Optimal Destabilisation of Cartels

Ludwig von Auer (Universität Trier)
Tu Anh Pham (Universität Trier)

March 20, 2017 (file: wp59v33.tex)

Abstract: The literature on cartel stability sidelines antitrust policy, whereas the literature on antitrust policy tends to neglect stability issues. The present paper attempts to connect these two interrelated aspects in the context of the quantity leadership model. In this model the cartel is the Stackelberg quantity leader and the fringe firms are in Cournot competition with respect to the residual demand. We extend this model by an antitrust authority that decides on its own investigative effort and on the size of the fine that cartel members have to pay when they are detected. For testifying cartel members a leniency program is implemented. We incorporate into our framework that these three instruments of antitrust policy are not costless for society. Our model demonstrates that an effective antitrust policy exploits the inherent instability of cartels. We derive an optimal antitrust policy that usually reduces the size of the cartel, but never eliminates it.

JEL-Classification: L0, L1

Keywords: antitrust, stability, Cournot fringe, oligopoly, leniency.

\footnote{We have had the benefit of many useful comments from Andranik Stepanyan.}
1 Introduction

Competitive markets are regarded as a crucial mechanism for creating proper conditions for economic growth and prosperity, because they encourage companies to innovate and to provide the consumers with their preferred products at lowest possible prices. However, there is a long history of anti-competitive agreements in which firms collude to raise market prices above competitive levels. Such agreements inflict short-term and long-term costs on society. Therefore, the formation of cartels is considered to be illegal in most developed economies. Competition law and antitrust authorities are important elements in combating anti-competitive practices.

Economists have made considerable progress in better understanding the formation of cartels and the effectiveness of antitrust policies. Oligopoly markets have emerged as the basic framework for studying the formation and stability of cartels. Two different oligopoly frameworks dominate this literature: The price leadership model developed by d’Aspremont et al. (1983) and the quantity leadership model propagated by Shafer (1995). In the price leadership model, the cartel is the Stackelberg price leader. The fringe firms take this price as given and set their quantities such that price equals marginal cost. In the quantity leadership model, the cartel is the Stackelberg quantity leader and the fringe firms are in Cournot competition with respect to the residual demand. The price leadership model with its perfectly competitive fringe might fit industries with a large number of competing firms, whereas the quantity leadership model with its Cournot fringe might be more suitable for markets with a more limited number of firms.

Both types of leadership models center on the concept of stability. Since the work of d’Aspremont et al. (1983) one distinguishes between internal and external stability. Internal stability requires that no member of the cartel wants to become a fringe firm, whereas external stability requires that no fringe firm wants to become a cartel member.

The leadership models have inspired further work on the conditions for the successful formation and stability of cartels. For example, Donsimoni (1985), Donsimoni et al. (1986), and Prokop (1999) utilize the price leadership model, whereas Shafer’s (1995) analysis and subsequent studies by Konishi and Lin (1999) and Zu et al. (2012) are based on the quantity leadership model. In neither type of leadership model, antitrust policy is an issue. Instead, these studies focus on the formal conditions for the existence and uniqueness of a stable cartel.

The leadership models have been criticized for assuming that within the cartel all members comply with the agreement, as if it were a legally enforceable contract (e.g., Escrivuela-Vilar, 2009, p. 138). Since almost all cartels are illegal, however, such enforcement is unrealistic. Accordingly, a different strand of literature has emerged that addresses the issue of compliance within the cartel. Following Friedman (1971), this literature studies repeated oligopoly games. Most of this work identifies grim-trigger strategies such that compliance is a subgame perfect Nash equilibrium. When such grim-trigger strategies exist, the cartel is said to be sustainable, that is, all firms comply with the collusive agreement. Motta and Polo (2003), Spagnolo (2004), and many subsequent studies have demonstrated that these repeated oligopoly games are well suited to study the effects
of antitrust policy on the compliance of cartel members. However, in these models only two basic market outcomes can arise. Either the antitrust policy successfully deters the firms from establishing a sustainable cartel or the antitrust policy completely fails and a perfect cartel arises, that is, all firms become members of the cartel. Missing in this type of analysis is a third case that, in the real world, appears to be the most frequent one: A sustainable cartel competes against fringe firms. Without this third case, stability is not an issue.

In short, the literature on cartel stability sidelines antitrust policy, whereas the repeated oligopoly games of antitrust policy neglect stability issues. However, there is a strong interdependency between cartel stability and an effective antitrust policy. Therefore, the first contribution of the present paper is to introduce a model that elaborates how an antitrust authority can exploit for its policy the inherent instability of cartels.

For this purpose one could try to develop a hybrid model that combines the leadership framework with a repeated oligopoly game. In fact, Escrihuela-Villar (2009, pp. 139-140) has proposed such a hybrid model. It supplements the quantity leadership model by a repeated oligopoly game à la Friedman (1971, pp. 4-5) which examines the sustainability of a given cartel. It turns out that, with a sufficiently large social discount factor, the stable cartel of the quantity leadership model is always sustainable. In other words, for the stable cartels of the quantity leadership model the sustainability issue can be neglected. Therefore, we have decided to avoid the repeated oligopoly framework. Instead we incorporate into the quantity leadership model an antitrust authority that tries to detect cartels and to prosecute its members.

A second contribution of our paper concerns the menu of policies available to the antitrust authority. The repeated oligopoly games described above mainly focus on the role of fines and leniency programs in the detection of cartels. The most direct way to increase the probability of detection, however, is an intensification of the antitrust authority’s own investigative efforts. Moreover, such efforts are indispensable for a functioning leniency program. Without the authority’s investigations, each cartel member would know that a detection is highly unlikely and that also the other cartel members know this. This knowledge, in turn, would destroy the incentives for a cartel member to turn itself in.

Somewhat surprisingly, the authority’s own investigative effort has rarely been addressed in the literature. An early exception is the repeated oligopoly game developed by Spagnolo (2004). It features an antitrust authority that can decide on its own effort, on the size of the fine, and on the extent of leniency. However, that study is concerned with a cartel’s sustainability rather than with its stability. In contrast, our own model emphasizes the connection between stability and an antitrust authority that can decide on its investigative effort and on the appropriate size of the fine that detected cartels have to pay. Furthermore, we assume that a leniency program exists that offers testifying firms a discount. The probability of detection increases with the authority’s effort, with the size of the discount, and with the number of cartel members.

In the tradition of the law enforcement literature initiated by Becker (1968), the operations of the antitrust authority come at a social cost (e.g., Spagnolo, 2004). This cost must be considered in the design of an optimal antitrust policy. We demonstrate how
such an optimal antitrust policy can be derived.

From a welfare perspective, the antitrust authority should attempt to reduce the size of cartels relative to the case without any antitrust policy. Therefore, the antitrust authority is interested in policies that create internal instability. Creating external instability is not a sensible policy option, because it would increase the size of the cartels. We show that in the context of our extended quantity leadership model the optimal antitrust policy reduces the size of the cartel, but never eliminates it.

This paper proceeds as follows. Section 2 summarizes the original quantity leadership model. In Section 3 we extend this model by an antitrust authority. How an optimal antitrust policy can be derived is outlined in Section 4. Section 5 applies this approach and discusses the underlying economics. Concluding remarks are contained in Section 6.

2 The Original Quantity Leadership Model

Daughety (1990, pp. 1232-33) introduces a model that combines Stackelberg and Cournot competition. The inverse demand function is \( P = a - bQ \), where \( P \) is the market price and \( Q \) is the aggregate quantity produced. The industry consists of a finite number of \( n \geq 2 \) identical firms, some of which are independent Stackelberg leaders and the remaining ones independent Stackelberg followers. Only integer numbers of firms are considered. All \( n \) firms have a constant marginal cost equal to \( c \). Among each other, the Stackelberg leaders act like Cournot players. The same is true for the Stackelberg followers. Daughety uses this model as a starting point for analyzing the welfare effects of horizontal mergers.

When only one Stackelberg leader exists, Daughety’s model simplifies to the quantity leadership model that Shaffer (1995, p. 745) adopts for studying the stability of cartels. In that model, the Stackelberg leader is a cartel comprising a group of symmetric firms that coordinate their output decisions in a binding manner. Brito and Catalão-Lopes (2011, p. 3) summarize the justifications for assuming that the cartel acts as a leader.

When \( n_C \in (1, \ldots, n) \) of the \( n \) firms form a cartel, then the number of fringe firms is \( n_F = n - n_C \). Of course, the model’s results do not depend on whether they are presented in terms of \( n_F \) or \( n_C \). For algebraic reasons we prefer the \( n_F \)-version. A second justification for using the \( n_F \)-version is the fact that in the quantity leadership model the cartel acts as one firm, regardless of \( n_C \).

After the \( (n-n_F) \) firms of the cartel have collectively determined their profit maximizing joint output \( Q_C \), each fringe firm determines its profit maximizing output \( q_F \). In other words, the cartel acts as a Stackelberg leader, whereas the group of fringe firms is the Stackelberg follower. Each fringe firm does not only consider the cartel’s output \( Q_C \) as given, but also the aggregate output of all other fringe firms, \( Q_{-F} \). Therefore, the output of each fringe firm, \( q_F \), is determined by the Cournot-Nash equilibrium concept.

Shaffer (1995, p. 745) showed that, with \( n \in (1, \ldots, n) \) and \( n_F \in (0, \ldots, n-1) \), the equilibrium is given by the total output

\[
Q = Q_C + n_F q_F = \frac{a - c}{b} \frac{2n_F + 1}{2n_F + 2}
\]
and the market price
\[ P = c + \frac{a - c}{2(n_F + 1)}. \] (2)

The profit of each cartel member is
\[ \pi_C(n_F) = \frac{(a - c)^2}{4b(n_F + 1)(n - n_F)}. \] (3)

and the profit of each fringe firm is
\[ \pi_F(n_F) = \frac{(a - c)^2}{4b(n_F + 1)^2}. \] (4)

We assume that, given the choice between cartel membership and fringe, the firm always chooses the option with the larger profit. However, if both options generate exactly the same profit, the firm prefers the legal status of a fringe firm to the illegal status of a cartel member. In other words, to stay in a cartel, the profit there must be strictly larger than the profit earned after becoming a fringe firm. This is the condition for internal stability. Formally, a cartel with \((n - n_F)\) members is internally stable, only if
\[ \pi_C(n_F) > \pi_F(n_F + 1). \] (5)

This is identical to the definition in Escrihuela-Villar (2009, p. 140). Note that in the original definition of d’Aspremont et al. (1983, p. 21) and most subsequent papers a weakly larger profit is sufficient for internal stability.

External stability requires that no fringe firm has an incentive to become a member of the cartel. This is satisfied, when the profit of the fringe firm is at least as large as the profit the firm would earn after entering the cartel. Formally, external stability of a cartel with \((n - n_F)\) members requires that
\[ \pi_C(n_F - 1) \leq \pi_F(n_F). \] (6)

A perfect cartel \((n_F = 0)\) is always externally stable, because no fringe firm exists that could enter the cartel. However, it is easy to show that for \(n > 4\), perfect cartels are never internally stable.

When both, internal and external stability are satisfied, the cartel is said to be stable (d’Aspremont et al., 1983, p. 21). When \(n\) is even (uneven), a stable cartel has exactly two (three) more members than the fringe (Shaffer, 1995, p. 746, Proposition 4) and the profit of each cartel member is smaller than that of each fringe firm (p. 747, Proposition 5). However, both, the members of the stable cartel and the fringe firms earn a higher profit than in the standard Cournot oligopoly without cartel (p. 748, Proposition 6). Note that in the standard Cournot oligopoly the market price is
\[ P = c + \frac{a - c}{n + 1}. \] (7)
and the profit of each firm is

\[ \pi_F(n) = \frac{(a - c)^2}{b(n + 1)^2} \]  \hspace{1cm} (8)

Due to the first-mover advantage, a single-firm cartel \((n_F = n - 1)\) always prefers the cartel situation to the standard Cournot oligopoly. This can be easily checked by comparing the profits in (3) with those in (8). Also the welfare arising from a market with a single-firm cartel is larger than the welfare arising from a standard Cournot oligopoly. This beneficial “timing effect” (Brito and Catalão-Lopes, 2011, p. 2) was already discussed in Daughety (1990, p. 1233).

Equation (2) reveals that the market price falls as the number of fringe firms increases. Therefore, the welfare generated by a market with a single-firm cartel is larger than the welfare generated by markets with larger cartels.

The latter two findings have important implications for the introduction of an antitrust policy. Such a policy should attempt to destabilize the existing stable cartel by providing incentives for the cartel members to become fringe firms. In other words, the policy should change the condition for internal stability. However, the policy should stop short of eliminating the cartel, because the ensuing Cournot competition yields a larger market price than the price arising from a single-firm cartel. If antitrust policy were costless, the largest welfare would arise from a policy that leads to a stable single-firm cartel. In the real world, however, antitrust authorities cause social costs. In the following section we integrate the antitrust authority and the associated social costs into the quantity leadership model.

3 Model Extension

We extend the original quantity leadership model by an antitrust authority that attempts to uncover the cartel. When the cartel is detected it must pay a fine. However, cartel members that cooperate with the authorities receive a discount.

3.1 Antitrust Authority

Let \(p \in [0,1]\) denote the probability of detection. It depends not only on the number of cartel members, but also on the policy implemented by the antitrust authority. To capture all these aspects we specify the probability of detection in the following way:

\[ p(n, n_F, e, d) = g(n - n_F) \cdot h(e) \cdot k(d) \]  \hspace{1cm} (9)

The first factor, \(g(n - n_F) \in [0,1]\), takes care of the fact that larger cartels are more likely to be detected than smaller ones. Therefore, \(g(n - n_F)\) has a positive first order derivative with respect to \((n - n_F)\) and a negative one with respect to \(n_F\). Furthermore, we assume that \(g(n - n_F)\) is concave and therefore approaching 1 from below.

The second factor of the probability of detection is \(h(e) \in [0,1]\), where \(e\) denotes the antitrust authority’s effort level. The authority is part of the government and invests its
complete endowment in the detection of cartels. The larger the endowment, the larger the effort and therefore the probability of detection. The marginal increase in the probability of detection falls with \( e \). That is, \( h(e) \) is also a continuous strictly increasing concave function that approaches 1 from below.

The antitrust authority runs a leniency program that offers a reduced fine to cartel firms that inform the authority about the cartel. There is ample evidence that such programs increase the probability of detecting cartels (e.g., Aubert et al., 2006, p. 1242; Brenner, 2009, pp. 642-644). In anticipation of being detected, cartel firms may apply for leniency by providing evidence of a cartel agreement. Furthermore, each cartel member may worry that some other member applies for leniency and, because of that worry, applies itself. Harrington (2013, pp. 2-3) denotes these two effects as the prosecution effect and the pre-emption effect, respectively. The third factor on the right hand side of Equation (9), \( k(d) \in [0, 1] \), captures these effects. The expected value of the fine is \( pf (1 - \mu d) \), where \( d \geq 0 \) is the percentage by which the fine is reduced for a testifying firm and \( \mu > 0 \) is the share of cartel firms eligible for reductions. We do not include \( \mu \) in \( k(d) \), because its effect on the probability of detection is ambiguous. An increase of \( \mu \) raises the number of potential candidates that testify and get a discount. Due the prosecution effect of discounts, this expansion may increase the probability of detection. At the same time, the increase in \( \mu \) may weaken the pre-emptive effect of discounts thereby reducing the probability of detection. It is unclear which effect dominates. For simplicity, we assume that the two effects cancel. Note also that our specification allows for discounts \( d > 1 \), that is, for rewards. Also \( k(d) \) is a continuous concave function that is strictly increasing and approaching 1 from below.

However, \( k(d) \) differs from \( g(n - n_F) \) and \( h(e) \) in one important respect. When no cartel exists \( (n_F = n) \), the probability of detection, \( p \), should be 0. Therefore, we need \( g(0) = 0 \). Similarly, we should have \( h(0) = 0 \), because without any effort on the side of the antitrust authority, the prosecution effect and the pre-emptive effect of discounts do not exist and also the size of the cartel is irrelevant for the probability of detection. However, when the reduction for testifying cartel members is lowered to \( d = 0 \), this should not imply that the probability of detection falls to 0. The antitrust authority must be able to detect an existing cartel through its own investigative effort \( e \). Therefore, we need \( 0 < k(0) \leq 1 \).

In Appendix A we sketch out a rather general class of functions that can capture all of the above features and has some additional attractive properties. We use this class of functions for all three factors determining the probability of detection. More specifically we assume that

\[
g(n - n_F) = \frac{n - n_F}{[(n - n_F)^2 + 1]^{1/2}} \quad (10)
\]

\[
h(e) = \frac{e}{(e^2 + 1)^{1/2}} \quad (11)
\]

\[
k(d) = \frac{d + \rho}{[(d + \rho)^2 + 1]^{1/2}} \quad (12)
\]
with $\rho > 0$.

If effort, $e$, were costless, the authority could choose an infinitely large effort level such that the factor $b(e)$ approaches 1. Clearly, this is not what we observe in reality. Authorities are limited in their investigative efforts by the number and competence of its staff and by its technical equipment. If society wishes for a larger effort of its antitrust authority, it must provide the resources necessary to hire more and better staff and to purchase more effective equipment. In our model, we capture this aspect by a monotonically increasing cost function, $c(e)$, that attaches a cost to each effort level. The cost must be financed by the public. Therefore, it is a social cost.

When the antitrust authority uncovers a cartel, it can penalize the cartel with a fine $f \geq 0$. The antitrust authority must fix the size of the fine in advance. This decision is the second policy instrument of the antitrust authority. It has a strong incentive to choose very large fines, because this reduces the expected profits from cartel membership. However, excessive fines induce a social cost, because they violate the principle of proportional justice and may increase the risk of convicting innocent firms (e.g., Allain et al., 2015). Governments have reacted to these concerns by setting upper bounds to legally admissible fines. We formalize the social cost of fines by a monotonically increasing cost function, $z(f)$.

Leniency programs cause similar social costs. The public may dislike the idea that testifying firms that have broken the law can get away with a discount or, even worse, are rewarded. Lenient treatment of guilty firms may undermine the general respect for the law and may encourage unlawful behaviour. This is captured by a monotonically increasing cost function, $r(\mu d)$. This specification says that the public incorporates into its dislike not only the percentage of the reduction, $d$, but also the share of cartel members eligible for this reduction, $\mu$.

In sum, our model extends the quantity leadership model by an antitrust authority that imposes a rather stylized, but comprehensive antitrust policy. The effort $e$ and the fine $f$ are chosen by the antitrust authority. The discount $d$ and the share of eligible cartel members $\mu$, however, are exogenously given, due to legal restrictions. The probability of detection is carefully specified. It increases with the number of cartel members, $n - n_F$, with the authority’s effort, $e$, and with the reduction in the fine for testifying cartel members, $d$. Furthermore, our model introduces the cost functions $c(e)$, $z(f)$, and $r(\mu d)$ to account for the social costs arising from the antitrust policy.

### 3.2 Stability

Since the cartel’s profit maximizing output, $Q_C$, does not depend on the probability of detection, $p$, the results (1), (2), and (4) remain valid. However, each cartel member may have to pay the fine $f$ and a share $\mu$ of these members may get a reduction of $d$ percent.\(^2\)

\(^2\)If $\delta = 0.1$ and the number of cartel members is $n - n_F = 10$, then exactly one member is randomly drawn that testifies and receives the reduction. If $\delta = 0.1$ and $n - n_F = 5$, again one member is randomly drawn and that member has a 50% chance to testify and to receive the reduction.
Therefore, result (3) must be modified. The expected profit of each cartel member is

\[ E[\pi_C(n_F)] = \pi_C(n_F) - pf(1 - \mu d) \]

\[ = \frac{(a - c)^2}{4b(n_F + 1)(n - n_F)} - g(n - n_F)h(e)k(d)f(1 - \mu d). \]  
(13)

Accordingly, the formal condition for internal stability becomes

\[ E[\pi_C(n_F)] > \pi_F(n_F + 1). \]  
(14)

The condition’s interpretation, however, is unaltered: As long as the expected profit from being a cartel member is larger than the profit from becoming a fringe firm, the cartel membership is preferred. However, if the two profits are identical, the firm chooses the riskless and law-abiding option, that is, the fringe status.

Inserting the profit definitions (4) and (13) in inequality (14) yields

\[ A < T(n_F), \]  
(15)

with

\[ T(n_F) = \frac{(a - c)^2}{4b} \frac{[(n - n_F)^2 + 1]^{1/2} [n_F^2 + 2 - (n_F + 1)(n - n_F)]}{(n_F + 1)(n - n_F)^2(n_F + 2)^2} \]  
(16)

and

\[ A = h(e)k(d)f(1 - \mu d). \]  
(17)

The value of \( A \) depends on the three punishment variables \( f, d, \) and \( \mu \), and on the effort \( e \), but not on \( n \) and \( n_F \). Increases in the effort, \( e \), and the fine, \( f \), raise the value of \( A \). The impact of the discount \( d \) on the value of \( A \) is ambiguous, since it increases the probability of detection, but lowers the expected value of the fine.

The value of \( T(n_F) \) depends on \( n \) and \( n_F \) as well as on the market volume \((a - c)/b\). Furthermore, we obtain the following result.

**Lemma 1** The function \( T(n_F) \) defined by (16) increases in \( n_F \). For \( T(n_F) \geq 0 \), the function \( T(n_F) \) decreases in \( n \).

**PROOF:** See Appendix B.

External stability requires that

\[ E[\pi_C(n_F - 1)] \leq \pi_F(n_F). \]

Inserting the profit definitions (4) and (14) yields

\[ T(n_F) \leq A \]  
(18)

with

\[ T(n_F) = \frac{(a - c)^2}{4b} \frac{[(n - n_F + 1)^2 + 1]^{1/2} [n_F(2n_F + 1) - n + 1]}{n_F(n - n_F + 1)^2(n_F + 1)^2}. \]  
(19)

The weak inequality in (18) reflects the fact that with \( E[\pi_C(n_F - 1)] = \pi_F(n_F) \) a firm always prefers the riskless and legal status of the fringe firm.

For later proofs the following result is useful.
Lemma 2 The function $T(n_F)$ defined by (19) increases in $n_F$ and decreases in $n$.

Proof: See Appendix C.

When conditions (15) and (18) are simultaneously satisfied, the cartel is stable:

$$T(n_F) \leq A < \overline{T}(n_F).$$  \hspace{1cm} (20)

For a market to have a stable cartel, the number of fringe firms, $n_F$, must be such that the given $A$-value satisfies inequalities (20). Therefore, $T(n_F)$ and $\overline{T}(n_F)$ can be viewed as two thresholds. In contrast to these thresholds, the $A$-term does not depend on $n_F$ and $n$. Instead it represents the antitrust policy $(e, f, d, \mu)$. By modifying the values of the policy instruments, the antitrust authority can modify the $A$-value, and therefore, the size of stable cartels. Reducing the effort or the fine would lower the stable $n_F$-value (enlarge the cartel), whereas increasing the effort or the fine would increase the stable $n_F$-value (shrink the cartel).

Note that

$$T(n_F + 1) = \overline{T}(n_F).$$  \hspace{1cm} (21)

Therefore, the condition for stability, (20), can be written in the form

$$T(n_F) \leq A < T(n_F + 1).$$  \hspace{1cm} (22)

This has an important implication:

Theorem 1 If for given parameters $n, a, b,$ and $c$, an antitrust policy $(e, f, d, \mu)$ generates a stable cartel, then this cartel is unique.

Proof: By Lemma 2, the sequence of intervals $[T(0), T(1)), [T(1), T(2)), \ldots, [T(n - 1), T(n))]$ is connected, but mutually exclusive. Therefore, for each given $A$-value, only one $n_F$-value satisfying (22) can exist.

3.3 Welfare

Welfare is defined here as the sum of consumer and producer rent minus the social cost caused by the antitrust policy. When the antitrust policy leads to a stable cartel with $n_F \in (0, \ldots, n - 1)$, the sum of consumer and producer rent is given by

$$(a - c) Q - 0.5(a - P)Q = \frac{(a - c)^2 (2n_F + 1) (2n_F + 3)}{8b (n_F + 1)^2},$$  \hspace{1cm} (23)

where the values of $Q$ and $P$ are defined by (1) and (2).

The implemented antitrust policy also determines the values of the three components of the social cost function: $c(e), z(f),$ and $r(\mu d)$. As described in Section 3.1, these three components are strictly increasing functions. We assume that the social cost function is

$$s(c(e), z(f), r(\mu d)) = (ae + \beta f + \gamma \mu d)^{1+\delta},$$  \hspace{1cm} (24)

10
with $\alpha, \beta, \gamma, \delta \geq 0$. The parameter $\alpha$ can be interpreted as an indicator of the authority’s economic efficiency. The parameters $\beta$ and $\gamma$ determine the damage to the rule of law arising from the fine and the discount. The degree of convexity in the social cost function is determined by the parameter $\delta$.

Welfare does not depend on the budgetary effects of the fines and discounts, because these are of purely redistributonal nature. Combining expressions (23) and (24), the welfare function is

$$ W(e, f) = \frac{(a - c)^2 (2n_F + 1)(2n_F + 3)}{8b (n_F + 1)^2} - (\alpha e + \beta f + \gamma \mu d)^{1+\delta}, \quad (25) $$

where $n_F$ is determined via condition (22). We want to find, for given $\mu$ and $d$, the antitrust authority’s policy $(e, f)$ that generates the maximum welfare. This policy is denoted as the authority’s optimal antitrust policy, $(e^*, f^*)$.

## 4 Optimal Antitrust Policy

### 4.1 Cartel Prevention and Passive Antitrust Policy

The authority could choose a policy $(e, f)$, such that $A \geq T(n - 1)$. This policy would prevent the formation of any stable cartel. Therefore, a symmetric Cournot oligopoly would arise. However, this is not the outcome the antitrust authority should pursue.

**Theorem 2** The symmetric Cournot oligopoly would generate a smaller welfare than a Stackelberg market with a stable single-firm cartel.

**Proof:** We know that the price (2) associated with $n_F = n - 1$ is smaller than the price (7) associated with $n_F = n$. As a consequence, the sum of consumer and producer rent is larger for $n_F = n - 1$ than for $n_F = n$. From Lemma 1 we know that $T(n_F)$ is increasing in $n_F$. Therefore, a policy that eliminates the single-firm cartel must increase the value of $A$ from $T(n - 1) \leq A < T(n - 1)$ to $A \geq T(n - 1)$. From (17) and the definition of $h(e)$ we know that $\partial A / \partial e > 0$ and $\partial A / \partial f > 0$. Therefore, the increase in $A$ required to eliminate the single-firm cartel necessitates larger values of $e$ and $f$. These larger values translate into a higher social cost. The increase in social cost and the fall in the sum of consumer and producer rent result in a lower welfare.

Theorem 2 says that an antitrust policy leading to $n_F = n$ cannot be optimal. Therefore, we can restrict our attention to policies consistent with stable cartels such that $n_F \in (0, \ldots, n - 1)$.

Relationships (17) and (22) imply that for $(e, f) = (0, 0)$ the condition for stability becomes

$$ T(n_F) \leq 0 < T(n_F + 1). \quad (26) $$

Only one $n_F$-value exists that satisfies this condition. This is the $n_F$-value arising from doing nothing on the side of the antitrust authority. We denote this value as $n_F^{\text{min}}$, because
a sensible antitrust policy \((e, f)\) must lead to a fringe that has at least \(n_F^{\text{min}}\) firms. A policy \((e, f)\) that leads to \(A < T(n_F^{\text{min}}) < 0\), and therefore to \(n_F < n_F^{\text{min}}\), would cause social costs and at the same time it would increase the size of the cartel and therefore reduce the sum of consumer and producer rent. Clearly, this would not be a sensible policy. Therefore, we can confine our search for the optimal antitrust policy to those policies \((e, f)\) that lead to \(n_F \in (n_F^{\text{min}}, \ldots, n-1)\).

To find the optimal antitrust policy \((e^*, f^*)\), we pursue the following three stage procedure:

1. Find \(n_F^{\text{min}}\).

2. Derive for each given \(n_F \in (n_F^{\text{min}}, \ldots, n-1)\) the antitrust policy that minimizes the social cost \((ae + \beta f + \gamma \mu d)^{1+\delta}\).

3. For each of these cost minimizing antitrust policies, compute the resulting welfare. The policy that generates the largest welfare is the optimal antitrust policy \((e^*, f^*)\).

In the following, we describe these three stages in more detail. An illustrating example is provided in Section 5.

### 4.2 Finding \(n_F^{\text{min}}\)

We know that the passive antitrust policy \((e, f) = (0, 0)\) leads to \(n_F^{\text{min}}\). This is the only \(n_F\)-value that satisfies both inequalities in (26). The sign of \(T(n_F)\) depends on the sign of the term \(n_F (2n_F + 1 - n) + 1\) in expression (19). Therefore, the left inequality of (26) – the external stability condition – gives

\[
n - n_F \geq n_F + 1 + \frac{1}{n_F}.
\]

This says that in the absence of an active antitrust policy the smallest externally stable cartel, and therefore the only stable cartel, has at least two more members than fringe firms exist. Therefore, \(n_F^{\text{min}}\) is the largest integer for which the condition \(n - n_F^{\text{min}} \geq n_F^{\text{min}} + 2\) is satisfied. Rearranging this condition gives \(n_F^{\text{min}} \leq (n - 2) / 2\). Therefore,

\[
n_F^{\text{min}} = \begin{cases} 
\frac{(n - 2)}{2} & \text{for even } n \\
\frac{(n - 3)}{2} & \text{for uneven } n.
\end{cases}
\] (27)

This is just a reformulation of Shaffer’s (1995, p. 746) Proposition 4. The antitrust authority can restrict its search for the optimal antitrust policy \((e^*, f^*)\) to policies that lead to \(n_F \in (n_F^{\text{min}}, \ldots, n-1)\), where \(n_F^{\text{min}}\) is defined by (27).
4.3 Computing the Cost Minimizing Policies

We know that the passive policy \((e, f) = (0, 0)\) leads to \(n_F = n_F^{\text{min}}\), that is, to a stable cartel with \(n - n_F^{\text{min}}\) members.

**Theorem 3** Among all antitrust policies leading to a stable cartel with \(n - n_F^{\text{min}}\) members, the passive policy \((e, f) = (0, 0)\) is the cost minimizing policy \((e^*_n, f^*_n)\).

**Proof:** The policy \((e, f) = (0, 0)\), leads to \(A = 0\). Inequalities (26) imply that any other policy leading to \(T(n_F) \leq A < 0\) or to \(0 < A < T(n_F + 1)\) would generate the same stable cartel, and therefore, the same sum of consumer and producer rent. However, it would cause a larger social cost.

Suppose that the antitrust authority wants to generate a stable cartel with exactly \((n - n_F)\) members, where \(n_F \in \{n_F^{\text{min}} + 1, \ldots, n - 1\}\). From (22) we know that this outcome requires an active antitrust policy \((e, f)\) such that the resulting \(A\)-value defined by (17) falls into the interval \([T(n_F), T(n_F + 1)]\). An infinite number of policies exist that satisfy this condition. Since all of these policies lead to the same given \(n_F\)-value, the equilibrium quantity, \(Q\), the equilibrium price, \(P\), and therefore, the consumer rent and the producer rent are the same for all these policies. However, the social cost is not.

Lower \(A\)-values allow for lower values of \(e\), \(f\), and \(d\), and therefore, for lower social cost. To exploit this property, the antitrust authority should choose a policy that generates the lowest \(A\)-value consistent with a stable cartel with \((n - n_F)\) members. In other words, for each given \(n_F\)-value the antitrust authority should decide for a policy such that \(A\) reaches the lower bound of its admissible interval:

\[
A = T(n_F) .
\]

In this situation each member of a cartel with \(n - (n_F - 1)\) firms would earn the same expected profit as it would earn after becoming a fringe firm. Therefore, each cartel member would prefer the riskless and legal status of the fringe firm. One of the cartel members would leave the cartel, reducing the number of cartel members to \((n - n_F)\). In the new situation no incentive remains for leaving the cartel.

Inserting (17), Equation (28) can be written in the form

\[
h(e) k(d) f (1 - \mu d) - T(n_F) = 0 .
\]

Still, an infinite number of policies \((e, f)\) satisfy condition (29). Therefore, the authority should choose the policy that causes the lowest social cost:

\[
\min (\alpha e + \beta f + \gamma \mu d) \quad \text{subject to (29)} .
\]

Note that the solution to this minimization problem is independent from \(\delta\), the parameter determining the convexity of the social cost function.
Theorem 4 For each \( n_F \in (n_F^{min} + 1, \ldots, n - 1) \), the cost minimizing policy \((e_{n_F}^*, f_{n_F}^*)\) that leads to a stable cartel with \((n - n_F)\) members, is given by the following two equations:

\[
e_{n_F}^* = \left[ \frac{\beta (d + \rho)^2 + 1}{\alpha (d + \rho)} \frac{T(n_F)}{(1 - \mu d)} \right]^{1/2},
\]

\[
f_{n_F}^* = \frac{\alpha}{\beta} e_{n_F}^* \left( e_{n_F}^* \right)^2 + 1 \right)^{1/2}.
\]

Proof: We can transform the constrained minimization problem (30) into an unconstrained one. Solving (29) for \( f \) yields

\[
f = \frac{T(n_F)}{h(e) k(d) (1 - \mu d)}.
\]

Inserting the right hand side of (33), in the social cost term, \( \alpha e + \beta f + \gamma d \), gives

\[
\alpha e + [h(e)]^{-1} \beta \frac{T(n_F)}{k(d) (1 - \mu d)} = \gamma d.
\]

Differentiating expression (34) with respect to \( e \) and setting the result equal to 0 yields

\[
\frac{[h(e)]^2}{h'(e)} = \beta \frac{T(n_F)}{\alpha k(d) (1 - \mu d)}.
\]

Inserting expressions (12), (11), and its derivative in (35) gives the optimal effort (31).

Expression (11) gives

\[
h(e_{n_F}^*) = \frac{e_{n_F}^*}{\left( e_{n_F}^* \right)^2 + 1}^{1/2}.
\]

We know from Equation (31) that

\[
\left( e_{n_F}^* \right)^2 = \frac{T(n_F)}{\alpha k(d) (1 - \mu d)}.
\]

Inserting (36) and (37) in (33) yields the optimal fine (32).

To summarize, Theorem 3 defines the cost minimizing policy for \( n_F = n_F^{min} \). This policy is \((e_{n_F}^{min*}, f_{n_F}^{min*}) = (0, 0)\). For every \( n_F \in (n_F^{min} + 1, \ldots, n - 1) \), Theorem 4 defines the cost minimizing policy \((e_{n_F}^*, f_{n_F}^*)\).

4.4 Selecting the Optimal Antitrust Policy

First, we insert \( n_F = n_F^{min} \) and \((e, f) = (0, 0)\) in welfare function (25) and compute the resulting welfare level. This is the welfare level arising from the passive antitrust policy. Then we consider the active antitrust policies. The welfare levels corresponding to each
integer \( n_F \in (n_F^{\text{min}} + 1, \ldots, n - 1) \) are calculated. For this purpose we insert the value \( n_F \) and the associated cost minimizing policy \((e_{n_F}^*, f_{n_F}^*)\) in the welfare function (25). We get a set of welfare levels. From this set we select the maximum value. If this welfare is smaller than the one generated by the passive policy, the optimal fringe size is \( n_F^* = n_F^{\text{min}} \). However, if the welfare from the best active antitrust policy is larger than the one from the passive policy, then the optimal fringe size is \( n_F^* > n_F^{\text{min}} \). The cost minimizing antitrust policy leading to the stable cartel with \((n - n_F^*)\) members is the optimal antitrust policy \((e^*, f^*)\).

5 Application and Discussion

In Section 4 the optimal antitrust policy was derived. Next we present a numerical illustration and we discuss the economics behind our results.

5.1 A Simple Application

For the numerical illustration the following values are assigned to the parameters: \( n = 10, a = 100, b = 1, c = 10, \alpha = 0.05, \beta = 0.05, \gamma = 0.05, \delta = 0.5, \mu = 0.1, d = 1, \) and \( \rho = 1 \).

Stage 1: The lowest relevant number of fringe firms, \( n_F^{\text{min}} \), is obtained from (27): \( n_F^{\text{min}} = 4 \). Therefore, we can confine the analysis to \( n_F \in \{4, \ldots, 9\} \).

Stage 2: We know that the cost minimizing effort level and fine for generating a stable cartel with \( n - n_F^{\text{min}} = 10 - 4 = 6 \) members, is the passive policy \((e, f) = (0, 0)\). This policy is listed in the second and third column of Table 1 in the line corresponding to \( n_F = 4 \).

Reducing the size of the cartel below six requires an active antitrust policy that makes the cartel internally instable and transforms it into a five-firm cartel that competes with five fringe firms \((n_F = 5)\). This policy must be such that \( A = T(5) \). Inserting in (19) all parameters gives \( T(n_F) = 11.405 \). Inserting this result together with the parameter values in Equation (31) gives the cost minimizing effort \( e_5^* = 3.76 \). Inserting everything in (32) yields the cost minimizing fine \( f_5^* = 14.66 \). The cost minimizing policies for \( n_F = 6 \) to \( n_F = 9 \) are calculated in an analogous manner. The results are listed in the second and third column of Table 1.

Stage 3: We compute for all cost minimizing policies the resulting welfare and select the policy that leads to the largest welfare. How large is the welfare arising from the passive policy \((e, f) = (0, 0)\)? Inserting the parameter values together with \( n_F = 4, e = 0, \) and \( f = 0 \) in the welfare function (25), yields the welfare level \( W = 4009.5 \) listed in the fourth column of Table 1 in line \( n_F = 4 \). This welfare level is the reference for any active antitrust policy. A policy that generates a welfare level below 4009.5 cannot be an optimal antitrust policy.
<table>
<thead>
<tr>
<th>$n_F$</th>
<th>$e^*_n$</th>
<th>$f^*_n$</th>
<th>$W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>4009.5</td>
</tr>
<tr>
<td>5</td>
<td>3.76</td>
<td>14.66</td>
<td>4021.0</td>
</tr>
<tr>
<td>6</td>
<td>5.76</td>
<td>33.65</td>
<td>4026.6</td>
</tr>
<tr>
<td>7</td>
<td>7.22</td>
<td>52.59</td>
<td>4029.0</td>
</tr>
<tr>
<td>8</td>
<td>8.82</td>
<td>78.25</td>
<td>4028.4</td>
</tr>
<tr>
<td>9</td>
<td>11.32</td>
<td>128.62</td>
<td>4021.3</td>
</tr>
</tbody>
</table>

Table 1: Identifying the Optimal Anti-Trust Policy.

The welfare level associated with $n_F = 5$ is obtained from inserting the parameter values together with $n_F = 5$, $e = 3.76$ and $f = 14.66$ in the welfare function (25). This gives $W = 4021.0$ which is larger than the welfare from the passive policy. However, the maximum welfare is obtained from a policy that leads to $n_F = 7$: $(e^*, f^*) = (7.22, 52.59)$. This solution corresponds to a three-firm cartel. Reducing the size of the cartel even further would require a larger effort and a higher fine. However, the increase in social cost would outweigh the gain in the sum of consumer and producer rent.

5.2 Discussion

A sufficient increase in the effort level, $e$, and the fine, $f$, increases the size of the fringe from $n_F$ to $n_F + 1$. The additional fringe firm has a positive and a negative welfare effect. It increases competition and therefore the sum of consumer and producer rent (positive “rent effect”), but it increases also the social cost of the antitrust policy (negative “cost effect”).

When $n_F$ increases by 1, the positive rent effect is

\[
\frac{(a - c)^2}{8b} \left( \frac{2n_F + 3}{n_F^2 + 3n_F + 2} \right)^2 > 0.
\]

For all positive values of $n_F$, this expression is positive and falling in $n_F$. In other words, the positive rent effect falls as $n_F$ rises. More formally, the positive rent effect is a decreasing convex function with a positive range and the domain given by the integers $n_F \in (0, 1, \ldots, N)$.

At the same time, the increase in $n_F$ increases the threshold $T(n_F)$ in (31). A larger effort is required and, by (32), also a larger fine. This increases the social cost. This change in social cost is the negative cost effect. Also the cost effect is a function in $n_F$. However, this function is much more complex than the rent effect (38).

We can confine our attention to the integers $n_F \in (n_F^{\text{min}}, \ldots, n - 1)$. If already at $n_F^{\text{min}}$ the negative cost effect outweighs the positive rent effect, then $n_F^* = n_F^{\text{min}}$ and the passive policy $(e, f) = (0, 0)$ is optimal. However, if at $n_F^{\text{min}}$ the positive rent effect outweighs the
negative cost effect, then an active antitrust policy should be implemented that increases the number of fringe firms above $n_F^{\text{min}}$. The expansion of the fringe leads to an overall welfare gain as long as the positive rent effect dominates the negative cost effect. If even for $n_F = n - 2$ the positive rent effect is larger than the negative cost effect, the size of the fringe should be increased to $n_F^* = n - 1$. However, another expansion to $n_F = n$ would cause a welfare loss, because the Stackelberg competition would be replaced by Cournot competition (see Theorem 2).

The model’s parameters affect the relative strength of the positive rent effect and the negative cost effect. Obviously, larger values of the cost parameters $\alpha, \beta, \gamma,$ and $\delta$ increase the strength of the negative cost effect, thus reducing the size of the optimal fringe, $n_F^*$. The impact of the market volume, $(a - c)/b$, is less obvious. From (38) it can be seen that, for every given $n_F$, a larger market volume increases the rent effect. However, it increases also the cost effect, because the threshold $T(n_F)$ increases with the market volume. Numerical simulations show that the market volume and the optimal fringe size, $n_F^*$, are negatively correlated. This says that the larger the market volume, the larger is the decline in the positive rent effect and the increase in the negative cost effect caused by an increase in $n_F$. As a result, in markets with a large volume the optimal fringe size tends to be smaller than in a market with a small volume.

How does the number of firms, $n$, affect the optimal share of fringe firms $n_F^*/n$? Numerical simulations show that for small values of $n$ the ratio $n_F^*/n$ increases with $n$. However, for larger values of $n$ the ratio starts falling.

The previous discussion has focused on the positive rent and negative cost effect of expanding the fringe and how these two effects depend on the model’s parameters. Changes in these parameters may lead to a new optimal fringe size, $n_F^*$, and therefore to a new welfare level. What happens when the parameter changes are insufficient to change the optimal fringe size, $n_F^*$? Then the sum of consumer and producer rent remains unchanged. However, the cost minimizing antitrust policy and therefore the social cost would change even without a change in $n_F^*$. This, in turn, would affect the welfare level. Therefore, we take a closer look at the impact of the parameters on the cost minimizing antitrust policies $(e_{n_F}^*, f_{n_F}^*)$ determined by Equations (31) and (32). These policies depend on the values of all parameters except $\gamma$ and $\delta$.

**Theorem 5** For all $n_F \in (n_F^{\text{min}} + 1, \ldots, n - 1)$, the cost minimizing effort $e_{n_F}^*$ and fine $f_{n_F}^*$ defined by Equations (31) and (32) increase with the market volume, $(a - c)/b$, and with the number of fringe firms, $n_F$, and they decrease with the total number of firms, $n$. Furthermore, $e_{n_F}^*$ increases and $f_{n_F}^*$ decreases with $\beta/\alpha$.

**Proof:** The impact of $(a - c)/b$ on $e_{n_F}^*$ is obvious from Equation (19). We know that for $n_F \in (n_F^{\text{min}} + 1, \ldots, n - 1)$ the term $T(n_F)$ is positive. From Lemma 2 we know that $T(n_F)$ increases in $n_F$, but decreases in $n$. Since $f_{n_F}^*$ is an increasing function of $e_{n_F}^*$, all these arguments carry over to $f_{n_F}^*$. It is obvious that an increase in $\beta/\alpha$ increases $e_{n_F}^*$. Inserting (31) in (32) and squaring both sides reveals that $\beta/\alpha$ has a decreasing effect on $f_{n_F}^*$. 

17
Most of the results of Theorem 5 have a straightforward intuitive explanation. An increase in $\alpha/\beta$ says that the social cost of the effort increases relative to the social cost of the fine. Therefore, the policy instrument effort is substituted by the policy instrument fine.

Less obvious is the positive impact of the market volume, $(a - c)/b$, on $e^*_{n_F}$ and $f^*_{n_F}$. We know that $e^*_{n_F}$ and $f^*_{n_F}$ must be such that the difference between the profit of a member of an undetected cartel with $(n_F - 1)$ members, $\pi_C(n_F - 1)$, and the profit from leaving the cartel and becoming a fringe firm, $\pi_F(n_F)$, must be just equal to the expected fine of the cartel member, $pf(1 - \mu d)$. When the market volume increases, the difference $\pi_C(n_F - 1) - \pi_F(n_F)$ increases. To balance this increase the expected fine must increase, too. This necessitates a larger effort and fine.

For given $n_F$, the cost minimizing effort $e^*_{n_F}$ and fine $f^*_{n_F}$ falls when the number of firms, $n$, increases. Note that for given $n_F$ the increase in $n$ is an enlargement of the cartel. We know from Equation (9) that larger cartels increase the probability of detection and that this enlargement can substitute the authority’s effort. The lower effort, in turn, corresponds to a lower fine (Theorem 4).

6 Conclusions

This paper demonstrated that a very close relationship exists between an antitrust authority’s policy and the stability of cartels. This relationship has been neglected in the existing literature. Studies such as Shaffer (1995), Konishi and Lin (1999), as well as Zu et al. (2012) verify that the quantity leadership model is well suited for the analysis of stable cartels. Antitrust policy, however, is not an issue in these studies. Conversely, studies of antitrust policy are typically embedded in repeated oligopoly games with the two market outcomes “no cartel” or “perfect cartel”. These studies are concerned with sustainability. Stability is not an issue.

Therefore, the present paper extended the quantity leadership model by an antitrust authority that can choose its own antitrust policy. The authority can use its policy instruments to make cartels internally instable. Such a policy, however, is costly. We showed that the authority is able to completely prevent the emergence of a stable cartel. However, such a policy causes too large social costs. Instead the antitrust authority should merely reduce the size of the cartel until the resulting gains in the sum of consumer and producer rent (the rent effect) no longer outweighs the resulting increase in social cost (the cost effect). We showed how such an optimal antitrust policy can be derived.

Furthermore we studied how changes in the market environment affect the optimal antitrust policy. For example, an increase in the total number of firms increases welfare via its cushioning effect on social cost. When this effect is sufficiently strong, some of the new firms will become fringe firms, increasing competition and raising welfare even further. The same beneficial effects arise from increasing the efficacy and/or efficiency of the antitrust authority’s investigative efforts or from reducing the harm caused by excessive fines and leniency programs.
We have proposed a stylized but comprehensive formal representation of antitrust policy. We took some care in specifying the probability of detection as a function of the number of cartel members, the authority’s investigative effort, and its fine and leniency policy. In future research it would be useful to refine our formalization even further.

Appendix A

The probability of detection specified by (9) is the product of three functions that all belong to the same general class of functions. This class is defined by

$$y(x) = \frac{x + \rho}{(x + \rho)^{\theta} + 1}.$$  

with $\theta > 0$ and $\rho \geq 0$. The first order derivative is

$$y'(x) = y(x) \frac{1}{(x + \rho)[(x + \rho)^{\theta} + 1]} > 0. \tag{39}$$

The second order derivative is

$$y''(x) = -y(x) \left( \frac{(x + \rho)^{\theta} + \theta}{(x + \rho)^2 [(x + \rho)^{\theta} + 1]^2} \right) < 0.\quad \text{(39)}$$

Furthermore

$$\lim_{x \to \infty} y(x) = 1$$

and

$$y(0) = \frac{\rho}{(\rho^\theta + 1)^{1/\theta}}.$$  

Therefore, the parameter $\rho$ determines the minimum value that the function $y(x)$ can take. For $\rho = 0$ we get $y(0) = 0$.

Equation (39) implies that

$$\frac{y'(x)}{[y(x)]^{\delta+2}} = \left( \frac{(x + \rho)^{\theta} + 1}{(x + \rho)^{\delta+2}} \right).$$

For $\theta = 1$, $\delta = 0$, and $\rho = 0$ this expression simplifies to

$$\frac{y'(x)}{[y(x)]^2} = \frac{1}{x^2}.$$
Appendix B

Proof of Lemma 1: (16) can be written as

\[ T(n_F) = \frac{(a - c)^2}{4b} \cdot \frac{I(n_F)}{g(n - n_F)} \]

with

\[ I(n_F) = \frac{2n_F^2 + (5 - n)n_F + 4 - n}{(n_F + 1)(n - n_F)(n_F + 2)^2}. \]  (40)

Using the quotient rule for the differentiation of (40) with respect to \( n_F \), gives the positive denominator

\[ (n_F + 1)^2(n - n_F)^2(n_F + 2)^4 > 0 \]

and the numerator

\[ (4n_F + 5 - n)(n_F + 1)(n - n_F)(n_F + 2)^2 \]
\[ - [2n_F^2 + (5 - n)n_F + 4 - n](n_F + 2)^2(n - n_F) \]  (41)
\[ - [2n_F^2 + (5 - n)n_F + 4 - n]2(n_F + 2)(n_F + 1)(n - n_F) \]  (42)
\[ + [2n_F^2 + (5 - n)n_F + 4 - n](n_F + 1)(n_F + 2)^2. \]  (43)
\[ + [2n_F^2 + (5 - n)n_F + 4 - n]2(n_F + 2)(n_F + 1)(n - n_F) \]  (44)

The expression in line (41) is equal to

\[ (2n_F + 5 - n)(n_F + 1)(n - n_F)(n_F + 2)^2 + 2n_F(n_F + 1)(n - n_F)(n_F + 2)^2. \]  (45)

We add to the first summand of (45) the expression in line (42) and obtain

\[ (2n_F + 1)(n - n_F)(n_F + 2)^2 > 0. \]

Next we add to the second summand of (45) the expressions in lines (43) and (44), factor out \((n_F + 2)(n_F + 1)\), simplify the remaining term to get

\[ n_F [4n_F^2 + 5 (3 - n) n_F + 2n^2 - 11n + 22] + 2n^2 - 10n + 8 \]
\[ = \left( 2n_F + \frac{5}{4} (3 - n) \right)^2 + \frac{7}{16} (3 - n)^2 + 2n^2 - 9n + 12 > 0 \]

Therefore, the expression in lines (41) to (44) is positive and so is the derivative of \( I(n_F) \) with respect to \( n_F \):

\[ I(n_F + 1) > I(n_F). \]

In addition, \( g(n - n_F) \) is increasing in \( n - n_F \), and therefore, decreasing in \( n_F \):

\[ g(n - (n_F + 1)) < g(n - n_F). \]

Therefore, we get

\[ \frac{I(n_F + 1)}{g(n - (n_F + 1))} > \frac{I(n_F)}{g(n - n_F)}. \]
which is identical to $T(n_F + 1) > T(n_F)$.

Let $N$ denote the numerator and $D$ the denominator of (16). We know that $\partial g(n - n_F)/\partial n > 0$ and $D > 0$. The derivative of (16) with respect to $n$ is

$$\frac{(-n_F - 1)D - (n_F + 1)(n_F + 2)^2 [(\partial g(n - n_F)/\partial n)(n - n_F) + g(n - n_F)] N}{D^2}$$

$$= \frac{-(n_F + 1) - (n_F + 1)(n_F + 2)^2 [(\partial g(n - n_F)/\partial n)(n - n_F) + g(n - n_F)] T(n_F)}{D}.$$

For $T(n_F) \geq 0$, this derivative is negative.

## Appendix C

**Proof of Lemma 2:** Equation (19) can be written as

$$T(n_F) = \frac{(a - c)^2 [(n - n_F + 1)^2 + 1]^{1/2}}{4b} \frac{n_F(2n_F + 1 - n) + 1}{n_F(n - n_F + 1)(n_F + 1)^2}.$$  \hspace{1cm} (46)

Differentiating the second and third fraction on the right hand side of (46) with respect to $n_F$ gives

$$\frac{1}{[(n - n_F + 1)^2 + 1]^{1/2}(n - n_F + 1)^2} > 0$$  \hspace{1cm} (47)

and

$$\frac{2n^2n_F^2 - 5nn_F^3 + nn_F^2 - 3nn_F - n + 4n_F^4 + n_F^3 + 5n_F^2 - n_F - 1}{n_F^2(n_F + 1)^3(n - n_F + 1)^2}.$$  \hspace{1cm} (48)

The denominator of (48) is positive. Therefore, the numerator determines the sign of the derivative. This numerator can be written in the following form:

$$n_F^2 \left(2n_F + \frac{1 - 5n}{4} \right)^2 + \left[\frac{7n^2 + 26n + 79}{16}n_F^2 - (3n + 1)n_F - n - 1 \right].$$

The first summand is positive. For $n_F \geq 1$, the second summand is at least as large as

$$\frac{7n^2 + 26n + 79}{16}n_F - (3n + 1)n_F - n - 1 = \frac{7n^2 - 22n + 63}{16}n_F - n - 1.$$

This term is at least as large as

$$\frac{7n^2 - 22n + 63}{16} - n - 1 = \frac{7n^2 - 38n + 47}{16} = \frac{(7n - 10)(n - 4) + 7}{16} > 0 \quad \text{for all} \ n \geq 4.$$

Since (47) and (48) are both positive, the product rule yields a positive derivative of the right hand side of (46) with respect to $n_F$.  

21
The derivatives of the second and third fraction on the right hand side of (46) with respect to \( n \) are

\[
1 \leq \frac{1}{[(n - n_F + 1)^2 + 1]^{1/2}(n - n_F + 1)^2} < 0
\]

and

\[
\frac{-n_F (n_F + 1)^4}{(n_F (n - n_F + 1) (n_F + 1)^2)^2} < 0.
\]

Using the product rule yields a negative derivative of the right hand side of (46) with respect to \( n \).

\[\blacksquare\]

**Literature**


