EconomiX
http://economix.fr

Document de Travail
Working Paper
2011-10

Do followers really matter in Stackelberg competition?

Ludovic A. Julien
Olivier Musy
Aurélien W. Saïdi

UMR 7235

Université de Paris Ouest Nanterre La Défense
(bâtiments T et G)
200, Avenue de la République
92001 NANTERRE CEDEX

Tél et Fax : 33.(0)1.40.97.59.07
Email : nasam.zaroualete@u-paris10.fr

université Paris Ouest Nanterre La Défense

Do followers really matter in Stackelberg competition?☆

Ludovic A. Julien[a,b,c], Olivier Musy[b], Aurélien W. Saïdi[d]

[a]*LEG, Université de Bourgogne*
[b]*EconomiX, Université Paris Ouest – Nanterre - La Défense*
[c]*IRES, Université Catholique de Louvain*
[d]*Department of Information & Operations Management, ESCP Europe*

Abstract

In this note, we consider a generalized $T-$stage Stackelberg oligopoly. We provide a proof and an interpretation that under the two necessary and sufficient conditions of linear aggregate demand and identical constant marginal costs, followers do not matter for leaders. Leaders act as *rational myopic agents*, voluntarily ignoring the number of followers and remaining stages, thereby behaving as Cournotian oligopolists. Strategies of incumbent firms are invariant to entry of new cohorts. Their profits can be studied by the way of two discount factors: the first impacting markup and the second impacting output supply. Some implications in terms of welfare and convergence toward competitive equilibrium are derived.

Keywords: Leader's markup discount factor, linear economy, follower's output discount factor, myopic behavior.
JEL classification: L13, L20

1. Introduction

The Stackelberg (1934) sequential oligopoly model is often put forward in the literature as a useful alternative to the static Cournot (1838) model. [1] Firms sequentially choose the quantities to produce and take into account the impact of their own decisions on the decisions of firms playing later. The basic model has been extended in order to integrate a larger number of stages and/or players than in the original model (Boyer and Moreaux (1986), Sherali (1984) or Watt (2002)). An interest of such a structure, which we call a *hierarchy*,[2] is to introduce heterogeneity between firms according to their place in the decision process. Notable implications concern welfare (Watt (2002)), merging (Daughety (1990) and Heywood and McGinty (2007, 2008)) and profits (Etro (2008)).

Under the two standard assumptions of linear market demand and identical constant marginal costs, this sequential process exhibits an interesting property: the $T-$stage Stackelberg equilibrium strategies coincide with the equilibrium strategies

☆We are grateful to C. Bidard for his comments and remarks.

Email addresses: ludovic.julien@u-paris10.fr (Ludovic A. Julien), omusy@u-paris10.fr (Olivier Musy), asaidi@escpeurope.eu (Aurélien W. Saïdi)

[1]We do not consider oligopoly competition on prices.

[2]This denomination comes from Boyer and Moreaux (1986).

obtained in a succession of T monopolists exploiting the residual demand. This feature of the linear economy has been noticed by Boyer and Moreaux (1986), Anderson and Engers (1992) and Pal and Sarkar (2001). Watt (2002) has extended the framework to integrate multiplayer cohorts and has shown that firms behave as oligopolists in a succession of Cournot games.[3]

The purpose of this note is twofold. First, we generalize the existing results and provide a formal proof for the following property: the two assumptions specified above are not only sufficient but also necessary for the sequential structure of the Stackelberg model to simply reduce to a succession of Cournot games. By contrast with the relevant literature, we also provide an explanation for such a result by showing that the rational behavior of leaders consists in voluntarily ignoring any information about their followers when solving their optimization program. A change in the number of remaining stages or followers results in a scale factor affecting negatively the objective profit function of a leader, but does not distort the relative profits associated with two different strategies, and thereby, the optimal strategy. The behavior of firms can be qualified as "rationally myopic": it is optimal for them to only focus on the decisions of the previous and current players and ignore the available information about their followers.[4] This property holds in and outside the equilibrium.

Second, our approach enables us to explore some properties concerning strategic interactions in this Stackelberg structure. The model exhibits two discount factors for any cohort. We call the first the *leader's markup discount factor*. It represents the reduction of a leader's markup associated with the existence of its followers in the hierarchy. Its value differs from one cohort to another, depending negatively on the number of remaining cohorts and corresponding players. This factor also brings into light the consequences of free entry for incumbents. We call the second factor the *follower's output discount factor*. It represents the decrease in optimal quantities for a follower resulting from a contraction of the residual demand when playing latter in the hierarchy. It is a share of the first cohort production, whether optimal or not. For any follower, this share decreases when going further in the sequence. Its value depends negatively on the number of leading cohorts and on the number of corresponding players (because residual demand decreases with this parameter). Both factors measure the profit reduction of a firm within the hierarchy.

The paper is presented as follows. In section 2, we present the model. Section 3 analyzes the behavior of the firms under the Stackelberg structure, and introduces the two kinds of discount factors. Section 4 derives welfare implications and free entry analysis. In section 5, we conclude.

2. The Model

Consider one homogeneous good produced by n firms which oligopolistically compete in a hierarchical framework. There are T stages of decisions indexed by t,

[3] Julien and Musy (2011) also show that considering various degrees of competition between the followers does not affect Watt (2002)'s results.

[4] It does not imply that a firm has no advantage in being a leader since its position in the sequential process does affect its profits.

$t \in [\![1, T]\!]$. Each stage embodies one cohort and is associated with a level of decision. The whole set of cohorts represents a hierarchy. Cohort t is populated by n_t firms, with $\sum_t n_t = n$. The distribution of the firms within each cohort is assumed to be observable and exogenous.[5] This latter assumption notably implies that position of firms and timing of moves are given. are given.[6]

A firm i which belongs to cohort t has to decide strategically (simultaneously with firms of the same cohort, and sequentially among the hierarchy) its level of output denoted by x_t^i. The aggregate output of cohort t is denoted $X_t = \sum_{i=1}^{n_t} x_t^i$, where x_t^i stands for firm i's output within cohort t. In addition, $X_t^{-i} = \sum_{-i} x_t^{-i}$ will denote the production of all firms belonging to cohort t but i.

The inverse market demand function specifies the market price p as a function of aggregate output X, with $X = \sum_{t=1}^{T} X_t$, and is denoted by $p(X)$. We assume that $p(.)$ is continuous and twice differentiable, with $\frac{dp(X)}{dX} < 0$. The cost function of any firm i which belongs to cohort t, is denoted by $\phi_t^i(.)$. It is a continuous and twice differentiable function with $\frac{d\phi_t^i(x_t^i)}{dx_t^i} > 0$ and $\frac{d^2\phi_t^i(x_t^i)}{d\left(x_t^i\right)^2} \geq 0$.

The n_t firms which belong to cohort t, behave as followers with respect to all firms of cohort τ, $\tau \in [\![1, t-1]\!]$, whose strategies are taken as given. However, they behave as Stackelberg leaders toward all firms of cohort τ, $\tau \in [\![t+1, T]\!]$. They consider the best-response functions of all firms belonging to these cohorts as functions of their strategies. Therefore the profit of firm i which belongs to cohort t may be written:

$$\pi_t^i(x_t^i) = p\left(\sum_{\tau=1}^{t-1} X_\tau + \sum_{\tau=t}^{T} X_\tau\right) x_t^i - \phi_t^i(x_t^i), \qquad i = 1, ..., n_t. \tag{1}$$

3. Rational myopic behavior and Stackelberg competition

3.1. Assumptions and graphical interpretation

In this section, we assume a linear market demand function

$$p(X) = a - bX, \qquad a, b > 0 \tag{H1}$$

and constant and identical marginal costs[7]

$$\phi_t^i(x_t^i) = cx_t^i, \qquad i = 1, ..., n_t \ \text{ and } \ t = 1, ..., T. \tag{H2}$$

These two assumptions are standard in the literature on oligopoly analysis (see Daughety (1990), Carlton and Perloff (1994), Vives (1999), among others).

Consider two successive stages, say $t-1$ and t. Let $p\left(\sum_{\tau=1}^{t} X_\tau\right)$ be the market price when cohort t enters the market while each cohort τ ($\tau < t$) produces a quantity

[5] The standard Stackelberg duopoly prevails when $T = 2$ and $n_1 = n_2 = 1$.

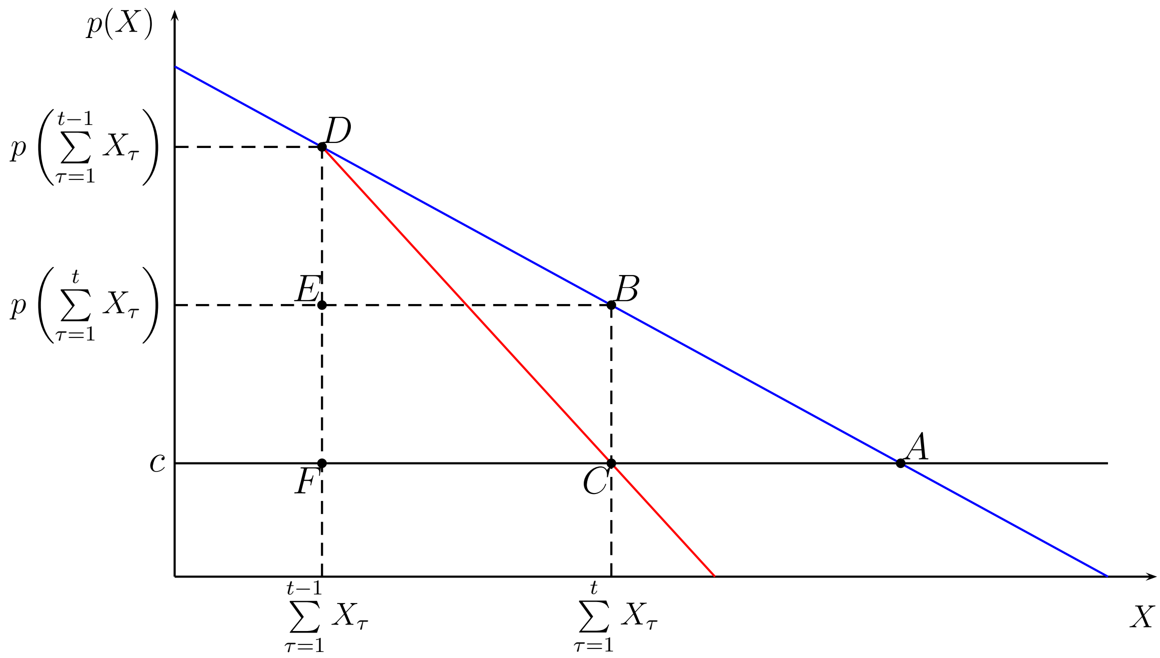
[6] We therefore do not question the way a specific firm could or should become a leader (see Anderson and Engers (1992), Amir and Grilo (1999), Matsumura (1999)).

[7] We therefore do not envisage the case where one firm is more efficient than another within the economy (see Pal and Sarkar (2001)).

of output X_τ. No particular value is given to X_τ, except that it must generate a non-negative profit. However, we assume that any leading cohort $\tau < t$ expects firms of cohort t (or more) to act rationally and symmetrically. As in the standard literature, they maximize their profits for any quantity $\sum_{\tau=1}^{t-1} X_\tau$ produced by their predecessors.

The behavior of firms can be qualified as "rationally myopic" when it is optimal for them to focus on the decision of the previous players, and thereby to ignore the information about their followers. In this case, the rational choice of firms is depicted in figure 1.

Figure 1: Rational myopia



In this Figure, we illustrate the behavior of cohort t when acting as a myopic cohort. Firms of this cohort behave as oligopolists on the residual demand left by firms of cohort $\tau < t$, and do not take into consideration firms playing after.

The purpose of the current section consists in proving that the following property holds if and only if (H1) and (H2) are satisfied.

Property 1. *While rational myopia is an optimal behavior for the last cohort, it is also optimal for any cohort in the economy, whatever its rank within the hierarchy.*

3.2. Sufficiency

Demonstrating property 1 requires to exhibit the link between leaders and followers' profits.

Lemma 1. *Let* $\gamma_t \equiv \prod_{\tau=t+1}^{T} \frac{1}{1+n_\tau}$ *be the leader's markup discount factor. Under (H1) and (H2), the markup earned by a cohort t firm, $t < T$, in a $T-$cohort economy is a constant share $\gamma_t < 1$ of the markup it earns in a $t-$cohort economy for any given vector of outputs $(X_1, ..., X_{t-1})$ produced by the previous cohorts:*

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \gamma_t \left[p\left(\sum_{\tau=1}^{t} X_\tau\right) - c\right] \qquad \textit{for } t < T. \tag{2}$$

Proof. See Appendix A. ■

Notice that under condition H2, the markup is always equal across cohorts. The discount factor γ_t measures the dependence of market power on the number of followers. It represents the reduction of markup of any leader due to the presence of the additional cohorts $t+1$ to T. It affects less intensively the market power of the last cohorts in the sequence since they face a reduced number of followers. Market power shrinks as t tends to infinity. This case will be discussed in Section 4.

The existence of cohort τ equally impacts by a coefficient $1/(1+n_\tau)$ the markup expected by a leader t $(t < \tau)$ in a $t-$stage economy, whatever the quantities produced by the first t cohorts (this results directly from H1 and H2).

Corollary 1. *For any strategy x_t^i, the profit obtained by a cohort$-t$ firm in the sequential $T-$stage structure is a constant share of the myopic profit (of the $t-$stage economy):*

$$\pi_t^i(x_t^i) = \left[p\left(\sum_{\tau=1}^{T} X_\tau\right) - c\right] x_t^i = \gamma_t \left[p\left(\sum_{\tau=1}^{t} X_\tau\right) - c\right] x_t^i \tag{3}$$

Proof. This corollary directly results from Lemma 1. ■

In other words, each cohort can behave *as if* there were no following cohorts behind it since it earns a constant share of the profit realized in an oligopoly structure market where it represents the last cohort, whatever the aggregate output $\sum_{\tau=1}^{t-1} X_\tau$ produced by the leaders. Provided that cohort$-t$ firms maximize their profit for any vector of strategies $(X_1, ..., X_{t-1})$, cohort$-\tau$ leaders $(\tau < t)$ act as oligopolists, ignoring the following cohorts, that only discount the value of their profits without changing the nature of the maximization program.

Lemma 2. *Let* $\eta_{t-h,t} \equiv \prod_{\tau=t-h+1}^{t} \frac{1}{1+n_\tau}$ *be the follower's output discount factor. Under (H1) and (H2), the output of a firm i in cohort $t \le T$ can be expressed as a share of the output produced by a firm playing previously and belonging to cohort $t-h$ for $h \in [\![1, t-1]\!]$, that is:*

$$x_t = \eta_{t-h,t} x_{t-h}.$$

Proof. See Appendix B. ■

The follower's output discount factor represents the contraction of output resulting from playing later in the hierarchy. It indicates the share of cohort $t-h$'s output which is optimal for cohort t to produce.

From the previous lemmas, the following proposition can be stated:

Proposition 1. *If conditions (H1) and (H2) hold, firms in any cohort behave as rational myopic agents and the T-stage Stackelberg linear economy reduces to a succession of staggered static problems in which firms compete oligopolistically on residual demands.*

Proof. The proposition directly ensues from Lemmas 1 and 2. ■

Maximizing the right-hand side of Equation (3) (sequential structure program) is tantamount to maximize the left-hand side of Equation (3) (myopic program) since γ_t is a constant term. In the linear economy, strategies of firms do not depend on the number of firms playing after, which equally impact the profit associated to each strategy. As a consequence the optimal strategies and the equilibrium strategies remain unchanged whatever the number of stages and the number of followers in the sequential structure.

The literature only covers the similarity of the equilibrium strategies in both the $T-$stage Stackelberg linear model and the succession of staggered static problems but does not provide any explanation for this coincidence (see Boyer and Moreaux (1986), Anderson and Engers (1992) and Watt (2002)).

Corollary 2. *The equilibrium strategy of cohort 1-firms may thus be obtained from the myopic profit maximization:*

$$x_1 = \frac{1}{1+n_1}\frac{a-c}{b} \equiv \eta_1 X^*,$$

where $X^* = (a-c)/b$ is equal to the perfect competition aggregate output. We then deduce the equilibrium strategy of any firm i in cohort t, $t \in [\![1, T]\!]$:

$$x_t = \frac{\eta_{1,t}}{1+n_1}\frac{a-c}{b} \equiv \eta_t X^* \qquad \textit{with} \qquad \eta_t \equiv \prod_{\tau=1}^{t} \frac{1}{1+n_\tau}$$

Notice that in the equilibrium, each firm of cohort t produces a share η_t of the perfect competition equilibrium output.

Corollary 3. *The equilibrium price and equilibrium profits are given by:*

$$p = c + (a-c)\prod_{\tau=1}^{T}\frac{1}{1+n_\tau},$$

$$\pi_t^i = \frac{(a-c)^2}{b}\prod_{\tau=1}^{t}\frac{1}{(1+n_\tau)^2}\prod_{\tau=t+1}^{T}\frac{1}{1+n_\tau} \qquad t = 1, ..., T.$$

De Quinto and Watt (2003) use a similar term to η_t to analyze welfare through market power and mergers. In our approach, we investigate the issue of welfare through a comparison with perfect competition representing the maximizing global surplus benchmark case.

Remark 1. *When leaders of any cohort t, $t \in [\![1, T-1]\!]$, ignore the true number of followers in cohort $\tau > t$ or form incorrect estimates about it, the effective equilibrium strategies and payoffs coincide with those obtained under perfect information regarding the number of followers.*

Remark 2. *The leader's markup discount factor is equivalent to a taxation of profits. The tax rate τ_t differs from one firm to another:*

$$\tau_t = 1 - \gamma_t.$$

3.3. Necessity

Lemma 3. *Under condition (H1), Property 1 occurs only if condition (H2) is satisfied. Under condition (H2) and $c > 0$, Property 1 occurs only if condition (H1) is satisfied.*

Proof. See Appendix C and Appendix D. ■

It is then possible to state the following proposition.

Proposition 2. *In the T-stage economy, firms behave as Cournotian oligopolists only if (H1) and (H2) are satisfied.*

Proof. The proposition directly ensues from Lemmas 1 to 3. ■

In the literature, neither the necessity of these conditions nor the fact that the leaders effectively behave as Cournotian oligopolists have ever been demonstrated. The only results in the literature feature that the equilibrium strategies in a $T-$stage model coincides with the equilibrium strategies that would be obtained with staggered static problems where firms either compete monopolistically (as in Boyer and Moreaux (1986)) or oligopolistically (as in Watt (2002)) on residual demand. This paper provides a formal and general proof.

4. Implications for convergence and welfare

Social welfare is maximized under perfect competition, that is when aggregate output is equal to X^*. Let ω be the index of social welfare. This index, included between 0 and 1 (maximum welfare), is measured by the sum of the shares $n_\tau \eta_\tau$:

$$\omega = \sum_{\tau=1}^{T} n_\tau \eta_\tau = 1 - \frac{1}{(1+n_1)(1+n_2)...(1+n_T)} = 1 - \eta_{1,T}.$$

It can be asserted from Corollary 2 that the aggregate equilibrium output in the model is given by ωX^*.

Lemma 4. *When the number of firms becomes arbitrarily large, either vertically (when T tends to infinity) or horizontally (when n_τ tends to infinity), the oligopoly equilibrium output converges toward the competitive equilibrium output.*

Proof. Immediate from $\lim_{T\to\infty} \sum_{\tau=1}^{T} n_\tau \eta_\tau = 1$ and $\lim_{n_\tau\to\infty} \sum_{\tau=1}^{T} n_\tau \eta_\tau = 1$. ∎

Convergence toward perfect competition is then achieved through an increase in the number of cohorts and/or in the number of firms in any cohort. A specific case of vertical convergence can be found in Boyer and Moreaux (1986) for $n_\tau = 1$, $t \in [\![1, T]\!]$.

From the previous lemma we know that welfare can be improved by increasing the number of firms. When the number of firms is fixed, welfare can be modified modified when firms are displaced in the decision sequence, either by enlarging the hierarchy or changing the size of existing cohorts.

Lemma 5. *For any given number of firms, a displacement of any firm results in a higher welfare gain when enlarging the hierarchy rather than modifying the size of an existing cohort.*

Proof. Assume a move of a cohort-t firm within the hierarchy. Let ω_1 be the social welfare index when this moves enlarges the hierarchy (adding a cohort $T+1$) and ω_2 be the same index when it modifies the size of an existing cohort (say t').

$$\begin{aligned} \omega_1 &= 1 - \frac{1+n_t}{(1+n_t-1)(1+n_{T+1})} \prod_{\tau=1}^{T} \frac{1}{1+n_t} \quad \text{with} \quad n_{T+1} = 1 \\ \omega_2 &= 1 - \frac{(1+n_t)(1+n_{t'})}{(1+n_t-1)(1+n_{t'}+1)} \prod_{\tau=1}^{T} \frac{1}{1+n_t}. \end{aligned}$$

Since $\frac{1}{2+n_{t'}} < \frac{1}{2}$ for any $n_t > 0$ then $\omega_1 > \omega_2$. ∎

For a constant number of firms, adding new cohorts is always welfare improving. Said differently, introducing position-based asymmetries is welfare enhancing. It echoes and generalizes the result of Daughety (1990), which is restricted to $T = 2$.

When both the number of stages and the number of firms are fixed, the following lemma shows how to improve welfare.

Lemma 6. *For a fixed number of firms and cohorts, welfare improves as long as firms are relocated between cohorts until the difference of sizes between any two cohorts is at most equal to 1. For each relocation, welfare enhancement is greater when the firm is moved from the largest to the smallest cohort.*

Proof. See Appendix E. ∎

It can now be stressed the assumptions upon which positions of firms do not matter for social welfare, i.e. are invariant to specific modifications in the decision process. This property is called hierarchy neutrality.

Lemma 7. *The linear economy, defined by assumptions (H1) and (H2), is hierarchy neutral when relocation of firms consists of switching the whole cohorts within the hierarchy: this relocation does not affect social welfare.*

Proof. Immediate: switching n_t and $n_{t'}$ backward or forward in $\eta_{1,T}$ does not change the value of ω. ■

From the preceding lemmas, one can state the following proposition relative to the link between welfare and the structure of the economy.

Proposition 3. *In this linear economy, maximizing social welfare can be achieved through two ways:*
(i) As a priority, by enlarging the hierarchy.
(ii) Then, by successively relocating firms from the most to the less populated cohort until equalizing the size of all cohorts.

Proof. Proposition 3 ensues from Lemmas 4 to 7. ■

This proposition could also be used to analyze how entry affects welfare. If new firms enter the economy, the increase in welfare is greater if new cohorts are created rather than if those firms integrate existing cohorts.

5. Conclusion

The paper investigates a hierarchic T-stage oligopoly model. It states that followers do not matter in the linear case, i.e. under constant identical marginal costs and linear demand. This means that at any stage each firm behaves as a Cournotian oligopolist on residual demand. In addition, the two discount factors presented in this paper enable us to characterize to fully characterize the market outcome of the linear economy, especially in terms of strategies and welfare.

The rational myopic behavior of firms could be interpreted as a situation in which firms compete under imperfect information on the number of stages and players. Since strategies are invariant to the number of followers, this framework is also suited to analyze free entry. The follower's output discount factor determines the quantities produced by new firms while the leader's markup discount factor captures the impact of additional competitors on incumbents' profits.

6. References

Amir R. and I. Grilo (1999), "Stackelberg versus Cournot equilibrium", *Games and Economic Behavior*, 26, pp. 1-21.

Anderson S.P. and M. Engers (1992), "Stackelberg versus Cournot equilibrium", *International Journal of Industrial Organization*, 10, pp. 127-135.

Boyer M. and M. Moreaux (1986), "Perfect competition as the limit of a hierarchical market game", *Economics Letters*, 22, pp. 115-118.

Carlton D. and J.M. Perloff (1994), *Modern industrial organization*, New-York: HarperCollins.

Daughety A. (1990), "Beneficial concentration", *American Economic Review*, 80, pp. 1231-1237.

De Quinto J. and R. Watt (2003), "Some simple graphical interpretations of the Herfindahl-Hirshman index", mimeo.

Etro F. (2008), "Stackelberg competition with endogenous entry", *Economic Journal*, 118, pp. 1670-1697.

Heywood J. S. and M. McGinty (2007), "Mergers among leaders and mergers among followers", *Economics Bulletin*, 12, pp. 1-7.

Heywood J. S. and M. McGinty (2008), "Leading and merging: convex costs, Stackelberg, and the merger paradox", *Southern Economic Journal*, 74, pp. 879-893.

Julien L.A. and O. Musy (2011), "A generalized oligopoly model with conjectural variations", forthcoming *Metroeconomica.*

Matsumura T. (1999), "Quantity-setting oligopoly with endogenous sequencing", *International Journal of Industrial Organization*, 17, pp. 289-296.

Pal D. and Sarkar (2001), "A Stackelberg oligopoly with nonidentical firms", *Bulletin of Economic Research*, 53(2), pp. 127-135.

Sherali H.D. (1984), "A multiple leader Stackelberg model and analysis", *Operation Research*, 32, pp. 390-404.

Vives X. (1999), *Oligopoly pricing. Old ideas and new tools*, MIT Press, Cambridge, Massachussets.

Watt R. (2002), "A generalized oligopoly model", *Metroeconomica*, 53, pp. 46-55.

Appendix A. Proof of Lemma 1

The proof is by backward induction and structured in three steps.

Step 1: property (2) is true for cohort $t = T-1$ (with $T > 1$).
The inverse demand function faced by firms (blue line) is defined by:

$$p(X) = a - bX \quad \text{with} \quad X = \sum_{\tau=1}^{T} X_\tau, \tag{H1}$$

where $X_\tau = \sum_{i=1}^{n_\tau} x_\tau^i \geq 0$ is the aggregate production of cohort τ. For any quantity of output X_{T-1} produced by cohort $T-1$, the resulting residual demand faced by followers of cohort T is:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) = \hat{a}_{T-1} - bX_T \quad \text{with} \quad \hat{a}_t \equiv a - b\sum_{\tau=1}^{t} X_\tau,$$

where $\hat{a}_{T-1}$ is considered as given by followers. Geometrically, followers must select a couple (X, p) on the segment $[D, A]$.

When acting symmetrically, the associated marginal revenue of cohort$-T$ firms (red line) is defined by:[8]

$$R_m(X_T) = \hat{a}_{T-1} - b\frac{1+n_T}{n_T}X_T.$$

Considering the following derivatives:

$$\begin{aligned} \frac{\partial R_m}{\partial X_T}(X_T) &= \frac{DF}{CF} = -b\frac{1+n_T}{n_T} \\ \frac{\partial p}{\partial X_T}(X_T) &= \frac{DF}{AF} = -b, \end{aligned}$$

it comes that:

$$CF = \frac{n_T}{1+n_T}AF, \quad \text{or equivalently} \quad AC = \frac{1}{1+n_T}AF.$$

Finally, applying Thales' theorem to triangles ABC and ADF leads to:

$$BC = \frac{1}{1+n_T}DF = EF. \tag{A.1}$$

[8]This function is derived from the total revenue of a follower i :
$RT(x_T^i) = \left[\hat{a}_{T-1} - b\sum_{k=1}^{n_T} x_T^k\right] x_T^i$. The symmetric behavior assumed for followers yields: $x_T^i = x_T$ for all $i \in [1, n_T]$ and $X_T = n_T x_T$.

Actually, EF is the markup of a leader after the entrance of the last cohort, while DF is the markup of a leader before the entrance of cohort T. Equation (A.1) can be rewritten as:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \frac{1}{1+n_T}\left[p\left(\sum_{\tau=1}^{T-1} X_\tau\right) - c\right].$$

Step 2: assume property (2) is true for any cohort $t = T-h$ $(1 \leq h \leq T-2)$ then it is true for cohort $T-h-1$.

If property (2) holds for cohort $T-h$ then:

$$\left[p\left(\sum_{\tau=1}^{T} X_\tau\right) - c\right] x^i_{T-h} = \gamma_{T-h}\left[p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c\right] x^i_{T-h} \qquad \text{with } \gamma_{T-h} \equiv \prod_{\tau=T-h+1}^{T} \frac{1}{1+n_\tau}$$

Thus, maximizing firm i's profit is tantamount to maximize the myopic profit defined as follows:

$$\max_{x^i_{T-h}} \left[p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c\right] x^i_{T-h}.$$

When firms of cohort$-(T-h)$ act symmetrically, the myopic marginal revenue (red line) is defined by:[9]

$$\tilde{R}_m(X_{T-h}) = \hat{a}_{T-h-1} - b\frac{1+n_{T-h}}{n_{T-h}} X_{T-h}.$$

In the same way as in step 1, it can be shown that:

$$p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c = \frac{1}{1+n_{T-h}}\left[p\left(\sum_{\tau=1}^{T-h-1} X_\tau\right) - c\right].$$

By assumption, the following property is satisfied:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c = \gamma_{T-h}\left[p\left(\sum_{\tau=1}^{T-h} X_\tau\right) - c\right].$$

We deduce from the two previous equations that:

$$\begin{aligned} p\left(\sum_{\tau=1}^{T} X_\tau\right) - c &= \frac{\gamma_{T-h}}{1+n_{T-h}}\left[p\left(\sum_{\tau=1}^{T-h-1} X_\tau\right) - c\right] \\ &= \gamma_{T-h-1}\left[p\left(\sum_{\tau=1}^{T-h-1} X_\tau\right) - c\right] \end{aligned}$$

Step 3: from steps 1 and 2 we conclude by backward induction that property (2) is true for any cohort t (with $1 \leq t \leq T-1$).

[9]The associated marginal revenue is: $R_m(X_{T-h}) = \tilde{R}_m(X_{T-h}) + (1-\gamma_{T-h})c$.

Appendix B. Proof of Lemma 2

Applying Thales' theorem to triangles ABC and ADF leads to:

$$CF = \frac{n_t}{1+n_t} AF. \tag{B.1}$$

Actually, CF is the optimal output produced by cohort t, that is X_t, while AF is the maximal quantities cohort t can produce to generate non-negative profit (equal to the difference between the perfect competition equilibrium supply and the output already produced by the previous cohorts). The property above can be rewritten as:

$$X_t = \frac{n_t}{1+n_t}(X_t + AC), \quad \text{or equivalently} \quad AC = \frac{X_t}{n_t} = x_t.$$

Notice that AC is also the maximal quantities cohort $t+1$ can produce to generate non-negative profit. Then, property (B.1) applied to cohorts t and $t+1$ becomes:

$$X_{t+1} = \frac{n_{t+1}}{1+n_{t+1}} AC, \quad \text{leading to} \quad \frac{X_{t+1}}{n_{t+1}} = x_{t+1} = \frac{1}{1+n_{t+1}} x_t.$$

By backward induction, it turns out that:

$$x_t = \eta_{1,t} x_1, \quad \text{where} \quad \eta_{1,t} \equiv \prod_{\tau=2}^{t} \frac{1}{1+n_\tau}.$$

Appendix C. Proof of Lemma 3

We intend to prove that condition (H2)

$$\phi_t(x_t) = c x_t \quad \text{for any } t \in [\![1, T]\!] \tag{H2}$$

is a necessary condition for properties (2) and (3) to hold.

Assume that all firms act as Cournotian oligopolists on the residual demand. The firm j of cohort t then determines x_t^j such that it maximizes its myopic profit function $\tilde{\pi}_t^j$:

$$\tilde{\pi}_t^j(x_t^j) = \left[\hat{a}_{t-1} - b \sum_{i=1}^{n_t} x_t^i \right] x_t^j - \phi_t(x_t^j),$$

where ϕ_t is a non-decreasing function.

The reaction function of cohort-t firms, when behaving symmetrically, is implicitly defined by:

$$\hat{a}_{t-1} - b(1+n_t)x_t - \phi_t'(x_t) = 0 \quad \text{for any } t > 1.$$

When maximizing its profit, a cohort-τ firm ($\tau < t$) substitutes to X_t the aggregate reaction function of cohort t. Strategy of a cohort-τ firm is then independent of n_t provided X_t does not contain any n_t term. If so x_t is a function of n_t, i.e. $x_t = X_t/(1+n_t)$.

The aggregate reaction function of cohort t can be rewritten as:

$$X_t = \frac{\hat{a}_{t-1} - \phi_t'(x_t)}{b} \quad \text{for any } t > 1.$$

It does not depend on n_t provided $\phi_t'(x_t)$ does not contain any n_t term. Since x_t is a function of n_t, the marginal cost is independent of x_t, that is constant:

$$\phi_t(x_t) = c_t x_t \quad \text{for any } \; t \in [\![1, T]\!].$$

The reaction function of a cohort-T firm can now be explicitly defined:

$$x_T = \frac{a - c_T - b \sum_{\tau=1}^{T-1} X_\tau}{b(1 + n_T)},$$

and the profit function of cohort$-(T-1)$ firms is:

$$\pi_{T-1}(x_{T-1}) = \frac{1}{1 + n_T} \left[\hat{a}_{T-1} - bX_{T-1} - c_{T-1} - n_T(c_{T-1} - c_T)\right] x_{T-1}$$

For cohort$-(T-1)$ firms to behave as Cournotian oligopolists, the profit function must be such that:

$$\pi_{T-1}(x_{T-1}) = \frac{1}{1 + n_T} \tilde{\pi}_{T-1}(x_{T-1}).$$

This is the case if and only if:

$$c_{T-1} = c_T.$$

Using backward induction, it can be proved that for cohort$-t$ firms, $t < T-1$, to behave as Cournotian oligopolists, the following properties must hold:

$$c_1 = c_2 = ... = c_{T-1} = c_T.$$

Thus, when the demand function is linear and cost functions are non-decreasing, a necessary condition for firms of any cohort to behave as Cournotian competitors on residual demand requires condition (H2).

Appendix D. Proof of Lemma 4

Under condition (H2), firms of cohort $T-1$ act as Cournotian competitors on residual demand if and only if:

$$\left[p\left(\sum_{\tau=1}^{T-1} X_\tau + X_T\right) - c\right] x_t = \gamma_t \left[p\left(\sum_{\tau=1}^{T-1} X_\tau\right) - c\right] x_t$$

for a coefficient $\gamma_t > 0$.

It requires X_T (the aggregate reaction function of cohort T) to be linear in $\sum_{\tau=1}^{T-1} X_\tau$ and to be a function of c, independent of $\sum_{\tau=1}^{T-1} X_\tau$:

$$X_T = \alpha_1 + \alpha_2 \sum_{\tau=1}^{T-1} X_\tau + g(c),$$

where $\alpha_1 > 0$ and $-1 < \alpha_2 < 0$ (strategies are substitutes) are constant terms, and $g(0) = 0$ and $g'(c) \leq 0$ on the ground of economic plausibility.

Adding $\sum_{\tau=1}^{T-1} X_\tau$ to each side of the equation above and rearranging yields:

$$\sum_{\tau=1}^{T-1} X_\tau = -\frac{\alpha_1}{1+\alpha_2} + \frac{1}{1+\alpha_2} \sum_{\tau=1}^{T} X_\tau - \frac{g(c)}{1+\alpha_2}.$$

The first-order condition for profit maximization of cohort-T firms is:

$$p\left(\sum_{\tau=1}^{T} X_\tau\right) - c + p'\left(\sum_{\tau=1}^{T} X_\tau\right) x_T = 0$$

that is:

$$-\frac{p\left(\sum\limits_{\tau=1}^{T} X_\tau\right)}{p'\left(\sum\limits_{\tau=1}^{T} X_\tau\right)} + \frac{c}{p'\left(\sum\limits_{\tau=1}^{T} X_\tau\right)} = \frac{\alpha_1}{n_T} + \frac{\alpha_2}{n_T} \sum_{\tau=1}^{T-1} X_\tau + \frac{g(c)}{n_T}.$$

or:

$$-\frac{p\left(\sum\limits_{\tau=1}^{T} X_\tau\right)}{p'\left(\sum\limits_{\tau=1}^{T} X_\tau\right)} + \frac{c}{p'\left(\sum\limits_{\tau=1}^{T} X_\tau\right)} = \mu_1 + \mu_2 \sum_{\tau=1}^{T} X_\tau + \mu_3 g(c), \tag{D.1}$$

with:

$$\begin{aligned}
\mu_1 &= \frac{\alpha_1}{n_T(1+\alpha_2)} > 0 \\
\mu_2 &= \frac{\alpha_2}{n_T(1+\alpha_2)} < 0 \\
\mu_3 &= \frac{1}{n_T(1+\alpha_2)} > 0
\end{aligned}$$

By identification, the previous equality is satisfied provided p' is independent of $\sum\limits_{\tau=1}^{T} X_\tau$ or $c = 0$.

Case 1

When $c = 0$, equation (D.1) becomes:

$$-\frac{p'\left(\sum_{\tau=1}^{T} X_\tau\right)}{p\left(\sum_{\tau=1}^{T} X_\tau\right)} = \frac{1}{\mu_1 + \mu_2 \sum_{\tau=1}^{T} X_\tau}.$$

By integrating both sides of the equation:

$$-\ln p\left(\sum_{\tau=1}^{T} X_\tau\right) = \frac{1}{\mu_2}\ln\left[\mu_1 + \mu_2\sum_{\tau=1}^{T} X_\tau\right] - \ln\mu_4 \quad \text{with } \mu_4 > 0,$$

that is:

$$p(X) = \mu_4\left[\mu_1 + \mu_2\sum_{\tau=1}^{T} X_\tau\right]^{-1/\mu_2} \equiv \left[a - b\sum_{\tau=1}^{T} X_\tau\right]^{\mu} \quad \text{with } a, b, \mu > 0.$$

A specific case is the linear demand function with $\mu = 1$. It is worth noting however that another class of non-linear functions is compatible with the property that firms behave as Cournotian oligopolist on residual demand when $c = 0$.

Case 2

When $c \neq 0$, identifying equation (D.1) implies:

$$\frac{c}{p'\left(\sum_{\tau=1}^{T} X_\tau\right)} = \mu_3 g(c),$$

where μ_3 is a constant and g is independent of $\sum_{\tau=1}^{T} X_\tau$. The derivative of p must then be a constant term. Thus:

$$p(X) = a - b\sum_{\tau=1}^{T} X_\tau \quad \text{with } a, b > 0.$$

The linear demand function is the only class of functions consistent with the property that firms act as Cournotian oligopolist on residual demand when $c \neq 0$.

Appendix E. Proof of Lemma 6

Maximizing the welfare index ω is tantamount to maximize:

$$\begin{cases} \max\limits_{n_\tau} & \prod\limits_{\tau=1}^{T} 1 + n_\tau \\ s.t. & \sum\limits_{\tau=1}^{T} n_\tau = n \end{cases}$$

Substituting n_T by $n - \sum_{\tau=1}^{T-1} n_\tau$ into the objective function, deriving with respect to n_τ, $\tau \in [\![1, T-1]\!]$, and assuming the $n_\tau' s$ are infinitely divisible yields the following first-order conditions:

$$n - \sum_{\tau=1}^{T-1} n_\tau = n_t \quad t \in [\![1, n_{T-1}]\!]$$

or equivalently (in addition with the definition of n_T):

$$n_t = n/T \equiv \bar{n} \quad t \in [\![1, n_{T-1}]\!].$$

Notice that for this value of n_τ the omitted constraint $0 \leq n_\tau \leq n$ is satisfied for any $\tau \in [\![1, T]\!]$.

At $(\bar{n}, ..., \bar{n})$, the $T \times T$ Hessian matrix M is

$$M = -(1+\bar{n})^{T-2} \begin{bmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{bmatrix} - (1+\bar{n})^{T-2} I,$$

where I is the identity matrix. The eigenvalues of M are $-(T+1)(1+\bar{n})^{T-2}$ and $-(1+\bar{n})^{T-2}$ (the associated eigenspaces have dimension 1 and $T-1$ respectively).

Matrix M is then negative definite:

- The unique solution to the first-order conditions is a global maximum when n/T is an integer.
- There are multiple optima when n/T is not an integer. Due to the strict concavity of the objective function, these optima must be as close as possible to the hypothetical solution above. In other words they must minimize the distances $\left|n_\tau - \frac{n}{T}\right|$, for $\tau = 1, ..., T$, such that $\sum_{\tau=1}^{T} n_\tau = n$. The minimum value of these distances is 1 and can be obtained as follows.

  Let $m < T$ be an integer such that $(n-m)/T = \lfloor n/T \rfloor$. An optimum is such that there are $T-m$ cohorts populated by $\lfloor n/T \rfloor$ firms and the other m cohorts by $\lfloor n/T \rfloor + 1$ firms. The number of combinations of m cohorts out of T defines the number of optima.

  Notice that the most populated cohorts embody one more firm than the less populated cohorts.

When n/T is not an integer and for given values of the n_τ's, the more efficient way to get closer to the hypothetical optimal as one firm is relocated consists in reducing the largest distance, e.g. $\left|n_t - \frac{n}{T}\right|$. Without loss of generality, assume that the difference $\left(n_t - \frac{n}{T}\right)$ is positive. Then, it is not efficient for social welfare to relocate the firm in a cohort t' with $n_{t'} > \bar{n}$ since this move decreases $\left|n_t - \frac{n}{T}\right|$ but increases $\left|n_{t'} - \frac{n}{T}\right|$. The firm must be relocated in a cohort τ such that $n_\tau < \bar{n}$. Within this set of cohorts, the cohort with the largest $\left|n_\tau - \frac{n}{T}\right|$ will be selected since the reduction of the distance is the highest.

The argument is similar when the difference $\left(n_t - \frac{n}{T}\right)$ is negative. As a conclusion, social welfare is more efficiently improved when relocating a firm from the most to the less populated cohort.