilibrium in pure strategies is a special case of an equilibrium in mixed strategies, (23) even applies for the (rare) cases where jointly maximized profits are at least as large as deviation profits ($\pi_{jpi} \geq \pi_{di}$).

If firms' optimal values of j_i are known, a firm i can be simulated to join the cartel if a random number generator draws a number from a uniform distribution in the interval $[0;1]$ which is smaller than or equal to j_i . The task of finding optimal participation-probabilities is not trivial. This is because maximizing (23) requires firms to form expectations about the likeliness of cartel-detection by the competition authority P and about the likeliness of a price-war P_{pw} . In the following, firms are assumed to be well-informed about the competition authority's activities. Therefore, they are assumed to know P for sure, where P is an input parameter to this simulation model. The process of forming expectations about the probability of a price war P_{pw} is described in section 3.2. Finally, firms must form expectations about their profits in competition π_{ci} , when remaining a fringe firm π_{di} , and when joining a cartel π_i . This task is complex since π_{ci} and π_{di} depend on the size and the composition of the cartel which again depend on the optimal values of firms' joining-probabilities j_i . The process of finding optimal j_i 's is described in sections 3.1 and 3.3.

3 DETERMINING NASH-EQUILIBRIA OF ASYMMETRIC GAMES

In this section, I show that a Nash-equilibrium of the above cartel-formation game for asymmetric firms can be found numerically by a Differential Evolution (DE) stochastic optimization heuristic. This algorithm, as described in section 3.3, iteratively converges to this Nash-equilibrium by minimizing an objective function. This function is described in section 3.1 and relies on the idea that all firms choose a participation probability such as to maximize their expected present value of

profits taking into account the strategies of all other firms. Since these profits depend on the likeliness of a price war, firms' expectation-making process concerning the price war probability is outlined in section 3.2. Section 3.4 provides an economic interpretation of the factors that influence firms' decision whether to join a cartel or to remain a fringe firm.

3.1 Nash-Equilibria as the Minimum of an Objective Function

In an industry with n firms and two strategies (collude or non-collude) 2^n market outcomes may arise. $n+1$ of these market structures are competitive since at maximum one firm participates in the cartel. The challenge is determining the optimal market structure by selecting participation probabilities. This is done by jointly maximizing the expected present value of profits (23) for all firms. Doing this optimization jointly ensures that the outcome constitutes a Nash-equilibrium, i.e. no firm would find it profitable to deviate from this mixed strategy. Beck et al. (2007) show how the Nash equilibrium of a game can be determined by using a stochastic optimization algorithm. Accordingly, the vector \mathbf{j}_{opt} of optimal participation probabilities shall be found as the minimum of some objective function $of(\mathbf{j})$. For this cartel-formation problem the following objective function is chosen. It aggregates the expected present value of profits over the mixed strategies for all firms.

$$of(\mathbf{j}) = \frac{1}{n} \cdot \sum_i \left[\frac{E(V_{max,i}) - H_i(\mathbf{j})}{E(V_{max,i})} \right]^2 \quad (24)$$

The idea underlying equation (24) is the following. No firm can make an expected present value of profits higher than $E(V_{max,i})$. This is the present value of profits it makes when all other firms form the cartel and it remains the only firm in the fringe. Therefore, it is each firm i 's objective to individually choose a participation probability j_i such that the difference between its expected present value of profits $H_i(\mathbf{j})$ and $E(V_{max,i})$ is minimized. Therefore, finding the Nash-equilibrium of the cartel formation game is equivalent to finding the joint minimum of this difference. Three additional features are added to the objective function. (i) Dividing by $E(V_{max,i})$ normalizes the term in brackets to the interval $[0;1]$ so that large firms are not assigned a higher weight in (24) than small firms. (ii) Squaring puts more weight on firms for whom a good j_i has not been found, yet. (iii) Dividing by the number of firms n normalizes the values of $of(\mathbf{j})$ to the interval $[0;1]$. This allows for a better comparison of objective function values for differently sized industries.

The expected present value of profits when just firm i is in the fringe can be expressed as

$$E(V_{max,i}) = \frac{(1 - P_{pw}) \cdot \pi_{max,i} + P_{pw} \cdot \pi_{ci} + \delta \cdot P \cdot \pi_{ci} / (1 - \delta)}{1 - (1 - P) \cdot \delta} \quad (25)$$

Therefore, the objective function can also be written as

$$of(\mathbf{j}) = \frac{1}{n} \cdot \sum_i \left[\frac{(1 - P_{pw}) \cdot [\pi_{max,i} - j_i \cdot \pi_{jpi} - (1 - j_i) \cdot \pi_{di}]}{(1 - P_{pw}) \cdot \pi_{max,i} + P_{pw} \cdot \pi_{ci} + \delta \cdot P \cdot \pi_{ci} / (1 - \delta)} \right]^2 \quad (26)$$

Please note that the cartel detection probability P only enters the objective function via the constant term $\delta \cdot P \cdot \pi_{ci} / (1 - \delta)$. Thus, in this model without fines the effectiveness of the competition authority does not impact firms' participation probabilities.

$\pi_{jpi} = \pi_{jpi}(\mathbf{j})$ and $\pi_{di} = \pi_{di}(\mathbf{j})$ are functions of all firms' participation probabilities. Therefore, computing the objective function value is not a simple task since theoretically for any combination of j_i 's $\in]0;1[$ any combination of cartel-firms and fringe-firms can arise. Therefore, three assumptions are made. First, the expected size of the cartel $E(m)$, with m being the number of cartelists, is defined as the rounded sum of joining probabilities. Second, π_{jpi} and π_{di} are computed for every firm i based on the assumption that the $m-1$ firms (excluding firm i) with the highest joining-probabilities form the cartel. i 's profits are computed both, when it joins the cartel (making π_{jpi} in periods without a price war) and when it remains independent (making π_{di}). Third, given the known industry characteristics and using $E(m)$ as a proxy for m , the price war probability P_{pw} is obtained from the cumulative density function of a standard-Normal distribution based on the results of the Probit-regression as shown in section 3.2.

The assumptions on cartel-size and cartel-composition are of particular importance. Firms can only attain deviation profits $\pi_{max,i}$ when the expected cartel encompasses all firms, i.e. $E(m)=n$, and firm i then decides to refrain. For symmetric firms $j_i = j_j \forall i,j$ holds, such that $\sum_i j_i = n \cdot j_i$. Since the expected cartel size is the rounded sum of participation probabilities, $n \cdot j = n - 0.5$ holds. Hence, for symmetric firms the optimal participation probability is

$$j_{sym}^* = \frac{n - 0.5}{n} \quad (27)$$

Any participation probability below j_{sym}^* would prevent the perfect cartel from being established. Any larger participation probability would lower expected profits (23) by putting too much weight on cartel profits. This is undesirable since cartel profits are smaller than deviation profits. The probability J that a cartel of size m is established in an industry with n symmetric firms is given by the following binomial distribution.

$$J(m) = \binom{n}{m} \cdot j_{sym}^{*m} \cdot (1 - j_{sym}^*)^{n-m} \quad (28)$$

This yields the, at the first sight, paradoxical result that a perfect cartel is more likely to be formed

in larger industries. At the second sight, this result can well be explained. The number of non-perfect cartel-combinations is higher in large industries than in small industries. Therefore, each firm must make a higher commitment, i.e. announcing a higher j^*_{sym} , in order to having a reasonable chance to ever make π_{max} .

3.2 Assessing the Probability of a Price War

Equation (23) indicates that firms' expected present value of profits is a function of the price war probability P_{pw} . This probability determines how often cartelists revert to a competitive equilibrium, which affects the profitability of the cartel. The basic idea, that is promoted here, is that firms use information about the occurrence of price wars in other collusive industries for forming an expectation about the price war probability in their own industry.

For doing so the following conditions must be satisfied. First, firms must know about collusion in other industries. Second, firms must be able to observe the characteristics of these industries that are decisive for cartelization. These are (i) supply-side characteristics (e.g. the number of firms, the degree of product homogeneity, firms' discount rate), (ii) demand-side characteristics, (iii) firms' production technology (e.g. cost parameters), and (iv) cartel characteristics (e.g. the number of cartelists, the competition authority's detection probability). Third, firms must observe whether and when a price war occurred. Such information can e.g. be obtained from competition authorities' decisions in cartel cases or from scientific journal articles on actual cartels.

Based on these observations firms can form expectations about P_{pw} by running a Probit-regression of the variable PW (with $PW = 1$ if a price war occurs and $PW = 0$ otherwise) on the above characteristics (i)-(iv) as explanatory variables. Based on the results of this Probit-regression firms find a proxy of P_{pw} for their specific industry which may then be used as an input in finding the optimal participation-probabilities.

This process of expectation-making is illustrated in the following. The above oligopoly model is used to randomly generate 50,000 different collusive industries.⁵ These industries' characteristics are drawn randomly from the intervals that are presented in Table 1. If firms' interest requirements r exceed the minimum competitive return on investment $\min_i((p_{ci}-c_i)/c_i)$ firms would exit the industry. Since entry and exit is beyond the scope of the above simulation model, none of such outcomes enters the below calculations. Cartelists and fringe firms are determined at random. Treating the cartel formation game as a game in mixed strategies there is a finite probability that the

5 The below Probit-regression has also been done for 5,000 simulated industries which does not strongly affect the regression coefficients and significance levels.

chosen cartel would be observed. As described in section 2.2, price wars are elements of the equilibrium-strategy for stabilizing cartels. Therefore, the observed industry-outcomes are equilibrium-outcomes. The minimum cartel-size is three firms, since deviation from a cartel with two members would give competitive profits and, thus, is not profitable.

	n	m	v	μ	a_1	a_3	a_4	a_5	P	r
Lower boundary	4	3	50	0	0.05	0	5	0.05	0.05	0.05
Upper boundary	20	n	150	100	0.9	1	15	0.15	0.4	0.25

Table 1: Intervals of Industry Characteristics

A price war is indicated in 0.1% of these collusive industries. This percentage might be used as a naïve proxy for P_{pw} . More accurate predictions can be derived from a Probit-regression⁶ whose results are presented in Table 2.

	c	n	m	μ	$\ln(\mu)$	a_1	a_3	r	P
Probit- β	-26.168	-1.719	2.238	-0.119	2.080	4.430	-1.831	11.682	31.394

Table 2: Parameter Values and Probit Regression

The McFadden- R^2 of the regression is 82.1%. The coefficients of variables v , a_4 , and a_5 were found to be statistically insignificant and, thus, were removed from the regression. This does not affect the remaining coefficients much but gives better information criteria. All other coefficients are statistically significant (based on a z-test) at the 1%-significance level. The results of this Probit-regression give a first set of findings.

- A.1. *Supply-side characteristics* – The price war propability P_{pw} rises (A1.1.) if firms are more impatient (higher r). (A.1.2.) The model also supports the finding of Ross (1992) that the transition from quite differentiated products to moderately homogenous products decreases cartel stability. This is because deviating firms win a larger share of the market as in the case of more heterogeneous products which increases the gain from deviation ($\pi_{di} - \pi_i$) as presented in equation (17). Moreover, Ross (1992) finds that a transition from moderately homogenous products to quite homogenous products increases cartel stability. This is because greater product homogeneity decreases price war profits and, thus, raises the punishment for deviation. The price war probability is lower in larger industries. This finding is commonly expected by economists (Davies and Olczak 2008: 176). However, it is

⁶ The results from a Logit-model are quite similar to those of the Probit-model. Hence, there is no reason for favoring the one or the other distribution.

not entirely obvious why P_{pw} decreases in the number of firms in the industry. Larger industries are characterized by a higher intensity of competition. Thus, they are characterized by both lower jointly maximized profits π_{jpi} and lower deviation-profits π_{di} . Therefore, the effect of n on the cartel-stability condition (17) can theoretically be ambiguous.

- A.2. *Firms' production technology* – The propensity to engage in a price war is higher (A.2.1.) if firms' costs are more asymmetric (lower a_3).
- A.3. *Cartel characteristics* – A price war is more likely (A.3.1.) if the cartel is larger (higher m) and, thus, harder to coordinate (see e.g. Kühn (2008: 115) or Levenstein and Suslow (2006: 58)), and (A.3.2.) if the antitrust authority is more effective (higher P).

3.3 Determining Equilibria by a Differential Evolution Heuristic Search Algorithm

Beck et al. (2007) show that the minimum of an objective function like (24) can be found by a stochastic search algorithm. The basic idea of such an algorithm is to start at a random point in the search space (i.e. a candidate vector of participation probabilities) and then converge towards the minimum of the objective function by iteratively examining and refining further points in the search space. The selection of these further candidate solutions is done on basis of two ideas. The first idea is, that the optimum should have some similarity to features of good points that have been examined throughout the search. The second idea is that by concentrating on these good, known candidate solutions one might miss further good points. Therefore, the search of an optimum should be carried out to some extent randomly in the vicinity of good candidate solutions. There are two alternatives to performing a stochastic search, either using a complete enumeration algorithm or a deterministic search algorithm.

A complete enumeration algorithm evaluates all possible candidate solutions. Such a search would be very costly in terms of computation time and, thus, is not favorable. Suppose that in the above case participation probabilities⁷ could only take integer values in the interval $[0;100]$, the number of firms was only four, and it would take about one minute to evaluate 50,000 candidate solutions. In this case, it would take a computer about 1.5 days to examine all 100^4 possible candidate solutions and find the solution to this simple problem when using a complete enumeration algorithm. It is shown below that a stochastic search algorithm – by stochastically generating and iteratively refining candidate solutions – does the same job in much shorter time (about one

⁷ Participation probabilities are measured in percentage points.

minute⁸) at the precision of floating numbers. Alternatively, for objective functions with a single optimum deterministic methods such as e.g. the method of steepest descent might be faster in converging to the optimum than a heuristic search method. However, the shape of the objective function is unknown for this n -dimensional problem. It is only known that the objective function makes discrete jumps when the expected size of the cartel changes. However, for a given cartel-size its shape is unknown such that one is advised to use a heuristic search method. In particular, I find that a Differential Evolution (DE) algorithm as proposed by Storn and Price (1997) works quite well in finding the global minimum of objective function (24). In the following, I provide an intuitive description of DE. A technical description of the DE-implementation as used in this paper is provided in appendix A.2.

DE belongs to the group of evolutionary algorithms. These algorithms consider different candidate solutions, i.e. vectors of participation probabilities \mathbf{j} , to be the DNA of individuals within a population. The members of a population may mate and pass on some part of their DNA (i.e. the participation probabilities of some firms) to their offsprings. This process is called crossover. In case of DE, an offspring is generated from four parents. First, the scaled (by a predefined scaling factor F) difference of two candidate solutions' participation probabilities is added to the vector of participation probabilities of a third individual. This generates a so called mutant vector. Second, some elements (here: firms' participation probabilities) of a fourth (so called target) vector are replaced with some predefined probability (i.e. the crossover rate CR) by the corresponding elements of the mutant vector. Mutation ensures that genetic diversity is not restricted to the DNA, i.e. the participation probabilities, in the starting population. If the objective function value of the generated offspring (the so-called trial vector) is lower (i.e. better) than that of the target vector (the parent), the offspring enters the new generation. Otherwise, the parent enters the new generation. This process of mutation and recombination is repeated for a predefined number G of generations. Since only the fittest individuals enter a new generation the algorithm converges towards the global minimum of the objective function. Such convergence does not mean that the algorithm attains the global optimum perfectly in each run. However, it is shown below that the DE algorithm, that is used here, arrives at the optimum very often with a high level of accuracy.

This implementation of DE is challenged by a Threshold Accepting (TA) algorithm. It is found that DE gives solutions with lower, i.e. better, objective function values than TA. TA as initially proposed by Dueck and Scheuer (1990) is chosen as a relevant alternative to DE since it is a variant of simulated annealing (SA). SA is found by Vorobeychik and Wellmann (2008: 1055) to be a good “general-purpose Nash equilibrium approximation technique[s] for infinite games”. Both

8 All evaluations were done on a PC running with 32-bit Windows Vista on a Intel Core2-architecture (3.00 Mhz) and 4 GB RAM. All programs were executed in Matlab version 7.7.0.

algorithms are run for 50 industries with symmetric firms and 50 industries with asymmetric firms. These industries' properties are randomly generated within the bounds provided by Table 1. Comparability of results is ensured by equalizing the number of candidate vectors ($= 50.000$) that are evaluated throughout the optimization. For DE the crossover and the scaling factor are chosen as $CR = F = 0.5^9$. This gives a second set of findings.

- B.1. In the case of asymmetric firms, DE is superior to TA by finding better solutions in 38 out of 50 cases. In these cases, objective function values that are obtained by DE are on average 41.87% lower than those obtained by TA.
- B.2. In the case of symmetric firms, DE dominates TA by finding better solutions in 36 out of 50 cases. In these cases, objective function values that are obtained by DE are on average 4.34% lower than those obtained by TA.

In section 3.1 it is argued that the minimum of objective function (24) constitutes a Nash-equilibrium of the cartel formation game. For symmetric firms it is shown that the optimal participation probabilities are $j^*_{opt} = (n-0.5)/n$. For being effective the DE-algorithm must (i) converge to this optimum solution and (ii) must do so irrespective of its starting conditions. In appendix A.3 it is shown that the DE-algorithm satisfies these two conditions. This is done by running the algorithm on ten randomly generated industries using five different starting populations. The industry characteristics are drawn randomly within the bounds provided by Table 1. The algorithm is run with 5,000 generations and $CR = F = 0.5$. This gives a third set of findings.

- C.1. The DE-algorithm converges to the optimum Nash-equilibrium in at least three out of the five runs. Some small variation in participation probabilities is economically insignificant and may be attributed to the stochastic nature of the search.

As a consequence, the DE-algorithm may be believed to converge to the optimum Nash-equilibrium of the participation-game for *asymmetric* firms, too. Therefore, it is run with the above parameter settings five times on ten randomly generated asymmetric industries.

- C.2. For these asymmetric industries the DE-algorithm converges to the same specific optimum for all industries. For seven industries, it converges to this optimum in all five restarts. Thus, the algorithm tends to work even better on asymmetric problems than on symmetric ones.

9 Ad-hoc evaluations indicate that results obtained by DE with different parameter settings are not miles apart. The focus of this paper is on the economic interpretation of results rather than obtaining them most efficiently. It is left to further research to fine-tune the search algorithm such as to obtain results with the same precision but fewer iterations, i.e. less computation time.

- C.3. Optimization results tend to be better for simpler optimization problems such as smaller industries. For more delicate problem sets, variation in results can be increased by raising the number of generations.

A more detailed description of this experiment's results is provided in appendix A.3.

3.4 The Economics of the Formation Equilibrium

The above techniques are used to derive knowledge on firms' decision to join a cartel. One prominent finding in this section is that large firms have a higher willingness to form a cartel than small firms. This finding is generated by econometrically relating the participation probabilities of 200 simulated industries to these industries' characteristics. These industries' characteristics are drawn randomly within the bounds provided by Table 1. Participation probabilities are obtained by running the DE algorithm (with 2,500 generations and $CR = F = 0.5$) on each industry. Since it is found that the significance-levels of estimated coefficients somewhat vary with the composition of the sample, standard errors and critical t-values are generated by a moving blocks bootstrap method.¹⁰ The idea of bootstrapping-methods is to construct multiple samples (here: 20,000) by randomly drawing 200 industries with replacement from the above sample. Then the regression is done for each of these 20,000 subsamples. This gives an empirically derived standard error as well as a distribution of t-values for each coefficient, which may then be used for statistical inference. More details on this method and the regression output are provided in appendix A.4. Evaluating the regression's results gives a fourth set of findings.

- D.1. Firms are quite likely to collude with a median participation probability of about 96%.
- D.2. Firms' participation decision depends most prominently on how much their marginal costs differ from average marginal costs in the industry. Firms with below-average marginal costs are found to have an above-average participation-probability. Firms with above-average marginal costs are found to have a below-average participation-probability. In this model, firms with below-average costs are also large firms, while firms with above-average marginal costs are small. The intuition underlying this finding is the following. When a large firm joins a cartel, it contributes much to raising prices. In comparison, staying independent is little profitable since – without the large firm – the other cartel members cannot raise the market price by (that) much. For a small firm, profitability effects are different. A small firm that joins a cartel will not raise the market price much. However, remaining independent is

¹⁰ For an introduction to this estimation method see e.g. Chernick (2008).

comparably profitable since the firm benefits from a high market price *and* an increased output by undercutting cartelists. Putting it more technically, solving equation (23) (i.e. a firm's) expected present value of profits for its participation probability gives

$$j_i = \frac{1 - H_i(\mathbf{j})/E(V_{di})}{1 - E(V_i)/E(V_{di})} \quad (29)$$

The ratio of the expected present value of cartel profits relative to the expected present value of fringe profits $E(V_i)/E(V_{di}) \in [0;1]$ is smaller for inefficient, small firms than for efficient, larger firms. Therefore, the participation probability j_i is c.p. higher for larger firms than for smaller firms.

- D.3. Collusion is less likely in industries that are characterized by market power even in competition (i.e. high values of the Herfindahl-Hirschmann index). In such industries the cartel generates smaller additional profits than in more competitive industries.
- D.4. The expectation to form a perfect cartel (i.e. all firms collude) provides an incentive for firms to join the cartel. However, larger cartels (relative to the number of firms in the industry) are not necessarily desirable since this raises the likeliness of a price war.

4 EVALUATION OF EMPIRICAL METHODS IN COMPETITION POLICY

The above methods are used to construct panel-data of a collusive industry for $T = 100$ periods and $n = 9$ firms. The cartel is formed among all firms but firms 3 and 8. The cartel is lived from period 21-100 without any price war periods. In the following, it is shown that the simulated data patterns are sufficiently reasonable such that the cartel can be detected by econometric screening methods like the tests for exchangeability and conditional independence proposed by Bajari and Ye (2003). Moreover, it is shown that simulated data of this type can be used to advance empirical methods applied in competition policy. In particular, Bajari and Ye's (2003) test for conditional independence – although originally proposed for auction markets – is made applicable for the detection of cartels in non-auction markets. For maintaining clarity the results are presented for two cartelists (1 and 2) and two fringe-firms (3 and 8) only.

4.1 Testing for Exchangeability

Bajari and Ye (2003) provide a screening method for collusion, i.e. their test for exchangeability. The basic idea underlying this screening method is to test the null hypothesis that firms' behavior is

consistent with a competitive model. In particular, it is tested if firms make their pricing decision independently. If this H_0 must be rejected competition authorities should conduct further examinations. Bajari and Ye's (2003) approach requires estimating a pricing-equation, i.e. regressing firms' prices p_i on a set of possibly explanatory variables X_i such as demand-side and cost-side variables.

$$p_i = X_i \beta_i + \epsilon_i \quad (30)$$

For the above simulated data a panel-regression of prices is run so that firm-specific coefficients are obtained for a constant, marginal costs, and the industry's Herfindahl-Hirschmann-Index (HHI) as computed from each firm's stake in industry-revenue. The results¹¹ are presented in Table 3. All estimated coefficients are statistically significant at the 1% level of confidence. In accordance with economic theory, cartel-firms tend to pass on marginal cost increases by less than fringe-firms. Moreover, an increase in concentration causes cartelists' prices to rise more strongly than fringe-firms' prices. This can be explained by the umbrella-effect, which says that competitive firms benefit from expanding price *and* quantity under the cartel's price umbrella.

firm	1 (cartel)	2 (cartel)	3 (fringe)	8 (fringe)
constant	-49.63 (0.51)***	-49.44 (0.51)***	-22.08 (0.51)***	-22.44 (0.51)***
marginal costs	0.9263 (0.0053)***	0.9318 (0.0054)***	0.9708 (0.0056)***	0.9894 (0.0058)***
HHI	460.45 (4.30)***	461.35 (4.32)***	210.77 (4.30)***	213.93 (4.33)***
adjusted R ² = 99.71%				

Table 3: Pricing Equation

A regression of the above form is used to detect collusion. This can be done by comparing firms' pricing behavior. Under competition, firm i is supposed to react on changes in the competitive environment (e.g. cost shocks) in the same way as any other firm j . Hence, if firm i was put in the same situation as firm j , i would act perfectly like j . Firms' reactions are *exchangeable* if

$$H_0: \beta_{ik} = \beta_{jk} \quad \forall i, j, i \neq j \quad (31)$$

and for all explanatory variables k . Such testing can be done using an F-test for all combinations of the four above firms. In the case of the above industry, for all firm-pairs the nullhypothesis of exchangeability must be rejected with a p-value of 0.0%. This finding is an artifact that is caused by the little noise in the simulated data, so that estimated coefficients' standard errors are quite low. Therefore, even similar looking pricing patterns must be accepted to differ when applying statistical tests. When the nullhypothesis is tested for each pair of coefficients separately – rather than jointly

11 Standard errors are given in parentheses. *, **, *** indicate that results are significant at the 10%, 5%, 1% significance level.

for firms' entire set of coefficients – one may well discriminate cartelists and competitive firms.

In principle, Bajari and Ye's (2003) test for exchangeability is shown to be a valid tool for indicating collusion. However, when taking the test literally, i.e. testing the similarity of coefficients jointly, would find all firms to behave differently. A remarkable fact is, that this failure is caused by the goodness of the above data and the estimated equation. In reality, one must neither expect marginal costs to be known nor firms' prices to be affected by such a limited number of variables. In reality, economic marginal costs can at best be approximated by average variable costs from accounting. This imperfectness of data attenuates the above problem but may cause that cartel-firms and fringe-firms cannot be discriminated because coefficients' standard errors are too large. Therefore, one can say that a test for exchangeability alone is no sufficient indication of collusion.

4.2 Testing for Conditional Independence

A second test for collusion, which is based on a pricing-regression of the above form, is a test of *conditional independence* (Bajari and Ye 2003). This test is based on the observation that in case of collusion the above regression is misspecified so that (some part of) the effect of collusion shows up in the error term ε_i . Thus, one should expect the error terms of colluding firms to be correlated to an above-normal extent. Such a correlation among error terms ε_i and ε_j can be measured by their correlation coefficient ρ_{ij} . Bajari and Ye's (2003) original concept deals with bid rigging in procurement auctions. In this case, conditional independence is given if one must reject

$$H_0: \rho_{ij}=0 \quad , \quad (32)$$

i.e. the null hypothesis of competition. Bajari and Ye (2003: 981) suggest to use the Fisher-test to test the hypothesis (32). This test is described in greater detail in appendix A.5.

In imperfect goods markets, as the one modeled here, the nature of competition differs from competition in auctions. Repeated interaction in goods markets reveals information about e.g. the pricing behavior of a firm's competitors which, in turn, affects its own pricing decision. I.e. if some firm raises its price for whatever reason, its competitors will readjust their prices accordingly. In the case of price fixing in imperfect goods markets, the above null hypothesis of complete independence of firms' behavior seems inappropriate. Therefore, one might rather want to test against a *normal* level of conditional dependence that is prevalent in oligopolistic competition. This normal level of independence can be found by running the above pricing-regression for the 20 competitive periods. The correlation coefficients of the error terms obtained from this regression lie in the range from [-0.404;0.975]. Therefore, one may only suspect that two firms collude if the correlation coefficient of their error terms from a pricing regression (including the supposedly cartelized periods) is significantly higher than the maximum level of conditional independence

under competition. The nullhypothesis can then be formulated as follows.

$$H_0: \rho_{ij} = 0.9750 \quad (33)$$

Results of a Z-test of the estimated collusive correlation coefficients against this null hypothesis of maximum normal dependence are presented in Table 4.

firm	1 (cartel)	2 (cartel)	3 (fringe)	8 (fringe)
1		0.9955***	0.9171***	0.7042***
2			0.9285***	0.7596***
3				0.8571***
8				

Table 4: Z-Test on Correlation Coefficients

All firm-pairs but the two cartelists are found to be below-normally dependent at the 1%-significance level. The two cartelists are above-normally dependent at the 1%-level of significance. With the standard nullhypothesis (32), which is appropriate for auction markets, one would have found all firm-pairs to be unnaturally dependent.

Consequently, Bajari and Ye's (2003) test for conditional independence is a powerful tool for complementing their test of exchangeability. For non-auction markets it is necessary to test firms' level of dependence against some normal, competitive level of dependence. Such a normal level can be found in two ways. On the one hand, it can be determined from a pricing-regression if one knows that the periods used have been competitive. On the other hand, it may be a feasible way to run an industry simulation like the one proposed here and obtain competitive correlation coefficients from the generated data.

5 CONCLUSION

In this paper I show how to simulate oligopolistic industries with differentiated products and asymmetric firms. A particular feature of this simulation model is that firms may endogenously decide whether they prefer to compete or to collude. Both these features, asymmetric firms and endogenous cartel formation, are elements that have not been explored extensively elsewhere, yet.

Firms' decision whether to form a cartel is modeled as a game in mixed strategies. This modeling assumption is made since in most cases it is more profitable for each individual firm to have others form the cartel and remain independent than joining the cartel. Thus, in pure strategies no cartel is formed. Giving firms' the chance to refrain from the cartel after negotiations about its formation widens their set of strategies. This greater variety of strategies (i.e. this mixed strategy) is the decisive element that motivates firms to meet at the bargaining table. It is shown that firms' optimal set of participation strategies is the one that jointly maximizes their expected present value

of profits. This Nash-equilibrium of the participation game can be found as the minimum of an objective function. The idea behind this function is to choose optimization strategies such as to minimize the difference between firms' expected profits and the maximum attainable profit over all possible outcomes of the game. Each firm obtains its individual maximum profit in a situation where all other firms form the cartel and only itself remains an independent fringe firm.

It is shown that the Nash-equilibrium of the game is characterized by a set of participation probabilities that (in expectation) motivates all firms to join the cartel. For symmetric firms the optimum participation probability is identical for all firms. Therefore, participation probabilities can easily be expressed analytically as an increasing function in the number of firms in the industry. The result of participation probabilities rising in the number of firms is counterintuitive only at first sight. At the second sight, it is clear that in an industry with many firms the number of non-perfect cartels is also large. Therefore, each firm must make a high commitment, i.e. announcing a high participation probability, in order to increase its chance for being the only fringe firm. For asymmetric firms the optimum cannot be expressed analytically but must be found numerically. I find that the optimum can be found by stochastic optimization algorithms at reasonable computational cost. In particular, a Differential Evolution algorithm is found to give more accurate results than a Threshold Accepting algorithm. Econometrically analyzing optimization results yields the prominent finding that large firms have a higher probability of joining the cartel than small firms. The intuition underlying this result is that a large fringe firm constrains collusive prices more than a small fringe firm. Therefore, participation in a cartel is above-average profitable for large firms as compared to remaining a fringe firm.

In making their participation-decision, firms must form expectations about the profitability of the cartel. These profits depend on industry characteristics, cartel-size, and the probability that a price war will occur in some period. In the paper, it is argued that firms can form expectations about the probability of a price war in their industry by (econometrically) analyzing other collusive industries. Doing such an analysis for simulated industries shows that large cartels with impatient firms are more instable. The probability of a price war is found to be U-shaped in the degree of products' homogeneity. One of the most important factors for cartel instability is an effective competition authority. However, despite this destabilizing effect on existing cartels, even effective competition authorities will not deter the formation of new cartels unless they impose perceptible fines on colluders.

Data, that was generated by this simulation model, is used to evaluate empirical methods for cartel detection. I find that Bajari and Ye's (2003) tests for *exchangeability* and *conditional independence* work well for detecting cartels. Additionally, I show that the test for conditional

independence can be extended to the analysis of collusion in imperfect goods markets by assuming some *normal* level of oligopolistic firm-interdependence. Such a normal level can e.g. be found from the above simulation model by calibrating it to the specifications of an industry of interest. Finding normal levels of firm-interdependence may be considered an interesting field for future research. One can summarize that data generated from the above simulation model is a valid and valuable ingredient for examining empirical methods for e.g. the detection of cartels.

The above simulation model provides a starting point for both the numerics-based theoretical analysis of cartels with asymmetric firms as well as the evaluation and advancement of empirical methods. From a *theoretical side*, further research should be devoted to broadening the variety of events that is covered by the model. Here, one may think of allowing the competition authority to set fines, with colluders being able to apply for leniency, and consumers having the opportunity to claim damages. Entry and exit of firms may be endogenized. Also, demand shocks and/or business cycles might be modeled. From the *technical side*, more research needs to be done in the area of determining Nash-equilibria by heuristic optimization methods. In particular, further fine-tuning of Differential Evolution or even different optimization techniques may further improve both the effectiveness and efficiency of the search for Nash-equilibria. From the *empirical side*, further econometric methods should be examined and evaluated, that are used in competition policy. Moreover, it may be worthwhile to examine whether the above finding of revenue-based market shares being superior to quantity-based market shares can also be validated for real data.

Summarizing one finds that simulation analyses of the above type are a valuable tool for researching collusion. On the one hand, numerical simulation methods are sufficiently flexible to model and theoretically examine the effects and interactions of a variety of influential factors. Examining these effects analytically may be much more complex (if being possible at all) than doing numerical analyses. On the other hand, the generated data is valuable in evaluating the power of empirical methods that are used in competition policy.

REFERENCES

- [1] Abrantes-Metz, R.M. and Froeb, L.M. and Geweke, J.F. and Taylor, C.T. (2006). "A variance screen for collusion." *International Journal of Industrial Organization*. Vol. 24, pp. 467-486
- [2] Aguirregabiria, V. and Mira, P. (2008). „Structural Estimation of Games when the Data Come from Multiple Equilibria.“ Working Paper.
<http://www.cireq.umontreal.ca/activites/papiers/08-09aguirregabiria.pdf>
- [3] Athey, S. and Bagwell, K. and Sanchirico, C. (2004). „Collusion and Price Rigidity.“ *Review of Economic Studies*. Vol. 71, pp. 317-349
- [4] Bajari, P. and Ye, L. (2003). "Deciding Between Competition and Collusion." *The Review of Economics and Statistics*. Vol. 85 No. 4, pp. 971-989
- [5] Beck, E.V. and Cherkaoui, R. and Minoia, A. and Ernst, D. (2007) "Nash equilibrium as the minimum of a function. Application to electricity markets with large number of actors." *Power Tech, 2007 IEEE Lausanne*, pp. 837-842
- [6] Bos, I. and Harrington, J.E. (2010). "Endogenous cartel formation with heterogeneous firms." *The RAND Journal of Economics*. Vol. 41 No. 1, pp. 92-117
- [7] Budzinski, O. and Ruhmer, I. (2008). "Merger Simulation in Competition Policy: A Survey." *MAGKS Joint Discussion Paper No. 07-2008*. http://www.uni-marburg.de/fb02/makro/forschung/magkspapers/index_html%28magks%29
- [8] Cameron, A.C. and Trivedi, P.K. (2005). "Microeconometrics – Methods and Applications." Cambridge University Press, 3rd printing 2006
- [9] Chernick, M.R. (2008). "Bootstrap Methods A Guide for Practitioners and Researchers." John Wiley & Sons Hoboken, 2nd editions
- [10] Compte, O. and Jenny, F. and Rey, P. (2002). "Capacity Constraints, Mergers and Collusion." *European Economic Review*. Vol. 46, pp. 1-29
- [11] Cramton, P. C. and Palfrey, T.R. (1990). "Cartel Enforcement with Uncertainty about Costs." *International Economic Review*. Vol. 31 No. 1, pp. 17-47
- [12] D'Aspremont, C. and Jacquemin, A. and Gabszewicz, J.J. and Weymark, J.A. (1983). "On the Stability of Collusive Price Leadership." *The Canadian Journal of Economics*. Vol. 16 No. 1, pp. 17-25

- [13] Davies, S. and Olczak, M. (2008). "Tacit versus Overt Collusion Firm Asymmetries and Numbers: What's the Evidence?" *Competition Policy International*. Vol. 4 No. 2, pp. 175-200
- [14] Deneckere, R. and Davidson, C. (1985). "Incentives to Form Coalitions with Bertrand Competition." *The RAND Journal of Economics*. Vol. 16 No. 4, pp. 473-486
- [15] Diamantoudi, E. (2005). "Stable cartels revisited." *Economic Theory*. Vol. 26, pp. 907-921
- [16] Dueck, G. and Scheuer, T. (1990). "Threshold Accepting: A General Purpose Optimization Algorithm Appearing Superior to Simulated Annealing." *Journal of Computational Physics*. Vol. 90, pp. 161-175
- [17] Fershtman, C. and Pakes, A. (2000). "A dynamic oligopoly with collusion and price wars." *The RAND Journal of Economics*. Vol. 31 No. 2, pp. 207-236
- [18] Friedman, J.W. (1971). "A Non-cooperative Equilibrium for Supergames." *Review of Economic Studies*. Vol. 38, pp. 1-12
- [19] Green, E.J. and Porter, R.H. (1984). "Noncooperative Collusion under Imperfect Price Information." *Econometrica*. Vol. 52 No. 1, pp. 87-100
- [20] Haltiwanger, J. and Harrington, J.E. (1991). "The Impact of Cyclical Demand Movements on Collusive Behavior." *The RAND Journal of Economics*. Vol. 22 No. 1, pp. 89-106
- [21] Harrington, J.E. (2008). "Detecting Cartels." in *Handbook of Antitrust Economics*. P. Buccirosi ed., The MIT Press, 2008
- [22] Hart, S. and Kurz, M. (1983). "Formation of Coalitions." *Econometrica*. Vol. 51 No. 4, pp. 1047-1064
- [23] Hinloopen, J. (2006). "Internal cartel stability with time-dependent detection probabilities." *International Journal of Industrial Organization*. Vol. 24, pp. 1213-1229
- [24] Johnston, J. and DiNardo, J. (2007). "Econometric Methods." McGraw-Hill International Editions, 4th edition
- [25] Kühn, K.U. (2008). "The coordinated Effects of Mergers." in *Handbook of Antitrust Economics*. P. Buccirosi ed., The MIT Press, 2008
- [26] Kuipers, J. and Olaizola, N. (2008). "A dynamic approach to cartel formation." *International Journal of Game Theory*. Vol. 37, pp. 397-408
- [27] Levenstein, M.C. And Suslow, V.Y. (2006). "What Determines Cartel Success?" *Journal of Economic Literature*. Vol. 44, pp. 43-95

- [28] Lofaro, A. (1999). "When Imperfect Collusion Is Profitable." *Journal of Economics*. Vol. 70 No. 3, pp. 235-259
- [29] Lyra, M. and Paha, J. and Paterlini, S. and Winker, P. (2009). "Optimization heuristics for determining internal rating grading scales." forthcoming in *Computational Statistics and Data Analysis*, pp. 1-14
- [30] Miklos-Thal, J. (2008). "Optimal Collusion under Cost Asymmetry." MPRA Paper No. 11044. <http://mpra.ub.uni-muenchen.de/11044/>
- [31] Motta, M. (2004). "Competition Policy – Theory and Practice." Cambridge University Press, 8th printing 2007
- [32] Olaizola, N. (2007). "Cartel formation and managerial incentives." *Spanish Economic Review*. Vol. 9, pp. 219-236
- [33] Porter, R.H. and Zona, J.D. (1993). "Detection of Bid Rigging in Procurement Auctions." *Journal of Political Economy*. Vol. 101 No. 3, pp. 518-538
- [34] Porter, R.H. and Zona, J.D. (1999). "Ohio school milk markets: an analysis of bidding." *RAND Journal of Economics*. Vol. 30 No. 2, pp. 263-288
- [35] Prokop, J. (1999). "Process of dominant-cartel formation." *International Journal of Industrial Organization*. Vol. 17, pp. 241-257
- [36] Quantitative Micro Software (2007). "Eviews 6 User's Guide II." Quantitative Micro Software, Irvine CA
- [37] Ross, T.W. (1992). "cartel stability and product differentiation." *International Journal of Industrial Organization*. Vol. 10, pp. 1-13
- [38] Rotemberg, J.J. and Saloner, G. (1986). "A Supergame-Theoretic Model of Price Wars during Booms." *The American Economic Review*. Vol. 76 No. 3, pp. 390-407
- [39] Selten, R. (1973). "A Simple Model of Imperfect Competition, where 4 Are Few and 6 Are Many." *International Journal of Game Theory*. Vol. 2 No. 1, pp. 141-201
- [40] Shubik, M. and Levitan, R. (1980). "Market Structure and Behavior." Harvard University Press, Cambridge
- [41] Spagnolo, G. (2008). "Leniency and Whistleblowers in Antitrust." in *Handbook of Antitrust Economics*. P. Buccirossi ed., The MIT Press, 2008
- [42] Spulber, D.F. (1995). "Bertrand Competition When Rivals' Costs Are Unknown." *The Journal of Industrial Economics*. Vol. 43 No. 1, pp. 1-11

- [43] Stigler, G.J. (1950). "Monopoly and Oligopoly by Merger." *The American Economic Review*. Vol. 40 No.2 Papers & Proceedings, pp. 23-24
- [44] Stigler, G.J. (1964). "A Theory of Oligopoly." *The Journal of Political Economy*. Vol. 72 No. 1, pp. 44-61
- [45] Storn, R. and Price, K. (1997). "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces." *Journal of Global Optimization*. Vol. 11, pp. 341-359
- [46] Thoron, S. (1998). "Formation of a Coalition-Proof Stable Cartel." *The Canadian Journal of Economics*. Vol. 31 No. 1, pp. 63-76
- [47] van Dijk, T. and Verboven, F. (2005). "Quantification of damages" forthcoming chapter for *Issues in Competition Law and Policy*. ABA Publications in Antitrust. Ed. W. Dale Collins. <http://www.econ.kuleuven.be/public/ndbad83/Frank/Papers/Van%20Dijk%20&%20Verboven,%202006.pdf>
- [48] Vasconcelos, H. (2005). "Tacit Collusion, Cost Asymmetries, and Mergers." *The RAND Journal of Economics*. Vol. 36 No. 1, pp. 39-62
- [49] Vorobeychik, Y. and Wellmann, M.P. (2008). "Stochastic search methods for nash equilibrium approximation in simulation-based games." *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - Volume 2*, pp. 1055-1062
- [50] Winker, P. (2001). "Optimization Heuristics in Econometrics – Applications of Threshold Accepting." John Wiley & Sons Ltd., Chichester

APPENDIX

A.1 The Basic Model in Non-Vector Notation

In the following, some functions are presented in scalar-notation that were presented in section 2.1 in vector-notation.

Utility function (1) can also be written as

$$U(q_1, \dots, q_i, \dots, q_n) = v \sum_{j=1}^n q_j - \frac{n}{2 \cdot (1 + \mu)} \left[\sum_{j=1}^n q_j^2 + \frac{\mu}{n} \left(\sum_{j=1}^n q_j \right)^2 \right] \quad (34)$$

so that the (inverse) demand function for a single firm i is given by

$$p_i = v - \frac{1}{1 + \mu} \left(n q_i + \mu \sum_{j=1}^n q_j \right) \quad (35)$$

$$q_i = \frac{1}{n} \left[v - p_i (1 + \mu) + \frac{\mu}{n} \sum_{j=1}^n p_j \right] \quad (36)$$

Equilibrium-prices for a non-colluding firm i can be written as

$$p_i = \frac{nv + (n + n\mu - \mu)c_i + \mu \sum_{j=1}^n p_j}{2n + 2\mu n - \mu} \quad (37)$$

While equilibrium-prices for a colluding firm i are

$$p_i = \frac{nv + (n + n\mu - \mu)c_i - \mu \left[\left(\sum_{j=1}^m c_j \right) - c_i \right] + \mu \left[\left(\sum_{j=1}^n p_j \right) + \left(\sum_{j=1}^m p_j \right) - p_i \right]}{2n + 2\mu n - \mu} \quad (38)$$

A.2 The Differential Evolution Algorithm

The Differential Evolution (DE) algorithm that is used in this paper is a variant of DE as proposed by Storn and Price (1997). The features of this specific algorithm are detailed in the following.

Each candidate solution is characterized by its vector of participation probabilities \mathbf{j} . The size of a population NP is predefined to be five times the number of firms n in the industry of interest. The number of generations G is selected by the researcher as well as the crossover probability CR and the scaling factor F . As is mentioned in footnote 9, ad-hoc evaluations indicate that results obtained by DE with different parameter settings are not miles apart. The focus of this paper is on economically meaningful, i.e. precise, results rather than obtaining them most efficiently. Therefore, the number of generations is generally set as relatively high levels. It is left to further research to fine-tune the search algorithm such as to obtain results with the same precision but fewer iterations, i.e. less computation time. A pseudocode of DE is provided below.

The search is started (1:) by generating a population GG_c of candidate vectors \mathbf{j}_g whose elements, i.e. firms' participation probabilities, are randomly chosen within the interval $[0;1]$. Then, the fitness (i.e. the objective function value) is computed for all candidates. Each individual \mathbf{j}_g now conceives an offspring by mating with three other members \mathbf{j}_{i1} , \mathbf{j}_{i2} , and \mathbf{j}_{i3} (with $g \neq i1 \neq i2 \neq i3$) of the population. This is done by, first generating a mutant vector \mathbf{j}_m (7:) from \mathbf{j}_{i1} , \mathbf{j}_{i2} , and \mathbf{j}_{i3} according to equation (39).

$$\mathbf{j}_m = \mathbf{j}_{i1} + \phi \cdot (\mathbf{j}_{i2} - \mathbf{j}_{i3}) \quad (39)$$

The factor ϕ controls the amplification of the differential variation. In Storn and Price (1997: 344) ϕ is a constant. Here, it is a random number that is determined from the interval $[0.9F; 1.1F]$ with $E(\phi) = F$. This somewhat altered version of DE enables the algorithm to generate slight changes in the differential variation. Therefore, the elements in the mutant vector are not restricted to the combinations of participation probabilities that exist in the current population.

In a second step, a trial vector \mathbf{j}_{ng} is generated (11:) by replacing participation probabilities in the target vector \mathbf{j}_g by the corresponding elements in the mutant vector \mathbf{j}_m . Each element is replaced with probability CR . Putting it in technical terms, the replacement is done if a random number $\in [0;1]$ is generated that is smaller than CR . The algorithm is designed such that $\mathbf{j}_g \neq \mathbf{j}_{ng}$ is ensured. The trial vector enters the following population GG_n if its fitness is better (i.e. its objective function value is lower) than that of the (old) target vector. Otherwise, the target vector \mathbf{j}_g enters GG_n . These steps are repeated for all candidate vectors in a population over G generations. Finally, the solution with the lowest objective function value is returned.

Algorithm 2: Pseudocode for DE

```

1:   Generate at random  $GG_c$  and compute  $of(\mathbf{j}_g)$  with  $\mathbf{j}_g \in GG_c$ 
2:   for  $\gamma = 1$  to  $G$ 
3:     if  $\mathbf{j}_g = \mathbf{j}_{-g} \quad \forall \mathbf{j}_g, \mathbf{j}_{-g} \in GG_c$ 
4:        $\mathbf{j}_i = \text{rand}(nx1)$ 
5:     end if
6:     for  $g = 1$  to  $NP$ 
7:        $\mathbf{j}_m = \mathbf{j}_{i1} + \phi (\mathbf{j}_{i2} - \mathbf{j}_{i3})$ 
8:       if  $\mathbf{j}_g = \mathbf{j}_m$ 
9:          $\mathbf{j}_m = \text{mean}(\mathbf{j})$ 
10:      end if
11:      generate  $\mathbf{j}_{ng}$  by combining  $\mathbf{j}_g$  and  $\mathbf{j}_m$ 
12:      if  $of(\mathbf{j}_{ng}) < of(\mathbf{j}_g)$ 
13:         $\mathbf{j}_{ng}$  enters the new population  $GG_n$ 
14:      else
15:         $\mathbf{j}_g$  enters the new population  $GG_n$ 
16:      end if
17:    end for
18:     $GG_c = GG_n$ 
19:    determine  $\mathbf{j}_{best}, \mathbf{j}_{worst} \in GG_c$ 
20:    generate  $\mathbf{j}_{mean} = \text{ones}(nx1) \cdot \text{mean}(\mathbf{j}_{best})$ 
21:    if  $of(\mathbf{j}_{mean}) < of(\mathbf{j}_{worst})$ 
22:      replace  $\mathbf{j}_{worst}$  by  $\mathbf{j}_{mean}$ 
23:    end if
24:  end for

```

It turns out that the effectiveness of the algorithm can be increased by selecting \mathbf{j}_{i1} , \mathbf{j}_{i2} , and \mathbf{j}_{i3} according to their fitness. Let \mathbf{j}_{rr} denote the candidate vectors in the current population GG_c with the index $rr = 1, \dots, NP$ denoting their rank. I.e. $\mathbf{j}_{rr=1}$ is the candidate vector with the lowest fitness. Then the probability PS_{rr} of selecting \mathbf{j}_{rr} as one of \mathbf{j}_{i1} , \mathbf{j}_{i2} , or \mathbf{j}_{i3} is denoted by equation (40).

$$PS_{rr} = rr / \sum_{\rho=1}^{NP} \rho \quad (40)$$

If the target vector and the mutant vector are identical they can only produce an identical offspring. In this

case, the mutant vector \mathbf{j}_m is replaced by the (firmwise) mean of participation probabilities of all candidate vectors in the current generation (9:). If all candidate solutions in the current generation are identical the first candidate vector is replaced (4:) by a vector whose elements are randomly generated from the interval [0;1].

In case of symmetric firms candidate solutions with identical participation probabilities are needed. DE's task to generate such candidate solutions is alleviated by the following routine. In every new generation, the candidate vector \mathbf{j}_{worst} with the lowest fitness is replaced by a candidate vector \mathbf{j}_{mean} whose elements equal the mean participation probability of the vector with the highest fitness \mathbf{j}_{best} . The new vector \mathbf{j}_{mean} replaces \mathbf{j}_{worst} only (22:) if its objective function value is better than that of \mathbf{j}_{worst} .

A.3 Evaluating Optimization Outcomes

In section 2.3 the joint (over all firms) maximum of equation (23) is proposed to constitute a Nash-equilibrium. For symmetric firms the optimum participation-probabilities can be derived analytically as shown in equation (27). In this section it is shown that the DE-algorithm converges reliably to this symmetric Nash-equilibrium. Therefore, it may be believed to also converge to a Nash-equilibrium in the case of asymmetric firms.

In order to show this, the DE algorithm is run five times on 20 randomly generated industries. Industries 1-10 are characterized by symmetric firms where firms in industries 11-20 are asymmetric. The algorithm is run with 5,000 generations and $CR = F = 0.5$. The population size is five times the number of firms. The results of the optimization for the symmetric industries are presented in Table 5 which provides the number of firms in an industry and the number of times the algorithm converges to the same solution. Moreover, the minimum, the median, and the maximum variation coefficient – i.e. the standard deviation of participation probabilities $\sigma_{j,i}$ divided by the mean participation probability \bar{j}_i computed for each firm i – is displayed besides the mean participation probability. Moreover, the variation coefficient, the mean objective function value $\overline{of(\mathbf{j})}$ and its standard deviation σ_{of} are presented. An interpretation of these results is provided in section 3.3.

industry		# runs to the same solution	participation probabilities				objective function values			
	# firms		$\min(\sigma_{j,i}/\bar{j}_i)$	$\text{median}(\sigma_{j,i}/\bar{j}_i)$	$\max(\sigma_{j,i}/\bar{j}_i)$	\bar{j}_i	$\sigma_{of}/\overline{of(\mathbf{j})}$	$\overline{of(\mathbf{j})}$	σ_{of}	
symmetric	1	8	5/5	0.0002	0.0006	0.0010	0.9375	0.0000	0.0767	0.0000
	2	13	4/5	0.0303	0.0377	0.0403	0.9462	0.2139	0.0172	0.0037
	3	6	5/5	0.0000	0.0000	0.0000	0.9167	0.0000	0.0566	0.0000
	4	14	3/5	0.0345	0.0418	0.0486	0.9357	0.3644	0.1199	0.0437
	5	15	3/5	0.0284	0.0403	0.0437	0.9400	0.4104	0.1958	0.0804
	6	10	4/5	0.0470	0.0481	0.0512	0.9300	0.5316	0.0378	0.0201
	7	13	5/5	0.0006	0.0019	0.0028	0.9615	0.0000	0.0141	0.0000
	8	13	4/5	0.0259	0.0348	0.0532	0.9462	2.2029	0.0506	0.1114
	9	15	4/5	0.0227	0.0316	0.0389	0.9533	0.3340	0.1216	0.0406
	10	15	3/5	0.0281	0.0419	0.0494	0.9400	0.7508	0.1109	0.0832
asymmetric	11	9	5/5	0.0004	0.0011	0.0015	0.9444	0.0000	0.0003	0.0000
	12	6	5/5	0.0000	0.0009	0.0030	0.9167	0.0000	0.1332	0.0000
	13	11	5/5	0.0000	0.0000	0.0001	0.9545	0.0000	0.0189	0.0000
	14	7	5/5	0.0002	0.0005	0.0007	0.9286	0.0000	0.0057	0.0000
	15	12	4/5	0.0106	0.0376	0.2728	0.9417	0.2911	0.2545	0.0741
	16	4	5/5	0.0012	0.0019	0.0026	0.8750	0.0000	0.0574	0.0000
	17	15	3/5	0.0346	0.0392	0.0472	0.9400	0.3559	0.1605	0.0571
	18	6	5/5	0.0002	0.0003	0.0006	0.9167	0.0000	0.0011	0.0000
	19	17	5/5	0.0000	0.0000	0.0001	0.9706	0.0000	0.1032	0.0000
	20	12	4/5	0.0308	0.0380	0.0534	0.9417	0.2829	0.0358	0.0101

Table 5: Optimization Evaluation

A.4 Bootstrap Estimation of Participation Probabilities

In this section, participation probabilities of firms in 200 randomly generated industries are regressed on a set of possibly explanatory variables. First, a standard OLS-regression is run, whose results are presented in Table 6. It is found that some coefficients vary quite strongly with the composition of the sample. In this case, standard statistical inference methods, which are based on asymptotic theory, may not be applicable in this finite sample. Therefore, and second, a bootstrap procedure is implemented that provides estimates of standard errors and critical t-values based on the sample properties and, thus, allows for more accurate inferences. The 200 industries' properties are determined randomly in the bounds provided by Table 1. Participation probabilities are obtained by running DE on each industry with 2,500 generations and $CR = F = 0.5$.

The idea of the bootstrap is to run the regression B times (here $B = 20,000$) for different samples that are generated from the above initial sample with 200 industries. This gives a distribution of values for each coefficient from which e.g. its standard error can be computed. Using a moving block bootstrap (Chernick 2008: 104), the B new samples are generated by randomly drawing 200 industries with replacement from the original sample. By drawing blocks of firms (= industries) rather than firms themselves, I treat industries as independent but allow for dependence of firms within industries. Running the below regression for each sample (indicated by index b) not only gives 20,000 values for the regression coefficients $\hat{\phi}_b^*$, that can be used to calculate its standard error $s_{\hat{\phi}_b^*}$. One also gets 20,000 t-values

$$t_b^* = (\hat{\phi}_b^* - \hat{\phi}) / s_{\hat{\phi}_b^*} \tag{41}$$

that are located around the original estimate $\hat{\phi}$. The nullhypothesis $H_0: \phi=0$ may now be rejected at confidence-level α if the test-statistic $(\hat{\phi}-0)/s_{\hat{\phi}}$ lies outside the range defined by the lower $\alpha/2$ and upper $\alpha/2$ quantiles of the ordered test statistics t^* . These intervals are provided in Table 6 for the 1%, 5%, and 10% confidence-levels.

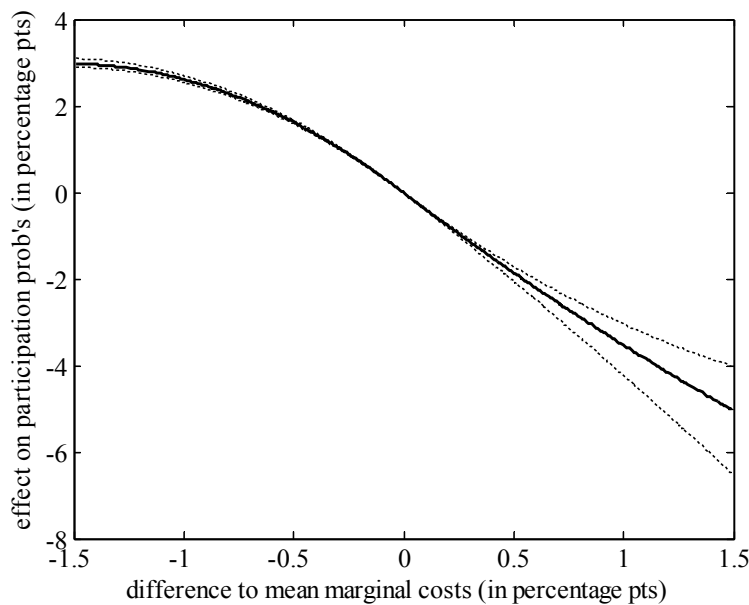


Figure 1: Scatter Plot - Change in Participation Probabilities

The results are interpreted in section 3.4. Participation probabilities j , the percentage deviation from mean marginal costs mc_dev , the detection probability P , the interest rate r , the variation coefficient of marginal costs (for each industry) σ_c/c , and the relative cartel size m/n are defined in percentage points i.e. in the interval $[0;100]$. The Herfindahl-Hirschmann index is defined in the interval $[0;1]$. Participation probabilities lower than 80 are outliers that are supposedly caused by DE's inability to always find the optimum solution for all firms. This effect is captured by a dummy variable. A further dummy variable captures the effect that participation probabilities are censored and cannot take values larger than 100. Figure 1 presents a plot that relates statistically significant changes in participation probabilities to the relative deviation of firms' marginal costs from mean marginal costs. The dotted lines are the respective upper and lower 95%-confidence bands.

dependent variable		participation probabilities (j)								
regressors	coefficient	t-values		bootstrap critical t-values						
		OLS	Bootstrap	1%		5%		10%		
constant	103.47	20.129 ***	15.589 ***	-1.127	1.101	-0.856	0.813	-0.712	0.680	
dummy for $j = 1$	3.476	11.916 ***	9.942 ***	-0.090	0.094	-0.068	0.070	-0.057	0.058	
dummy for $j < 0.8$	-25.493	-44.950 ***	-9.095 ***	-0.678	0.896	-0.535	0.671	-0.452	0.555	
dummy perfect cartel	5.078	10.476 ***	9.220 ***	-0.158	0.155	-0.119	0.118	-0.099	0.098	
deviation from \bar{c}	mc_dev	-3.877	-14.588 ***	-6.568 ***	-0.244	0.227	-0.188	0.167	-0.159	0.137
$(mc_dev > 0)^2$		0.342	2.412 ***	0.775 *	-1.743	1.159	-1.333	0.879	-1.109	0.749
$(mc_dev < 0)^2$		-1.263	-9.341 ***	-4.991 ***	-0.107	0.165	-0.083	0.117	-0.071	0.094
variation coeff. of c	σ_c/c	1.395	5.485 ***	4.449 ***	-0.157	0.209	-0.119	0.158	-0.099	0.134
degree of homogeneity	μ	0.010	1.927 **	1.167 ***	-0.019	0.018	-0.014	0.014	-0.012	0.012
cost parameters	a_1	0.399	0.463	0.377	-7.376	7.342	-5.540	5.612	-4.579	4.642
	a_3	0.418	0.974	0.942	-1.307	1.178	-0.977	0.863	-0.811	0.730
	a_4	0.008	0.255	0.196	-0.541	0.513	-0.413	0.392	-0.348	0.330
	a_5	1.312	0.377	0.344	-29.77	27.78	-22.57	20.84	-19.07	17.51
detection probability	P	-0.004	-0.411	-0.369 ***	-0.080	0.073	-0.060	0.057	-0.049	0.047
discount rate	r	-0.021	-1.141	-1.076 ***	-0.048	0.049	-0.036	0.036	-0.030	0.030
HHI		-26.521	-6.295 ***	-5.792 ***	-2.395	1.747	-1.714	1.385	-1.402	1.188
relative cartel size	m/n	-0.114	-2.152 **	-1.694 ***	-0.105	0.102	-0.081	0.075	-0.067	0.064
R^2	65.12%									
\bar{R}^2	65.36%									

Table 6: Participation Probabilities

A.5 Testing for Conditional Independence

Bajari and Ye (2003: 981) suggest to use the Fisher-test to test hypothesis (32) (respectively (33)) of conditional independence of firms' behavior. For being able to do statistical inference testing the correlation coefficient ρ_{ij} must be transformed into a variable Z whose distribution is known. Using Fisher's Z -transformation ρ_{ij} can be transformed into a normally distributed variable Z

$$Z = \frac{1}{2} \ln \left(\frac{1 + \rho_{ij}}{1 - \rho_{ij}} \right) \quad (42)$$

with expected value

$$\mu_Z = \frac{1}{2} \ln \left(\frac{1 + \rho_{ij}}{1 - \rho_{ij}} \right) + \frac{\rho_{ij}}{2(T-1)} \quad (43)$$

(i.e. $\mu_{H0} = 0$ under (32)) and standard deviation $\sigma_Z = (T-3)^{-1/2}$ with T being the length of the time-series. (32) must be rejected at confidence-level α if

$$|(Z - \mu_{H0}) \cdot \sqrt{T-3}| > u_{1-\alpha/2} \quad , \quad (44)$$

where $u_{1-\alpha/2}$ is the $(1-\alpha/2)$ -quantile of the standard-normal distribution.