Journal of Economics (2023) 140:181–207
https://doi.org/10.1007/s00712-023-00835-w

Imperfect collusion in monitored markets with free entry

Ludwig von Auer[1] · **Tu Anh Pham**[1]

Received: 6 July 2023 / Accepted: 13 July 2023 / Published online: 7 August 2023
© The Author(s) 2023

Abstract
Surveys of antitrust cases reveal that colluding firms usually (1) attempt to minimise the risk of prosecution, (2) achieve merely imperfect levels of collusion, (3) compete against some independently acting firms, and (4) adjust to market entries and exits. In contrast, existing oligopoly models neglect some of the four listed stylised facts and, thus, overlook important interdependencies between them. Therefore, the present paper develops a general quantity leadership model that simultaneously accommodates all four stylised facts. The model is a three-stage game in which each firm must make three consecutive decisions: market entry or not, collusion or not, and output quantity. The framework is augmented by an antitrust authority that ensures free market access. In addition, the antitrust authority may directly obstruct collusion and it may threaten prosecution. The results of this study indicate that the latter two instruments are rather ineffective.

Keywords Cartel · Game · Leadership · Oligopoly · Stability · Sustainability

JEL Classification L13 · L41

1 Introduction

Surveys of antitrust cases reveal many interesting facets of collusive behaviour.[1] Four stylised facts stand out. The colluding firms (1) introduce elaborate arrangements to minimise the risk of being caught and punished by antitrust authorities, (2)

[1] See, for example, Harrington (2006) or Levenstein and Suslow (2006).

A former version of this study was presented at a research seminar of BETA at Université de Strasbourg. Helpful comments and suggestions from participants are gratefully acknowledged. The detailed advice of an anonymous referee greatly improved the presentation. We also thank Xenia Matschke, Matthew Mitchel, and Volker Schulz for their valuable comments on an earlier draft.

✉ Ludwig von Auer
vonauer@uni-trier.de

[1] Universität Trier, Trier, Germany

struggle to enforce their agreement and, therefore, develop sophisticated means to improve compliance, (3) usually compete against some independently acting firms, and (4) operate on markets characterised by occasional market entries and exits. The four listed empirical facts translate into four features that a comprehensive model of collusive behaviour should accommodate.

The first empirical fact (exposure to antitrust surveillance) is the *antitrust* issue of the colluding firms. The most common instruments of antitrust policy are (a) ensuring free market access, (b) making collusion more difficult, and (c) discouraging collusion through law enforcement. To compare the efficacy of these instruments, a model of collusive behaviour must accommodate them in an appropriate form.

The antitrust surveillance forces the colluding firms to renounce legally binding contracts and to operate in secrecy. This limits the ability of the firms to coordinate their actions and to prevent new competitors from entering the market. Therefore, colluding firms develop other means of enforceable coordination. This is the second empirical fact (incomplete internal compliance). It highlights the issue of *output sustainability*. A mutual agreement on output is sustainable if no firm wants to deviate from this agreement. This is the first condition of a comprehensive oligopoly model's equilibrium solution.

The third empirical fact (competition from independent firms) raises the issue of *status stability*. This type of stability requires that no firm wants to change its status, that is, no colluding firm has the incentive to become an independently acting firm and, at the same time, no independently acting firm has the incentive to join the group of colluding firms.[2] This is the second condition of the equilibrium solution.

The fourth empirical fact (market entries and exits) emphasises a problem that is labelled here as *market stability*. A market is regarded as stable, if no firm wants to enter or exit it. This is the third condition of the equilibrium solution.

To the best of our knowledge, none of the existing oligopoly models that analyse output sustainability or status stability (few consider both) also include the issues of antitrust surveillance and market stability. However, such a comprehensive approach is necessary to capture the subtle interdependence between these issues and to compare the efficacy of the three instruments of antitrust policy. Therefore, the overall contribution of the present paper is the joint analysis of all four issues within one comprehensive model. The results of the model indicate that ensuring free market access is a very effective remedy against collusion, while making collusion more difficult and discouraging collusion through law enforcement provides no significant additional welfare gains.

The model is based upon the leadership approach, that is, the colluding firms appropriate the Stackelberg leadership and the other firms act as independent Stackelberg followers. The studies of d'Aspremont et al. (1983), Donsimoni (1985), Donsimoni et al. (1986), and Prokop (1999) utilise the price leadership model. The

[2] Studies of illegal collusion often denote the group of colluding firms as "the cartel" and the independent competitors as "fringe firms". Since the framework developed in the present study is general enough to allow for legal (or tacit) as well as illegal collusion, it avoids this terminology.

quantity leadership model is applied by Shaffer (1995), Lofaro (1999), Konishi and Lin (1999), Zu et al. (2012), Auer and Pham (2021), and by the model of the present study.[3] Martin (1990) considers both variants. All listed studies neglect the problem of market stability. Instead, the primary topic of the leadership approach is status stability. The problem of output sustainability is usually evaded by simply assuming that the leader acts like a single firm even though it would be profitable for an individual firm to deviate from the agreement.[4]

Output sustainability is extensively studied in supergames (repeated oligopoly games) with grim-trigger strategies or some other strategies that form a subgame-perfect Nash equilibrium. This strand of literature has been pioneered by Friedman (1971). Output sustainability is exclusively ensured by the threat of penalties imposed by the other colluding firms. Other motivations for compliance (e.g., mutual trust and/or transparency) are not considered. Supergames that simultaneously tackle status stability and penalty-enforced output sustainability include Escrihuela-Villar (2008, 2009), Bos (2009), and Bos and Harrington (2010). Bos and Harrington (2015) augment the analysis by an antitrust policy. Same as the existing leadership models, these supergames do not deal with the problem of market stability.

The supergame approach is a dynamic framework of imperfect competition. The simplified time dimension of the present study's leadership approach can be viewed as a "reduced-form" representation in the sense that the controversies and complexities associated with discounting a distant future are avoided, while the economic aspects of the collusion's fragility and the relationship between the colluding firms and their independent rivals are fully preserved. The reduced form creates the possibility to design a comprehensive oligopoly model that simultaneously addresses all four features of collusive behaviour and, therefore, their interdependence.[5]

The model is developed as a game with three stages and solved by backward induction. The paper is organised accordingly. Section 2 introduces the last stage of the game. At this stage, the number of operating firms and their individual status (colluding leader or independent follower) are given. For each firm, the profit-maximising output is derived. It satisfies output sustainability, that is, no firm wants to deviate from its output decision. The second stage is presented in Sect. 3. The firms can choose their preferred status and output, though the number of operating firms is still given. The solution satisfies output sustainability and status stability, that is, no firm has the desire to change its status and output. In Sect. 4, the first stage of the game is added. The firms decide whether they want to enter the market. The unique subgame-perfect equilibrium is derived. It satisfies output

[3] Some important results derived by Martin (1990) and Shaffer (1995) coincide with findings presented in Selten's (1973) pioneering study.

[4] A notable exception is Lofaro (1999).

[5] A less closely related strand of literature deals with common ownership or cross-ownership. Such arrangements facilitate collusion by increasing mutual transparency and by reducing the incentives for aggressive competitive behaviour. For a brief review see, for example, López and Vives (2019, pp. 2399-2400).

sustainability, status stability, and market stability. Furthermore, the implications for the design of an effective antitrust policy are discussed. Concluding remarks are contained in Sect. 5.

2 Output sustainability

This section is devoted to the final stage of the game. Thus, the number of operating firms and their individual status has been determined on the previous two stages.

2.1 Overview

The inverse demand function for a homogeneous product is $P = a - bQ$, where P is the market price and Q is the aggregate quantity produced. The industry consists of a given finite number of $n \geq 1$ identical firms. Only integer numbers of firms are considered. All n firms have a constant marginal cost equal to c and a positive fixed cost that can represent an entry cost or a cost of production.

Without loss of generality, we can change the currency and the units in which output is measured (e.g., Selten 1973, p. 144). An original unit of output is equivalent to $b/(a-c)$ new units of output and a unit of the original currency is equivalent to $b/(a-c)^2$ units of the new currency. With this normalization, the new values of the parameters a, b, and c yield $a - c = 1$ and $b = 1$. Thus, the market volume (perfect competition output), $(a-c)/b$, is normalised to 1. The normalised fixed cost is denoted by $f \leq 1/9$. The upper bound is the gross profit (revenues minus variable cost) of Cournot duopolists.

Of the n operating firms, a given group of $k \in (2, \ldots, n)$ firms appropriates the role of the Stackelberg leader. The group of leaders competes against $(n-k)$ Stackelberg followers.[6] They consider the leaders' output as given and compete on the residual demand.

Given the reaction function of the followers and the level of collusion of the leaders, the latter determine their profit maximising joint *sustainable output* Q_K (details in Sect. 2.4). The output Q_K is sustainable in the sense that no firm has an incentive to deviate from its output decision. Inserting the output Q_K into the reaction function of the followers yields the aggregated output of the followers Q_F, the total output $(Q = Q_K + Q_F)$, the profit of the $(n-k)$ followers, and the profit of the k leaders.

For each industry size, n, and number of leaders, k, such a sustainable equilibrium can be derived. Only in Sect. 3, the firms can choose whether they want to join the group of leaders or prefer the status of an independent follower, that is, we add the issue of status stability and the number of leaders, k, becomes endogenous. The issue of market stability is introduced in Sect. 4 through the endogeneity of the number of operating firms, n.

[6] Huck et al. (2007) provide some experimental evidence that firms that cooperate in a binding manner show leadership behaviour, whereas the remaining firms exhibit follower behaviour. Brito and Catalão-Lopes (2011, p. 3) summarise the justifications for assuming that colluding firms act as a leader.

2.2 Reaction function of followers

The followers consider the joint output of the leaders as given and compete on the residual demand. The followers recognise the interdependence of their individual quantity decisions. Therefore, quantity leadership models assume that the followers are in Cournot competition to each other.

The profit of a follower is

$$\pi_F = P(Q)q_F - cq_F - f = \left(1 - Q_K - Q_{-F}\right)q_F - q_F^2 - f \ ,$$

where q_F is the follower's output and Q_{-F} is the aggregate output of all other followers. Each follower considers Q_{-F} and the leaders' output, Q_K, as given. Exploiting the symmetry of the followers, their profit maximising total output is

$$Q_F = (n-k)\frac{1 - Q_K}{n - k + 1} \ . \tag{1}$$

This is the reaction function of the followers, taking Q_K as given.

Inserting this result in the demand function yields

$$P(Q) - c = \frac{1 - Q_K}{n - k + 1} \ . \tag{2}$$

This mark-up incorporates the profit maximising reaction of the $(n-k)$ followers to the leaders' output Q_K. The mark-up does not depend on the process by which Q_K is determined. In our model, the leaders opt for a sustainable output. What is sustainable depends on the leaders' level of collusion.

2.3 Antitrust authority

To address the antitrust issue, the model includes an antitrust authority that monitors the market and attempts to detect and punish illegal collusive behaviour. Let $\Pr \in [0, 1]$ denote the leaders' perception of the probability of a successful conviction. Such a conviction requires not only effective market surveillance but also success in court. Only if the conviction is successful, the leaders must pay a fine. The fine is proportional to the leaders' gross profit. The factor of proportion is denoted by $\phi \in [0\,,\, 1/\Pr]$.

Thus, the gross profit of a detected leader is $(P-c)q_K(1-\phi)$ and a leader's expected profit is

$$\begin{aligned} E(\pi_K) &= (1 - \Pr)(P-c)q_K + \Pr(P-c)q_K(1-\phi) - f \\ &= (P-c)q_K(1-p) - f \ , \end{aligned} \tag{3}$$

with $p = \Pr \cdot \phi \in [0, 1]$ denoting the antitrust policy's *rigour*. Analogously, the expected average profit of the other leaders is

$$E\left(\bar{\pi}_{-K}\right) = (P-c)\bar{q}_{-K}(1-p) - f \ , \tag{4}$$

where $\bar{q}_{-K}$ is the average output of the other leaders.

In the U.S. antitrust policy and the EU competition policy, collusion is not necessarily illegal. Only explicit (or formal) collusion is prosecuted, while tacit collusion is not (e.g., Martin 2006, p. 1300). In practice, the distinction between explicit and tacit collusion is difficult. For the courts, it is easier to observe the market price P than the n firms' total output Q. If Q is equal to the Cournot output, $n/(n+1)$, the observable market price will be indistinguishable from the market price arising in Cournot competition. Therefore, this Cournot price may serve as a reference. When the market price falls below the Cournot price (that is, $Q > n/(n+1)$), the courts will be more inclined to consider the behaviour of the firms as unsuspicious. In the extreme case, the probability of a successful conviction, Pr, becomes 0, and so does the antitrust authority's rigour, $p = \Pr \cdot \phi$. If the market price rises above the Cournot price ($Q < n/(n+1)$), the courts become less generous and the antitrust authority can increase p.[7]

2.4 Imperfect collusion

A single firm cannot proclaim itself as Stackelberg leader. To appropriate the leadership position requires collusion and this involves at least two firms: $k \geq 2$.[8] In quantity leadership models, the leaders' collusion has two levels. The basic level is the appropriation of the Stackelberg leadership and the upper level is the coordination of its output decisions.

The standard quantity leadership model assumes that perfect collusion prevails, that is, not only the cooperation at the basic level of collusion (appropriation of Stackelberg leadership) is perfectly smooth but also at the upper level of collusion (coordination of quantities). The leaders act as if they were the subsidiaries of a company that determines its joint profit maximising quantity without worrying that individual subsidiaries may deviate from it (that is, produce a larger quantity).[9]

As the standard model, our own quantity leadership model assumes that the colluding firms successfully appropriate the Stackelberg leadership. However, our model generalises the standard model with respect to the upper level of collusion. More specifically, it allows for all possible degrees of *quantity coordination*, that is, from completely ineffective coordination of the output decisions (within their group, the leaders act like perfect competitors) to perfect coordination of these decisions (the leaders act like a merged firm).

Perfect quantity coordination is rarely feasible. However, antitrust proceedings reveal that firms are impressively innovative in establishing mechanisms that ensure at least imperfect coordination. The achieved degree of coordination varies widely.

[7] This becomes relevant only in Sect. 4.3 where we present a detailed analysis of the impact of rigour, p, on the subgame-perfect equilibrium and on output and welfare.

[8] This condition requires that $n \geq 2$. In the monopoly case ($n = 1$) collusion is redundant. The monopolist's gross profit is $\pi = 1/4$ and total output is $Q = 1/2$.

[9] The case of perfectly colluding leaders competing against independent Cournot competitors is explored in the quantity leadership models of Martin (1990, pp. 9-12, 1993, pp. 98-100), Shaffer (1995, pp. 744-749) as well as Auer and Pham (2021). Note that perfect quantity coordination is not sufficient to establish a monopoly. The latter also requires that no independent competitor remains.

It depends on aspects such as the antitrust authority's diligence and resources, the nature of the judicial system, the type of product, the number of firms, the market size, the cost function, the degree of product differentiation, or the potential for regional separation. Equally important are the design of the collusive agreement (market strategy, trust building, surveillance mechanism, and system of internal sanctions), behavioural fundamentals (e.g., performance-linked payment systems and personal relationships between top level managers of different firms), and institutional structures (e.g., industry federation and ownership arrangements).[10]

To accommodate in our model the broadest possible spectrum of institutional arrangements and behavioural assumptions, we modify a game-theoretic concept that Cyert and DeGroot (1973, p. 25) coined as the "coefficient of cooperation"-approach.[11] In our modified approach, we define the *coefficient of coordination* $\lambda \in [0, 1]$. Instead of its own expected profit, $E(\pi_K)$, each leader maximises the following *compound profit*:

$$\hat{\pi}_K = E(\pi_K) + (k\lambda - 1)E(\bar{\pi}_{-K}) \ . \tag{5}$$

In this objective function, the expected average profit's weight, $(k\lambda - 1)$, is a strictly positive monotonic transformation of the coefficient of coordination, λ.[12] When the quantity coordination of the group of leaders is completely dysfunctional, the coefficient of coordination is $\lambda = 0$ and the weight becomes (-1): $\hat{\pi}_K = E(\pi_K) - E(\bar{\pi}_{-K})$. Each group member maximises the difference between its own expected profit and the expected average profit of the other leaders, that is, only relative performance matters. For $\lambda = 1/k$, the weight is 0: $\hat{\pi}_K = E(\pi_K)$. Each leader maximises only its own expected profit. The performance of the other leaders is irrelevant. For $\lambda = 1$, the weight is $(k - 1)$: $\hat{\pi}_K = E(\pi_K) + (k - 1)E(\bar{\pi}_{-K})$. Each leader maximises the group's expected total profit, that is, only the group's joint performance matters.

The coefficient of coordination, λ, is a measure of the *level* of collusion and does not discriminate against any *means* of coordination. More specifically, the admissible means include not only mutual surveillance and penalties (as in the supergame literature) but also regular communication, transparency, and other trust-building measures (as documented in many real-world cartel cases).

[10] For example, Matsumura and Matsushima (2012) and Matsumura et al. (2013) point out that managers often do not only care about their own firm's profit (the Cournot-Nash case) but also about their firm's profit relative to the profit of the competing firms. The perfect competition outcome arises when each firm cares only about its relative profit.

[11] Martin (1993, pp. 30-31) and Escrihuela-Villar (2015, pp. 476-479) demonstrate the close correspondence between the coefficient of cooperation and Bowley's (1924) concept of conjectural variation.

[12] Edgeworth (1881, p. 53, fn. 1) proposes a similar specification of the objective function and calls the term $(k\lambda - 1)$ the "coefficient of effective sympathy". Instead of the weight $(k\lambda - 1)$, Lofaro (1999) and López and Vives (2019) use in their objective functions the weight $\lambda(k - 1)$. The latter study is concerned with the consequences of partial cross-ownership, while the former study introduces an oligopoly model of imperfect collusion that resembles the last two stages of our own three-stage game. However, Lofaro's model provides no closed form solution and neither the antitrust issue nor the issue of market stability are considered.

Inserting Eqs. (2), (3), and (4) into a leader's compound profit defined by Eq. (5) yields

$$\hat{\pi}_K = (1-p)\frac{1-q_K-(k-1)\bar{q}_{-K}}{n-k+1}\left[q_K+(k\lambda-1)\bar{q}_{-K}\right]-k\lambda f\ . \tag{6}$$

2.5 Profit maximising sustainable quantities

Each leader maximises its compound profit, $\hat{\pi}_K$, and takes the average output of the other leaders, $\bar{q}_{-K}$, as given.[13] Differentiating Eq. (6) with respect to q_K leads to the first-order condition

$$[k(1+\lambda)-2]\bar{q}_{-K}+2q_K=1\ .$$

Since all leaders are identical, a symmetric solution arises. Substituting $\bar{q}_{-K}$ by q_K yields the leaders' sustainable total output

$$Q_K = kq_K = \frac{1}{1+\lambda}\ . \tag{7}$$

It is independent of the number of leaders, k, and the number of firms on the market, n. With this output, no leader has an incentive to deviate from its own output decision. Equation (7) reveals that $\partial Q_K/\partial\lambda<0$.

Since the status of each firm is given, each leader considers the risk of being prosecuted by the antitrust authority as an unavoidable cost. Thus, q_K is independent of the antitrust authority's rigour, $p=\Pr\cdot\phi$.

Inserting Eq. (7) in Eq. (1) gives the sustainable total output of the followers:

$$Q_F=(n-k)q_F=\frac{n-k}{n-k+1}\ \frac{\lambda}{1+\lambda}\ . \tag{8}$$

Therefore, $\partial Q_F/\partial\lambda>0$ and $Q_F<Q_K$. When $\lambda=0$, the total output of the leaders is equal to the market volume ($Q=1$) and the followers produce no output.

Equations (7) and (8) define the sustainable equilibrium for given values of n and k. Thus, total (sustainable) output is

$$Q=Q_K+Q_F=\frac{n-k+1/(1+\lambda)}{n-k+1}\ . \tag{9}$$

Q increases with the number of followers, $\partial Q/\partial(n-k)>0$, and decreases when the level of coordination increases, $\partial Q/\partial\lambda<0$. When n and k simultaneously change by the same number (e.g., a leader leaves the market), Q is not affected. The antitrust policy's rigour, p, has no direct effect on Q.

[13] From a formal game-theoretic perspective, this behaviour of the individual members of the cartel could be considered as "non-cooperative". Note, however, that for $\lambda>1/k$ this "non-cooperative" behaviour is quite "cooperative" because each member maximises the compound profit, $\hat{\pi}_K$, instead of its own expected profit, $E(\pi_K)$. For $\lambda<1/k$, the behaviour takes on a "destructive" component.

Equation (9) is a very general expression of total output, because it covers the equilibrium output of all standard oligopoly models that use quantities as the strategic variable. For $k = n$ and $\lambda = 1/n$, the total output of the standard Cournot model arises: $Q = n/(n+1)$. For $k = 1$ (which we have ruled out) and $\lambda = 1$, Eq. (9) gives $Q = (2n-1)/(2n)$ which is the total output of the standard Stackelberg model. The monopoly output, $Q = 1/2$, arises for $k = n$ and $\lambda = 1$, while the market volume arising on perfectly competitive markets ($Q = 1$) is obtained for $\lambda = 0$.

2.6 Sustainable profits

Inserting Eq. (7) in Eq. (2), the mark-up simplifies to

$$P(Q) - c = \frac{\lambda}{(n-k+1)(1+\lambda)} \qquad (= 1 - Q)\ . \tag{10}$$

It increases with the level of coordination, λ.

Inserting $q_K = 1/[(1+\lambda)k]$ and Eq. (10) in Eq. (3) yields the sustainable expected profit of each leader:

$$E[\pi_K(k)] = \left(\frac{\lambda}{1+\lambda}\right)^2 \frac{\Omega}{(n-k+1)k} - f\ , \tag{11}$$

where $\Omega = (1-p)/\lambda$. The notation $E[\pi_K(k)]$ emphasises that the expected profit depends on the number of leaders, k. Similarly, inserting $q_F = \lambda/[(1+\lambda)(n-k+1)]$ and Eq. (10) in $\pi_F(k) = (P-c)q_F - f$, gives the sustainable profit of each follower:

$$\pi_F(k) = \left(\frac{\lambda}{1+\lambda}\right)^2 \frac{1}{(n-k+1)^2} - f\ . \tag{12}$$

Both, $E[\pi_K(k)]$ and $\pi_F(k)$ are increasing with the coefficient of coordination, λ.

Since $\lambda^2\Omega = \lambda(1-p)$, profit function (11) gives for $p = 1$ or $\lambda = 0$ negative expected profits. Thus, in the remainder of this study, we restrict our attention to the case $\lambda > 0$ and $p < 1$. For given p, an increase in the level of coordination, λ, reduces the output of the leaders and, therefore, increases the residual demand available for the followers. As a consequence, not only the leaders but also the $(n-k)$ followers benefit from improved coordination among the leaders. The optimal situation for the leaders and followers is a perfectly colluding group of leaders ($\lambda = 1$).

The profit functions (11) and (12) directly yield the following findings:

Lemma 1 *The sustainable profit of a follower exceeds the expected sustainable profit of a leader, if and only if* $k > \Omega(n+1)/(\Omega+1)$.

Furthermore, we can derive the following result:

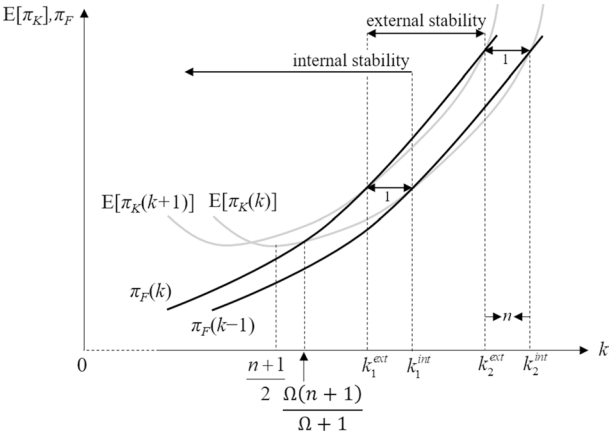

Fig. 1 Profits as a function of the number of leaders

Lemma 2 *The profit function of a leader,* $E[\pi_K(k)]$*, is convex with the minimum profit at* $k = (n+1)/2$*. The profit function of a follower,* $\pi_F(k)$*, is convex, too, but with the minimum profit at* $k = 0$*.*

Proof See Appendix A.

The graphical implications of Lemmas 1 and 2 are depicted in Fig. 1. It shows the $E[\pi_K(k)]$-curve, the $\pi_F(k)$ -curve and their intersection. All other elements of Fig. 1 will be explained shortly.

In the derivation of the sustainable total output defined in Eq. (9), the status of the firms was fixed during the previous stages of the game, that is, n and k were given. As a consequence, the solution was independent of the antitrust policy's rigour, p. In the next section, each firm decides on its own status, that is, k becomes endogenous. This is the penultimate stage of the game.

3 Status stability

Output sustainability merely ensures that no cheating occurs within the group of leaders. It does not preclude situations in which a leader wants to become a follower (violation of internal stability; k decreases by one) or a follower wants to become a leader (violation of external stability; k increases by one). In other words, the sustainable solution defined by Eq. (9) is not necessarily stable. For given n, we derive the unique integer value k which ensures that the solution defined by Eq. (9) is not only sustainable but also stable.

3.1 Definition of status stability and important thresholds

Following Selten (1973) and d'Aspremont et al. (1983), the stability of a group of k leading firms requires that both the condition for internal stability,

$$E[\pi_K(k)] > \pi_F(k-1)\,, \tag{13}$$

and the condition for external stability,

$$E[\pi_K(k+1)] \leq \pi_F(k)\,, \tag{14}$$

are satisfied.[14] The concept of stability implicitly assumes that only one firm at a time can change its status. The only exception is the case of $k = 2$. Since a single firm cannot appropriate the leadership position, the internal stability of a group of two leaders requires that its expected profit is larger than the expected profit that arises when no leader exists, that is, larger than the profit in the standard Cournot model: $E[\pi_K(2)] > 1/(n+1)^2 - f$.[15]

Let k_1^{ext} denote the smallest k-value at which $E[\pi_K(k+1)] \leq \pi_F(k)$. It defines the minimum k-value for external stability (see Fig. 1). Similarly, let k_1^{int} denote the smallest k-value at which $E[\pi_K(k)] \leq \pi_F(k-1)$. All k -values smaller than k_1^{int} ensure that internal stability prevails. The formulas for the compilation of k_1^{ext} and k_1^{int} are derived in Sect. 3.2. Figure 1 reveals that $k_1^{\text{int}} = k_1^{\text{ext}} + 1$. Therefore, in the interval $[k_1^{\text{ext}}, k_1^{\text{int}})$ only a single integer exists. We denote this integer by k^*. It is the *unique* integer value of k that satisfies both, the internal stability condition (13) and the external stability condition (14). In other words, k^* is the unique equilibrium number of leaders that ensures output sustainability and status stability.

3.2 Equilibrium analysis

It is helpful to distinguish between the case of a complete group of leaders ($k = n$) and the case of an incomplete group of leaders ($n - k \geq 1$). When a complete group arises, the sustainable expected profit of each leader is $E[\pi_K(n)] = \Omega\lambda^2/[n(1+\lambda)^2] - f$ and the external stability condition is redundant because no follower exists. The internal stability condition of a complete group of leaders is $E[\pi_K(n)] > \pi_F(n-1)$. If one leader leaves this group, this leader becomes the first follower. The associated profit is $\pi_F(n-1) = \lambda^2/[4(1+\lambda)^2] - f$. Therefore, we directly obtain the following result:

Theorem 1 *For $n < 4\Omega$, a complete group of leaders is internally stable, that is, the equilibrium is a group of $k^* = n$ leaders.*

[14] In his formulation of stability, Selten (1973, pp. 179-181) denotes the colluding firms as the "participators" while the followers are called the "non-participators". Note that in d'Aspremont et al. (1983) the internal stability condition (13) has a weak inequality sign, while the external stability condition (14) has a strict inequality sign. We prefer our own definition because it implies that firms prefer the status of a (possibly prosecuted) leader only if the profit is strictly larger than that associated with the legal status of a follower.

[15] Conversely, when $k = 0$, external stability requires that $E[\pi_K(2)] \leq 1/(n+1)^2 - f$.

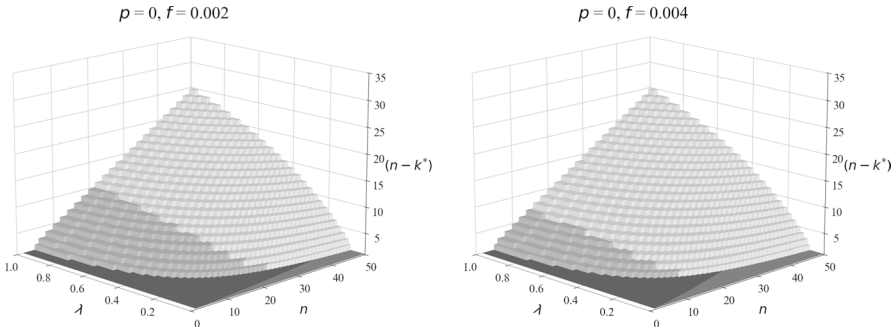

Fig. 2 The Equilibrium Values $(n - k^*)$ As a Function of the Total Number of Firms, n, and the Degree of Quantity Coordination, λ, for $p = 0$ and $f = 0.002$ as well as $f = 0.004$

When n is at least as large as 4Ω, an incomplete group of leaders arises. It is characterised by the following finding:[16]

Theorem 2 *For $n \geq 4\Omega$, the sustainable and stable equilibrium number of leaders, k^*, is either 0 (Cournot competition) or it is the unique integer satisfying the condition $1 < k_1^{ext} \leq k^* < k_1^{int} \leq n$, where $k_1^{int} = k_1^{ext} + 1$ and*

$$k_1^{ext} = \frac{n - 1 + 2\Omega(n+1) - z}{2(\Omega + 1)}\,, \tag{15}$$

with

$$z = \sqrt{(n+1)^2 - 4\Omega(n+2)}\,. \tag{16}$$

Proof See Appendix B

Corollary 1 *For $n \geq 4\Omega$, the first-order derivatives of k_1^{ext} with respect to λ, n, and p yield*

$$\frac{\partial k_1^{ext}}{\partial \lambda} < 0\,, \qquad \frac{\partial k_1^{ext}}{\partial n} > 0\,, \qquad \text{and} \qquad \frac{\partial k_1^{ext}}{\partial p} < 0\,.$$

Proof See Appendix C.

The numerical implications of Theorems 1 and 2, and Corollary 1 are illustrated in Fig. 2. It comprises two diagrams. Both correspond to $p = 0$, while the f-values differ. Both diagrams look like a big flight of stairs. The height of each step

[16] In Appendix B (proof of Theorem 2) it is shown that the largest k -value consistent with external stability is $k_2^{\text{ext}} = k_1^{\text{ext}} + z/(\Omega + 1)$ and that all k-values larger than $k_2^{\text{int}} = k_1^{\text{int}} + z/(\Omega + 1)$ satisfy the internal stability condition. However, it is also shown that, for $n \geq 4\Omega$, the integer in the interval $[k_2^{\text{ext}}, k_2^{\text{int}})$ is n (this case is depicted in Fig. 1). Then, the threshold $k_2^{\text{int}} (> n)$ is irrelevant for the formal analysis.

(measured from the bottom of the diagram) shows, for the given p-f-combination, the equilibrium number of followers, $(n - k^*)$, corresponding to the respective values of λ and n. The dark area in front of the steps represents the λ-n-combinations leading to a complete group of leaders.[17]

For given n, the equilibrium number of followers *increases* with the cartel's degree of quantity coordination, λ. It is worthwhile to explain this somewhat counterintuitive relationship. Strengthened coordination, λ, reduces each leader's output, q_K, and total output, Q, but increases the output of each follower, q_F; see Eq. (9). As a consequence, the mark-up, $(P - c)$, increases. The increase of $(P - c)$ in conjunction with the increase of q_F and the reduction of q_K implies that, for given k, the increase of λ raises the profit of each follower by more than the expected profit of each leader. If this difference is sufficiently large, the internal stability condition may no longer hold and one leader may want to become a follower, that is, k^* falls and $(n - k^*)$ increases. This reasoning confirms our earlier claim that important interdependencies between the sustainability issue and the stability issue exist.

In contrast to all existing quantity leadership models, we were able to derive a closed solution for the sustainable and stable equilibrium. As a consequence, we can proceed to study the consequences of market entries and exits. That is, n becomes endogenous. We consider this as an important step forward because ensuring free market access is a viable antitrust policy of its own and previous models were unable to compare its effectiveness to that of surveillance, prosecution, and punishment of the colluding firms.

4 Market stability and antitrust policy

In Sects. 2 and 3 we analysed the subgame formed by the last two stages of our three-stage game. For this subgame, we derived the equilibrium (n, k^*). It specifies for each given number of firms, n, the equilibrium size of the cartel, k^*. This equilibrium satisfies *output sustainability* as well as *status stability*.

The *antitrust policy* issue is related to the parameters p and λ. The parameter p indicates the rigour of the implemented antitrust policy, while the parameter λ reflects the level of quantity coordination within the group of leaders. The antitrust authority can influence the values of these two parameters. For example, it can introduce stricter surveillance that complicates the leaders' coordination. However, the policy portfolio of antitrust authorities is not limited to a change of the parameters p and λ. Another important policy option is to ensure free market access (e.g., effective measures against predatory pricing and against firms that penalise clients that order from new entrants).

When free market access is ensured, new firms can enter the market. Therefore, the entry decision is the first stage of our three-stage game. In its entry deliberations, each firm anticipates its status decision (the second stage, discussed in Sect. 3) and its profit maximising output decision (the final stage, discussed in Sect. 2). In other

[17] The meaning of the different shades of grey will be described in Sect. 4.1.

words, the total number of firms, n, becomes endogenous raising the issue of *market stability*. A market is stable if no firm wants to enter or exit the market. A firm wants to enter the market if and only if this is weakly profitable. Strictly unprofitable firms leave the market.

In Sect. 4.1, we derive the subgame-perfect equilibrium of our three-stage game as a function of the remaining parameters λ, p, and f. Furthermore, we show that an antitrust policy that ensures free market access raises the level of welfare. In Sects. 4.2 and 4.3, we investigate the other two antitrust policy options: the reduction of the degree of quantity coordination, λ, and the increase of the antitrust policy's rigour, p. Our analysis examines whether these additional measures further increase the level of welfare.

Even though policy measures exist that possibly affect the barriers to entry as well as the parameters p and λ (e.g., a more generous leniency program or more effective protection of employees that become whistle-blowers), we keep the analysis of these three policy options separate from each other. This allows us to identify the individual contributions of each option.

With the endogeneity of n comes a minor modification in notation. Henceforth, the expected profit of a leader is denoted by $E\left[\pi_K(n,k)\right]$ instead of $E\left[\pi_K(k)\right]$. The profits of a follower are represented by $\pi_F(n,k)$ instead of $\pi_F(k)$.

4.1 Antitrust policy I: free market access

The first stage of our game is the firms' entry decision. In the following, we derive the subgame-perfect equilibrium of our three-stage game. We denote this equilibrium by (n^{**}, k^{**}).

Let $(n+1, k^{*\prime})$ denote the equilibrium of the subgame formed by the last two stages when $(n+1)$ firms operate on the market. Since we assumed that the normalised fixed cost is not larger than the profit of a Cournot duopolist ($f \leq 1/9$), a market with only two firms ($n = 2$) is always profitable. If for $n = 2$ we have

$$E\left[\pi_K(n+1, k^{*\prime})\right] \geq 0 \qquad \text{and} \qquad \pi_F(n^* + 1, k^{*\prime}) \geq 0\,, \tag{17}$$

a third firm enters the market raising the number of firms to $n = 3$. As long as conditions (17) are satisfied, additional firms enter the market. This process stops only when an additional firm would trigger negative profits. Thus, the subgame-perfect equilibrium (n^{**}, k^{**}) is characterised by the following conditions:

$$E\left[\pi_K(n^{**}, k^{**})\right] \geq 0 \qquad \text{and} \qquad \pi_F(n^{**}, k^{**}) \geq 0\,, \tag{18}$$

but

$$E\left[\pi_K(n^{**} + 1, k^{**\prime})\right] < 0 \qquad \text{and/or} \qquad \pi_F(n^{**} + 1, k^{**\prime}) < 0\,, \tag{19}$$

where $(n^{**} + 1, k^{**\prime})$ denotes the equilibrium for the subgame formed by the last two stages of the game when $(n^{**} + 1)$ firms operate on the market.

These considerations are illustrated in Fig. 2 (p. 16). One can see that for each given λ-value the height of the stairs is non-decreasing in n. For example, we know from Theorem 1 that for the parameter values $\lambda = 0.4$, $f = 0.002$, and $p = 0$ (left diagram of Fig. 2), the maximum size of a complete group of leaders is $n = 9$. When n increases to 10, the entering firm does not want to join the group of leaders but prefers to become the first follower. The expected profit of the leaders is still positive. Therefore, further firms enter until $n = 22$ firms operate on the market. From Eq. (15) of Theorem 2 we know that the corresponding equilibrium number of leaders is $k^* = 17$, while the remaining five firms are followers. From the profit function (3) it can be seen that the leaders' expected profit is still positive. If an additional firm entered the market, the corresponding equilibrium number of leaders would increase to $k^* = 18$. Inserting $n = 23$ and $k^* = 18$ in the profit function (3) reveals that the expected profit of the leaders would become negative. In the left diagram of Fig. 2, this transition from profitability to loss-making is highlighted by the changeover to a lighter shade of grey.

In sum, a market with the parameter values $\lambda = 0.4$, $f = 0.002$, and $p = 0$ can support up to 22 firms, 17 of which form the group of leaders. Therefore, $(n^{**}, k^{**}) = (22, 17)$ is the subgame-perfect equilibrium. It satisfies not only output sustainability and status stability but also market stability. For each λ-p-f-combination the subgame-perfect equilibrium (n^{**}, k^{**}) can be derived. It is defined by the market stability conditions (18) and (19) in conjunction with Theorems 1 and 2.

What are the welfare consequences of the antitrust authority's free market access policy? We define welfare as the sum of consumer and producer rent. This sum is equal to $(Q - 0.5Q^2)$, where the value of total output, Q, is defined by Eq. (9). Recall that the perfect competition output is 1. For Q-values smaller than 1, welfare and total output, Q, are positively correlated. Therefore, we can confine the welfare analysis to an analysis of total output, Q. From Eq. (9) we know that total output depends only on the degree of coordination, λ, and the equilibrium number of followers, $(n^{**} - k^{**})$. Therefore, for given λ, in Fig. 2 a higher step represents a higher welfare.

Positive profits induce new firms to enter the market. However, the output of the group of leaders is $Q_K = \lambda/(1 + \lambda)$, regardless of the number of leaders. Thus, only those new firms that become followers increase total output and, therefore, welfare.

Figure 2 illustrates the welfare effect of the free market access policy. Returning to the example with $\lambda = 0.4$, $f = 0.002$, and $p = 0$ (left diagram in Fig. 2), the equilibrium corresponding to $n = 9$ is a complete group of leading firms: $(n, k^*) = (9, 9)$. Equation (9) implies that the associated total output is $Q = 0.714$. Free market access induces thirteen firms to enter the market, five of which become followers: $(n^{**}, k^{**}) = (22, 17)$. These followers raise total output to $Q = 0.952$. This increase represents a considerable welfare gain.

In many models, the fixed cost f is interpreted as a cost of market entry. The larger f, the larger the barriers to entry. Figure 2 illustrates the welfare consequences of a change of f. Comparing the two diagrams reveals that decreasing the entry cost, f, from 0.004 (right diagram) to 0.002 (left diagram) does not affect the flight of stairs but shifts the borderline between profitable and unprofitable short-run equilibria

upwards, that is, away from the origin ($\lambda = 0$ and $n = 0$). Thus, for each given λ-p -combination, a reduction of f increases the subgame-perfect equilibrium number of followers and, therefore, output and welfare. This merely reinforces our previous result: a reduction of the barriers of entry (here the entry cost f) increases welfare.

4.2 Antitrust policy II: obstructing collusion

The antitrust authority could complement its free market access policy by a more rigorous obstruction of collusive behaviour. For example, it could improve its surveillance of the industry federation and its protection of whistle-blowers. Such measures are likely to reduce the level of quantity coordination, λ.

We know from Eq. (9) that, for given n and k, a reduction of λ generates an increase in total output, Q. This is the direct effect of λ on Q. However, with free market access, there is also a less obvious indirect effect. A sufficiently strong reduction of λ also changes the subgame-perfect equilibrium (n^{**}, k^{**}). Somewhat paradoxically, the leaders' deteriorating quantity coordination tends to *raise* the number of leaders, k^{**}, and to *increase* the number of operating firms, n^{**} (see Corollary 1). Since the k^{**}-increasing effect dominates the n^{**}-increasing effect, the number of followers, $(n^{**} - k^{**})$, decreases and so does total output, Q.

Before discussing the logic behind this detrimental indirect effect, it is worthwhile to re-consider Fig. 2. It illustrates the negative relationship between λ and the subgame-perfect equilibrium number of firms, n^{**}. For each given λ-value, the equilibrium value n^{**} is indicated by the (last) changeover between the darker and lighter shade of grey. This boundary point separates the sustainable and stable equilibria with positive profits from those with negative profits. The combination of all boundary points shows the relationship between λ and n^{**}. A decrease in the degree of quantity coordination, λ, leads to a moderate increase of n^{**} until a complete group of leaders is reached.

Why does a deteriorating coordination of an incomplete group of leaders raise k^{**} and, to a lesser extent, also the number of operating firms n^{**}? Fig. 3 illustrates the answer. As in Fig. 2, the parameter values are $f = 0.002$ and $p = 0$. The graph shows the impact of λ on the subgame-perfect equilibrium (n^{**}, k^{**}) and the associated output Q. For $\lambda = 1$, the subgame-perfect equilibrium is $(n_1^{**}, k_1^{**}) = (21, 12)$ and total output is $Q = 0.9500$; see the right-hand side of Fig. 3. When λ is gradually reduced and reaches 0.8547, external stability is violated and one of the nine followers becomes a leader. The new subgame-perfect equilibrium is $(n_1^{**}, k_1^{**} + 1) = (21, 13)$. From Eq. (9) it follows that the changeover of the follower reduces total output. In Fig. 3, this fall in total output is shown by the kink at $\lambda = 0.8547$.

Figure 3 shows the typical relationship between λ and Q.[18] A reduction of λ leads to a zigzag pattern of declining output levels, Q, until some minimum output is reached. Comprehensive numerical analysis reveals that this general result is largely independent of the underlying p-f-combination. In Fig. 3, the minimum output is

[18] A detailed description and discussion of Fig. 3 is provided in Appendix D.

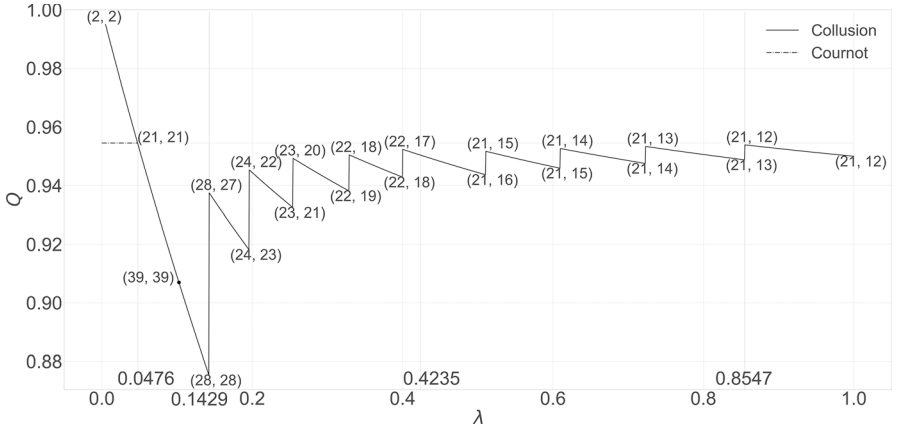

Fig. 3 The Total Output, Q, and the Subgame-Perfect Equilibria, (n^{**}, k^{**}), Corresponding to Different Values of λ

at $\lambda = 0.1429$. This is the λ-value at which the last follower joins the cartel. At that moment, the group of leaders becomes complete with 28 members and the external stability condition becomes redundant. The resulting output is considerably smaller than the output produced by a group of leaders that perfectly coordinate their quantities ($\lambda = 1$) plus the output of the competing followers.

When λ falls to $1/21 = 0.04762$, the size of the complete group of leaders falls to 21 members, that is, $\lambda = 1/k$. Therefore, the group's degree of quantity coordination is equivalent to the behaviour of 21 firms that are in Cournot competition with each other (see Sect. 2.5). The associated output is $Q = 0.955$. Remarkably, this output is only slightly larger than the output related to the market configuration with perfect quantity coordination ($\lambda = 1$) and free market access. The following theorem shows that this is a general result valid for all *p-f*-combinations.

Theorem 3 *If* (n^{**}, k^{**}) *with total output* Q' *is the subgame-perfect equilibrium corresponding to* $\lambda = 1$, *then* (n^{**}, n^{**}) *with total output* Q'' *is the subgame-perfect equilibrium corresponding to* $\lambda = 1/n^{**}$, *where*

$$1 < \frac{Q''}{Q'} < 1 + \frac{2}{(n+1)(n-2)} \,. \tag{20}$$

Proof See Appendix E.

The welfare and policy implications of the preceding discussions (Sects. 4.1 and 4.2) are rather obvious. The policy instrument of ensuring free market access is effective. Attempts to further increase welfare by obstructing collusion are largely futile.

4.3 Antitrust policy III: prosecution and punishment

The third antitrust policy instrument is the prosecution and punishment of the colluding leaders. This can be captured by an increase in the antitrust policy's rigour, $p = \Pr \cdot \phi$. Formally, such an increase can be accomplished by a more severe punishment, ϕ, or by a higher probability of a successful conviction, Pr. Possible means of increasing Pr include more efficient auditing, an expanded leniency program, or changes in the judicial system.

However, the scope for increasing p is limited. One limitation (not considered in the present paper) is the budget of the antitrust authority. The other limitation is the definition of illegal collusion. As pointed out in Sect. 2.3, when the n^{**} operating firms produce a total output that is not smaller than the Cournot output, $n^{**}/(n^{**}+1)$, courts tend to consider the market as unsuspicious. In the extreme, the probability of a successful conviction, Pr, approximates 0, and so does rigour, p. Then, raising p is not a viable policy option. However, when the total output of the n^{**} operating firms falls below the Cournot output, the antitrust authority might be able to implement a policy with $p > 0$.

To identify the welfare consequences of an increase of p, it again suffices to study the effect of p on total output, Q. A formal analysis is rather tedious because various cases and subcases can arise. However, a numerical analysis is straightforward. It reveals a clear pattern.

When the fixed cost, f, is small, many firms operate on the market. A continuous increase in p may induce leaders to exit the market but this does not lead to a continuous increase in the number of followers ($n^{**} - k^{**}$). When an increase of p "generates" a new follower, an additional increase of p usually reverts this increment because a follower exits the market. In other words, the number of followers fluctuates within a small range and so does total output, Q.

This is illustrated in Fig. 4 which connects to our previous example. The level of coordination is $\lambda = 0.4$ and the fixed cost is $f = 0.002$. The rigour, p, ranges from 0 to 1. The number of followers fluctuates between four and five. Accordingly, total output, Q, fluctuates between 0.942 and 0.953.

The dotted line in Fig. 4 indicates the output that would arise in Cournot competition with n^{**} firms. As p increases, total output, Q, remains consistently above this dotted line and the courts are increasingly inclined to consider the leaders' collusion as legal. Thus, from a legal point of view, these p-values may not be feasible. In other words, at some level, raising p becomes self-defeating.[19]

In sum, also the effectiveness of the third antitrust policy instrument (prosecution and punishment) is limited. It certainly does not reach the effectiveness of the first instrument (ensuring free market access).

[19] Fig. 4 illustrates the results related to the fixed cost $f = 0.002$. When the fixed cost is considerably higher (e.g., $f = 0.05$), an increase in p can cause a perceptible increase in Q. However, these events have only very limited practical relevance because they occur only in the range of p-values that are probably not feasible from a legal point of view.

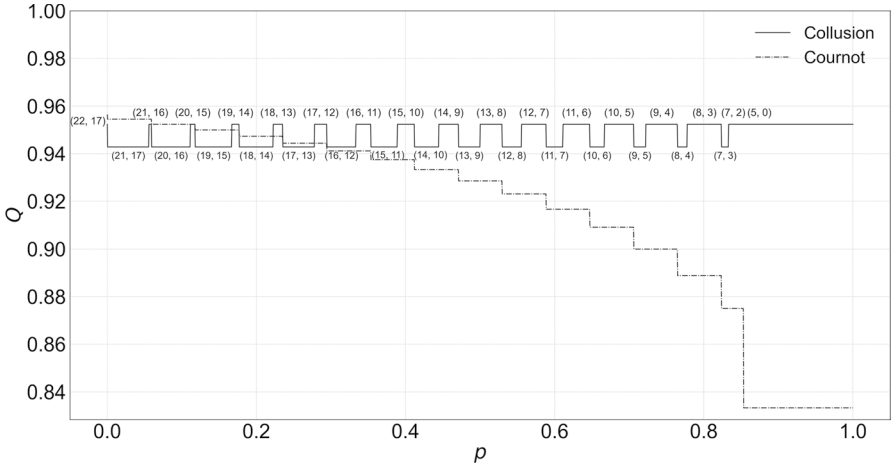

Fig. 4 The Total Output, Q, and the Subgame-Perfect Equilibria, (n^{**}, k^{**}), Corresponding to Different Levels of Rigour, p

5 Concluding remarks

In the present paper, we developed a three-stage game that represents a general oligopoly model of imperfect collusion. All standard oligopoly models are special cases of this general model. We derived the equilibrium number of firms operating on the market, the associated number of colluding firms, and the associated total output. The model addresses the interplay between the four core issues of cartel theory: output sustainability, status stability, market stability, and antitrust policy.

For example, an antitrust policy of obstructing collusion does not lead to the desired goal. Although the deteriorated collusion leads to a higher output of the colluding firms (the leaders), it harms the profit of the other firms (the followers). As a consequence, some of the followers either become leaders or leave the market. The associated reduction in the followers' output overcompensates the increase in the output of the leaders. Thus, a lower level of collusion results in a lower total output and, therefore, welfare.

An increase in the antitrust policy's rigour causes only minor changes in total output. A larger increase can arise only for high levels of rigour. However, these higher levels are unattainable because the associated total output is larger than the Cournot output and the leaders' collusion could not be classified as illegal.

According to our model, the most effective antitrust policy is the removal of entry and exit barriers. This recommendation echoes a key result of the theory of contestable markets. However, that theory has been criticised for its assumption that incumbent firms can only react to market entries with delay. Our own model reveals that this controversial assumption is not necessary to emphasise the relevance of free market access for a successful competition policy.

Appendix A

Proof of Lemma 2: The first- and second-order derivatives of $E[\pi_K(k)]$ with respect to k are

$$\frac{\partial E[\pi_K(k)]}{\partial k} = \frac{\Omega\lambda^2}{(1+\lambda)^2} \frac{2k-(n+1)}{(n-k+1)^2k^2} \tag{21}$$

$$\frac{\partial^2 E[\pi_K(k)]}{\partial k^2} = \frac{\Omega\lambda^2}{(1+\lambda)^2} \frac{2(n^2+2n-3nk+3k^2-3k+1)}{(n-k+1)^3k^3} . \tag{22}$$

The denominators of the second quotients on the right-hand side are positive. The numerator of the second quotient on the right-hand side of (22) can be expressed in the form

$$2\left(n-\frac{3k}{2}\right)^2 + \frac{3}{2}(k-2)^2 + 4(n-1) > 0 .$$

Therefore, $\partial^2 E[\pi_K(k)]/\partial k^2 > 0$. At $k = (n+1)/2$, the numerator of (21) is equal to 0. Therefore, $\partial E[\pi_K(k)]/\partial k > 0$, for all $k > (n+1)/2$. The first- and second-order derivatives of $\pi_F(k)$ with respect to k are

$$\frac{\partial \pi_F(k)}{\partial k} = \left(\frac{\lambda}{1+\lambda}\right)^2 \frac{2}{(n-k+1)^3} > 0$$
$$\frac{\partial^2 \pi_F(k)}{\partial k^2} = \left(\frac{\lambda}{1+\lambda}\right)^2 \frac{6}{(n-k+1)^4} > 0 .$$

□

Appendix B

Proof of Theorem 2: External stability requires that $E[\pi_K(k+1)] \leq \pi_F(k)$. Using the profit functions (12) and (11), this condition yields

$$\frac{(n-k)(k+1)}{(n-k+1)^2} \geq \Omega.$$

This inequality yields the following convex quadratic function:

$$(\Omega+1)k^2 - (\Omega 2(n+1) - 1 + n)k + \Omega(n+1)^2 - n \leq 0.$$

Setting the left-hand side equal to 0, gives the two solutions

$$k_1^{ext} = \frac{n-1+2\Omega(n+1)-z}{2(\Omega+1)} \tag{23}$$

$$k_2^{ext} = \frac{n - 1 + 2\Omega(n+1) + z}{2(\Omega + 1)}, \tag{24}$$

with z being defined by (16):

$$z = \sqrt{(n+1)^2 - 4\Omega(n+2)}.$$

This square root is defined, if and only if

$$n^2 + 2n + 1 \geq 4\Omega n + 8\Omega \ . \tag{25}$$

By definition, $n \geq 4\Omega$. Thus, we have $n^2 > 4\Omega n$ and $2n \geq 8\Omega$. As a consequence, for all $n \geq 4\Omega$, the term represented by z is defined and external stability holds for all $k \in [k_1^{ext}, k_2^{ext}]$. Expressions (23) and (24) imply that $k_2^{ext} = k_1^{ext} + z/(\Omega + 1)$.

For $k_1^{ext} \leq 1$, the smallest integer k^* satisfying $k_1^{ext} \leq k^*$ is 1. However, a single firm cannot proclaim itself as a leader. Thus, by definition, $k_1^{ext} \leq 1$ yields $k^* = 0$.

Next, we show that $k_2^{ext} \leq n$. Using (16) and (24), this inequality can be rearranged to

$$\sqrt{(n+1)^2 - 4\Omega(n+2)} \leq n - 2\Omega + 1 \ . \tag{26}$$

Since $n \geq 4\Omega$, both sides of (26) are defined and positive. Taking squares on both sides of (26) and simplifying yields

$$0 \leq 4\Omega(\Omega + 1)$$

which is always satisfied.

Internal stability requires that $E[\pi_K(k)] > \pi_F(k-1)$. Using the profit functions (12) and (11), this condition yields

$$\frac{(n-k+1)k}{(n-k+2)^2} < \Omega \ . \tag{27}$$

For $p = 1$, we get $\Omega = 0$ and internal stability is always violated. Inequality (27) gives the following convex quadratic function:

$$(\Omega + 1)k^2 - (1 + 4\Omega + n + 2n\Omega)k + \Omega\left(4n + n^2 + 4\right) > 0.$$

Setting the left-hand side equal to 0, gives the two solutions

$$k_1^{int} = \frac{n + 1 + 2\Omega(n+2) - z}{2(\Omega + 1)} \tag{28}$$

$$k_2^{int} = \frac{n + 1 + 2\Omega(n+2) + z}{2(\Omega + 1)}\,, \tag{29}$$

where $k_2^{int} = k_1^{int} + z/(\Omega + 1)$.

Next, we show that $k_2^{int} > n$ which implies that the threshold k_2^{int} is irrelevant for any stability considerations and that internal stability holds for all $k < k_1^{int}$. Rearranging the inequality $k_2^{int} > n$ gives:

$$n - 4\Omega - 1 < \sqrt{(n+1)^2 - 4\Omega(n+2)}\,. \tag{30}$$

For $n \geq 4\Omega$, the right-hand side of (30) is defined and positive. If also $n < 4\Omega + 1$, (30) is satisfied because its left-hand side becomes negative. If $n \geq 4\Omega + 1$, squaring both sides of (30) and simplifying yields $-4(\Omega + 1)(n - 4\Omega) < 0$, which is true.

Finally, we show that $k_1^{int} \leq n$. This inequality can be rearranged to

$$\sqrt{(n+1)^2 - 4\Omega(n+2)} \geq -n + 4\Omega + 1\,. \tag{31}$$

For $n \geq 4\Omega$, the left-hand side of (31) is defined and positive. If also $n > 4\Omega + 1$, the right-hand side of (31) is negative and the inequality is satisfied. If $n \leq 4\Omega + 1$, squaring both sides of (31) and simplifying yields $(n - 4\Omega)(1 + \Omega) \geq 0$, which is true.

In sum, for $n \geq 4\Omega$, status stability arises for the unique integer k^* in the interval $[k_1^{ext}, k_1^{int})$. □

Appendix C

Proof of Corollary 1: Since $\Omega = (1 - p)/\lambda$, we get

$$\frac{\partial \Omega}{\partial \lambda} = -\frac{1-p}{\lambda^2} < 0 \quad \text{and} \quad \frac{\partial \Omega}{\partial p} = \frac{-1}{\lambda} < 0\,.$$

Furthermore, the derivative of k_1^{int} with respect to Ω is

$$\frac{\partial k_1^{int}}{\partial \Omega} = \frac{2n + 3z + 4\Omega + 2n\Omega + nz + z^2 + 4}{z2(\Omega + 1)^2} > 0\,.$$

Thus, the derivatives of k_1^{int} with respect to λ and p are

$$\frac{\partial k_1^{int}}{\partial \lambda} = \frac{\partial k_1^{int}}{\partial \Omega}\frac{\partial \Omega}{\partial \lambda} < 0 \quad \text{and} \quad \frac{\partial k_1^{int}}{\partial p} = \frac{\partial k_1^{int}}{\partial \Omega}\frac{\partial \Omega}{\partial p} < 0$$

The derivative of k_1^{int} with respect to n is

$$\frac{\partial k_1^{int}}{\partial n} = \frac{(2\Omega+1)z - (n+1-2\Omega)}{2(\Omega+1)z} .$$

The denominator is defined and positive. Defining

$$X = (2\Omega+1)z \qquad (>0)$$
$$Y = n+1-2\Omega \qquad (>0) ,$$

the numerator can be expressed as

$$X - Y = \frac{(X-Y)(X+Y)}{X+Y} = \frac{X^2-Y^2}{X+Y}.$$

The denominator of this quotient is positive. The numerator gives

$$\begin{aligned} X^2 - Y^2 &= (4\Omega^2+4\Omega+1)z^2 - \left[(n+1)^2 - 4\Omega(n+2) + 4\Omega + 4\Omega^2\right] \\ &= 4\Omega(\Omega+1)\left(z^2-1\right) \\ &= 4\Omega(\Omega+1)(n+2)(n-4\Omega) . \end{aligned}$$

This expression is positive because $n > 4\Omega$. Thus, $\partial k_1^{int}/\partial n > 0$.

Since $k_1^{ext} = k_1^{int} - 1$, we get

$$\frac{\partial k_1^{ext}}{\partial \lambda} = \frac{\partial k_1^{int}}{\partial \lambda} < 0 , \qquad \frac{\partial k_1^{ext}}{\partial n} = \frac{\partial k_1^{int}}{\partial n} > 0, \qquad \text{and} \qquad \frac{\partial k_1^{ext}}{\partial p} = \frac{\partial k_1^{int}}{\partial p} < 0 .$$

□

Appendix D

Figure 3 shows the typical relationship between λ and Q. There are several interesting aspects that are not directly visible in this graph. In the main text, it was pointed out that a sufficient reduction of λ usually causes a violation of external stability such that a follower joins the group of leaders. In the following, this effect is denoted as case 1. For example, when λ falls to 0.7213, the new subgame-perfect equilibrium is (21, 14) instead of (21, 13). Note that the resulting output is *lower* than with perfect quantity coordination ($\lambda = 1$).

Besides case 1, three other cases exist. When λ reaches the value 0.6735, external stability is again violated and one follower joins the group of leaders. However, the resulting increase in the profits of the remaining followers is so large that a new

firm enters the market and becomes a follower. This is case 2. The new subgame-perfect equilibrium is (22, 15). Since the number of followers is the same as in the subgame-perfect equilibrium (21, 14), output is not affected. Therefore, in Fig. 3 no kink arises at $\lambda = 0.6735$.

The different welfare consequences of cases 1 and 2 reinforce our claim that status stability and market stability should be studied together. Additional support for this claim arises when λ falls to 0.6655. At this λ-level, external stability is not violated, but the leaders' profits become negative and one leader leaves the market. Thus, the profits of the remaining leaders increase, while the profits of the followers remain unchanged. Therefore, external stability must be examined. In the present situation ($\lambda = 0.6655$) it is still satisfied. Therefore, welfare remains unchanged. This is case 3. It reverses the effects of case 2 on n and k.[20]

When $\lambda = 0.1429$, the last follower joins the group of colluding leaders. Since the profits are positive, a new firm may consider to enter the market and to join the group of leaders. However, it will abstain from an entry because it anticipates that this entry would lead to $n \geq 4\Omega$ causing a violation of internal stability; see Theorem 1. One leader would become the first follower and, as a consequence, the leaders' profits would become negative.

Therefore, when $\lambda = 0.1429$, the internal stability condition, $n < 4\Omega$, is the binding restriction. As λ falls further, this restriction is relaxed. New firms enter the market and join the complete group of leaders until it has 39 members (see Fig. 3). Note that the output of the leaders' increases as λ falls and that this output is independent of the number of leaders. Therefore, no kink arises to the left of $\lambda = 0.1429$ and reductions of λ reduce the aggregate profit of the leaders. New leaders aggravate this effect.

When λ reaches the value 0.0961, the binding restriction is no longer internal stability but profitability. Further reductions of λ make the existing leaders unprofitable. Therefore, leaders are forced to leave the market. When λ falls to $1/21 = 0.04762$, the number of leaders falls to 21 and the group's degree of quantity coordination is equivalent to the behaviour of 21 firms that are in Cournot competition with each other.

Appendix E

Proof of Theorem 3: When the level of collusion of a group of n leaders is $\lambda = 1/n$, the n firms are in Cournot competition and no followers exist. Each leader's expected profit is $E[\pi_K(n,n)] = (1-p)/(n+1)^2 - f$; see profit function (11). Let n_C denote the *unique* integer n satisfying the two conditions

[20] If the exit induced increase in the leaders' profits triggered a violation of external stability, the number of leaders would return to its original value, while the number of followers and, therefore, welfare, would fall. This would be case 4.

$$\text{A:}\quad \frac{1-p}{\left(n_C+1\right)^2}-f\geq 0 \qquad \text{and} \qquad \text{B:}\quad \frac{1-p}{\left(n_C+2\right)^2}-f<0.$$

Thus, for leaders that are in Cournot competition, n_C is the maximum integer n with non-negative expected profits.

When $k(\leq n)$ firms form a perfectly colluding group of leaders ($\lambda=1$) and compete against $(n-k)$ followers, the expected profit of each leader is $E[\pi_K(n,k)]=(1-p)/[4(n-k+1)k]-f$; see profit function (11). Let (n^{**},k^{**}) denote the corresponding subgame-perfect equilibrium. If a new firm entered the market and became a leader, output and price would not be affected, but expected profits would become negative (otherwise, the firm would have joined the group of leaders before). Then, n^{**} is the *unique* integer satisfying the following two conditions:

$$\text{C:}\quad \frac{1-p}{4(n^{**}-k^{**}+1)k^{**}}-f\geq 0 \quad \text{and} \quad \text{D:}\quad \frac{1-p}{4(n^{**}-k^{**}+1)(k^{**}+1)}-f<0.$$

We have to prove that, for given p and f, the unique n_C-integer satisfying conditions A and B is always equal to the unique n^{**}-integer satisfying conditions C and D, and *vice versa*.

Because

$$(n+1)^2-4(n-k+1)k=(n+1-2k)^2\geq 0,$$

the denominator in A is at least as large as the denominator in C (when $n_C=n^{**}$). Therefore, the left hand-side of A is smaller than the left-hand side of C. Thus, all n-values that satisfy A also satisfy C (A $\Rightarrow$ C). Similarly,

$$(n+2)^2-4(n-k+1)(k+1)=(n-2k)^2\geq 0.$$

Therefore, the left-hand side of B is always smaller than the left hand-side of D and all n-values that satisfy D also satisfy B (D $\Rightarrow$ B). Since n_C is the unique integer satisfying conditions A and B, and n^{**} is the unique integer satisfying conditions C and D, simultaneous satisfaction of the logical relationships A $\Rightarrow$ C and D $\Rightarrow$ B requires that $n_C=n^{**}$.

The total output corresponding to $(n_C,n_C)=(n^{**},n^{**})$ and $\lambda=1/n^{**}$ is $Q''=n^{**}/(n^{**}+1)$ while the total output corresponding to (n^{**},k^{**}) and $\lambda=1$ is $Q'=(n^{**}-k^{**}+1/2)/(n^{**}-k^{**}+1)$. The ratio of the two output levels is

$$\frac{Q''}{Q'}=1+\frac{2k-n-1}{(n+1)[n-(2k-n-1)]}\,. \tag{32}$$

Lemma 2 implies that $2k-n-1>0$. From Proposition 2 of Shaffer (1995, p. 746) we know that, for $\lambda=1$ and $p=0$, internal stability requires that $2k-n-1<2$. Our Corollary 1 implies that, for given n, the value of k_1^{int} falls as p increases. Therefore, the quotient on the right-hand side of Eq. (32) is always positive, but smaller than $2/[(n+1)(n-2)]$. □

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Auer L von, Pham TA (2021) Optimal destabilization of cartels. J Regul Econ 59(2):175–192
Bos I, Harrington JE Jr (2015) Competition policy and cartel size. Rand J Econ 56(1):133–153
Bos I, Harrington JE Jr (2010) Endogenous cartel formation with heterogeneous firms. Int Econ Rev 41(1):92–117
Bos I (2009) Incomplete cartels and antitrust policy, Doctoral dissertation, PhD thesis, Tinbergen Institute
Bowley AL (1924) Mathematical groundwork of economics. California Press, Oxford
Brito D, Catalão-Lopes M (2011) Small fish become big fish: mergers in Stackelberg markets revisited. BE J Econ Anal Policy 11(1):24
Cyert RM, DeGroot MH (1973) An analysis of cooperation and learning in a duopoly context. Am Econ Rev 63(1):24–37
d'Aspremont C, Jacquemin A, Gabszewicz JJ & Weymark JA (1983) On the stability of collusive price leadership. Can J Econ 16(1):17–25
Donsimoni MP, Economides N, Polenachakis H (1986) Stable cartels. Int Econ Rev 27(2):317–327
Donsimoni MP (1985) Stable heterogeneous cartels. Int J Ind Organ 3(4):451–467
Edgeworth FY (1881) Mathematical psychics. C. Kegan Paul & Co, London
Escrihuela-Villar M (2015) A note on the equivalence of the conjectural variations solution and the coefficient of cooperation. BE J Theor Econ 15(2):473–480
Escrihuela-Villar M (2009) A note on cartel stability and endogenous sequencing with tacit collusion. J Econ 96(2):137–147
Escrihuela-Villar M (2008) On endogenous cartel size under tacit collusion. Investig Econ 32(3):325–338
Friedman JW (1971) A non-cooperative equilibrium for supergames. Rev Econ Stud 38(1):1–12
Harrington JE Jr (2006) How do cartels operate? Found Trends Microecon 2(1):1–105
Huck S, Konrad KA, Müller W, Normann H-T (2007) The merger paradox and why aspiration levels let it fail in the laboratory. Econ J 117:1073–1095
Konishi H, Lin P (1999) Stable cartels with a Cournot fringe in a symmetric oligopoly. KEIO Econ Stud 36(2):1–10
Levenstein MC, Suslow VY (2006) What determines cartel success? J Econ Lit 44(1):43–95
Lofaro A (1999) When imperfect collusion is profitable. J Econ 70(3):235–259
López ÁL, Vives X (2019) Overlapping ownership, R&D spillovers, and antitrust policy. J Polit Econ 127(5):2394–2437
Martin S (2006) Competition policy, collusion, and tacit collusion. Int J Ind Organ 24:1299–1332
Martin S (1993) Advanced industrial economics. Blackwell, Oxford
Martin S (1990) Fringe size and cartel stability. EUI Working Papers in Economics, No. 90/16
Matsumura T, Matsushima N, Cato S (2013) Competitiveness and R&D competition revisited. Econ Model 31:541–547
Matsumura T, Matsushima N (2012) Competitiveness and stability of collusive behaviour. Bull Econ Res 64:22–31
Prokop J (1999) Process of dominant cartel formation. Int J Ind Organ 17(2):241–257

Selten R (1973) A simple model of imperfect competition, where 4 are few and 6 are many. Int J Game Theory 2:141–201
Shaffer S (1995) Stable cartels with a Cournot fringe. South Econ J 61:744–754
Zu L, Zhang J, Wang S (2012) The size of stable cartels: an analytical approach. Int J Ind Organ 30(2):217–222

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.