Version 6.3 – November 24, 2010

ENDOGENOUS CARTEL FORMATION WITH HETEROGENEOUS FIRMS AND DIFFERENTIATED PRODUCTS

Johannes Paha*

ABSTRACT

Real-world collusive agreements do not necessarily include all firms that are active in the cartelized industry. Often, cartels are formed by asymmetric firms offering products that are not completely homogenous. This article endogenizes the process of cartel formation in a simulation model where firms are asymmetric in marginal costs and offer differentiated products. After a cartel is formed, it may be destabilized if the evolution of costs produces a sufficiently asymmetric cost-structure. In this case, the cartelists use price wars to stabilize the collusive agreement. Cartels arise as the outcome of a dynamic formation game in mixed strategies. This game is complex because of firms' asymmetry and the multitude of possible cartel-compositions. I show that the Nash-equilibrium of this game can be obtained efficiently by a Differential Evolution stochastic optimization algorithm. It turns out that large firms gain more from colluding than small firms. Nevertheless, large firms choose a lower probability of joining the cartel.

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, the 3rd RNIC conference, the CRESSE conference '10, and several workshops. I would like to thank all participants of these conferences and workshops – especially my discussants – for their valuable comments. A special thanks goes to Iwan Bos, Georg Götz, Joe Harrington, Peter Winker, my colleagues in Giessen, and numerous researchers who read the paper and gave me advise on it.

1 INTRODUCTION

The Belgian beer cartel is one example of a collusive agreement in a market where asymmetric firms offer a differentiated product. While much literature on collusion focuses on industries with symmetric firms and homogeneous products, this article proposes a model of collusion in markets with asymmetric firms and differentiated products. An important element in this model is the endogenous choice of firms whether to form a cartel. I find that – in terms of profits – large firms gain more from colluding than small firms. Interestingly, large firms optimally choose a low probability of cartel-participation. The reason for this is a strategic one. By choosing a low participation-probability large firms reduce the chance that small firms free-ride on the cartel. Additionally, I find that greater product-homogeneity raises the internal stability of a cartel whose combined market share is small. However, homogeneity decreases cartel-stability when the combined market share of the cartel is high.

These results are of particular importance as they are derived from a comprehensive but fairly standard model. The model is comprehensive because it allows for different degrees of product-differentiation and cost-asymmetry as well as a wide variety in the number of firms in the industry. The model is standard because it relies on well-known theories regarding product differentiation (Shubik and Levitan 1980), cartel stability (Friedman 1971, Rotemberg and Saloner 1986), and cartel formation (Prokop 1999).

Firms are assumed to produce differentiated goods (Shubik and Levitan 1980) at constant marginal costs that are firm-specific and evolve over time according to a random walk (Harrington 2008). By using a cost function of this type, I provide a fairly general model. Thus, its results regarding cartel-formation and cartel-stability neither depend on assumptions about dis-economies of scale (see, e.g., Donsimoni 1985, Vasconcelos 2005) nor on assumptions about the production capacity of the firms (see, e.g., Compte et al. 2002, Bos and Harrington 2010). These parameters are important in previous research. Colluding firms are assumed to maximize profits jointly. Therefore, my model does not require an assumption on the allocation of demand among cartelists as is the case in the model of Bos and Harrington (2010: 97).

The focus of this paper lies on endogenous cartel formation. At an exogenously given point in time, each firm in an industry may endogenously decide whether it participates in a newly established cartel or whether it prefers to remain in the competitive fringe. This decision is a function of firms' expected present value of profits both when joining the cartel and when remaining in the competitive fringe. As such, cartel formation is a function of the (expected) stability of the cartel. I assume that firms play a grim trigger strategy (Friedman 1971) in order to prevent deviations from the cartel. Allowing marginal costs to evolve over time changes the incentive of firms whether to deviate from a cartel. In particular, a deviation from a previously stable cartel may become profitable. In this case, firms are assumed to precautionarily lower the collusive price in order to make the deviation unprofitable (Rotemberg and Saloner 1986). In particular, firms are assumed to revert to the one-shot Nash-equilibrium prices. This trigger strategy generates time series of firms' prices that resemble those of a Green and Porter (1984)-type strategy. In contrast to the literature mentioned, price changes are not triggered by demand shocks but by observable cost shocks.

Section 2.3 presents a model of a cartel's endogenous formation. Previous research (see, e.g., Selten 1973, D'Aspremont et al. 1983, Diamantoudi 2005, and Kuipers and Olaizola 2008) shows that it can be profitable for firms to form a cartel when either the number of firms in the industry and/or the number of firms in the cartel is sufficiently small. However, in industries where both the overall number of firms and the number of cartelists is large, it can be more profitable to be a competitive fringe-firm rather than being a cartelist. This is because fringe-firms can expand their output under the price-umbrella of the cartel, while cartelists must reduce output in order to raise the market price. Prokop (1999) shows that even in industries with many firms a large cartel can arise as a Nash-equilibrium when firms play a mixed strategy in the formation game. I use this equilibrium-concept for modeling cartel-formation.

The assumption of cost-asymmetry practically prevents analytic solutions of the cartel formation game. Therefore, I resort to a new strand in the literature and propose to solve for the Nash-equilibrium of this complex game by means of heuristic optimization (Beck et al. 2007, Contreras et al. 2004, Vorobeychik and Wellmann 2008). In particular, I contribute to this literature by introducing a so-called objective function whose global minimum coincides with the mixed-strategy Nash-equilibrium of the cartel-formation game. I show that this Nash-equilibrium is attained fast and accurately by a Differential Evolution search heuristic (Storn and Price 1997). Using optimization techniques for solving the model requires its numerical parametrization.

Therefore, section 3.1 is concerned with formulating a parametric function for marginal costs. In particular, marginal costs are assumed to follow a random walk (Harrington 2008). This is especially convenient with regard to the theoretical model because firms' expectation about future marginal costs is their current level. This assumption keeps the structure of the model easy. I argue that firms take into account the stability of a cartel when deciding about their participation probability. A cartel is rather ineffective if cartelists frequently engage in price wars, i.e. if they frequently revert to the competitive equilibrium in order to prevent deviations from the cartel. Therefore, section 3.2 proposes a method for determining a numerical value of industries' price war probabilities and analyzes the determinants of cartelists' incentive to deviate from the collusive agreement.

To my knowledge, I am the first to analyze the effect of product differentiation on cartel stability in the presence of a competitive fringe. In the context of all-inclusive cartels, Ross (1992: 4) identifies two effects. “First, the segmentation of the markets through differentiation means that the gains from cheating are less. [... Second], the Bertrand-Nash punishments are less severe with differentiation.” In a model with two firms and a quadratic utility function, he finds a u-shaped relation between product homogeneity and cartel-stability. When products are very differentiated the gains from cheating are low relative to the punishment. When products are rather homogenous the punishment is high relative to the gain from cheating. In both cases, deviations from the collusive agreement are prevented quite effectively. Only for intermediate levels of product differentiation the gain from deviating is high relative to the punishment. This makes deviations more likely. My model shows that the relative strength of these effects also depends on the size of the competitive fringe. A strong competitive fringe constrains the deviation profit of a cartelist the more strongly the less differentiated products are. In this case, greater product homogeneity makes deviations less likely. If the competitive fringe is small this constraining effect is small, too. In this case, greater product homogeneity makes deviations more likely.

In section 4.1, I present the Differential Evolution stochastic search heuristic that is used to determine the Nash-equilibrium of the cartel formation game. The application of search heuristics for solving game theoretic problems is rather new (see, e.g., Beck et al. (2007), Contreras et al. (2004), and Vorobeychik and Wellmann (2008)). Methodologically, I contribute to this literature by introducing a distance measure for the formation-game in mixed strategies.

Section 4.2 analyzes the optimal participation decision of firms. In line with previous research (see, e.g., Donsimoni 1985, Bos and Harrington 2010) I find that large, cost-efficient firms gain more from colluding relative to remaining in the competitive fringe than small, cost-inefficient firms. A small firm “finds it optimal not to join the cartel [...] because the effects of its membership on price is trivial but, at the same time, it experiences a nontrivial reduction in its output” (Bos and Harrington 2010: 93). In contrast, by joining the cartel a large firm raises the level of market prices much even for a small relative reduction of quantity.

In addition to this research, I find that this effect makes small firms join the cartel with a higher probability than large firms. The reason for this is a strategic one. All firms are assumed to correctly anticipate (for themselves and all other firms) the effect on profits of joining the cartel or remaining in the fringe. If large firms joined the cartel with a high probability, small firms would prefer to remain in the fringe and free-ride on the cartel. This is disadvantageous for the cartel-firms. Therefore, large firms reduce their probability of joining the cartel. By reducing the expected size of the cartel, this lowers the profits of all firms both in the cartel and in the fringe. Now, being in the cartel becomes relatively more profitable as compared to being in the fringe even for small firms. Hence, they raise their participation probability. To summarize this main finding of my paper: By choosing low participation probabilities large firms punish and to some degree prevent free-riding behavior of small firms.

The numerical nature of the model adds a further benefit. The model allows for generating data on firms' prices, output quantities, and profits. In a companion paper (Paha 2010) I show that such data can be used to evaluate and advance empirical methods used in the detection and prosecution of cartels.

The remainder of the article is structured as follows. Section 2 presents the theoretical model. In particular, section 2.1 presents the basic model and its one-shot Nash equilibrium. Section 2.2 elaborates on the stability of cartels in the simulated industries. Section 2.3 describes the process of cartel formation. Section 3 is concerned with parameterizing the model so that it can be solved by a stochastic search heuristic. Section 3.1 formulates a parametric function for marginal costs. Section 3.2 numerically determines and analyzes cartels' likeliness of ending up in a price war. Section 4.1 shows how the Nash-equilibrium of the cartel-formation process can be obtained by a stochastic search algorithm. In section 4.2, I analyze this equilibrium. Section 5 concludes.

2 THE ECONOMIC MODEL

In this section, I present the economic theory underlying my model and its basic assumptions. The one-shot Nash-equilibrium 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 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 One-Shot Nash-Equilibrium

Stigler (1964: 45) proposes that homogeneous product markets are rather the exception than the rule. Therefore, cartels also may take place in (mildly) differentiated products markets. One example for such collusive agreement is the cartel in the Belgian beer market (1993-1998), where one would hardly argue that beer is a homogeneous product. A second example is the recent cartel of bathroom equipment manufacturers. In this cartel, firms colluded in differentiated products such as sinks, taps and fittings.

Assumption 1: To model these industries, my simulation model is based on the following utility function for differentiated products of a representative agent (Shubik and Levitan 1980).

Vectors and matrices are denoted in bold.

$$V = q_0 + U(\boldsymbol{q}) = q_0 + v\,\boldsymbol{q}'\boldsymbol{\iota} - \frac{n}{2(1+\mu)}\left[\boldsymbol{q}'\boldsymbol{q} + \frac{\mu}{n}\left(\boldsymbol{q}'\boldsymbol{\iota}\right)^2\right] \tag{1}$$

In this function q_0 is the outside option of the consumers. $\boldsymbol{q}$ is a (n×1)-vector whose elements are the quantities q_i of n products. Each product is produced by exactly one firm. Hence, there are n firms in the industry.

Assumption 2: The number of firms is assumed to be 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×1)-vector where each element takes a value of 1. v is a positive parameter and $\mu \in [0, \infty)$ represents the degree of substitutability[1] of the n products.

Consumers maximize utility subject to their budget constraint. Taking the outside option as numéraire, i.e. setting $p_0 = 1$, yields optimality condition (2).

$$\boldsymbol{p} = \frac{dU}{d\boldsymbol{q}'} \tag{2}$$

This gives the set of inverse demand functions whose right-hand side equals $dU/d\boldsymbol{q}'$.

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

The system of corresponding demand functions can be written as:

$$\boldsymbol{q} = \frac{1}{n}\left[v\boldsymbol{\iota} - \boldsymbol{p}(1+\mu) + \frac{\mu}{n}(\boldsymbol{p}'\boldsymbol{\iota})\boldsymbol{\iota}\right] \tag{4}$$

1 For $\mu = \infty$ goods are perfect substitutes. For $\mu = 0$ goods are independent. As this paper is interested in analyzing (imperfect) substitutes, μ is set at values greater than 0.

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 yields the reaction function of firm *i*.

$$p_i = \frac{c_i}{2} + \frac{n\nu + \mu \sum_{(j=1)\setminus i}^{n} p_j}{2\cdot(n+n\mu-\mu)} \tag{5}$$

Solving for the vector of competitive equilibrium-prices gives.

$$\boldsymbol{p} = \left(\boldsymbol{I}(2\mathrm{n}+2\mu n-\mu) - \mu\,\boldsymbol{\iota}\boldsymbol{\iota}'\right)^{-1}\cdot\left(\boldsymbol{\iota} n\nu + (n+n\mu-\mu)\right)\cdot\boldsymbol{c} \quad . \tag{6}$$

Firms 1 to *m* (with *m≤n*) may reduce competition by forming a cartel. If they do so, they maximize profits jointly (Stigler 1964: 44, 45). In this case, prices rise and equation (6) becomes

$$\boldsymbol{p} = \left[\boldsymbol{I}(2\mathrm{n}+2\mu n-\mu) - \mu(\boldsymbol{\iota}\boldsymbol{\iota}'+\boldsymbol{A})\right]^{-1}\cdot\left[\boldsymbol{\iota} n\nu + \left((n+n\mu-\mu)-\mu\boldsymbol{A}\right)\cdot\boldsymbol{c}\right] \tag{7}$$

where $\boldsymbol{A}$ is a *(n×n)*-matrix of the form shown in (8). The non-zero elements in the upper left part are of dimension *(m×m)* and stand for the effect of the joint profit maximization in the cartel.

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & & & & & \\ 1 & 1 & 0 & & & & & \\ \vdots & & & \ddots & & & & \vdots \\ 1 & & & & 0 & & & \\ 0 & & & & \cdots & & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & & & & \cdots & & & 0 \end{pmatrix} \tag{8}$$

Please note that equation (6) is a special case of equation (7) with $m = 0$, i.e. no collusion occurs.

> *Assumption 3*: Marginal costs have two features. (i) They are firm-specific which makes firms asymmetric. (ii) Cost-shocks are assumed to occur in every period such that marginal costs follow a random walk (Harrington 2008: 241).

Firm-specificity allows to analyze differences in firms' probabilities of participation in a cartel. The occurrence of cost-shocks introduces dynamics to the simulation model. A firm-specific evolution of costs not only affects the productive efficiency of the firms but also their competitiveness relative to each other. Such dynamics may create incentives to deviate from a collusive agreement that were not present upon the time of its establishment. This triggers price wars in order to mitigate the deviation incentive of the firms. Such price wars affect the profitability of the cartel. The exact parametrization of marginal costs in terms of the model's numerical nature is provided in section 3.1.

2.2 Cartel Stability

Prokop (1999: 248) proposes that firms play a game in two stages. “In the first stage, the firms simultaneously decide whether to join the cartel or remain in the competitive fringe.” In the second stage, firms set prices as described below. The game is solved recursively. Therefore, I start with defining the pricing decision of the firms and the resulting profits. I turn to the cartel formation game in section 2.3.

For the moment, I assume that the firms in the industry have established some cartel. Its composition is irrelevant for the discussions in this section. This assumption is relaxed in section 2.3. Concerning the cartel and the information structure of the model I make the following assumptions.

> *Assumption 4*: No more than one cartel exists at any time (Kuipers and Olaizola 2008: 407). The cartel does not necessarily have to consist of all n firms in the industry.

> *Assumption 5*: At the beginning of each period, firms observe their own marginal costs and those of their competitors.

> *Assumption 6*: It takes some time to detect a deviation from the cartel (see, e.g., Rotemberg and Saloner 1986). This time defines one period.

Assumption 4 is standard in the literature. Assumption 5 ensures that firms correctly anticipate their current, jointly maximized profits in the cartel π_{jpi}, their profits in competition π_{ci}, and the profits they make when deviating from the cartel π_{di}. According to Assumption 6, it takes some time for cartelists to establish evidence of a deviation by one of the cartel's members. During this time, cartel- and fringe-firms leave their prices at the collusive level (see equation (7)) while deviator i realizes his individual profit-maximum (π_{di}) at levels above his cartel-profits (π_{jpi}). This is done by lowering the price of good i according to the reaction function (5) of the deviator.

Cartels are called *internally unstable* (D'Aspremont et al. (1983: 21), Stigler (1964: 46)) if their members have an incentive to deviate from the collusive agreement. If the other firms were able to detect the deviation upon the time of its occurrence they would lower their prices and, thus, might render the deviation unprofitable. D'Aspremont et al. (1983) show that the anticipation of such immediate reactions may prevent firm i from deviating even in a one-shot game. Reactions of this type are beyond the scope of this paper.

> *Assumption 7*: Firms only consider the deviation-profits that they can make by being the sole deviator.

Theoretically, $d \in [1;m]$ cartel-firms can jointly deviate from a cartel with m members. This gives rise to potentially $\sum_{d=1}^{m} \binom{m}{d}$ new equilibria. Such joint deviations are difficult in two respects. (i) Allowing for joint deviations might create internally stable cartels. To see this, consider the extreme case where all cartelists defected from the collusive agreement by lowering prices at the same time. This contemporaneous deviation prevents the business steeling effect which makes deviations profitable. Only the price effect remains. Therefore, deviation-profits will be lower than collusive profits. (ii) There is no uncontested method to select one of these multiple equilibria. Therefore, I go with the more widespread solution (see, e.g., 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 stable if for any single firm there is no individual incentive to deviate from the cartel. In contrast, a cartel is called strongly internally stable if “for all possible coalitions of firms in the cartel there is no incentive to jointly defect from the cartel” (Olaizola 2007: 224). 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.

> *Assumption 8*: I assume that the cartel is stabilized by a grim trigger strategy (Friedman 1971) with price wars (Rotemberg and Saloner 1986).
>
> *Assumption 9*: I assume that the industry is infinitely lived. Respectively, the terminal date is uncertain.

For cartels of the above type, Friedman (1971) shows that deviations can be prevented if firms are sufficiently patient. This is the case if the cartelists credibly threaten potential deviators with reverting to the competitive equilibrium forever, once they have collected sufficient evidence for proving the deviation. In Assumption 6 the time for collecting this evidence is defined as one period. For deriving the below condition (12) for internal cartel stability, let the payoffs of the firms be defined as follows. The expected present value of competitive profits V_{ci} is given by equation (9).

$$V_{ci} = \pi_{ci} / (1-\delta) \tag{9}$$

This equation rests on Assumption 3 that marginal costs follow a random walk. Therefore, firms' best guess about costs and profits is their current values, i.e. the conditions $E(\pi_{jpi}) = \pi_{jpi}$, $E(\pi_{ci}) = \pi_{ci}$, and $E(\pi_{di}) = \pi_{di}$ apply. Future profits are discounted by a discount factor $\delta = 1/(1+r)$ with discount rate r. V_{di} defines the present value of deviation profits of firm i.

$$V_{di} = \pi_{di} + \delta \cdot V_{ci} \tag{10}$$

If cartelist i deviates from the collusive agreement, it enjoys deviation profits for one period only and is punished by a grim trigger strategy – i.e. reversion to the competitive equilibrium – in all subsequent periods (Friedman 1971). Equation (11) defines the present value of profits in the cartel V_{jpi}.

$$V_{jpi} = \pi_{jpi} + \delta[P \cdot V_{ci} + (1-P) \cdot V_{jpi}] \tag{11}$$

If firm i stays in the cartel, it makes profit π_i in the current period. At the beginning of each following period, the cartel may be discovered 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. Therefore, for future periods firm i expects to make the competitive present value of profits V_{ci} after a detection by the antitrust authority and the collusive present value of profits V_{jpi} otherwise. 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, standard condition (12) applies.

$$\frac{\pi_{di} - \pi_{jpi}}{\pi_{di} - \pi_{ci}} \leq \delta \cdot (1-P) \tag{12}$$

Cost shocks change the value of profits in equation (12) and may, thus, render previously stable cartels unstable. In this case, firms may stabilize the cartel by engaging in price wars. These are defined as periods where cartelists set prices in the range above or equal to the competitive price and below the jointly profit maximizing price. Such price wars have, e.g., been described by Green and Porter (1984), Rotemberg and Saloner (1986), and Haltiwanger and Harrington (1991). In particular, I rely on the strategy outlined by Rotemberg and Saloner (1986). While they assume observable, i.i.d. demand shocks, my model is concerned with observable, i.i.d. cost shocks. Therefore, their strategy can well be applied to my modeling structure.

Equation (13) shows the strategy that firms are supposed to pursue. The upper part shows that firms engage in joint profit maximization when neither cartelist has an incentive to deviate from the collusive agreement. The lower part applies if inequality (12) is not satisfied for at least one firm that, thus, has an incentive to deviate from the cartel. As costs are perfectly observable (see Assumption 5) this incentive is anticipated by the other cartelists. Therefore, they lower their prices until the incentive to deviate is eliminated for every firm.

$$\pi_i = \begin{cases} \pi_{jpi} & \text{if } \forall\, j \in [1,m] \;\; \dfrac{\pi_{dj} - \pi_{jpj}}{\pi_{dj} - \pi_{cj}} \leq \delta \cdot (1-P) \\ \pi_{ci} & \text{if } \exists\, j \in [1,m] \;\; \dfrac{\pi_{dj} - \pi_{jpj}}{\pi_{dj} - \pi_{cj}} > \delta \cdot (1-P) \end{cases} \tag{13}$$

In large industries with asymmetric firms and differentiated products it is challenging even for an omniscient researcher to determine the set of prices that equalize the left-hand side to the right-hand side of equation (12). 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. Therefore, I assume cartelists to overcome this problem by explicitly coordinating on the one stable solution that is known to every firm, i.e. setting competitive prices. Therefore, the strategy space of the firms is restricted to setting either jointly profit maximizing prices (making profit π_{jpi}) or engaging in a price war, i.e. setting competitive prices (making profit π_{ci}).

A simple pricing strategy of this form is also used by Fershtman and Pakes (2000: 213) for the same reason of computational complexity. 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.

The above strategy yields a self-enforcing internally stable cartel. Taking the model literally, cartels never break down unless they are discovered by the competition authority. In case of a perfectly ineffective competition authority, i.e. $P = 0$, a cartel would stay active forever once it is formed. However, even in this case it does not necessarily remain effective. If cartelists are sufficiently asymmetric, there is often an incentive for some firm to deviate. One observes cartelists to play the seemingly competitive price war-strategy. When costs evolve to become more symmetric, the cartel becomes effective again without any change in the composition of its members.[2]

2.3 Cartel Formation

After describing the second stage of the game, i.e. collusive price setting, I turn to the first stage, i.e. the decision of the firms whether to join the cartel or remain in the competitive fringe.

> *Assumption 10*: At some exogenously given point in time, firms meet to form a cartel. After its formation, the cartel is infinitely lived and perfectly stabilized by the pricing strategy described in section 2.2. After the establishment of cartel, there are

2 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 illegal. 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 because 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 as “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.

no re-negotiations about the composition of the cartel that might affect firms' current participation decision.

First, to motivate the assumption of an exogenous date of cartel formation consider that conferences and regular meetings organized by industry associations are exogenous to the cartel but provide the opportunity to set up a collusive agreement. Second, assuming that no re-negotiations about the cartel-composition occur is a plausible assumption for many cartels. It certainly does not apply to *all* cartels. However, for reasons of conciseness modeling a game with re-negotiations must be left to further research.

Each firm $i \in N$ has two pure strategies, $g_i = 0$ to stay in the competitive fringe and make an expected present value of profits $E(V_{fi}(\boldsymbol{g}_{-i}))$, and $g_i = 1$ to join the cartel and make an expected present value of profits $E(V_i(\boldsymbol{g}_{-i}))$. A cartel is formed if at least two firms decide to behave collusively. $\boldsymbol{g}_{-i}=(g_{1,}\cdots, g_{i-1}, g_{i+1}, \cdots, g_n)$ is an (n-1)-tuple whose elements display the strategies g_{-i} of the other firms. I denote the strategy space {0, 1} of firm i by G_i. A firm chooses the strategy g_i that maximizes its payoff function $h_i(\boldsymbol{g})$.

$$h_i(\boldsymbol{g})=h_i(\boldsymbol{g}_{-i}, g_i)=\begin{cases} E(V_i(\boldsymbol{g}_{-i})) & \text{if } g_i=1 \\ E(V_{fi}(\boldsymbol{g}_{-i})) & \text{if } g_i=0 \end{cases} \tag{14}$$

In the following, I define $E(V_{fi}(\boldsymbol{g}_{-i}))$ and $E(V_i(\boldsymbol{g}_{-i}))$ more closely. Equation (11) implies that the expected present value of cartel-profits $E(V_i(\boldsymbol{g}_{-i}))$ of firm i is a function of its current expected cartel-profit $\pi_i(\boldsymbol{g}_{-i})$.When no price war occurs, pricing strategy (13) implies that firm i makes the jointly maximized profit $\pi_{jpi}(\boldsymbol{g}_{-i})$. A price war occurs with probability P_{pw}. In this case, firm i only makes the competitive profit π_{ci}. Therefore,the expected cartel-profit $\pi_i(\boldsymbol{g}_{-i})$ of firm i is expressed as the weighted sum of the jointly maximized profit and the competitive profit.

$$\pi_i(\boldsymbol{g}_{-i})=(1-P_{pw})\cdot\pi_{jpi}(\boldsymbol{g}_{-i})+P_{pw}\cdot\pi_{ci} \tag{15}$$

By plugging condition (15) into equation (11) the expected present value of cartel-profits is found as

$$E(V_i(\boldsymbol{g}_{-i}))=\frac{(1-P_{pw}(\boldsymbol{g}_{-i},1))\cdot\pi_{jpi}(\boldsymbol{g}_{-i})+P_{pw}(\boldsymbol{g}_{-i},1)\cdot\pi_{ci}+\delta\cdot P\cdot\pi_{ci}/(1-\delta)}{1-(1-P)\cdot\delta} \quad . \tag{16}$$

$P_{pw}(\boldsymbol{g}_{-i},1)$ denotes the probability of a price-war when m-1 firms other than firm i form a cartel and firm i participates in this cartel. If firm i remains in the fringe, the probability of a price war is denoted by $P_{pw}(\boldsymbol{g}_{-i},0)$. In section 3.2 I show that price wars occur more frequently in large cartels and, thus, reduce the effectiveness of the cartel. Therefore, the effect of firm i's participation in the cartel on the likeliness of price wars is an important determinant of the firm's participation-decision.

Analogously to equation (16), the expected present value of profits in the fringe can be expressed as

$$E\left(V_{fi}(\boldsymbol{g}_{-i})\right)=\frac{(1-P_{pw}(\boldsymbol{g}_{-i},0))\cdot\pi_{fi}(\boldsymbol{g}_{-i})+P_{pw}(\boldsymbol{g}_{-i},0)\cdot\pi_{ci}+\delta\cdot P\cdot\pi_{ci}/(1-\delta)}{1-(1-P)\cdot\delta} \quad . \tag{17}$$

Equations (16) and (17) may be interpreted as follows. Firms make supra-competitive profits $\pi_i(\boldsymbol{g}_i)$ in the cartel and $\pi_{fi}(\boldsymbol{g}_i)$ in the fringe, as long as the cartel remains undiscovered by the competition authority, which occurs with probability 1-*P*, and as long as no price war occurs with probability P_{pw}.

Firm *i* considers collusion a dominant strategy and joins the cartel with probability $j_i = 1$ if it can make a higher expected present value of cartel-profits as when behaving competitively, i.e. condition (18) is satisfied for all possible cartel-combinations.

$$E\left(V_i(\boldsymbol{g}_{-i})\right)\geq E\left(V_{fi}(\boldsymbol{g}_{-i})\right)\forall\,\boldsymbol{g}_{-i} \tag{18}$$

If condition (18) is satisfied for all firms, an all-inclusive cartel is formed even if firms play a pure strategy. This is called an externally un̲stable cartel (D'Aspremont et al. 1983) as firms have an incentive to join the cartel irrespective of its size and composition. Selten (1973: 142) and, e.g., Prokop (1999: 253) show that condition (18) is satisfied in sufficiently small industries.

In many other cases, it can be more profitable for a firm to remain in the competitive fringe rather than joining the cartel. This is especially the case when the cartel encompasses a high proportion of the firms in the industry. The reason for this is that large cartels provide a positive externality to firms outside the cartel (see Deneckere and Davidson 1985, and Stigler 1950). Fringe firms may increase prices under the price umbrella of the cartel and *expand* quantity, while cartelists must *reduce* output in order to maintain a high price-level. Therefore, it can be more profitable for firms to stay outside the cartel than participating in the collusive agreement. Stigler[3] (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." However, a large fringe constrains the scope of the cartelists for raising prices. In this case, it might have been more profitable for the firm not to join the fringe but the cartel. In the latter case, the price-effect of participating in the cartel would have overcompensated the required reduction in quantity. Therefore, under the above assumptions, cartels of a certain size can be Nash-equilibria even if firms employ a pure strategy (for an example see Diamantoudi (2005: 909) or section 4.2 below).

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.

Such a Nash-equilibrium of a non-complete cartel is asymmetric even when firms are cost-symmetric. This is because fringe-firms can be shown to make higher equilibrium-profits than cartel-firms. Therefore, the decision of the firms whether to form a cartel is similar to Dixit and Shapiro's (1986) entry game of *n* firms in an industry that can only sustain *m* of them. Applying Dixit and Shapiro's (1986: 64) reasoning to the cartel-formation game means that firms engage in some vacillation: *I will regret joining the cartel if everyone else does too. But if other firms have the same fear and remain competitive we sacrifice profits and should rather form a (larger) cartel.* “The point at which each firm stops this process and plunges for one choice or the other will depend on many factors such as the organizational structure, the time available for decision, and even the moods of the executives. All these involve many chance elements, and modeling the outcome as a mixed strategy makes sense” (Dixit and Shapiro 1986: 64).

Assumption 11: Firms are assumed to play a mixed strategy when deciding about their participation in the cartel (Prokop 1999).

To further motivate this assumption, I start with describing the formal structure of the game and, then, elaborate on its economic interpretation. Denote the set of mixed strategies of firm i by $J_i = \{j_i: 0 \leq j_i \leq 1\}$, i.e. firm i joins the cartel with probability j_i and remains independent with probability $(1 - j_i)$. This builds up confidence among firms about the establishment of a cartel but leaves each firm with the chance to remain in the fringe. In such an environment a perfect cartel may even arise when condition (18) is not satisfied for all firms. When playing such a mixed strategy, each firm is interested in maximizing its expected payoff $H_i(\boldsymbol{j})$ (see equation (7) in Prokop (1999: 249)).

$$\begin{aligned} H_i(\boldsymbol{j}) \quad = \quad & j_i \cdot \left[\sum_{G \subseteq N-\{i\}} \prod_{-i \in G} j_{-i} \prod_{\substack{-i \notin G \\ -i \in N-\{i\}}} (1-j_{-i}) \cdot E\left(V_i(\boldsymbol{g_{-i}})\right) \right] \\ & + (1-j_i) \cdot \left[\sum_{G \subseteq N-\{i\}} \prod_{-i \in G} j_{-i} \prod_{\substack{-i \notin G \\ -i \in N-\{i\}}} (1-j_{-i}) \cdot E\left(V_{fi}(\boldsymbol{g_{-i}})\right) \right] \end{aligned} \qquad (19)$$

The upper bracket in (19) is the expected present value of cartel-profits of firm i, aggregated over all possible compositions of the cartel and weighted by their probability of occurrence. The lower bracket stands for i's weighted expected present value of fringe-profits. One can think of the mixed strategy in the following way. At a meeting of an industry association, the sales managers of the firms discuss the possibility of forming a cartel. They agree on participating in the cartel with probability j_i. However, the final decision is made by firms' CEOs, who do best to stick to the announced j_i's, when these are equilibrium outcomes.

In accordance with the above reasoning, the Nash equilibrium concept is applied to this game. A strategy n-tuple $\hat{\boldsymbol{j}}=\left(\hat{j}_{1,}\ldots,\hat{j}_{n}\right)$ is called a Nash equilibrium if for all firms i and for all $j_i \in J_i$ the condition

$$H_i(\hat{\boldsymbol{j}})\geq H_i(\hat{\boldsymbol{j}}_{-j},j_i) \tag{20}$$

applies where $\hat{\boldsymbol{j}}_{-j}=\left(\hat{j}_{1,}\ldots,\hat{j}_{j-1},\hat{j}_{j+1},\ldots,\hat{j}_{n}\right)$. Because an equilibrium in pure strategies is a special case of an equilibrium in mixed strategies, this condition even applies for the case of external instability, where firms have an individual incentive to join the cartel. In equilibrium, each firm chooses a participation probability such as to make the other firms indifferent between joining the cartel or remaining in the fringe. This is the case, when the weighted expected present values of cartel-profits and of fringe-profits are equalized for all firms.

$$\begin{aligned}&\left[\sum_{G\subseteq N-\{i\}}\prod_{-i\in G}j_{-i}\prod_{\substack{-i\notin G\\-i\in N-\{i\}}}(1-j_{-i})\cdot E\left(V_i(\boldsymbol{g}_{-i})\right)\right]\\=&\left[\sum_{G\subseteq N-\{i\}}\prod_{-i\in G}j_{-i}\prod_{\substack{-i\notin G\\-i\in N-\{i\}}}(1-j_{-i})\cdot E\left(V_{fi}(\boldsymbol{g}_{-i})\right)\right]\end{aligned} \tag{21}$$

By choosing a higher probability, a firm would forgive some of its chance to remain in the fringe and benefit from the positive cartel-externality on prices. By choosing a lower probability, the expected cartel would be too small to make remaining in the fringe a lucrative alternative. This is exactly the dilemma described by Dixit and Shapiro (1986: 64).

Equilibrium-condition (21) can be simplified by subtracting the right-hand side from the left-hand side. By plugging in expressions (16) and (17) for the expected values of profits, the equilibrium-condition can be stated in terms of the known stage-game profits. Multiplying the resulting function by the constant $(1-(1-P)\delta)$ does not affect the position of the function's root. Therefore, equation (22) is an alternative expression of the equilibrium-condition. By setting P_{pw}=0, equation (22) corresponds to equation (10) for symmetric firms without price wars in Prokop (1999: 250).

$$\begin{aligned}\psi_i(\boldsymbol{j}) \;=\; \sum_{G\subseteq N-\{i\}}\prod_{-i\in G}j_{-i}\prod_{\substack{-i\notin G\\-i\in N-\{i\}}}(1-j_{-i}) \cdot \Big\{ &\left[\pi_{jpi}(\boldsymbol{g}_{-i})-\pi_{fi}(\boldsymbol{g}_{-i})\right]+\\ P_{pw}(\boldsymbol{g}_{-i},1)\cdot &\left[\pi_{ci}-\pi_{jpi}(\boldsymbol{g}_{-i})\right]+\\ P_{pw}(\boldsymbol{g}_{-i},0)\cdot &\left[\pi_{fi}(\boldsymbol{g}_{-i})-\pi_{ci}\right]\Big\} \;=\; 0\end{aligned} \tag{22}$$

Result 1: The participation decision of a firm does not directly depend on the effectiveness of the competition authority P or on firms' discount factor δ.

In the Nash-equilibrium, condition (22) must apply for all firms. Therefore, the Nash-equilibrium is determined by the set of participation probabilities $\hat{\boldsymbol{j}}$ that constitutes the global minimum of a function Ψ as defined in equation (23).

$$\Psi(\boldsymbol{j})=\sum_i (1-\gamma_i)\cdot\psi_i^2 \tag{23}$$

In the Nash-equilibrium, i.e. the global minimum of Ψ, the condition $\Psi(\hat{\boldsymbol{j}})=0$ applies. To see this, consider that squaring ψ_i does not affect the root of function Ψ but prevents that positive values of ψ_i for firm i and negative values of ψ_{-i} for some firm $-i$ sum up to zero. Taking the sum of these squared terms ensures that a proposed solution $\hat{\boldsymbol{j}}$ constitutes a Nash-equilibrium for all firms. If some firm considers participation in the cartel a dominant strategy (i.e. equation (18) applies) one can pre-assign this firm a probability of $\hat{j}_i=1$. Firms of this type need not be considered in the calculation of other firms' optimal strategy. Therefore, an indicator-variable γ_i is defined that takes a value of 1 if some firm i considers cartel-participation a dominant strategy and a value of zero otherwise.

Defining this objective function is one of the main contributions of this article. This is because previous literature defines functions of this type for Nash-Cournot equilibria in pure strategies only (see, e.g., Beck et al. (2007), Contreras et al. (2004), Vorobeychik and Wellmann (2008) or Protopapas et al. (2010)). I argue that, from a computational point of view, determining a set of optimal participation probabilities is quite similar to the determination of a set of optimal output-quantities. Therefore, both problems can also be addressed by similar solution techniques as is shown in section 4.

3 PARAMETERIZING THE ECONOMIC MODEL

The main objective of this paper is to determine asymmetric firms' optimal probability of joining a cartel (see section 4). Prokop (1999: 252) analyzes the optimal participation probabilities of symmetric firms. His case is not too complex, as there is just one way for firms to be symmetric and only one optimal participation probability applies to all firms. Even this model is so difficult that it requires a solution by numerical simulation (Prokop 1999: 253). In contrast, firms may be asymmetric in infinitely many ways, which requires a determination of an optimal participation probability for each and every firm. Therefore, the asymmetric game also requires a numerical solution. Section 3.1 is concerned with describing the parametrization of firms' cost-function.

As can be seen from equation (22), the (expected) price war probability is an important determinant in firm i's decision of joining a cartel. When making their participation-decision firms trade-off two effects. First, by joining the cartel, firm *i* makes price wars more likely and lowers the share of jointly maximized profits in expected cartel-profits (see equation (15)). Second, by remaining in the fringe, firm *i* causes prices and profits to remain closer to their competitive levels. In section 3.2, I provide a method for numerically obtaining proxies of the price war probability for all possible cartel-combinations in particular industries. These proxies are used as input parameters for obtaining firms' optimal probabilities of participation in a cartel. This is shown in section 4. As the above model allows for analyzing the determinants of the occurrence of price wars, I additionally provide an economic analysis of these effects in section 3.2.

3.1 Parameterizing Marginal Costs

Assumption 3 states that marginal costs are (i) firm-specific and (ii) follow a random walk. A cost-function with these features is defined in this section. Marginal costs of firm *i*, i.e. $c_{i,t}$, are generated according to equation (24) in conjunction with conditions (25) and (26).

$$c_{i,t}=\begin{cases} a_1 \cdot v + a_{2i,t} \cdot s_t & \text{if } t=1 \\ c_{i,t-1} + a_{2i,t} \cdot s_t & \text{if } t>1 \end{cases} \tag{24}$$

$$\left.\begin{array}{lcl} a_1 & \in &]0;1[\\ a_{2i,t} & \sim & CN\left(\dfrac{a_3+1}{2}, \mu^{-2}, a_3, 1\right) \\ a_3 & \in & [0;1] \\ a_4 & \in &]0;1[\end{array}\right\} \tag{25}$$

$$s_t \in \begin{cases} \left[-a_4 \cdot (a_1 \cdot v); a_4 \cdot (a_1 \cdot v)\right] & \text{if } t=1 \\ \left[-a_4 \cdot \min_i (c_{i,t-1}); a_4 \cdot \min_i (c_{i,t-1})\right] & \text{if } t>1 \end{cases} \tag{26}$$

The base level of marginal costs, i.e. in the initial period $t = 1$, 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. First, the multiplicative, firm-specific technology-parameter $a_{2i,t}$ is drawn randomly from a censored normal distribution in the interval $[a_3;1]$. The expected value $E(a_{2i,t})$ is the mean of the interval $[a_3;1]$. Second, a realization of $a_{2i,t}$ is drawn in every period so that the asymmetry across firms changes over time. Third, the variance μ^{-2} of the technology-parameter $a_{2i,t}$ is assumed to depend on the product differentiation parameter μ.

Assumption 12: The variance of marginal costs is modeled to decrease in the degree of product homogeneity μ.

To motivate this assumption, consider that homogeneous products are produced by similar production technologies and, thus, at the same cost. If production costs differed, the less efficient firms would want to exit the industry. This process of convergence is assumed to have occurred in the past and, thus, is beyond the scope of the model. Product differentiation has two effects. On the one hand, differentiated products may require more diverse production technologies that cause differing costs of production. On the other hand, consumers' preference for differentiated products helps to sustain differences in production costs. The assumed distribution of $a_{2i,t}$ does not entirely prevent the generation of industries, where some firms would want to exit the industry. Such industries are discarded and do not enter the below analyses.

Cost shocks s_t are drawn randomly from a uniform distribution in the interval provided by equation (26). $a_4 \in [0;1]$ determines the amplitude of cost-shocks. Setting $a_4 = 0$ gives marginal costs that are symmetric across firms and constant over time. This interval ensures that marginal costs cannot become negative.

As in Harrington (2008: 241) marginal costs are assumed to follow a random walk in subsequent periods, i.e. $t > 1$. 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 may be considered a fluctuation in input prices. It is drawn randomly from a uniform distribution in the interval $\left[-a_4 \cdot min_i(c_{i,t-1}); a_4 \cdot min_i(c_{i,t-1})\right]$, where $min_i(c_{i,t-1})$ is the minimum (over all firms i) of last period's marginal costs. This ensures that marginal costs cannot become negative.

To summarize, the sign of the composite cost-shock $a_{2i,t}s_t$ is the same for all firms, because they produce substitute products with similar inputs. As these goods are imperfect substitutes, the production technology and, thus, the absolute value of the composite cost-shock differ across firms.

Using three fundamental input parameters, i.e. a_1, a_3, and a_4, for specifying the cost function must not be considered overly rich. This is because the application of three choice parameters is the most parsimonious way for generating a whole variety of different industries and simulate data on collusive and competitive behavior. This data may be used for evaluations of empirical methods that were proposed for the detection and prosecution of cartels. Paha (2010) is primarily concerned with such evaluations. With regard to firms' decision whether to participate in a cartel, only parameters a_1 and a_3 are found to have a decisive effect. Concerning the effectiveness of a cartel, it can be nicely seen from Table 2 below that a high level of marginal costs (a_1 is close to 1) and a more asymmetric evolution of costs (low values for a_3) reduces the effectiveness of a cartel by making price wars more likely. a_4 neither has a decisive effect on the decision of the firms whether to participate in a cartel nor on the effectiveness of a cartel.

3.2 Assessing the Probability of a Price War

In the above model, price wars arise naturally on the equilibrium-path of a cartel. They occur in order to prevent deviations from the collusive agreement. Cartels that are stabilized by frequent price wars may be considered rather ineffective. Therefore, let the ineffectiveness of a cartel be measured by its probability of entering a price war P_{pw}. Equation (15) illustrates that ineffective cartels (i.e. P_{pw} is high) do not generate much excess profits in comparison to a competitive situation. Therefore, the decision of firm i whether to join a cartel is a function of its collusive excess profits. This makes the participation decision depend on the price war probability P_{pw}, as can be seen from equation (22). Even the fringe-profits of a firm depend on the effectiveness of the cartel, because fringe-firms can only enjoy supra-competitive profits when the cartel is effective.

Consequently, when determining firms' optimal participation probabilities it is necessary to build expectations about the likely size of P_{pw}. Therefore, this section proceeds as follows. First, I present an approach for numerically determining price war probabilities in specific industries and cartels. These values are used in section 4 for determining firms' optimal probabilities of cartel-participation. Second, I analyze the economic determinants of P_{pw}. On the one hand, this is done by analyzing a sample-industry. On the other hand, I infer additional knowledge by econometrically analyzing the occurrence of price wars in 50,000 randomly generated industries.

I start with describing the method that is used to calculate a numerical proxy of P_{pw} for each possible cartel-composition in a particular industry. In particular, I calculate a numerical proxy of P_{pw} for each possible cartel-composition in an industry. For doing so, I use equation (24) to generate 200 vectors of firms' marginal costs. Then, I evaluate for each combination of marginal costs and each possible composition of cartels, whether some cartelist would find it profitable to deviate from the cartel. In an asymmetric industry 2^n different market structures can arise. Of those, n-1 are non-collusive because less than one firm decides to join the cartel. Each collusive market structure is assigned a price war probability P_{pw} equaling the frequency of observed price wars relative to the 200 simulated situations.

Using a number of 200 different combinations of marginal costs appears reasonable as this yields quite accurate predictions of P_{pw}. On the other hand, it leaves the computational burden at acceptable levels. To illustrate this trade-off between accuracy and computation-time, consider that for, e.g., n=10 firms one has to compute one competitive and 1013 collusive market outcomes. For each collusive market outcome, one needs to determine the incentive to deviate for on average about 5 cartelists. This is repeated 200 times. Therefore, one has to calculate about 1m equilibria of the stage-game only for approximating the price war probability P_{pw}.

After having presented the method for calculating a proxy of P_{pw}, I continue with a brief economic analysis of the determinants of price wars. In doing so, I resort to two types of analysis. First, I analyze firms' incentive to deviate from the collusive agreement in a particular collusive industry. This incentive is mitigated by price wars. Second, I infer additional knowledge from an econometric analysis of 50,000 different industries.

Cartelists engage in a price war if (at collusive prices) at least one cartel firm finds it profitable to deviate from the collusive agreement, i.e. inequality (12) is violated.

> *Result 2*: Inequality (12) indicates that a higher discount rate r and a higher probability of detection by the competition authority P c.p. destabilize cartels.

The profits at the left-hand side of condition (12) are complex terms. Theoretically, an analytical expression for these profits can be determined from the expressions of firms' prices and quantities (4)-(7). In practice, deriving an analytical expression of the left-hand side of stability condition (12) is far from trivial. However, further inferences on the determinants of the effectiveness of cartels can be drawn from the graphical analysis of a particular industry as is shown in the following.

Of particular importance is the effect of the relative cartel size m/n and product homogeneity μ. Figure 1 displays the effect of μ and m/n for an industry with ten symmetric firms and marginal costs c_i=v/4. On the vertical axis it maps the value of the left-hand side of inequality (12), i.e. the critical discount factor of the firms. The gray pane displays the right-hand side of (12), i.e. (1-P)/(1+r), for P=0.25 and r=0.10. Cartelists do not have an incentive to deviate from cartels where the gray pane lies above the critical discount factor. Interpreting Figure 1 yields the below results.

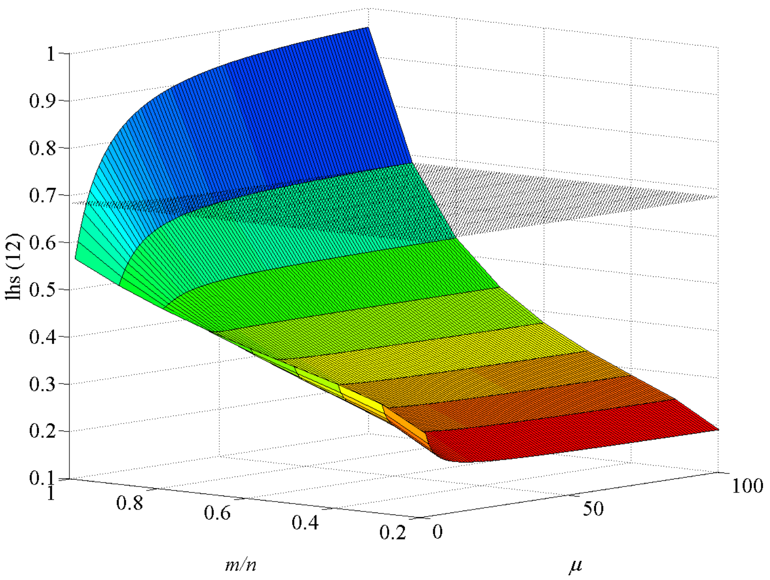

Figure 1: Condition (12) for ten symmetric firms

Result 3: When the cartel controls a *small* share of the market, greater product homogeneity *increases* the effectiveness of the cartel.

This is because in case of a small cartel the strong competitive fringe constrains the scope of the cartelists for raising prices. Therefore, neither the jointly maximized profits π_{jpi} of the firms nor their deviation profits π_{di} exceed competitive profits π_{ci} much. Greater product homogeneity lowers competitive profits and, thus, makes punishment harsher relative to the gain from deviating.

Result 4: When the cartel controls a *large* share of the market, greater product homogeneity *decreases* the effectiveness of the cartel.

Therefore, the absolute value of deviation profits is high because, in the case of a large cartel, the competitive fringe is weak. Moreover, the short-run gain from deviation ($\pi_{di} - \pi_{jpi}$) is higher for more homogeneous goods than for rather differentiated products. This is because firms win a larger share of the market when deviating from a cartel for more homogeneous products than by deviating from a cartel for quite differentiated products. The reason for this is the customers' preference for variety which cannot be overcome by simply lowering one's price.

These two effects have also been identified by Ross (1992) who, however, does not control for cartel-size. In analogy to Ross (1992), I find that the short-run gain from deviation is smaller for homogeneous goods than for mildly differentiated products. In contrast to Ross (1992), I find that this effect does not raise the stability of an all-inclusive cartel in a market for homogeneous products over its stability in a market for differentiated products. In my model, raising product homogeneity μ lowers the numerator of inequality (12) less strongly than the denominator.

Result 3 and Result 4 are in line with the below Probit-regression. A further result can be formulated for the impact of relative cartel-size on the effectiveness of the collusive agreement.

Result 5: Figure 1 shows that a price war is more likely if the cartel is larger (higher relative cartel-size *m/n*) and, thus, harder to coordinate (see, e.g., Kühn (2008: 115) or Levenstein and Suslow (2006: 58)).

Two additional results (Result 6 and Result 7) cannot easily be inferred from a graph like Figure 1. They are derived from a Probit-regression as shown in the following. Porter (1985) suggests to infer the probability of a price war P_{pw} from a Probit-regression that relates a variable PW (with $PW = 1$ if a price war occurred and $PW = 0$ otherwise) to relevant industry- and cartel-characteristics. Examples of 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 variables), and (iv) cartel characteristics (e.g., the fraction of cartelists, the probability of detection by the competition authority). The numerical nature of the above model makes it easy to apply such a regression approach for analyzing the determinants of cartel-stability.

I start with using the simulation model to randomly generate a cross-sectional dataset of 50,000 different collusive industries, i.e. I generate firm-level data at a single point in time.[4] Industries' characteristics are drawn randomly from the intervals given in Table 1.[5] At this point of the analysis, firms' optimal participation probabilities $\hat{j}$ have not yet been determined. Therefore, the identity of cartelists and fringe-firms is determined randomly with each firm being assigned a probability *m/n* of participating in the cartel. Treating the cartel formation game as a game in mixed strategies there is a finite probability that the 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. It is observed and stored in variable *PW* whether in the observed period a price war occurred in the observed industries (*PW* = 1) or not (*PW* = 0). A price war occurs in about 21% of the modeled cartels.

	n	m	v	μ	a_1	a_3	a_4	P	r
Lower boundary	3	2	50	0	0.05	0	0.05	0.05	0.05
Upper boundary	20	n	150	100	0.9	1	0.15	0.4	0.25

Table 1: Intervals of Industry Characteristics

After generating the data, *PW* is regressed on firm- and industry-characteristics (i)-(iv) using a Probit-regression. The results of this regression are presented in Table 2. The McFadden-R^2 of the regression is 82.31%. All coefficients are statistically significant (based on a z-test) at the 1%-significance level.

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

5 An economic explanation for the choice of these interval-borders is provided in Appendix B. Additionally, some literature (see, e.g., Li and Winker 2003) indicates that such a Monte Carlo approach for generating sample-industries might be outperformed by a quasi-Monte Carlo method where industries are generated deterministically. This is done for populating the space of industry characteristics more equally. Using such an advanced method appears not necessary in this context. This is because the results obtained when using the simpler Monte-Carlo method perfectly match the above theoretical results. Moreover, slight imbalances in the numerical value of the estimated coefficients do not affect the outcome of the model as these results are not used for any further calculations.

	constant	m/n	$(m/n)^2$	$(m/n)=1$	μ	$\ln(\mu)$	$(m/n)\ln(\mu)$	$\bar{c}/v$	$(\bar{c}/v)^2$	r	P	$\sigma_c/\bar{c}$	$\mu \cdot \sigma_c/\bar{c}$
Probit-β	-3.597	-22.371	19.739	0.995	-0.018	-0.283	1.782	-4.64	7.387	6.409	10.21	51.682	11.475

Table 2: Parameter Values and Probit Regression

Because the below analysis only concentrates on the sign of the estimated coefficients, I do not calculate marginal effects. The regression shows that the economic model (see section 2) in conjunction with the assumptions on marginal costs (see section 3.1) gives economically sensible results as it confirms the above results 2-5. Only the effect of the relative cartel-size on the effectiveness of the agreement cannot that nicely be seen from the above Probit-regression, as the coefficient of the size-variable m/n is negative. However, the coefficient of $(m/n)^2$ is positive so that the overall effect is positive for m/n>57%. This is because of the regression's inadequacy to capture the economic model's non-linearities. In addition, the Probit-regression provides two further findings that cannot easily be inferred from a graph like Figure 1.

> *Result 6*: Firms have a higher incentive to deviate from a cartel when the marginal costs of production are close to consumers' prohibitive price. This is measured by the ratio of average marginal costs $\bar{c}$ to variable v.
>
> *Result 7*: The propensity to engage in a price war is higher if the costs of the firms are more asymmetric. Asymmetry is measured by the variation-coefficient of marginal costs, i.e. the standard deviation of marginal costs σ_c divided by their mean $\bar{c}$.

4 SOLVING AND ANALYZING THE ASYMMETRIC FORMATION-GAME

In section 2.3, I show that the set of participation probabilities $\hat{\boldsymbol{j}}$, which globally minimizes the function $\Psi(\boldsymbol{j})$ (see equation (23)), coincides with the Nash-equilibrium of asymmetric firms' cartel-formation game. Determining $\hat{\boldsymbol{j}}$ is a non-trivial problem because an analytical expression of $\Psi(\boldsymbol{j})$ is quite complex and is hard to solve for $\hat{\boldsymbol{j}}$. Hence, Prokop (1999) determines the Nash-equilibria numerically even for the simpler problem of symmetric firms without price wars. Therefore, I also obtain numerical solutions for the cartel-formation game of asymmetric firms with price wars.

One possibility for determining the global minimum of $\Psi(\boldsymbol{j})$ is to evaluate all possible combinations of participation probabilities. This is a so-called complete enumeration algorithm. Such a search is very costly in terms of computation time and, thus, is not favorable. To see this, suppose that participation probabilities[6] could only take integer values in the interval [0;100], and

6 Participation probabilities are measured in percentage points.

the number of firms in the industry was ten. On a standard desktop computer[7] it takes about 1 minute to calculate the value of $\Psi(\boldsymbol{j})$ for 10^5 different $\boldsymbol{j}$-vectors. Therefore, it would take about 1.9bn years to calculate the value of $\Psi(\boldsymbol{j})$ for the 100^{10} possible combinations of participation probabilities and select $\hat{\boldsymbol{j}}$. This is quite some time.

In section 4.1, I show that the Nash-equilibrium of ten asymmetric firms' cartel formation-game is found by a Differential Evolution (DE) stochastic optimization heuristic in about 15 minutes at the precision of floating numbers. The idea of this algorithm is to start with a random set of different candidate vectors of participation probabilities and calculate the objective function values $\Psi(\boldsymbol{j})$ of these candidates. In a second step, new $\boldsymbol{j}$-vectors are generated by combining elements of the old ones. A new vector replaces an old one if the new $\boldsymbol{j}$-vector yields a better – i.e. lower – value of $\Psi(\boldsymbol{j})$. By concentrating on the evaluation of promising candidate vectors, the algorithm converges to the optimum of $\Psi(\boldsymbol{j})$. This causes the algorithm to be much faster than the complete enumeration.

In section 4.2, I show that large firms benefit more from collusion than small firms. However, small firms choose a higher probability of entering a cartel than large firms. At first sight, this result may seem counterintuitive. At second sight, the logic of this results becomes pretty clear, as it is driven by the strategic reasoning of the firms. The more firms stay in the competitive fringe, the lower are profits both in the cartel *and* in the fringe. Therefore, large firms choose a small participation probability. This imposes a threat on their small competitors, who anticipate the cartel and, thus, expected profits in the fringe to be small. That way, collusion becomes more profitable relative to staying in the fringe, and small firms choose a somewhat higher participation probability. In section 4.2, these points are illustrated in greater detail.

4.1 Determining Equilibria by a Differential Evolution Heuristic Search Algorithm

To my knowledge, Beck et al. (2007) provide the first contribution to determining Nash-equilibria of complex games by means of algorithmic optimization. I extend this literature by showing that the global minimum of equation (23) constitutes the Nash-equilibrium of the cartel-formation game and may be found by a stochastic search algorithm. As is standard in the optimization-literature, I refer to this function $\Psi(\boldsymbol{j})$ as the objective function.

The basic idea of a stochastic search algorithm is to start at a random point in the search space (i.e. a candidate vector of participation probabilities) and converge towards the minimum of

---

7 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. I will be happy to provide the Matlab-files on request. Please email me at johannes.paha@wirtschaft.uni-giessen.de.

the objective function. This is done 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, which 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 performed to some extent randomly in the vicinity of good candidate solutions. There are two alternatives to performing a stochastic search. These are using (i) a complete enumeration algorithm or (ii) a deterministic search algorithm.

A complete enumeration algorithm evaluates all possible candidate solutions. In the introduction to section 4, I argue that such a search is not favorable because it is very costly in terms of computation time. Therefore, the global optimum of the objective function should rather be obtained by a deterministic or stochastic optimization algorithm. The decisive advantage of stochastic search algorithms over deterministic search algorithms is their ability to overcome local optima in the search space. As the exact shape of the objective function is unknown for this n-dimensional problem, I use a Differential Evolution (DE) stochastic optimization algorithm (Storn and Price (1997) for determining the optimum of objective function (23). Below, I provide an intuitive description of DE. A more technical description of the DE-implementation as used in this article is provided in Appendix A.

Additionally, DE is advantageous as the generation of new candidate solutions is computationally quite efficient. Some prior evaluations indicate that a deterministic search algorithm, i.e. the method of steepest descent, often converges to the same optima as the below DE-algorithm. However, in the remaining cases the method of steepest descent has a tendency to end up in suboptimal corner-solutions (i.e. 0% or 100% participation probabilities for all firms). Moreover, for generating new candidate solutions it requires the calculation of the objective function's gradient-vector, which imposes some computational burden. This burden cannot necessarily be overcompensated by a reduction in the number of candidate solutions which must be evaluated for converging sufficiently close to the objective function's optimum. To summarize, DE appears to be more effective in finding the global optimum. Moreover, when fixing the computation time DE has a tendency to provide more accurate results than an algorithm searching for the steepest descent. Future research should be directed at (i) exploring the features of the objective function and (ii) fine-tuning a deterministic search method to the characteristics of this particular problem.

Additionally, within the group of stochastic search algorithms, DE seems to be suited particularly well to the determination of the global minimum of objective function (23). This result is attained by challenging DE by a Threshold Accepting (TA) algorithm. I find that DE yields 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 because 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".

After having explained the choice of a DE-algorithm, I intuitively describe the functioning of the algorithm. DE belongs to the group of evolutionary algorithms. These algorithms consider different candidate solutions, i.e. vectors of participation probabilities ***j***, to be the DNA of individuals within a population. The members of a population 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 difference of two candidate solutions' participation probabilities is scaled by a predefined scaling factor F. A so called mutant vector is generated by adding the scaled difference to the vector of participation probabilities of a third individual. Second, some elements (i.e. 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 of generations G. Because only the fittest individuals enter a new generation, the algorithm converges towards the global minimum of the objective function. Such a convergence does not mean that the algorithm perfectly attains the global optimum in every run. Nonetheless, it is shown below that the DE-algorithm arrives at the optimum both effectively and reliably. Appendix B provides details on this evaluation.

For being effective, the DE-algorithm must converge to the global minimum of objective function (23). For being reliable, it must do so irrespective of its starting conditions. DE is shown to satisfy these two criteria by evaluating its convergence in twenty randomly generated industries, using five different starting populations. Ten of these industries are characterized by symmetric firms. In the remaining ten industries, firms are asymmetric. The characteristics of the industries are drawn randomly within the bounds provided by Table 1. The number of firms n is restricted to the interval [11;13]. This keeps computation times at reasonable levels. The algorithm is run with CR=0.9, F=0.8, G=5,000 generations in the symmetric cases, and G=15,000 in the more complex, asymmetric cases.

For all ten symmetric industries, the DE algorithm converges to the same solution in all runs irrespective of its starting-population. These solutions are characterized by objective function values lower than 10^{-31}. As these values are sufficiently close to zero, a proposed solution may be considered the Nash-equilibrium of the cartel-formation game. As expected, the DE-algorithm returns (almost) identical participation probabilities for all symmetric firms. For further evaluating the quality of these solutions, I calculate the variation coefficient of firms' participation probabilities obtained in the five runs of the algorithm. None of these variation coefficients takes a value higher than 10^{-14}. This small variation is economically insignificant and may be attributed to the stochastic nature of the search.

For three out of the ten asymmetric industries, DE attains equally good results. In the seven remaining industries, DE converges to the same optimum in all runs but does not always exactly arrive there. To provide an example, the algorithm might return an optimal participation probability for a particular firm of 44.73% in one run and of 44.67% in another run. From the viewpoint of optimization-theory, this leaves room for further improvements of the algorithm. From the viewpoint of economics, the economic interpretation of these participation probabilities is practically the same. In summary, one finds that determining the optimum of objective function $\Psi(\boldsymbol{j})$ for asymmetric firms imposes a greater burden on DE as in the case of symmetric firms. This requires the researcher to increase the number of generations G and/or to fine-tune the parameter CR and F more thoroughly when applying DE to an asymmetric industry.

> *Result 8*: The Differential Evolution stochastic search algorithm is an appropriate means for determining the global minimum of objective function (23) and, thus, the Nash-equilibrium of the cartel-formation game.

4.2 The Economics of the Formation Equilibrium

This section analyzes the decision of firms whether to join a cartel. My research supports the hypothesis that small firms have a higher probability of joining the cartel although large firms have a higher benefit from collusion than small firms. First, I derive this hypothesis from the analysis of a particular industry. Second, I randomly generate 95 asymmetric industries and apply the DE-algorithm for determining firms' optimal participation probabilities. An econometric analysis of these participation probabilities confirms the hypothesis that small, inefficient firms choose higher optimal participation probabilities than large, efficient firms.

In order to illustrate the effect of asymmetries on firms' collusion-decision, I introduce a faint asymmetry in the sample-industry from section 3.2 and analyze the decision of these firms. The industry consists of n=10 firms. Firms 1-9 remain symmetric with marginal costs $c_l = v/4$. Firm 10 is assumed to be slightly less efficient with marginal costs $c_h = 1.02 \cdot c_l$. As this implies $p_h > p_l$, firm

10 sells less quantity and makes lower profits than firms 1-9. Hence, firm 10 is smaller than its more efficient competitors.

Below, I show that efficient firms prefer being in the cartel to being in the fringe even when the cartel is large. (However, it must not be too large.) When the cartel is large, inefficient firms often find it more profitable to be in the fringe than to be in the cartel. To see this, assume that firm i expects the other firms in the industry to behave according to the strategy-tuple $\boldsymbol{g}_{-i}$. Then, its profit in an effective cartel is $\pi_{jpi}(\boldsymbol{g}_{-i})$, and its profit when remaining in the fringe is $\pi_{fi}(\boldsymbol{g}_{-i})$. Firm i considers it individually profitable to participate in the cartel when condition (27) is satisfied.

$$\frac{\pi_{jpi}(\boldsymbol{g}_{-i})-\pi_{fi}(\boldsymbol{g}_{-i})}{\pi_{jpi}(\boldsymbol{g}_{-i})}\geq 0 \tag{27}$$

Figure 2 displays the value of the left-hand side of equation (27) for the efficient low-cost firm (solid lines) and the inefficient high-cost firm (dotted lines) for different degrees of product differentiation. The black lines apply for the case of μ=12, and the gray ones for μ=14. On the horizontal axis, I summarize the strategy-tuple $\boldsymbol{g}_{-i}$ by the number of cartelists m when firm i decides to participate in the cartel. Displaying the value of equation (27) for only up to seven cartelists allows for a reasonable scaling of the figure without limiting its economic interpretation. With regard to the composition of the cartel, I assume that the large, low-cost firm always colludes with the small high-cost firm and m-2 other low-cost firms.

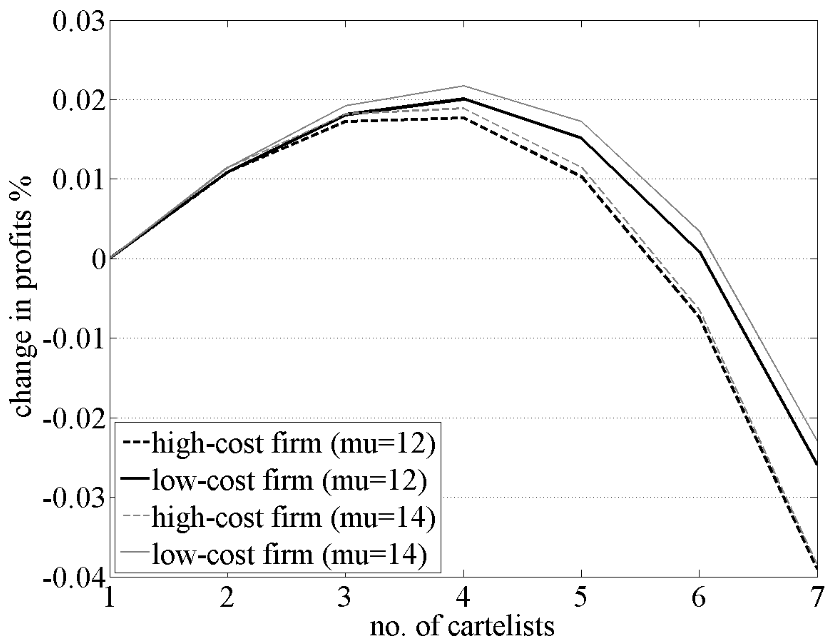

Figure 2: Incentive for joining the cartel

I start with analyzing the effects for μ=12 and comment on the differences of different degrees of product differentiation further below. When the number of cartelists is small (i.e. $m \leq 5$), both inefficient and efficient firms make higher profits in the cartel than in the fringe. In this case, the collusive profit of an efficient firm exceeds its profit in the fringe by 1.51%. The collusive profit of the inefficient firm exceeds its profit in the fringe by only 1.03%. The difference between these two values can be explained by the fact that the large, efficient firm achieves a particular increase in its price by a smaller relative reduction in its quantity than the small, inefficient firm. To see this, consider that the profit-maximizing increase in price when joining the cartel is about 6.4% for both firms. However, the efficient firm achieves the price-effect by reducing its output by only 22% while the inefficient firm must reduce its output by 23.2%. Therefore, collusion is more profitable for efficient firms.

When the cartel is large (i.e. $m \geq 7$), both firms prefer being in the fringe to being in the cartel. This is because a large cartel causes a perceptible increase in prices. At this elevated level of prices, consumers react more price-sensitive. Therefore, by joining a large cartel, a firm would have to reduce its output quite strongly in order to raise prices any further. In this case, it is more profitable for a firm to remain in the fringe and undercut the prices of the cartelists than colluding itself. In the above example, the inefficient firm finds it more profitable to be in the fringe when the cartel encompasses $m \geq 6$ firms. This is different for efficient firms who may affect market-prices more easily. An efficient prefers the fringe to the cartel only when the cartel is large. Given that the inefficient firm remains in the fringe, an inefficient firm finds it more profitable to be in the fringe when the cartel encompasses $m \geq 7$ firms. Based on the analysis of this particular industry, I formulate Hypothesis 1.

> *Hypothesis 1*: Collusion is more profitable for efficient firms than for inefficient firms. Efficient firms would want to form larger cartels than inefficient firms.

This hypothesis is perfectly in line with previous literature such as Bos and Harrington (2010) and Donsimoni (1985). While Bos and Harrington (2010) stress the effect of firms' size on their incentive to collude, Donsimoni (1985) analyzes the productive efficiency of the firms. I integrate both views, because in my model greater productive efficiency leads to a larger firm.

I proceed with analyzing the optimal cartel-participation probabilities of the firms in the above industry. It can be shown that a cartel of five efficient firms – i.e. the fringe consists of four efficient and one inefficient firm – constitutes a Nash-equilibrium of the cartel-formation game even when firms play pure strategies. Therefore, the efficient, symmetric firms would have to play asymmetric strategies, because some must stay in the fringe while others collude. Consequently, it is proposed in section 2.3 that firms play a mixed strategy. The DE-algorithm is used to determine the mixed-strategy Nash-equilibrium of the above industry, which is characterized by an optimal participation probability of $\hat{j}_h = 86.08\%$ for the small, high-cost firm and a probability of $\hat{j}_l = 39.08\%$ for the large, low-cost firm (and its look-alikes). Figure 3 provides a histogram showing the relative frequency of cartel-sizes given these participation probabilities.

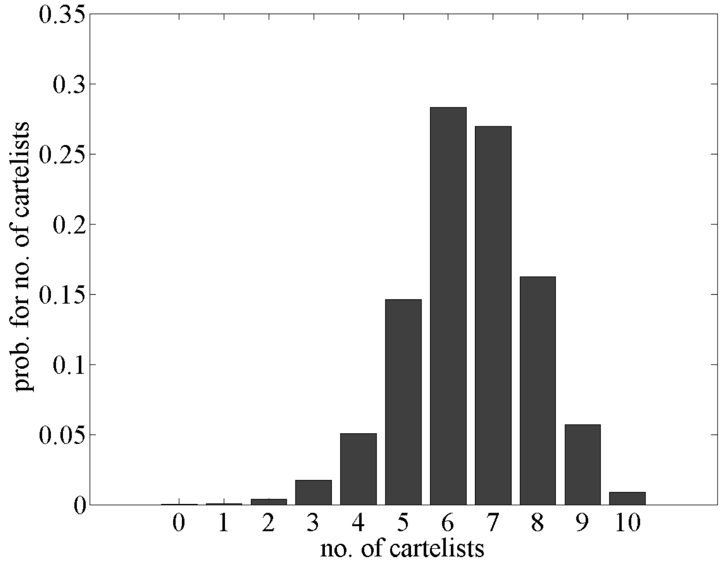

Figure 3: Histogram of the expected cartel size

Cartels with six members are found to be most frequent. Such cartels may also arise when firms play a pure strategy. However, playing a mixed strategy widens the set of cartels that may arise in this industry. This yields Result 9.

> *Result 9*: By playing a mixed-strategy in the cartel-formation game, firms can create larger cartels than by playing pure strategies.

To see the economic meaning of this result, suppose for a moment that small and large firms set identical participation probabilities $j_i = \hat{j}_h \forall i$. In expectation, this would generate relatively large cartels. Hypothesis 1 implies that, in this case, especially the small firm would (in expectation) find it profitable to remain in the fringe. This does not constitute a mixed-strategy Nash-equilibrium of the formation game. Therefore, the large firms lower their participation probability, which reduces the expected size of the cartel. This makes the fringe less desirable for the small firm, while the desirability of being in the cartel is raised. Graphically spoken, we move further to the left on the black, dotted line in Figure 2. This process takes place until all firms are made indifferent between joining the cartel or remaining in the fringe. Based on this analysis of a particular industry, I formulate a second hypothesis.

Hypothesis 2: Given Hypothesis 1, large firms choose a lower probability for joining a cartel than small firms.

Below in this section, I provide evidence that Hypothesis 2 also applies to industries other than the above example. Note that the two above hypotheses are interrelated and must apply together. Therefore, showing that Hypothesis 2 applies, means showing that Hypothesis 1 also applies.

Before demonstrating the generalization of the above hypotheses to other industries, I show that the qualitative interpretation of the hypotheses does not depend on the degree of product differentiation. In industries with less differentiated goods, the same effects apply but have a somewhat stronger impact. This can be seen from the gray lines in Figure 2. These represent the percentage increase in profits of being in the cartel relative to being in the fringe when the product-differentiation parameter takes a value of μ=14. In this case, the DE-algorithm returns an optimal participation probability of $\hat{j}_h = 96.33\%$ for the high-cost firm, while the low-cost firms choose $\hat{j}_l = 40.07\%$. These results allow for formulating two additional hypotheses.

Hypothesis 3: Collusion is more profitable in industries with similar products. Thus, firms choose higher participation probabilities in industries with homogeneous products than in industries with differentiated products.

Hypothesis 4: Product homogeneity increases the effect of cost-differences on the profits of the firms. This produces a greater spread in firms' participation probabilities.

Now, I turn to the second part of this section and show that the above hypotheses apply to more industries than the above sample-industry. The asymmetric and complex nature of the game prevents analytic solutions that would allow for a mathematical proof of the hypotheses. Therefore, I show that the above hypotheses apply to a large number of different industries, which firmly supports the validity of the hypotheses. In a first step, I generate 95 industries by randomly drawing their characteristics within the bounds provided by Table 1. The number of firms is determined in the interval from 11 to 13 firms. In a second step, I use the DE-algorithm (with *G*=10,000 generations, *CR*=0.9, and *F*=0.8) to obtain a participation probability for each firm in these 95 industries. In a third step, the participation probabilities of the firms are regressed on firm-, cartel- and industry-characteristics. This is done by using a standard OLS-regression in combination with a moving blocks bootstrap method.[8] The regression output and more details on the bootstrap-regression method are provided in Appendix B. With regard to its economic interpretation, the

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

regression provides the following central results.

Result 10: Firms with below-average marginal costs choose lower participation probabilities than firms with above-average marginal costs. This substantiates the above hypotheses 1 and 2.

Result 11: The effect of cost-dispersion on firms' participation probabilities becomes more pronounced the more similar the products of the firms are. This effect provides evidence for the validity of hypothesis 4.

Result 12: The propensity of firms to collude is higher in industries with homogenous products than in industries with differentiated products. This substantiates hypothesis 3.

Result 13: The participation decision of a firm depends indirectly, i.e. via the effect of the price war probability $P_{pw,}$ on the effectiveness of the competition authority P and on the discount rate r. Higher values of these variables imply lower participation probabilities.

5 CONCLUSION

In this article, I show how to simulate oligopolistic industries with differentiated products and asymmetric firms. A particular feature of this simulation model is that each firm may endogenously decide whether it prefers to compete or to collude. The decision about forming a cartel is made at an exogenously determined point in time. Both features, asymmetric firms and endogenous cartel formation, are elements that have not been explored extensively elsewhere.

The decision of the firms whether to form a cartel is modeled as a game in mixed strategies. This modeling assumption is made because in many cases it is more profitable for an individual firm to have others form the cartel and remain independent rather than joining the cartel. Thus, in pure strategies no cartel or only a small cartel is formed. Randomizing the participation decision of the firms widens their set of strategies. This greater variety of strategies (i.e. this mixed strategy) is a decisive element that motivates firms to form larger cartels than they would create when playing a pure strategy only.

In the Nash-equilibrium, every firm chooses a probability of participating in the cartel such as to make its rivals indifferent between joining the collusive agreement and remaining in the fringe. As a consequence, in equilibrium no firm has a systematic preference for staying in the fringe and, thus, free-riding on the cartel. Large firms, i.e. firms with below-average marginal costs, are found to choose lower participation probabilities than small, cost-inefficient firms. This result is interesting because in terms of profits large, efficient firms gain more from colluding (relative to staying in the fringe) than small, inefficient firms. This result on profits has also been shown by previous research. To my knowledge, I am the first to show that efficient firms should nonetheless choose a low probability of colluding. This is for strategic reasons. If the efficient firms chose high participation probabilities and, thus, showed a massive interest in colluding, the inefficient firms would find it individually profitable to stay in the competitive fringe and free-ride on the cartel's increased price-level. This behavior is harmful for cartelists. Hence, large firms choose a low participation probability for motivating their small rivals to join the cartel. The effect of dispersion in firms' participation probabilities becomes more pronounced when firms offer quite similar products.

These findings on the optimal participation probabilities are the central economic results of my paper. They are derived from a model that is, on the one hand, firmly grounded in economic theory while, on the other hand, being innovative with regard to the modeling of firm-size and cost-dynamics. First, firm-size is a result of cost-efficient production. Being productively efficient allows for setting lower prices and attracting more customers as compared to cost-inefficient firms. Thus, the model does not require an exogenous capacity-parameter for modeling firm-size. Second, as marginal costs are assumed to be subject to cost-shocks, they evolve over time. This dynamic structure leads to a variation of prices over time and triggers price wars in a cartel. In other models, this is modeled by assuming exogenous shocks to demand. Third, specifying marginal costs in the above way allows for numerical solutions of the cartel participation game. This is convenient as the asymmetry of the firms prevents an analytical solution of this game.

The Nash-equilibrium of the participation game is determined as the minimum of an objective function. Formulating this function is the main technical contribution of this paper. Moreover, I show that the minimum of the objective function can be obtained by a Differential Evolution stochastic optimization algorithm at reasonable computational cost. Econometrically analyzing optimization results yields the above finding that large, efficient firms have a lower probability of joining the cartel than small, inefficient firms.

The objective function shows that in the above model the participation decision of the firms does not directly depend on the effectiveness of the competition authority or the discount rate. First, this is because the relative profitability of being in the fringe – in comparison to participating in the cartel – is not affected by the value of the discount rate. Second, in this model without fines, an effective competition authority can only render the cartel ineffective but may not put the cartelists in a situation with profits below their competitive profits. Hence, the competition authority does not affect the profitability of being in the cartel relative to being in the fringe.

However, the participation decision of the firms indirectly depends on the discount rate and the effectiveness of the competition authority, because both parameters affect the probability that the cartelists engage in a price war. More frequent price wars make the cartel less effective and, thus, affect the participation decision of the firms. In the above model, I find that the probability of a price war rises in both firms' discount rate and the probability of detection by the competition authority. Consequently, both factors have an indirect negative effect on the firms' participation probabilities.

In addition to prior research, I find that cartels with a small combined market share are the more effective the more similar goods they offer. This is because a large competitive fringe strongly constrains the cartel's effectiveness when firms offer homogeneous goods. Therefore, a deviation from the small cartel is not very profitable, either. However, when the cartel controls a large share of the market it is least effective in industries for homogeneous goods. This is because a large cartel raises prices much. This makes a deviation quite profitable. Moreover, by deviating a cartelist wins a higher additional market share in markets for homogenous goods than in markets for differentiated goods.

Future research should be devoted to broadening the variety of effects 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 may be modeled. Additionally, one might relax the assumption of cartelists setting either competitive or jointly profit maximizing prices. It will be interesting to determine the set of maximum sustainable prices that *just* stabilizes the collusive agreement. Moreover, the model might be advanced in order to endogenize the time of cartelization, i.e. to predict *when* a cartel will be formed. In the existing model, cartel-firms do not have the chance to join the fringe after the establishment of the cartel. Similarly, fringe-firms may not belatedly join the cartel. Allowing for such changes in the cartel's structure may have an effect on the initial participation-decision of the firms. From the methodological side, more research needs to be done in the area of determining the Nash-equilibria of this game by optimization methods. In particular, further fine-tuning of Differential Evolution or other optimization techniques may improve the effectiveness and efficiency of the search for Nash-equilibria.

REFERENCES

[1] 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

[2] 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

[3] Chernick, M.R. (2008). “Bootstrap Methods A Guide for Practitioners and Researchers.” 2nd edition. New York: John Wiley & Sons

[4] Compte, O. and Jenny, F. and Rey, P. (2002). “Capacity Constraints, Mergers and Collusion.” European Economic Review. Vol. 46 No. 1, pp. 1-29

[5] Contreras, J. and Klusch, M. and Krawczyk, J. (2004). “Numerical Solutions to Nash-Cournot Equilibria in Coupled Constraint Electricity Markets.” IEEE Transactions on Power Systems. Vol. 19 No. 1, pp. 195-206

[6] 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

[7] 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

[8] 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

[9] Diamantoudi, E. (2005). “Stable cartels revisited.” Economic Theory. Vol. 26 No. 4, pp. 907-921

[10] Dixit, A. and Shapiro, C. (1986). “Entry Dynamics with Mixed Strategies.” in The Economics of Strategic Planning. Edited by L.G. Thomas. Lexington, MA: Lexington Books, 1986

[11] Donsimoni, M.P. (1985). “Stable Heterogeneous Cartels.” International Journal of Industrial Organization. Vol. 3 No. 4, pp. 451-467

[12] Dueck, G. and Scheuer, T. (1990). “Threshold Accepting: A General Purpose Optimization Algorithm Appearing Superior to Simulated Annealing.” Journal of Computational Physics. Vol. 90 No. 1, pp. 161-175

[13] 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

[14] Friedman, J.W. (1971). “A Non-cooperative Equilibrium for Supergames.” Review of Economic Studies. Vol. 38 No. 113, pp. 1-12

[15] Green, E.J. and Porter, R.H. (1984). “Noncooperative Collusion under Imperfect Price Information.” Econometrica. Vol. 52 No. 1, pp. 87-100

[16] 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

[17] Harrington, J.E. (2008). “Detecting Cartels.” in Handbook of Antitrust Economics. Edited by P. Buccirossi. Cambridge, MA: The MIT Press, 2008

[18] Hart, S. and Kurz, M. (1983). “Formation of Coalitions.” Econometrica. Vol. 51 No. 4, pp. 1047-1064

[19] Hinloopen, J. (2006). “Internal cartel stability with time-dependent detection probabilities.” International Journal of Industrial Organization. Vol. 24 No. 6, pp. 1213-1229

[20] Kühn, K.U. (2008). “The coordinated Effects of Mergers.” in Handbook of Antitrust Economics. Edited by P. Buccirossi. Cambridge, MA: The MIT Press, 2008

[21] Kuipers, J. and Olaizola, N. (2008). “A dynamic approach to cartel formation.” International Journal of Game Theory. Vol. 37 No. 3, pp. 397-408

[22] Levenstein, M.C. and Suslow, V.Y. (2006). “What Determines Cartel Success?” Journal of Economic Literature. Vol. 44 No. 1, pp. 43-95

[23] Li, J.X. and Winker, P. (2003). “Time Series Simulation with Quasi Monte Carlo Methods.” Computational Economics. Vol. 21 No. 1-2, pp. 23-43

[24] Olaizola, N. (2007). “Cartel formation and managerial incentives.” Spanish Economic Review. Vol. 9 No. 3, pp. 219-236

[25] Paha, J. (2010). “Empirical Methods in the Analysis of Collusion.” MAGKS Discussion Paper. No. 33-2010, pp. 1-32

[26] Porter, R.H. (1985). “On the Incidence and Duration of Price Wars.” The Journal of Industrial Economics. Vol. 33 No. 4, pp. 415-426

[27] Prokop, J. (1999). “Process of dominant-cartel formation.” International Journal of Industrial Organization. Vol. 17 No. 2, pp. 241-257

[28] Protopapas, M.K. and Battaglia, F. and Kosmatopoulos, E.B. (2010). “Coevolutionary Genetic Algorithms for Establishing Nash Equilibrium in Symmetric Cournot Games.” Advances in Decision Sciences. Vol. 2010, pp. 1-18

[29] Ross, T.W. (1992). “Cartel Stability and Product Differentiation.” International Journal of Industrial Organization. Vol. 10 No. 1, pp. 1-13

[30] Rotemberg, J.J. and Saloner, G. (1986). “A Supergame-Theoretic Model of Price Wars during Booms.” The Americam Economic Review. Vol. 76 No. 3, pp. 390-407

[31] 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

[32] Shubik, M. and Levitan, R. (1980). “Market Structure and Behavior.” Cambridge, MA: Harvard University Press, 1980

[33] Spagnolo, G. (2008). “Leniency and Whistleblowers in Antitrust.” in Handbook of Antitrust Economics. Edited by P. Buccirossi. Cambridge, MA: The MIT Press, 2008

[34] Stigler, G.J. (1950). “Monopoly and Oligopoly by Merger.” The American Economic Review. Vol. 40 No.2 Papers & Proceedings, pp. 23-24

[35] Stigler, G.J. (1964). “A Theory of Oligopoly.” The Journal of Political Economy. Vol. 72 No. 1, pp. 44-61

[36] 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 No. 4, pp. 341-359

[37] Vasconcelos, H. (2005). “Tacit Collusion, Cost Asymmetries, and Mergers.” The RAND Journal of Economics. Vol. 36 No. 1, pp. 39-62

[38] 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

APPENDICES

Appendix A The Differential Evolution Algorithm

The Differential Evolution (DE) algorithm that is used in this article 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 $\boldsymbol{j}$. As suggested by Rainer Storn[9] the size of a population NP is predefined to be 10 times the number of parameters, i.e. 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. Ad-hoc evaluations indicate that results obtained by DE with different parameter settings are not miles apart. The focus of this article is on economically meaningful results rather than obtaining them most efficiently. Therefore, the number of generations is generally set at relatively high levels. It is left to further research to fine-tune the search algorithm in order 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 $\boldsymbol{j_g}$ 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 $\boldsymbol{j_g}$ now conceives an offspring by mating with three other members $\boldsymbol{j_{i1}}$, $\boldsymbol{j_{i2}}$, and $\boldsymbol{j_{i3}}$ (with $g \neq i1 \neq i2 \neq i3$) of the population. This is done by, first generating a mutant vector $\boldsymbol{j_m}$ (7:) from $\boldsymbol{j_{i1}}$, $\boldsymbol{j_{i2}}$, and $\boldsymbol{j_{i3}}$ according to equation (28).

$$\boldsymbol{j_m} = \boldsymbol{j_{i1}} + \phi \cdot (\boldsymbol{j_{i2}} - \boldsymbol{j_{i3}}) \tag{28}$$

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 somewhat more diverse as in the case of a standard DE-algorithm.

In a second step, a trial vector $\boldsymbol{j_{ng}}$ is generated (11:) by replacing participation probabilities in the target vector $\boldsymbol{j_g}$ by the corresponding elements in the mutant vector $\boldsymbol{j_m}$. Each element is replaced with probability CR. Putting it in technical terms, the replacement is done if a random number is generated from the interval [0;1] that is smaller than CR. The algorithm is designed such that $\boldsymbol{j_g} \neq \boldsymbol{j_{ng}}$ 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 $\boldsymbol{j_g}$ 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.

It turns out that the effectiveness of the algorithm can be increased by selecting $\boldsymbol{j_{i1}}$, $\boldsymbol{j_{i2}}$, and $\boldsymbol{j_{i3}}$ according to their fitness. This is tested by comparing the performance (i.e. the obtained solution's objective function value when fixing the number of generations NP) of the above DE-algorithm with and without the fitness-based selection for several industries. To illustrate the fitness-based generation of offsprings, let $\boldsymbol{j_{rr}}$ denote the candidate vectors in the current population GG_c with the index rr = 1, ..., NP denoting their rank. As an example, $\boldsymbol{j_{rr}}$ with rr=1 is the candidate vector with the lowest fitness. Then the probability PS_{rr} of selecting $\boldsymbol{j_{rr}}$ as one of $\boldsymbol{j_{i1}}$, $\boldsymbol{j_{i2}}$, or $\boldsymbol{j_{i3}}$ is denoted by equation (29).

9 http://www.icsi.berkeley.edu/~storn/code.html#prac

$$PS_{rr} = rr / \sum_{\rho=1}^{NP} \rho \tag{29}$$

If the target vector and the mutant vector are identical they can only produce an identical offspring. In this case, the mutant vector $\boldsymbol{j_m}$ is replaced by the mean of every firm's participation probabilities across 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 $\boldsymbol{j}_{worst}$ with the lowest fitness is replaced by a candidate vector $\boldsymbol{j}_{mean}$ whose elements equal the mean participation probability of the vector with the highest fitness $\boldsymbol{j}_{best}$. The new vector $\boldsymbol{j}_{mean}$ replaces $\boldsymbol{j}_{worst}$ only (22:) if its objective function value is better than that of $\boldsymbol{j}_{worst}$.

Algorithm 2: Pseudocode for DE

1: Generate at random GG_c and compute $of(\boldsymbol{j}_g)$ with $\boldsymbol{j}_g \in GG_c$
2: **for** $\gamma = 1$ to G
3:   **if** $\boldsymbol{j}_g = \boldsymbol{j}_{-g} \quad \forall \quad \boldsymbol{j}_g, \boldsymbol{j}_{-g} \in GG_c$
4: $\boldsymbol{j}_1 = \text{rand}(n\text{x}1)$
5:   **end if**
6:   **for** $g = 1$ to NP
7: $\boldsymbol{j}_m = \boldsymbol{j}_{i1} + \phi\,(\boldsymbol{j}_{i2} - \boldsymbol{j}_{i3})$
8:   **if** $\boldsymbol{j}_g = \boldsymbol{j}_m$
9: $\boldsymbol{j}_m = \text{mean}(\boldsymbol{j})$
10:   **end if**
11:   generate $\boldsymbol{j}_{ng}$ by combining $\boldsymbol{j}_g$ and $\boldsymbol{j}_m$
12:   **if** $of(\boldsymbol{j}_{ng}) < of(\boldsymbol{j}_g)$
13: $\boldsymbol{j}_{ng}$ enters the new population GG_n
14:   **else**
15: $\boldsymbol{j}_g$ enters the new population GG_n
16:   **end if**
17:   **end for**
18: $GG_c = GG_n$
19:   determine $\boldsymbol{j}_{best}, \boldsymbol{j}_{worst} \in GG_c$
20:   generate $\boldsymbol{j}_{mean} = \text{ones}(n\text{x}1)\cdot\text{mean}(\boldsymbol{j}_{best})$
21:   **if** $of(\boldsymbol{j}_{mean}) < of(\boldsymbol{j}_{worst})$
22:   replace $\boldsymbol{j}_{worst}$ by $\boldsymbol{j}_{mean}$
23:   **end if**
24: **end for**

Appendix B Evaluating Optimization Outcomes and Bootstrap Estimation of Participation Probabilities

In section 2.3 the joint (over all firms) maximum of equation (19) is proposed to constitute a Nash-equilibrium that is found as the global minimum of equation (23). In this appendix, I show that the above DE-algorithm is appropriate to attain this Nash-Equilibrium for industries with either symmetric or asymmetric firms. Moreover, I analyze the determinants of firms' optimal participation probabilities. This is done based on a bootstrap-regression methodology.

In the following, industries are generated by randomly drawing values of the model's parameters within the bounds provided by Table 1. Therefore, an explanation for choosing these bounds is provided here. From the viewpoint of economic theory, the size of v is irrelevant as it only affects the scale of prices and quantities but has not impact on the ratio of profit-measures. Using μ_{upper}=100 as an upper bound is reasonable, as it suffices to give rather homogeneous goods. Choosing $a_1 \in [0.05,0.9]$ is reasonable because values below 0.05 would indicate that marginal costs are quite negligible. Such firms' production may be supposed to generate substantial fixed costs. These are beyond the scope of this model. If marginal costs were close to consumers' maximum willingness to pay (a_1>0.9) the entire market might break down when costs shocks drive marginal costs further upwards. Choosing $a_4 \in [0.05,0.15]$ yields economically meaningful, however, not unrealistically large cost shocks. Drawing P from the wide interval [0.05,0.4] reflects our lack of knowledge about the effectiveness of competition authorities. This is because one knows the number of discovered cartels but can hardly determine the number of undiscovered ones. The interval encloses the 15-20% detection probability that some studies suggest. Choosing $r \in [0.05,0.25]$ suggests that firms' discount rate is somewhere between the return of government bonds and some (ambitious) firms' target value of their return on equity. The number of firms is determined in the interval between 11 and 13 firms. The lower bound is chosen because in small industries firms always find it profitable to collude. The upper bound is determined such as to keep computation times at reasonable levels.

In order to show DE's appropriateness for finding the global minimum of equation (22), 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 CR=0.9, F=0.8, G=5,000 generations in the symmetric case and G=15,000 in the asymmetric case. The population size is ten times the number of firms. The optimization's results are presented in Table 3. The table 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 $\overline{j_i}$ computed for each firm i – is displayed besides the mean participation probability. Moreover, the variation coefficient, the mean objective function value $\overline{of(\boldsymbol{j})}$ and its standard deviation σ_{of} are presented. An interpretation of these results is provided in section 4.1.

industry		# runs to the	participation probabilities				objective function values		
	# firms	best solution	$min(\sigma_{j,i}/\bar{j}_i)$	$median(\sigma_{j,i}/\bar{j}_i)$	$max(\sigma_{j,i}/\bar{j}_i)$	$\bar{j}_i$	$\sigma_{of}/\overline{of(\boldsymbol{j})}$	$\overline{of(\boldsymbol{j})}$	σ_{of}
symmetric 1	12	5/5	1.55E-016	2.58E-016	3.74E-016	0.3999	0.3061	3.72E-034	1.14E-034
symmetric 2	12	5/5	2.60E-016	4.38E-016	7.85E-016	0.2499	0.3695	2.43E-034	8.99E-035
symmetric 3	13	5/5	2.33E-016	5.61E-016	1.00E-015	0.3566	0.1366	5.77E-033	7.88E-034
symmetric 4	13	5/5	3.81E-016	5.74E-016	1.24E-015	0.3716	0.4205	8.14E-036	3.42E-036
symmetric 5	12	5/5	1.20E-016	2.82E-016	7.08E-016	0.3998	0.3841	3.11E-036	1.19E-036
symmetric 6	12	5/5	1.57E-016	3.60E-016	7.45E-016	0.2499	0.4304	1.39E-033	6.00E-034
symmetric 7	13	5/5	1.93E-016	3.78E-016	6.83E-016	0.3521	0.1988	5.09E-034	1.01E-034
symmetric 8	12	5/5	9.89E-017	2.84E-016	5.19E-016	0.3967	0.3579	3.63E-034	1.30E-034
symmetric 9	12	5/5	1.79E-016	3.32E-016	5.55E-016	0.4091	0.1903	4.90E-037	9.33E-038
symmetric 10	12	5/5	1.99E-016	3.07E-016	6.79E-016	0.3944	0.2190	1.36E-035	2.98E-036
asymmetric 11	11	1/5	2.26E-003	4.53E-003	2.60E-002	0.4289	0.3317	1.49E-034	4.95E-035
asymmetric 12	11	1/5	1.67E-003	1.51E-002	4.18E-002	0.4206	0.3675	8.20E-033	3.01E-033
asymmetric 13	11	5/5	9.00E-017	2.12E-016	3.61E-016	0.4248	0.4238	3.45E-033	1.46E-033
asymmetric 14	13	5/5	2.46E-016	4.39E-016	6.61E-016	0.2756	0.1591	3.93E-033	6.25E-034
asymmetric 15	11	5/5	8.82E-017	1.90E-016	7.78E-016	0.4380	0.3069	9.26E-033	2.84E-033
asymmetric 16	11	1/5	5.46E-003	1.14E-002	3.31E-002	0.4144	0.2475	1.56E-032	3.86E-033
asymmetric 17	11	1/5	8.64E-001	1.18E+000	2.24E+000	0.2652	0.9284	2.12E-005	1.97E-005
asymmetric 18	12	4/5	1.46E-015	2.24E+000	2.24E+000	0.0517	2.2361	9.23E-005	2.06E-004
asymmetric 19	11	3/5	1.30E-001	1.92E-001	7.11E-001	0.3608	2.2319	1.13E-003	2.53E-003
asymmetric 20	12	1/5	1.40E-001	5.70E-001	8.87E-001	0.3969	2.2150	6.95E-005	1.54E-004

Table 3: Optimization Evaluation

In the following, participation probabilities of firms in 95 randomly generated industries are regressed on a set of possibly explanatory variables. I only regard industries where no firm has an incentive to employ a pure strategy. First, a standard OLS-regression is run, whose results are presented in Table 4. It is found that some coefficients vary somewhat 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 properties of the 95 industries are determined randomly in the bounds provided by Table 1. Again, the number of firms is determined in the interval between 11 and 13 firms. Participation probabilities are obtained by running DE on each industry with G=10,000 generations, CR=0.9, and F=0.8.

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 95 industries. This gives a distribution of values for each coefficient from which, e.g., its standard error can be computed. Using a moving blocks bootstrap (Chernick 2008: 104), the B new samples are generated by randomly drawing 95 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 yields 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

$$t_b^* = (\hat{\phi}_b^* - \hat{\phi}) / s_{\hat{\phi}_b^*} \tag{30}$$

that are located around the original estimate $\hat{\phi}$. The null hypothesis $H_0 : \phi = 0$ may now be rejected at confidence-level α if the test-statistic $(\hat{\phi} - 0)/s_{\hat{\phi}_b^*}$ lies outside the range defined by the lower α/2 and upper α/2 quantiles of the ordered test statistics t*. These intervals are provided in Table 4 for the 1%, 5%, and 10% confidence-levels.

The results are interpreted in section 4.2. Participation probabilities *j*, the percentage deviation from mean marginal costs *mc_dev*, the detection probability *P*, and the interest rate *r* are defined in percentage points, i.e. in the interval [0;100]. ($mc\_dev \geq 0$) is an indicator-variable that takes a value of 1, when the marginal costs of firm *i* are at or above the level of mean marginal costs in the industry, and a value of 0 otherwise. ($mc\_dev < 0$) is an indicator-variable that takes a value of 1, when the marginal costs of firm *i* are below the level of mean marginal costs in the industry, and a value of 0 otherwise.

dependent variable	participation probabilities (*j*)									
regressors		**coefficient**	**t-values**		**bootstrap critical t-values**					
			OLS	**Bootstrap**	**1%**		**5%**		**10%**	
constant		65.756	17.398 ***	17.641 ***	-0.60	0.51	-0.43	0.40	-0.36	0.33
			3.779	3.727						
$mc\_dev \cdot (mc\_dev \geq 0)$		6.175	3.462 ***	1.447	-10.98	3.22	-8.26	2.59	-6.88	2.25
			1.784	4.266						
$mc\_dev \cdot (mc\_dev < 0)$		8.089	4.653 ***	3.148 ***	-2.78	2.35	-1.83	1.74	-1.30	1.38
			1.738	2.570						
$\mu \cdot mc\_dev \cdot (mc\_dev \geq 0)$		1.188	11.566 ***	6.902 ***	-0.06	0.08	-0.04	0.06	-0.03	0.05
			0.103	0.172						
$\mu \cdot mc\_dev \cdot (mc\_dev < 0)$		0.553	5.499 ***	3.892 ***	-0.10	0.10	-0.07	0.07	-0.06	0.06
			0.100	0.142						
degree of homogeneity	μ	0.252	6.608 ***	5.604 ***	-0.02	0.02	-0.02	0.01	-0.01	0.01
			0.038	0.045						
	μ^2	-0.002	-5.253 ***	-4.653 ***	0.00	0.00	0.00	0.00	0.00	0.00
			0.000	0.000						
detection probability	P	-0.156	-6.334 ***	-5.440 ***	-0.01	0.01	-0.01	0.01	-0.01	0.01
			0.025	0.029						
discount rate	r	-0.077	-1.768 **	-1.603 ***	-0.07	0.08	-0.06	0.06	-0.05	0.05
			0.044	0.048						
number of firms	n	-2.395	-7.937 ***	-6.693 ***	-0.15	0.12	-0.11	0.10	-0.09	0.08
			0.302	0.358						
cost parameters	c/v	-0.062	-4.451 ***	-3.172 ***	-0.02	0.02	-0.01	0.01	-0.01	0.01
			0.014	0.020						
	a_3	1.589	1.947 **	1.625 **	-1.42	1.68	-1.05	1.31	-0.88	1.10
			0.816	0.978						
	a_4	-11.376	-1.291 *	-1.249	-19.96	17.74	-15.29	13.45	-12.83	11.03
			8.813	9.109						
R^2	41.76%									
$\bar{R}^2$	42.43%									

Table 4: Participation Probabilities