Rationalizability and Efficiency in an Asymmetric Cournot Oligopoly

Gabriel Desgranges, Stéphane Gauthier

Université Paris 1 – Panthéon - Sorbonne

Documents de Travail du Centre d'Economie de la Sorbonne

CES Working Papers

Rationalizability and Efficiency in an Asymmetric Cournot Oligopoly

Gabriel DESGRANGES, Stéphane GAUTHIER

2014.28

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/
ISSN : 1955-611X

Rationalizability and Efficiency in an Asymmetric Cournot Oligopoly*

Gabriel Desgranges† and Stéphane Gauthier‡

March 26, 2014

Abstract

We study rationalizable solutions in a linear asymmetric Cournot oligopoly. We show that symmetry across firms favors multiplicity of rationalizable solutions: A merger (implying a greater asymmetry across firms) makes out-of-equilibrium behavior less likely and should dampen 'coordination' volatility. The market structure maximizing consumers' surplus at a rationalizable solution is not always the competitive one: This may be a symmetric oligopoly with few firms. An empirical illustration to the airlines industry shows that a reallocation of 1% of market share from a small carrier to a larger one yields a 1.3% decrease in volatility, measured by the within carrier standard error of the number of passengers.

JEL codes: D43, D84, L40.

Keywords: competition policy, Cournot oligopoly, dominance solvability, efficiency, rationalizability, stability, airline industry.

*We thank Jess Benhabib, Gaetano Gaballo, Roger Guesnerie, Laurent Linnemer, Régis Renault, Jean-Philippe Tropéano and Thibault Vergé for useful discussions. Alexandra Belova, Philippe Gagnepain, Nicolas Jacquemet, and Antoine Terracol provided us generous guidance in the empirical illustration to the airline industry. The usual disclaimers apply.

†University of Cergy-Pontoise, France. gabriel.desgranges@u-cergy.fr

‡University of Paris 1 and Paris School of Economics, France. Stephane.gauthier@univ-paris1.fr

1 Introduction

The issue of rationalizability in a Cournot oligopoly with identical firms has received much attention in the literature (see, e.g., Bernheim, 1984; Basu, 1992; or Börgers and Janssen, 1995). Our paper considers the rationalizable outcomes in a linear asymmetric Cournot setup. We interpret multiplicity of rationalizable outcomes as favoring coordination failures. The main results of our paper show that identical firms are detrimental to the uniqueness of the rationalizable outcome (dominance solvability of the equilibrium): Firm's homogeneity actually favors efficiency in equilibrium, but it makes multiplicity of rationalizable outcomes (and thus coordination failures) more likely.

We consider the rationalizable outcomes in a linear Cournot oligopoly where the overall quantity of productive assets can be controlled (Perry and Porter, 1985). This allows us to separate the own effect of a change in the market size (measured by the total number of available productive assets) from the effect of a greater asset inequality given the overall quantity of assets. This setup naturally arises in competition policy when a competition regulator has to choose how production facilities should be allocated across competitors. Examples include airlines routes and airports (Borenstein, 1990), nuclear reactors in the power industry (Davis and Wolfram, 2011), hospitals in the health insurance market (Town et al., 2006) or water sources (Compte et al., 2002). In these examples, firms possibly differ according to the number of productive assets under their control, but the overall quantity of assets is given. A monopoly-like situation, where one firm holds most of the assets, is detrimental to efficiency in the equilibrium: The aggregate production is lower than in the equilibrium where facilities are equally shared across several competitors.

A unique rationalizable outcome is obtained when the spectral radius of the best-response map mapping a vector of individual productions to the vector of best-responses is less than one (Bernheim, 1984; Moulin, 1984). We show that this spectral radius increases when the total number of available productive assets increases: A large market size favors multiplicity of rationalizable outcomes.

Our main results concern reallocations of productive assets across firms. A reallocation of a given number of assets from a large firm (a firm which owns a high number of assets) to a smaller one also yields a higher spectral radius: Asymmetry across firms also favor multiplicity of rationalizable outcomes. In addition, in the case where there are multiple rationalizable outcomes, this same reallocation enlarges the set of rationalizable aggregate productions.

The intuition for these results proceeds as follows. The spectral radius of the best-response map is increasing in the slopes of the firms' reaction functions. These slopes are increasing and concave in the firms' number of assets. Concavity implies that, when productive assets are reallocated from a large firm to a smaller one, the increase in the slope of the reaction function of the smaller firm dominates the decrease in the slope of the larger one. The overall effect then favors multiplicity. Intuitively, this reallocation relaxes the capacity constraint of the smaller firm. The behavior of the smaller firm becomes less predictable. The behavior of the larger firm, reacting to its expectation about the production of the smaller firm, in turn becomes less predictable.

Since the equilibrium aggregate production decreases when the market gets closer to the monopoly situation (where one firm holds most of the available productive assets), our results show that a reallocation of assets whose goal is to reach more efficiency in the Cournot equilibrium may in fact result in a multiplicity of rationalizable outcomes.

These first results allow us to study the optimal allocation of assets among firms considering the rationalizable outcomes, rather than the equilibrium outcome only. We first solve for the distribution of assets maximizing the aggregate equilibrium production subject to the constraint that this equilibrium is the only rationalizable outcome. Of course, the equilibrium production reaches its maximum in the competitive case (corresponding to an infinite number of small identical firms). Thus, when the competitive equilibrium is the only rationalizable outcome, it is the solution we look for. Otherwise, we find that the solution is an oligopoly with few identical firms: the optimal allocation of assets now yields a non competitive market. Finally, we show that this same distribution of assets also maximizes the lowest aggregate production in a rationalizable outcome. In this sense, there is no room for an asset distribution which would give rise to multiple rationalizable outcomes.

Bernheim (1984), Basu (1992) and Börgers and Janssen (1995) study rationalizable outcomes in symmetric Cournot games. Guesnerie (1992) studies eductive stability (that coincides with uniqueness of rationalizable outcomes) in the competitive case, and Gaballo (2013) considers eductive stability in linear symmetric Cournot games. The closest paper to ours is Moulin (1984). Moulin (1984) provides a condition for local Cournot stability and shows that this same condition also governs elimination of non-best responses. Our paper considers global stability of the equilibrium (that amounts to take account of the entry decision of the firms). It provides a global characterization of the set of rationalizable outcomes in the presence of asymmetric firms. It also studies how this set relates to the asset distribution.

There are close links between the outcomes surviving an iterated elimitation process (dominance solvability, iterated weak dominance, etc.) and the outcomes stable under adaptive learning. See, among others, Milgrom and Roberts (1990), Guesnerie (1993), Marx (1999), Hommes and Wagener (2010) or Durieu, Solal and Tercieux (2011). Our paper does not adress this issue. However this literature suggests that our results may have close counterparts in adaptive learning.

In practice it is well known that there is no clear evidence that concentration of production facilities is associated with higher prices (Gugler et al., 2003). This lack of conclusive evidence is usually viewed as reflecting a trade-off between economies of scale and the ability of larger firms to exercise market power (Williamson, 1968; Perry and Porter, 1985; Farrell and Shapiro, 1990). Our theoretical analysis suggests that this lack of evidence may reflect different out-of-equilibrium rationalizable behaviors: A reallocation of production facilities changes the set of rationalizable outcomes. A merger (reducing possible out-of-equilibrium rationalizable behavior) should dampen market volatility by making easier for each firm to form accurate predictions about others' behavior. We illustrate this prediction by studying the impact of the merger between Delta Airlines and Northwest in 2008. We use the within carriers standard error of the number of passengers as a measure of 'coordination' volatility. We provide evidence that the change in market power implied by this merger has reduced 'coordination' volatility: A 1 percent transfer of market share from a small firm to a larger one would decrease volatility by 1.3 percent.

The paper is organized as follows. In Section 2, we briefly present the setup, and we show that equilibrium production increases following a reallocation of assets from a large firm to a smaller one. In Section 3, we give a necessary and sufficient condition for the equilibrium to be the only rationalizable outcome. In Section 4, we establish the trade-off between efficiency in equilibrium and uniqueness of the rationalizable outcome. In Section 5, we characterize the optimal distribution of assets. Finally the illustration to the airline industry is given in Section 6.

2 Setup

Following Perry and Porter (1985), we consider a single product model of Cournot competition with M firms and N units of a productive asset. Firm ℓ owns N_ℓ units of the asset, with N_ℓ decreasing in ℓ ($N_\ell \in \mathbb{R}_+$). The only source of heterogeneity across firms is the (exogenous) distribution of assets. Producing q_ℓ costs $C\left(q_\ell, N_\ell\right) = q_\ell^2/2\sigma N_\ell$ to firm ℓ ($\sigma > 0$). A possible interpretation of this cost function is to think of a unit of the asset as a plant, and to assume that producing q units of the good in one plant costs $q^2/2\sigma$. By convexity, a firm minimizes its overall cost by producing the same quantity in each plant.

Firm ℓ produces q_ℓ which maximizes $p(q_\ell + Q_{-\ell})q_\ell - C\left(q_\ell, N_\ell\right)$, where $Q_{-\ell}$ is the aggregate production of firms other than ℓ and $p\left(\cdot\right) = \delta_0 - \delta Q$ is the inverse demand function (where Q is the aggregate production, and $\delta, \delta_0 > 0$). The best response of firm ℓ is

$$R_\ell(Q_{-\ell}) = \begin{cases} q_\ell^m - b_\ell Q_{-\ell} & \text{if } Q_{-\ell} \leq \delta_0/\delta, \\ 0 & \text{if } Q_{-\ell} \geq \delta_0/\delta, \end{cases} \tag{1}$$

where $q_\ell^m = b_\ell \delta_0/\delta$ is the monopoly production of firm ℓ, and

$$b_\ell = \frac{\sigma\delta N_\ell}{2\delta\sigma N_\ell + 1}. \tag{2}$$

The linear/quadratic specification implies that the sensitivity parameter b_ℓ (the slope of the reaction function) is increasing and concave in N_ℓ.

A Cournot equilibrium is a vector (q_ℓ^*) such that $q_\ell^* = R_\ell(Q_{-\ell}^*)$ for every ℓ. Straightforwardly, there is a unique equilibrium. Let Q^* be the aggregate production in equilibrium. Since the equilibrium price $p(Q^*)$ is positive (otherwise no firm would be active in equilibrium) and the marginal cost tends to zero when the production tends to zero, it is always profitable for a firm to enter the market. Hence all the firms are active in equilibrium ($q_\ell^* > 0$ for every ℓ).

Our first result confirms that an equal distribution of assets across the firms yields the highest aggregate production in equilibrium, and thus the highest consumers' surplus.

Proposition 1. *A transfer of assets from firm h to firm s increases the aggregate output Q^* in the Cournot equilibrium if and only if $N_h > N_s$ (firm h is larger than firm s).*

Proof. The equilibrium aggregate production is

$$Q^* = \frac{S}{1+S}\frac{\delta_0}{\delta}, \quad \text{with } S = \sum_{\ell=1}^{M} \frac{b_\ell}{1-b_\ell}. \tag{3}$$

The production Q^* increases in S, and the ratio $b_\ell/(1-b_\ell)$ is increasing and concave in N_ℓ. A transfer of assets from a large to a small firm implies a lower $b_h/(1-b_h)$ and a higher $b_s/(1-b_s)$. By concavity, S increases, and so Q^* increases. □

This result is a particular case of Perry and Porter (1985) or Farrell and Shapiro (1990). It includes the case of a merger (a merger between firms s and h amounts to transfer all the assets of s to h). A corollary of Proposition 1 is that, when all the firms have the same number of assets, the aggregate output Q^* increases in the number M of firms (at a symmetric oligopoly with M firms, the transfer of all the assets of one firm to the others results into a symmetric oligopoly with $M-1$ firms, and Q^* decreases). Hence, Q^* and the consumers' surplus are maximized in a competitive equilibrium (an equilibrium with an infinite number of identical firms).

3 Dominance solvability of the equilibrium

An equilibrium is dominant solvable when it is the unique rationalizable outcome of the Cournot game. To study dominance solvability, we first define rationalizable strategies by the following elimination process.

Suppose first that the strategy set of every firm ℓ is $\left[q_\ell^{\inf}(0), q_\ell^{\sup}(0)\right) = [0, +\infty)$. Then, define iteratively (for all $t \geq 1$) the sequences $\left[q_\ell^{\inf}(t), q_\ell^{\sup}(t)\right]$ of sets of best responses of firm ℓ to the belief that the aggregate production of others is in $\left[Q_{-\ell}^{\inf}(t-1), Q_{-\ell}^{\sup}(t-1)\right]$, with $Q_{-\ell}^{\inf}(t-1) = \sum_{k\neq\ell} q_k^{\inf}(t-1)$ and $Q_{-\ell}^{\sup}(t-1) = \sum_{k\neq\ell} q_k^{\sup}(t-1)$. Strategic substitutabilities imply that

$$q_\ell^{\inf}(t) = R_\ell\left(Q_{-\ell}^{\sup}(t-1)\right), \quad \text{and } q_\ell^{\sup}(t) = R_\ell(Q_{-\ell}^{\inf}(t-1)). \tag{4}$$

These sequences are converging since $(q_\ell^{\inf}(t))$ increases in t, $(q_\ell^{\sup}(t))$ decreases in t, and they are bounded ($0 \leq q_\ell^{\inf}(t) \leq q_\ell^* \leq q_\ell^{\sup}(t) \leq q_\ell^m$ for all $t \geq 1$). Their limits, denoted $q_\ell^{\inf}$and $q_\ell^{\sup}$, are fixed points of the recursive system (4). The limit set $[q_\ell^{\inf}, q_\ell^{\sup}]$ is the set of rationalizable productions of firm ℓ.

Let us first adopt a local viewpoint. Local dominance solvability is defined as the uniqueness of the rationalizable outcome in a game where the strategy sets are restricted to a neighborhood of the equilibrium ($q_\ell^{\inf}(0)$ and $q_\ell^{\sup}(0)$ are close to the equilibrium q_ℓ^* for every firm). Local dominance solvability obtains whenever the recursive system (4) is locally contracting at the equilibrium. Since the system (4) is linear, this is equivalent to the spectral radius of the matrix

$$B = \begin{pmatrix} 0 & b_1 & \cdots & b_1 \\ b_2 & \ddots & & \vdots \\ \vdots & & \ddots & b_{M-1} \\ b_M & \cdots & b_M & 0 \end{pmatrix}$$

being less than 1.

Lemma 1. *The spectral radius of B is the unique positive root ρ of*

$$F(\rho) \equiv \sum_{\ell=1}^{M} \frac{b_\ell}{\rho + b_\ell} = 1. \tag{5}$$

We have $\rho < 1 \Leftrightarrow F(1) < 1$.

Proof. Let e be an eigenvalue of B, and v an associated eigenvector. Then, $ev = Bv$ yields

$$ev_\ell + b_\ell v_\ell = b_\ell \sum_{k=1}^{M} v_k \Leftrightarrow v_\ell = \frac{b_\ell}{e + b_\ell} \sum_{k=1}^{M} v_k \quad \text{for all } \ell.$$

Summing over ℓ implies that every eigenvalue e of B is such that

$$F(e) \equiv \sum_{\ell=1}^{M} \frac{b_\ell}{e + b_\ell} = 1.$$

For $e \geq 0$, the function F is continuous and decreasing. Moreover, $F(0) = n > 1 > 0 = F(+\infty)$. Hence, B admits a unique positive real eigenvalue. Since B is a positive matrix, it follows from Perron-Frobenius theorem that this positive real eigenvalue is the spectral radius ρ of B. That is, $F(\rho) = 1$ for $\rho > 0$. Finally, since F is decreasing, we have: $\rho < 1$ if and only if $F(1) < 1$. □

The inequality $F(1) < 1$ is the local condition found by Moulin (1984). We show below that it is also the condition for global dominance solvability of the equilibrium. This is done by investigating the set of rationalizable outcomes. Characterizations of this set have been obtained by Bernheim (1984), Basu (1991), Börgers and Janssen (1995) and Gaballo (2013) in the context of Cournot competition with identical firms (all the firms own the same number of assets). With identical firms, either the equilibrium is dominant solvable ($q_\ell^{\inf} = q_\ell^{\sup} = q_\ell^* \equiv q^*$ for all ℓ), or $[q_\ell^{\inf}, q_\ell^{\sup}] = [0, q^m]$ with $q^m = q_\ell^m$ for all ℓ. Indeed, by symmetry, either $q_\ell^{\inf} > 0$ for all ℓ or $q_\ell^{\inf} = 0$ for all ℓ. In the first case, the linear system (4) admits a unique fixed point which coincides with the equilibrium. In the latter case ($\forall \ell, q_\ell^{\inf} = 0$), $q_\ell^{\sup}$ is the best response of firm ℓ to others producing 0: This is the monopoly production q^m.

In our setup firms are heterogenous and it is no longer true that $q_\ell^{\inf} = 0$ for every ℓ when the equilibrium is not dominant solvable. The next result shows that the values of $q_\ell^{\inf}$ are ranked according to ℓ.

Lemma 2. *The bounds $q_\ell^{\inf}$ and $q_\ell^{\sup}$ are nonincreasing in ℓ. Furthermore, the lowest rationalizable production $q_\ell^{\inf}$ is 0 if and only if $\ell > \bar{\ell}$, where $\bar{\ell} \geq 0$ is the largest ℓ such that*

$$\sum_{k \leq \ell} \frac{b_k}{1+b_k} + \sum_{k > \ell} \frac{b_k}{1+b_\ell} < 1. \tag{6}$$

Proof. See in appendix. □

To get an intuition about the existence of the threshold $\bar{\ell}$, let us consider the first two steps of the iterative process of elimination of non best responses. In the first step, $q_\ell^{\sup}(1)$ is the monopoly production q_ℓ^m which is decreasing in ℓ (it is increasing in the number of assets). In the second step, $q_\ell^{\inf}(2)$ is the best response to $Q_{-\ell}^{\sup}(1)$ which is increasing in ℓ (small firms face a higher aggregate production of others than large firms). It follows that $q_\ell^{\inf}(2)$ is decreasing in ℓ and $q_\ell^{\inf}(2)$ is possibly 0 for ℓ large enough. The argument extends to every further step of the elimination process.

Given the threshold $\bar{\ell}$, we can characterize the rationalizable outcomes of a linear Cournot game.

Lemma 3. *The set of rationalizable aggregate productions is the interval $\left[Q^{\inf}, Q^{\sup}\right]$, where*

$$Q^{\inf} = \left(1 + \frac{c-a}{a^2 - c\,(c+e)}\right)\frac{\delta_0}{\delta}, \quad Q^{\sup} = Q^{\inf} + \frac{e}{a^2 - c\,(c+e)}\frac{\delta_0}{\delta}, \tag{7}$$

with

$$a = 1 + \sum_{\ell \leq \bar{\ell}} \frac{b_\ell^2}{1-b_\ell^2},\; c = \sum_{\ell \leq \bar{\ell}} \frac{b_\ell}{1-b_\ell^2} \text{ and } e = \sum_{\ell > \bar{\ell}} b_\ell.$$

Proof. See in appendix. The appendix also characterizes the set $\left[q_\ell^{\inf}, q_\ell^{\sup}\right]$ of rationalizable individual productions. □

Lemma 2 directly yields a necessary and sufficient condition for dominance solvability of the Cournot game. On the one hand, when all the firms are active ($\bar{\ell} = M$), $q_\ell^{\inf} = q_\ell^{\sup} = q_\ell^*$ for all ℓ since the equilibrium is the unique fixed point of the linear system (4). On the other hand, the Cournot equilibrium is not the only rationalizable outcome when some firms remain inactive ($\bar{\ell} < M$).

Proposition 2. *The Cournot equilibrium is globally dominant solvable (the unique rationalizable outcome) if and only if $\bar{\ell} = M$, or equivalently*

$$\Gamma \equiv \sum_{\ell} \frac{b_\ell}{1+b_\ell} < 1. \tag{8}$$

Dominance solvability is obtained whenever the total sensitivity of firms is low enough, an intuition similar to the one found in the competitive case (Guesnerie, 1992). Proposition 2 generalizes the local analysis by taking into account the decision of entry analyzed in Lemma 2. It shows that condition (8) also governs global dominance solvability, and thus it governs rationalizability of entry.

4 Rationalizability and asset distribution

We now relate the asset distribution to the set of rationalizable outcomes. A first approach consists in studying the variations of the spectral radius ρ w.r.t. the distribution of assets. The underlying idea is that a lower spectral radius favors dominance solvability. The spectral radius indeed reflects the speed of convergence of the sequences $q_\ell^{\inf}(t)$ and $q_\ell^{\sup}(t)$ toward their limits $q_\ell^{\inf}$and $q_\ell^{\sup}$. Desgranges and Ghosal (2010) provides a formal justification of the interpretation of ρ as a plausibility index for dominance solvability.

Proposition 3. *A transfer of assets from firm h to firm s increases the spectral radius ρ of B if and only if firm h is larger than firm s ($N_h > N_s$).*

Proof. Consider a transfer of $dN > 0$ assets from firm h to firm s, i.e., N_s increases by dN and N_h decreases by dN ($N_s < N_h$). The resulting change $d\rho$ in the spectral radius is obtained by differentiating (5):

$$F'(\rho)d\rho + \left[\frac{\partial}{\partial N_s}\left(\frac{b_s}{1+b_s}\right) - \frac{\partial}{\partial N_h}\left(\frac{b_h}{1+b_h}\right)\right] dN = 0.$$

Since the ratio $b_\ell/(1+b_\ell)$ is increasing and concave in N_ℓ, the term into brackets is positive. Since $F'(\rho) < 0$ for $\rho > 0$, we have $d\rho > 0$. □

Proposition 3 shows that introducing asymmetries across firms favors dominance solvability: For a given number of firms, the equilibrium is more likely to be dominant solvable when there are both large and small firms, rather than identical firms.

The effect of a transfer of assets from h to s is *a priori* ambiguous. Dominance solvability is favored by firms' inertia to changes in the (expected) production of others, i.e., a slope of the reaction functions close to 0. The (absolute value of the) slope is always increasing in the number of assets. Therefore the transfer of assets considered in Proposition 3 implies that the slope b_s of the reaction function of the smaller firm increases (which is detrimental to dominance solvability) while the slope b_h for the large firm decreases (which favors dominance solvability). The overall effect is made unambiguous by appealing to concavity of the ratio $b_\ell/\left(1+b_\ell\right)$ in the number of assets (so that the increase in b_s has a greater effect than the decrease in b_h). When every slope b_ℓ is increasing in the number of assets N_ℓ, this ratio is concave when b_ℓ is concave in the number of assets N_ℓ.[1]

[1] In a general nonlinear setup, Proposition 3 holds in the neighborhood of the equilibrium when these monotonicity and concavity properties are satisfied locally. Such properties rely on assumptions on 4th derivatives of cost and demand functions (the slope b_ℓ of the reaction function being characterized by second derivatives of these functions).

Propositions 1 and 3 highlight that any reallocation of assets which improves consumers' surplus in equilibrium makes less likely that this equilibrium is the unique rationalizable outcome.[2]

A second approach to the impact of assets reallocations bears on the set of rationalizable aggregate productions $\left[Q^{\text{inf}}, Q^{\text{sup}}\right]$. The difference $Q^{\text{sup}} - Q^{\text{inf}}$ can be viewed as a measure of the strategic uncertainty in the market. When the equilibrium is not dominant-solvable, multiplicity of rationalizable outcomes corresponds to the possible occurrence of fluctuations due to out-of-equilibrium beliefs: some 'coordination' volatility occurs following the process of elimination of non-best response strategies. The magnitude of this volatility can be measured by the size of the interval $\left[Q^{\text{inf}}, Q^{\text{sup}}\right]$ of rationalizable aggregate productions. Volatility is dampened when $\left[Q^{\text{inf}}, Q^{\text{sup}}\right]$ is a narrow interval around the Cournot equilibrium Q^*.

The next result is another version of the trade-off between efficiency and dominance solvability: A reallocation of assets which yields a higher aggregate production in equilibrium also enlarges the set of rationalizable outcomes.

Proposition 4. *Assume that the equilibrium is not dominant solvable ($Q^{\text{inf}} < Q^{\text{sup}}$). Consider an infinitesimal reallocation of assets from a large firm h to a smaller one s, $dN_h = -dN_s < 0$ ($N_s < N_h$). We have:*

$$d\left(Q^{\text{sup}} - Q^{\text{inf}}\right) > 0.$$

Proof. See in appendix. □

In order to grasp some intuition, consider again the iterative process (4). There, the production $q_\ell^{\text{sup}}(1)$ is the monopoly production q_ℓ^m which is increasing and concave in N_ℓ. The reallocation of assets from firm h to firm s implies that q_h^m decreases while q_s^m increases. The total effect is not ambiguous because of concavity: $q_s^m + q_h^m$ increases. Hence, for each firm $\ell \neq s$, the reallocation of assets implies an increase in $Q_{-\ell}^{\text{sup}}(1)$ so that $q_\ell^{\text{inf}}(2)$ decreases (firm s faces a smaller production $Q_{-s}^{\text{sup}}(1)$ but we show that this effect is dominated by the aggregate effect on all the other firms). The argument then extends to every further step of the iterative process.

5 Optimal asset distribution

This section characterizes the distribution of assets which maximizes consumers' surplus. When one considers the issue of multiplicity of rationalizable solutions, there are two ways of assessing consumers' surplus. One way is to maximize consumers' surplus in equilibrium under the additional constraint that the equilibrium is the only rationalizable solution. The other way is to maximize consumers' surplus at some (non-equilibrium) rationalizable

[2] An alternative assessment could refer to Γ characterized in Proposition 2. Proposition 3 actually holds true when the spectral radius ρ is replaced by Γ .

solution. Assuming risk aversion of the competition regulator whose goal is to maximize consumers' surplus leads to consider rationalizable solutions involving low production levels. We consider the polar case of the worst scenario where the lowest rationalizable aggregate production $Q^{\inf}$ occurs (this corresponds to an arbitrarily large risk aversion towards strategic uncertainty).

First, we characterize the distribution of assets which maximizes the aggregate production in a Cournot equilibrium subject to the constraint that this equilibrium is the unique rationalizable outcome. Second, we show that this same distribution also maximizes the lowest rationalizable aggregate production $Q^{\inf}$.

Consider first the distribution of assets (N_ℓ) and the number M of firms which maximize the equilibrium aggregate production Q^* defined in (3) subject to two constraints: the constraint of dominant solvability ($F(1) < 1$ in (8)), and the feasibility constraint

$$\sum_{\ell=1}^{M} N_\ell \leq N. \tag{9}$$

Proposition 5. *Any $((N_\ell), M)$ maximizing the aggregate equilibrium production Q^* given by (3) subject to the constraints (8) and (9) involves an equal sharing of productive assets: $N_\ell = N/M$ for all ℓ. Furthermore,*

- *if $\sigma\delta N < 1$, then the solution involves an infinite number of firms (competitive market), and (9) is the only binding constraint;*
- *if $\sigma\delta N \geq 1$, then the solution is a symmetric oligopoly with*

$$M^{**} = \frac{3\sigma\delta N}{\sigma\delta N - 1}$$

firms.[3] The aggregate production is

$$Q^{**} = \frac{1}{2}\frac{3\sigma\delta N}{1 + 2\sigma\delta N}\frac{\delta_0}{\delta} \tag{10}$$

Both constraints (8) and (9) are binding at the optimum.

Proof. See in appendix. □

By Proposition 5 the competitive equilibrium is the only rationalizable outcome when $\sigma\delta N < 1$. This is the condition found by Guesnerie (1992) in a setup where there is a continuum of size $N = 1$ of competitive firms. The competitive case with a unique rationalizable outcome can be viewed as a situation where a large number of identical firms share a small capacity of production.

[3] When M^{**} is not an integer, the solution is the largest integer below M^{**}.

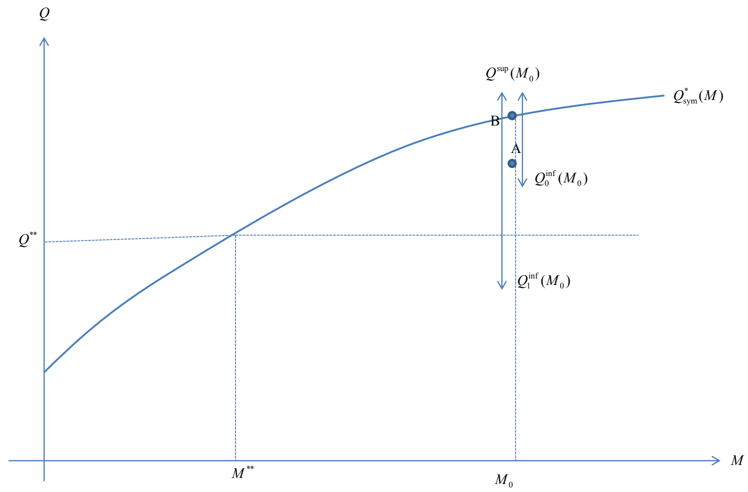

Figure 1: Optimal distribution of assets

When $\sigma\delta N \geq 1$, the competitive equilibrium is not dominant solvable. Uniqueness could be restored using two different kinds of policies. The first kind consists in allocating only a part of the existing production assets: the Cournot equilibrium becomes dominant solvable when only $1/\sigma\delta$ assets $(1/\sigma\delta < N)$ are allocated to a large number of firms, so that one goes back to the previous competitive situation. High competition then goes with production inefficiency, since some productive assets are not used. The second policy is to allocate the whole stock of assets to few firms, so that the market gets closer to a monopoly-like situation. Proposition 5 shows that this last policy is better for consumers' surplus.

At this stage, however, Proposition 5 only gives partial insights into the optimal distribution of assets. Indeed, when $\sigma\delta N \geq 1$, there might exist some distributions of assets such that the associated lowest rationalizable production Q^{inf} is greater than Q^{**} exhibited in Proposition 5. With such distributions, the consumers' surplus at a rationalizable outcome is necessarily greater than the highest surplus achievable at a dominant solvable equilibrium.

An illustration is given in Figure 1. The solid curve represents the equilibrium aggregate quantity when the productive assets are equally shared across firms. This quantity increases in the number M of firms and eventually coincides with the competitive equilibrium.

In Figure 1 the competitive equilibrium is not dominant solvable: By Proposition 5 there exists a finite threshold M^{**} such that the equilibrium is dominant solvable if and only if the number M of firms satisfies $M \leq M^{**}$. For all M, the highest equilibrium production obtains in the case of identical firms. Suppose that there are M_0 firms in the market, with $M_0 \geq M^{**}$. The equilibrium production in the case of identical firms is at point B. When the market structure is close to the monopoly situation, the equilibrium is the only rationalizable outcome. By Proposition 5 the aggregate production then stands below Q^{**}. Figure 1 depicts the case where the market structure is close the the symmetric case, with a set of rationalizable outcome being not reduced to the Cournot equilibrium. Point A corresponds to an equilibrium aggregate production where firms differ according to the number of productive assets they hold.

There are two possible cases for the set of rationalizable aggregate productions. In the first case rationalizable production is always higher than $Q_0^{\inf}(M_0)$ satisfying $Q_0^{\inf}(M_0) > Q^{**}$. In this case any rationalizable outcome in a market with M_0 firms dominates (in terms of consumers' surplus) the dominant solvable symmetric equilibrium with M^{**}. The optimal distribution of assets then involves a multiplicity of rationalizable outcomes. The other possibility where the lowest rationalizable aggregate production is $Q_1^{\inf}(M_0)$ is however theoretically possible too. Then, $Q^{**} > Q_1^{\inf}(M_0)$, so that a multiplicity of rationalizable outcomes can be detrimental to consumers' surplus.

The next result shows that the last configuration always prevails: for all $M > M^{**}$, the lowest rationalizable aggregate production is always below Q^{**}, as is $Q_1^{\inf}(M_0)$ in Figure 1.

Proposition 6. *There is no distribution of assets such that the lowest aggregate production $Q^{\inf}$ is greater than Q^{**}.*

Proof. See in appendix. □

This result implies that there is no asset distribution such that the competition regulator can hold for sure that one can achieve a surplus higher than in the Cournot-Nash equilibrium where the aggregate production is Q^{**} defined in Proposition 6. In this sense, one can not justify a policy which would give rise to multiple rationalizable outcomes.

In summary, this section raises the question of the choice by a competition regulator of the distribution of assets. If one relaxes the Nash equilibrium assumption and one considers rationalizable outcomes as possible outcomes, then one must wonder whether out-of-equilibrium outcomes may improve consumers' surplus. Looking at $Q^{\inf}$ as in Proposition 6 amounts to have a competition regulator whose risk aversion toward the strategic uncertainty is infinite (the regulator puts a high probability on worst aggregate productions in the case of multiplicity of rationalizable outcomes). Proposition 6 shows that sufficient risk aversion leads to select a distribution of assets which yields dominance solvability of the Cournot equilibrium. But sufficient risk aversion does not always recommend to pick out the competitive outcome: The optimal distribution of assets is an oligopolistic one when the production capacity is large ($\sigma\delta N > 1$).

6 An Illustration from the U.S. Airline Industry

Our analysis predicts that a merger dampens 'coordination' volatility by making out-of-equilibrium behavior less likely. We assess this prediction by considering the merger between Delta Air Lines (DL) and Northwest Airlines (NW) in the U.S. airline industry. This merger was announced on April 2008, approved by the Department of Justice on October 2008, and completed on January 2010. It took place over a period of high volatility, with the global recession, soaring fuel prices and H1N1 flu pandemic.

In this industry a route linking two cities can be viewed as a separate market, and carriers as producing an amount of passengers transported. The number of passengers transported in a given route by a given airline is limited by the time allocated by the airports to this airline under the form of landing slots. In the short/medium run, the distribution of the slots across airlines can be considered as given. These slots, or the corresponding seat capacity, are used as a proxy for the productive assets of our Cournot setup.

6.1 Data Description

Our data comes from the Airline Origin and Destination Survey, a quarterly 10% sample of airline tickets collected by the U.S. Department of Transportation. The data gives the origin and the destination airports, the ticketing carrier, the number of passengers and the airfare for about 5 millions observations per quarter.

A market (route) is defined as all flights between two U.S. cities, irrespective of the serving airports, intermediate transfer points and the direction of the flight path. The period of analysis comprises 18 quarters from 2006:3 to 2010:4, i.e., 9 quarters before and after the merger approval.[4]

There are $1,154$ (resp. $3,305$) routes from which DL (resp. NW) is absent every quarter during the whole period of analysis. The intersection of these two sets of routes yields a group of $1,099$ routes. These routes are considered to be not affected by the merger, and they are used as control group. The treatment group consists of all the routes where both DL and NW are active every quarter before the merger approval. This is the case in $2,934$ (resp. $1,702$) routes for DL (resp. NW). DL and NW were not competing directly on most routes: The intersection of these two sets of routes only comprises 839 routes.

The data is aggregated so that one observation gives the number of tickets per carrier, quarter and route. The final sample comprises $12,639$ observations, corresponding to $1,939$ routes linking 235 cities. The market share of DL (resp. NW) computed from 2006:3 to 2008:3 equals 11.3% (resp. 6.7%). Table 1 shows that routes in the treatment group involve a lower traffic volume and greater competition, with a lower Herfindahl index and more carriers active in 2006:3.

6.2 Variable Definitions

The main variable of interest is the variance within firms of the number of passengers transported. It measures how much a firm changes its output with respect to its own average output. It provides a proxy for 'coordination' volatility by reflecting the difficulties

[4]Data cleaning proceeds as in Kim and Singal (1993). We remove observations with missing carrier, with a zero fare or abnormal fares in the bottom and top 1% of observations, and tickets with a fare higher than 3 USD per mile. Following Ciliberto and Williams (2010) we only consider carriers with at least 20 reported passengers per quarter. This corresponds to an airline using at least a 20-seat plane at full capacity every week. Finally we neglect routes with less than 30 reported passengers per quarter, and routes where information is missing for some quarter during the period of analysis.

Table 1: Descriptive statistics

Variable	Control	Treatment
Number of routes	1,099	839
Number of cities	219	125
Number of passengers by route (2006:3)	3,071	1,194
Mean fare by route (2006:3)	205.2	226.5
Number of carriers by route (2006:3)	2.25	5.64
Herfindahl by route (2006:3)	0.7	0.3

faced by a carrier to set a suitable output level, i.e., to adjust suitably its production to others' behavior. This variance is our endogenous variable.

Let carriers be indexed by i, routes by j, quarters by $t = 1, \ldots, 18$ (2006:3 – 2010:4), and periods by p, with $p = 0$ before the merger announcement ($t \leq 7$) and $p = 1$ after the announcement. Let $\mathcal{P}_0$ (resp. $\mathcal{P}_1$) the set of quarters such that $p = 0$ (resp. $p = 1$). The average production of firm i in market j over period $\mathcal{P} \in \{\mathcal{P}_0, \mathcal{P}_1\}$ is

$$\bar{q}_{ij} = \frac{1}{(\#\mathcal{P})} \sum_{t \in \mathcal{P}} q_{ij}(t),$$

and our measure of 'coordination' volatility is

$$\sigma_w^2(i,j) = \frac{1}{(\#\mathcal{P}) - 1} \sum_{t \in \mathcal{P}} \left(q_{ij}(t) - \bar{q}_{ij}\right)^2 .$$

The set of regressors includes route and airline variables. All the route variables are computed from the first quarter of the period of analysis (2006:3). They consist of the number of passengers transported and the average fare for each route, the number of active carriers, and the Herfindahl index of the route. We also use the lowest distance between the two market endpoints and the total number of airports in the corresponding cities.

Airline variables include a company fixed effect, the fuel price per gallon paid by the company (the ratio between its domestic fuel cost in USD and its domestic fuel consumption in gallons) and the distance between the market endpoints and the closest hub of the company. The last two variables vary over time to control for changes in the production cost in a context of sharp increase in oil prices.

6.3 Merger and Market Volatility

Figure 2 shows the evolution of the (log of the) total number of passengers transported in the control and treatment group. Both follow a regular seasonal pattern, with a slight (not

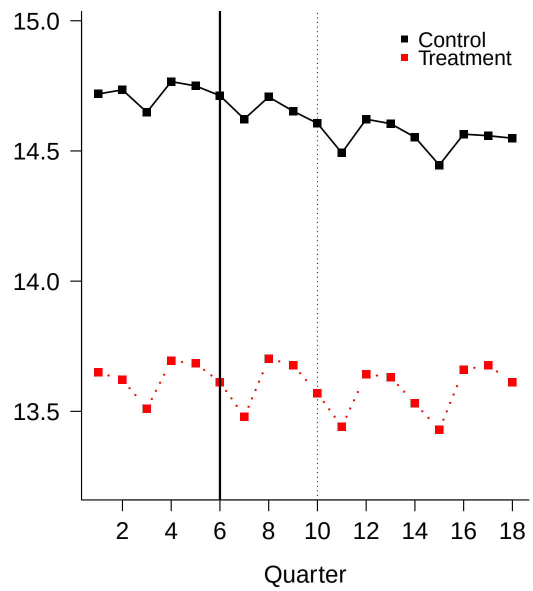

Figure 2: Number of passengers by group

statistically significant) decrease in the control group. The two vertical lines respectively indicate the time of the merger announcement (2008:2) and approval (2008:4).

We consider the model

$$\log(\sigma_{ij}(p)) = \text{cste} + r_j + c_i^0 + c_i^1(p) + \text{Period} + \text{Treatment} + (\text{Period} \times \text{Treatment}) + \epsilon_{ij}(p), \quad (11)$$

where r_j comprises the route variables, c_i^0 is a company fixed effect, and $c_i^1(p)$ includes the two company variables which change over time (fuel price and the distance to the closest hub of the company). The variable Period is 0 before the merger announcement, and 1 after the announcement. Treatment is a dummy which takes value 0 if the route belongs to the control group, and 1 otherwise. The interaction term (Period $\times$ Treatment) gives the impact of the merger onto the within standard error of the number of passengers.

Table 2 reports the results for three variants of the model (11). In each variant the period before the merger covers all the quarters from 2006:3 to the quarter preceding the announcement of the merger (2008:1), and the period after the merger begins on the merger announcement (2008:2). In the variant reported in Column 1 (resp., 2 and 3) the period after the merger ends at the merger approval in 2008:4 (resp., the merger completion in 2010:1, and the last quarter of 2010).

Route and company variables affect within volatility in a similar way in the three variants. Routes with a low traffic volume and a low fare display low volatility. This tends to be also the case for routes with a high number of competitors dominated by few large airlines. The company fixed effects indicate that large companies typically have a stabilizing effect, though the two merging companies DL and NW seem to encounter greater difficulties to set a stable output.

The merger occurs over a period of steep rise in the price of oil in 2008, and including the begining of the global crisis in 2008:3. The overall period of analysis has indeed seen a huge increase in the within standard error of the number of passengers. This effect dampens in 2010 but remains 58.3% higher than before the merger announcement.

One could have expected the merger to magnify market volatility: The market could have reached an equilibrium that is perturbed by the merger. The merger would then cause an increase in the standard error as the market transitions to the new equilibrium. This is not what happens: The short-run (three-quarter) impact of the merger is to reduce significantly the within standard error of the number of passengers transported. The merger is found to reduce 'coordination' volatility by 23.2% in the short-run. This large impact dampens over time: Volatility is only reduced by 8.6% during the whole completion period (Variant 2), which covers eight quarters after the announcement. The impact entirely looses significance after two years.

There is consequently a strong stabilizing effect of the merger in the short/medium run, in a context of high volatility. Following Kim and Singal (1993) one can interpret the result in Variant 2 as reflecting the own impact of a change in market power. Indeed they argue that the anticipation of the merger make DL and NW more cooperative from the announcement quarter, and that efficiency gains (synergies) are absent until reaching the completion: Only the change in market power matters during the period from announcement to completion. The impact measured in Variant 3 would instead mix changes in market power and synergies by covering about one year after the completion. A possible interpretation of our results is therefore that the change in market power due the merger stabilizes 'coordination' volatility by 8.6% while synergies eventually offset most of the effect of market power.

From the model in Variant 2, NW having 6.7% market share before the annoucement, we find that the own impact of a change in market power corresponding to a 1% transfer of market share is to reduce the within standard error of the output by about 1.3% ($\simeq 8.6/6.7$).

Table 2: Impact of the merger

	$\log(\sigma_w(i,j))$		
	2006:3–2008:1 *vs* 2008:2–2008:4	2006:3–2008:1 *vs* 2008:2–2010:1	2006:3–2008:1 *vs* 2008:2–2010:4
Constant	14.473***	10.538***	13.025***
	(1.063)	(1.168)	(1.282)
Route variables (2006:3)			
log(Nb of passengers)	0.761***	0.739***	0.728***
	(0.008)	(0.008)	(0.008)
log(Fare)	0.070**	0.063*	0.084**
	(0.035)	(0.035)	(0.035)
log(Nb of carriers)	−0.827***	−0.785***	−0.731***
	(0.032)	(0.032)	(0.032)
log(Herfindahl index)	−0.022	0.005	−0.006
	(0.015)	(0.016)	(0.016)
log(Market distance)	−0.018	0.005	−0.006
	(0.015)	(0.016)	(0.016)
log(Nb of airports)	0.013	0.010	0.002
	(0.015)	(0.015)	(0.015)
Company (*ref: American Airlines*)			
AS (Alaska Airlines)	−0.063	−0.079*	−0.101**
	(0.042)	(0.043)	(0.042)
CO (Continental Airlines)	−0.293***	−0.236***	−0.228***
	(0.025)	(0.025)	(0.025)
DL (Delta Air Lines)	0.097***	0.244***	0.231***
	(0.031)	(0.028)	(0.028)
FL (AirTran)	0.371***	0.349***	0.332***
	(0.041)	(0.041)	(0.040)
G4 (Allegiant Air)	1.467***	1.568***	1.595***
	(0.111)	(0.112)	(0.111)
HA (Hawaiian Airlines)	0.554***	0.614***	0.614***
	(0.113)	(0.118)	(0.120)
NW (Northwest Airlines)	0.039	0.085**	0.073**
	(0.032)	(0.033)	(0.034)
UA (United Airlines)	−0.229***	−0.174***	−0.209***
	(0.024)	(0.024)	(0.024)
US (US Airways)	−0.012	0.007	−0.013
	(0.024)	(0.025)	(0.025)
WN (Southwest Airlines)	0.221***	0.136***	0.186***
	(0.033)	(0.032)	(0.032)
YX (Midwest Airlines)	−0.099	−0.180***	−0.241***
	(0.066)	(0.064)	(0.061)
log(Distance to the closest hub)	−0.096***	−0.103***	−0.104***
	(0.003)	(0.003)	(0.003)
log(Fuel price)	1.143***	0.840***	1.030***
	(0.081)	(0.089)	(0.098)
Impact of the merger			
Period	0.838***	0.717***	0.583***
	(0.023)	(0.027)	(0.026)
Treatment	−0.213***	−0.285***	−0.314***
	(0.028)	(0.029)	(0.029)
Period × Treatment	−0.232***	−0.086***	−0.014
	(0.029)	(0.029)	(0.029)
Observations	11,770	12,397	12,639
R^2	0.697	0.647	0.632
Adjusted R^2	0.697	0.646	0.632
Residual Std. Error	0.714 (df = 11747)	0.748 (df = 12374)	0.759 (df = 12616)
F Statistic	1,229.037*** (df = 22; 11747)	1,030.666*** (df = 22; 12374)	985.608*** (df = 22; 12616)
Notes:	***Significant at the 1 percent level.		
	**Significant at the 5 percent level.		
	*Significant at the 10 percent level.		

References

[1] Basu, K., 1992, A characterization of the class of rationalizable equilibria of oligopoly games, Economics Letters 40, 187-191.

[2] Bernheim, B., 1984, Rationalizable strategic behavior, Econometrica 52, 1007-28.

[3] Borenstein, S., 1990, Airline mergers, airport dominance, and market power, American Economic Review 80(2), 400-404.

[4] Börgers, T. and M. Janssen, 1995, On the dominance solvability of large Cournot games, Games and Economic Behavior 8, 297-321.

[5] Ciliberto, F. and J. Williams, 2010, Limited Access to Airport Facilities and Market Power in the Airline Industry, Journal of Law and Economics 53 (2010), 467-495.

[6] Compte, O., F. Jenny and P. Rey, 2002, Collusion, mergers and capacity constraints, European Economic Review 46, 1-29.

[7] Davis, L. and C. Wolfram, 2011, Deregulation, consolidation, and efficiency: evidence from U.S. nuclear industry, NBER working paper 17341.

[8] Desgranges, G. and S. Ghosal, 2010, P-Stable Equilibrium : Definition and Some Properties, The Warwick Economics Research Paper Series (TWERPS) 952, University of Warwick, Department of Economics.

[9] Durieu, P., P. Solal and O. Tercieux, 2011, Adaptive learning and p-best response sets, International Journal of Game Theory 40, 735-747.

[10] Farrell, J. and C. Shapiro, 1990, Asset ownership and market structure in oligopoly, Rand Journal of Economics 21, 275-292.

[11] Gaballo, G., 2013, Market Power, Expectational Instability and Welfare, mimeo.

[12] Gugler, K., D.C. Mueller, B.B. Yurtoglu and C. Zulehner, 2003, The effects of mergers: an international comparison, International Journal of Industrial Organization 21, 625-653.

[13] Guesnerie, R., 1992, An exploration of the eductive justifications of the rational-expectations hypothesis, American Economic Review 82, 1254-78.

[14] Guesnerie, R., 1993, Theoretical tests of the rational expectations hypothesis in economic dynamical models, Journal of Economic Dynamics and Control 17, 847-864.

[15] Hommes, C. and F. Wagener, 2010, Does eductive stability imply evolutionary stability?, Journal of Economic Behavior and Organization 75, 25-39.

[16] Kim, E. and V. Singal, 1993, Mergers and Market Power: Evidence from the Airline Industry, American Economic Review 83, 549-569.

[17] Marx, L., 1999, Adaptive Learning and Iterated Weak Dominance, Games and Economic Behavior 26, 253-278.

[18] Milgrom, P., and J. Roberts, 1990, Rationalizability, learning, and equilibrium in games with strategic complementarities, Econometrica, 58, 12551277.

[19] Moulin, H., 1984, Dominance solvability and Cournot stability, Mathematical Social Science 7, 83-102.

[20] Perry, M. and M. Porter, 1985, Oligopoly and the incentive for horizontal merger, American Economic Review 75, 219-227.

[21] Town, R., D. Wholey, R. Feldman, L.R. Burns, 2006, The welfare consequences of hospital mergers, NBER working paper 12244.

[22] Williamson, O., 1968, Economies as and antitrust defense: The welfare trade-offs, American Economic Review, 58(1), 18-36.

7 Proof of Lemmas 2 and 3

Let L_0 be the set of values of ℓ such that $q_\ell^{\text{inf}} = 0$. The relations (4) give:

$$\begin{aligned}
\forall \ell \notin L_0 \quad &, \quad q_\ell^{\text{inf}} = b_\ell \left(\frac{\delta_0}{\delta} - (Q^{\text{sup}} - q_\ell^{\text{sup}}) \right), \\
\forall \ell \in L_0 \quad &, \quad q_\ell^{\text{inf}} = 0, \\
\forall \ell \quad &, \quad q_\ell^{\text{sup}} = b_\ell \left(\frac{\delta_0}{\delta} - (Q^{\text{inf}} - q_\ell^{\text{inf}}) \right).
\end{aligned}$$

Solving for q_ℓ^{inf} and q_ℓ^{sup} gives

$$\forall \ell \notin L_0 \quad , \quad q_\ell^{\text{inf}} = \frac{b_\ell^2}{1 - b_\ell^2} \left(\frac{\delta_0}{\delta} - Q^{\text{inf}} \right) + \frac{b_\ell}{1 - b_\ell^2} \left(\frac{\delta_0}{\delta} - Q^{\text{sup}} \right), \tag{12}$$

$$\forall \ell \notin L_0 \quad , \quad q_\ell^{\text{sup}} = \frac{b_\ell}{1 - b_\ell^2} \left(\frac{\delta_0}{\delta} - Q^{\text{inf}} \right) + \frac{b_\ell^2}{1 - b_\ell^2} \left(\frac{\delta_0}{\delta} - Q^{\text{sup}} \right), \tag{13}$$

$$\forall \ell \in L_0 \quad , \quad q_\ell^{\text{sup}} = b_\ell \left(\frac{\delta_0}{\delta} - Q^{\text{inf}} \right), \tag{14}$$

and $\forall \ell \in L_0$, $q_\ell^{\text{inf}} = 0$. Summing over ℓ gives a linear system in Q^{inf} and Q^{sup} whose solution is (7), namely:

$$Q^{\text{inf}} = \left(1 + \frac{c - a}{a^2 - c\,(c + e)} \right) \frac{\delta_0}{\delta}, \quad Q^{\text{sup}} = Q^{\text{inf}} + \frac{e}{a^2 - c\,(c + e)} \frac{\delta_0}{\delta},$$

with

$$a = 1 + \sum_{\ell \notin L_0} \frac{b_\ell^2}{1 - b_\ell^2}, \; c = \sum_{\ell \notin L_0} \frac{b_\ell}{1 - b_\ell^2} \text{ and } e = \sum_{\ell \in L_0} b_\ell.$$

For every $\ell \in L_0$, $q_\ell^{\text{inf}} = 0$ so that $\sum_{k \neq \ell} q_k^{\text{sup}} > \delta_0 / \delta$. Using (13), (14) and the expressions of Q^{inf} and Q^{sup}, this latter inequality is equivalent to:

$$\frac{(c - a)\,(b_m + 1) + e}{a^2 - c\,(e + c)} > 0. \tag{15}$$

Since $e \geq 0$ and

$$Q^{\text{sup}} - Q^{\text{inf}} = \frac{e}{a^2 - c\,(e + c)} \frac{\delta_0}{\delta} \geq 0, \tag{16}$$

it follows that $a^2 - c^2 - ce > 0$, so that the inequality (15) is equivalent to

$$(a - c)\,(b_m + 1) < e, \tag{17}$$

which is equivalent to (6) since

$$c - a = \sum_{k \notin L_0} \frac{b_k}{1 + b_k} - 1.$$

Hence, $q_\ell^{\text{inf}} = 0$ if and only if (6) does not hold true. Since the LHS of (6) is increasing in ℓ, there is a value $\bar{\ell}$ such that $q_\ell^{\text{inf}} = 0$ if and only if $\ell > \bar{\ell}$.

8 Proof of Proposition 4

Let $dN_h = -dN_s < 0$. By (2), b_ℓ is increasing and concave in N_ℓ. Hence we have $db_h < 0 < db_s$ and $db_h + db_s > 0$. Differentiating (16) gives:

$$dQ^{\text{sup}} - dQ^{\text{inf}} = \frac{\delta_0}{\delta} d\left(\frac{e}{a^2 - c(c+e)}\right). \tag{18}$$

We distinguish between 3 cases for the computation of $dQ^{\text{sup}} - dQ^{\text{inf}}$.

Case 1: $\bar{\ell} < h < s$. a and c remain constant and (18) writes:

$$dQ^{\text{sup}} - dQ^{\text{inf}} = \frac{\delta_0}{\delta} \frac{a^2 - c^2}{(a^2 - c(c+e))^2} de,$$

where $de = db_h + db_s > 0$. Since $a^2 - c^2 = (a-c)(a+c)$, simple algebra allows us to check that the above numerator is positive so that $dQ^{\text{sup}} - dQ^{\text{inf}} > 0$.

Case 2: $h < s \leq \bar{\ell}$. e remains constant and (18) writes:

$$dQ^{\text{sup}} - dQ^{\text{inf}} = -\frac{\delta_0}{\delta} \frac{e}{(a^2 - c(c+e))^2} (2ada - (2c+e)\,dc).$$

This has the same sign as $((2c+e)\,dc - 2ada)$. It is positive if and only if

$$\left(c + \frac{1}{2}e\right) \sum_{\ell \leq \bar{\ell}} d\left(\frac{b_\ell}{1 - b_\ell^2}\right) > a \sum_{\ell \leq \bar{\ell}} d\left(\frac{b_\ell^2}{1 - b_\ell^2}\right). \tag{19}$$

On the one hand, $\frac{b_\ell}{1+b_\ell}$ is increasing and concave in N_ℓ, which implies

$$\sum_{\ell \leq \bar{\ell}} d\left(\frac{b_\ell}{1 + b_\ell}\right) > 0.$$

This latter inequality rewrites

$$\sum_{\ell \leq \bar{\ell}} d\left(\frac{b_\ell}{1 - b_\ell^2}\right) > \sum_{\ell \leq \bar{\ell}} d\left(\frac{b_\ell^2}{1 - b_\ell^2}\right), \tag{20}$$

The LHS is positive since $\frac{b_\ell}{1-b_\ell^2}$ is shown to be increasing and concave in N_ℓ (but the RHS cannot be signed because $\frac{b_\ell^2}{1-b_\ell^2}$ is neither concave nor convex in N_ℓ). If the RHS is negative, then (20) implies that (19) holds true. If the RHS is positive, then rewriting (17) for $\bar{\ell}$ gives

$$a - c > \frac{e}{1 + b_{\bar{\ell}}}, \tag{21}$$

which implies $c + \frac{1}{2}e < a$. Combining this latter inequality with (20) proves that (19) holds true. This shows that $dQ^{\text{sup}} - dQ^{\text{inf}} > 0$.

Case 3: $h \leq \bar{\ell} < s$**.** (18) writes:

$$\begin{aligned} dQ^{\sup} - dQ^{\inf} &= -\frac{\delta_0}{\delta} \frac{e}{\left(a^2 - c\left(c+e\right)\right)^2} \left(2ada - \left(2c+e\right) dc\right) \\ &\quad + \frac{\delta_0}{\delta} \frac{a^2 - c^2}{\left(a^2 - c\left(c+e\right)\right)^2} de \end{aligned}$$

$dQ^{\sup} - dQ^{\inf}$ has the same sign as

$$\begin{aligned} & -e\left(2ada - \left(2c+e\right) dc\right) + \left(a^2 - c^2\right) de \\ = \ & e \frac{\left(2c+e\right)\left(1+b_h^2\right) - 4ab_h}{\left(1-b_h\right)^2 \left(1+b_h\right)^2} db_h + \left(a^2 - c^2\right) db_s \end{aligned}$$

Since $db_s > 0 > db_h$ and $db_h + db_s > 0$, the above expression is positive if

$$e \frac{\left(2c+e\right)\left(1+b_h^2\right) - 4ab_h}{\left(1-b_h\right)^2 \left(1+b_h\right)^2} < \left(a^2 - c^2\right) \tag{22}$$

Inequality (21) implies ($h \leq \bar{\ell}$):

$$a - c > \frac{e}{1+b_{\bar{\ell}}} \geq \frac{e}{1+b_h} > 0. \tag{23}$$

Using $\left(a^2 - c^2\right) = \left(a-c\right)\left(a+c\right)$ a sufficient condition for Inequality (22) is

$$e \frac{\left(2c+e\right)\left(1+b_h^2\right) - 4ab_h}{\left(1-b_h\right)^2 \left(1+b_h\right)^2} < \left(a+c\right) \frac{1}{1+b_h} e.$$

This rewrites:

$$\left(2c+e\right)\left(1+b_h^2\right) - 4ab_h < \left(a+c\right)\left(1-b_h\right)^2 \left(1+b_h\right).$$

Using again Inequality (23) ($e < \left(a-c\right)\left(1+b_h\right)$), a sufficient condition for the above inequality is

$$\left(2c + \left(a-c\right)\left(1+b_h\right)\right)\left(1+b_h^2\right) - 4ab_h < \left(a+c\right)\left(1-b_h\right)^2 \left(1+b_h\right).$$

This rewrites

$$\begin{aligned} 2ab_h\left(b_h - 1\right) &< -c2b_h^2\left(1-b_h\right), \\ a &> cb_h. \end{aligned}$$

Since $a - c > 0$, $a > c > cb_h$. This shows $dQ^{\sup} - dQ^{\inf} > 0$.

9 Proof of Proposition 5

We maximize Q^* subject to (8) and (9) in three steps. Since Q^* is increasing in S (see (3)), the optimization problem is to maximize S subject to (8) and (9).

Step 1. Consider the Lagrangian:

$$\sum_{\ell=1}^{M} \frac{\sigma\delta N_\ell}{\sigma\delta N_\ell + 1} + \mu\left(1 - \sum_{\ell=1}^{M} \frac{\sigma\delta N_\ell}{3\sigma\delta N_\ell + 1}\right) + \eta\left(N - \sum_{\ell=1}^{M} N_\ell\right).$$

It is the Lagrangian associated with the maximization problem for a given value of M. Any solution to the initial optimization problem satisfies the first-order conditions in N_ℓ associated with this Lagrangian. The first-order conditions in N_ℓ are:

$$\sigma\delta P(\sigma\delta N_\ell) = 0, \text{ for every } \ell,$$

where

$$P(x) = \frac{1}{(1+x)^2} - \mu\frac{1}{(1+3x)^2} - \frac{\eta}{\sigma\delta}.$$

Hence, the number of different firms (different values of N_ℓ) at a solution of the optimization problem equals the number of positive roots of P. Observe that

$$P'(x) = -\frac{2}{(1+x)^3} + \mu\frac{6}{(1+3x)^3}.$$

Since $P'(x) \geq 0$ rewrites

$$(3 - (3\mu)^{1/3})x \leq (3\mu)^{1/3} - 1,$$

P' can change its sign at most once. Hence, either P is monotonic or P admits one local extremum. It follows that P admits at most 2 positive roots: the solution to the optimization problem involves at most two types of firms.

Denote $i = 1, 2$ the type of a firm. Let M_i the number of firms of type i ($i = 1, 2$). Every type i firm uses N_i assets ($0 \leq N_1 \leq N_2$ w.l.o.g.).

Step 2. We maximize S for given N_1 and N_2 under the 2 constraints (8) and (9). S is linear in M_1 and M_2:

$$S = M_1\frac{\sigma\delta N_1}{\sigma\delta N_1 + 1} + M_2\frac{\sigma\delta N_2}{\sigma\delta N_2 + 1}.$$

The stability constraint (8) is linear:

$$M_1 \leq \frac{3\sigma\delta N_1 + 1}{\sigma\delta N_1} - \frac{N_2}{N_1}\frac{3\sigma\delta N_1 + 1}{3\sigma\delta N_2 + 1}M_2,$$

and the feasibility constraint (9) is linear too:

$$M_1 \leq \frac{N}{N_1} - \frac{N_2}{N_1}M_2.$$

The marginal rate of substitution of S is $\frac{\frac{\sigma\delta N_2}{\sigma\delta N_2+1}}{\frac{\sigma\delta N_1}{\sigma\delta N_1+1}}$, it lies between the slopes of the constraints:

$$\frac{N_2}{N_1}\frac{3\sigma\delta N_1+1}{3\sigma\delta N_2+1} < \frac{\frac{\sigma\delta N_2}{\sigma\delta N_2+1}}{\frac{\sigma\delta N_1}{\sigma\delta N_1+1}} < \frac{N_2}{N_1}.$$

Thus, we have 3 cases:

Case 1: $N_1 \geq \frac{\sigma\delta N-1}{3\sigma\delta}$**.** Then

$$\frac{3\sigma\delta N_1+1}{\sigma\delta N_1} \geq \frac{N}{N_1},$$

and the feasibility constraint is the only relevant constraint (i.e., feasibility implies stability). The solution is $M_2 = 0$, $M_1 = N/N_1$. The value of S is

$$\frac{\sigma\delta N}{\sigma\delta N_1+1}.$$

Case 2:$N_2 \leq \frac{\sigma\delta N-1}{3\sigma\delta}$**.** Then,

$$\frac{3\sigma\delta N_2+1}{\sigma\delta N_2} \leq \frac{N}{N_2},$$

and the stability constraint is the only relevant constraint (i.e., stability implies feasibility). The solution is $M_1 = 0$, $M_2 = \frac{3\sigma\delta N_2+1}{\sigma\delta N_2}$. The value of S is

$$\frac{3\sigma\delta N_2+1}{\sigma\delta N_2+1}.$$

Case 3: $N_1 < \frac{\sigma\delta N-1}{3\sigma\delta} < N_2$**.** The 2 constraints are relevant. The solution is at the unique intersection between the constraints, namely

$$\begin{cases} M_1 = \dfrac{1}{N_2-N_1}\dfrac{3\sigma\delta N_1+1}{3\sigma\delta N_1}\left(\dfrac{3\sigma\delta N_2+1}{\sigma\delta} - N\right) \\ M_2 = \dfrac{1}{N_2-N_1}\dfrac{3\sigma\delta N_2+1}{3\sigma\delta N_2}\left(N - \dfrac{3\sigma\delta N_1+1}{\sigma\delta}\right) \end{cases}$$

The value of S is (after some computations):

$$\frac{\sigma\delta N_1\left(1+3\sigma\delta N_2\right)+\sigma\delta N_2+\frac{2\sigma\delta N+1}{3}}{\left(\sigma\delta N_1+1\right)\left(\sigma\delta N_2+1\right)}.$$

Step 3. We solve for N_1 and N_2 maximizing S in each of the 3 above cases.

Case 1. Maximizing S amounts to minimize N_1. If $\sigma\delta N \leq 1$, then the solution is $N_1 = 0$ and $M_1 = +\infty$ ($M_2 = 0$) and the aggregate production is $\sigma\delta N$ (we are in the competitive case). If $\sigma\delta N > 1$, then the solution is $N_1 = \frac{\sigma\delta N-1}{3\sigma\delta}$ and

$$\begin{aligned} M_1 &= \frac{3\sigma\delta N}{\sigma\delta N - 1}, \\ S &= \frac{3\sigma\delta N}{\sigma\delta N + 2}. \end{aligned}$$

The aggregate production is

$$Q^* = \frac{S}{1+S}\frac{\delta_0}{\delta} = \frac{1}{2}\frac{3\sigma\delta N}{1+2\sigma\delta N}\frac{\delta_0}{\delta}.$$

Case 2. (this case requires $\sigma\delta N > 1$) Maximizing S amounts to maximize N_2. The solution is

$$N_2 = \frac{\sigma\delta N - 1}{3\sigma\delta},$$

this is the same solution as Case 1.

Case 3. (this case requires $\sigma\delta N > 1$). The derivatives of S w.r.t. N_1 and N_2 have the following signs:

$$\begin{aligned} \frac{\partial S}{\partial N_1} &\sim 3\sigma\delta N_2 - N\sigma\delta + 1 > 3\sigma\delta\frac{\sigma\delta N - 1}{3\sigma\delta} - N\sigma\delta + 1 = 0, \\ \frac{\partial S}{\partial N_2} &\sim 3\sigma\delta N_1 - N\sigma\delta + 1 < 3\sigma\delta\frac{\sigma\delta N - 1}{3\sigma\delta} - N\sigma\delta + 1 = 0. \end{aligned}$$

S is increasing in N_1 and decreasing in N_2. At the optimum,

$$N_1 = N_2 = \frac{\sigma\delta N - 1}{3\sigma\delta}.$$

This is again the same solution as Case 1.

Summing up the 3 cases:

- If $\sigma\delta N \leq 1$, then Case 1 is the only possible case and the solution is that N is divided equally across an infinite number of firms (competitive market).
- If $\sigma\delta N > 1$, then the 3 cases give the same solution: a symmetric oligopoly with $\frac{3\sigma\delta N}{\sigma\delta N-1}$ firms and where each firm owns the same number $\frac{\sigma\delta N-1}{3\sigma\delta}$ of assets. The aggregate production is $\frac{1}{2}\frac{3\sigma\delta N}{1+2\sigma\delta N}\frac{\delta}{\delta_0}$. One easily checks that both constraints (8) and (9) are binding.

10 Proof of Proposition 6

The expression (7) of $Q^{\inf}$ implies:

$$\frac{1}{1-\frac{\delta}{\delta_0}Q^{\inf}} = 1 + \sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1-b_\ell} + \frac{\left(\sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1-b_\ell^2}\right)\sum_{\ell>\bar{\ell}} b_\ell}{\sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1+b_\ell} - 1}.$$

By definition of $\bar{\ell}$ and Lemma 2, we have

$$\sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1+b_\ell} < \sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1+b_\ell} + \sum_{\ell>\bar{\ell}} \frac{b_\ell}{1+b_{\bar{\ell}}} \leq 1,$$

where the strict inequality comes from the equilibrium being unstable. Hence

$$\sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1+b_\ell} - 1 < 0,$$

and

$$\frac{1}{1-\frac{\delta}{\delta_0}Q^{\inf}} \leq 1 + \sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1-b_\ell}.$$

From (10), we have

$$\frac{1}{1-\frac{\delta}{\delta_0}Q^{**}} = 2\frac{1+2\sigma\delta N}{2+\sigma\delta N}.$$

Since every aggregate production ($Q^{\inf}$ or Q^{**}) is smaller than $\frac{\delta_0}{\delta}$, we have:

$$Q^{\inf} < Q^{**} \Leftrightarrow \frac{1}{1-\frac{\delta}{\delta_0}Q^{\inf}} < \frac{1}{1-\frac{\delta}{\delta_0}Q^{**}}.$$

A sufficient condition for this last inequality is

$$1 + \sum_{\ell\leq\bar{\ell}} \frac{b_\ell}{1-b_\ell} < 2\frac{1+2\sigma\delta N}{2+\sigma\delta N}. \tag{24}$$

Let $\alpha_\ell = N_\ell/N \geq 0$ (α_ℓ decreasing in ℓ and $\sum \alpha_\ell = 1$). Using the definition (2) of b_ℓ, (24) rewrites

$$\sum_{\ell\leq\bar{\ell}} \frac{\sigma\delta\alpha_\ell N}{\sigma\delta\alpha_\ell N + 1} < \frac{3\sigma\delta N}{\sigma\delta N + 2}. \tag{25}$$

Note that $f(\alpha_\ell) = \frac{\sigma\delta\alpha_\ell N}{\sigma\delta\alpha_\ell N+1}$ is concave in α_ℓ so that (Jensen inequality)

$$\sum_{\ell\leq\bar{\ell}} \frac{f(\alpha_\ell)}{\sum_{\ell\leq\bar{\ell}}\alpha_\ell} \leq f\left(\frac{\sum_{\ell\leq\bar{\ell}}\alpha_\ell}{\sum_{\ell\leq\bar{\ell}}\alpha_\ell}\right) = f(1).$$

This rewrites

$$\sum_{\ell \leq \bar{\ell}} f(\alpha_\ell) \leq \frac{\sigma\delta N}{\sigma\delta N + 1} \sum_{\ell \leq \bar{\ell}} \alpha_\ell$$

Since $\sum_{\ell \leq \bar{\ell}} \alpha_\ell \leq 1$ and

$$\frac{\sigma\delta N}{\sigma\delta N + 1} < \frac{3\sigma\delta N}{\sigma\delta N + 2},$$

this implies

$$\sum_{\ell \notin L_0} f(\alpha_\ell) < \frac{3\sigma\delta N}{\sigma\delta N + 2},$$

which shows that (25) holds true.