Short- and long-term optimality under sustainable threats in Contest Theory models of advertising and short-run competition

Reinoud Joosten, Rogier Harmelink, Thom Sparrius*
Financial Engineering, University of Twente
The Netherlands.
Email corresponding author: r.a.m.g.joosten@utwente.nl.

July 21, 2023

Abstract

We model advertising with effects on different time scales for a duopoly in imperfect substitutes using elements from Contest Theory. Firms addionally compete in a short run strategic variable, here price or quantity, allowing simultaneous or sequential decisions, or collusion in endogenously changing stage games. Strategic variables range from 'slow' (advertising), over 'moderate' (quantities) to 'fast' (prices).

We find feasible rewards and equilibria for the limiting average reward criterion. Uniqueness of equilibrium is not guaranteed, and we introduce two criteria which act as natural refinements. We impose stage-game rationality, i.e., the firms play optimally in each stage game. Furthermore, in establishing threats, we require that punishment is sustainable, i.e., the punisher must have nonnegative long term average own profits to avoid bankruptcy.

JEL-codes: C72, C73, L13, M31, M37.
Keywords: advertising, externalities, limiting average rewards, equilibria, long-term and stage-game optimality, sustainable threats

1 Introduction

We model and analyze strategic interaction over time in a duopoly in which advertising may have different types of effects in the time dimension. We engineer a series of connected games with joint frequency dependent stage payoffs which allows us to build complex relationships over time, and analyze them with techniques often used for infinitely repeated games.[1]

Conflict of Interest The authors have no involvement in any organization or entity with any financial or non-financial interest in this manuscript.

[1] Engineer as in Aumann [2008]. See Joosten *et al.* [2003] for early work leading up to the class of *ETP-ESP* games, and Joosten & Samuel [2020] for an overview.

Our paper is a continuation and extension of earlier efforts in applying (stochastic) game theory[2] to duopoly (cf., Joosten [2009, 2016b]). These contributions feature a combination of competitive advertising in the long run and strategic choice in an additional strategic variable in the short run. Joosten [2009[assumed each short-run outcome to result from a Cournot equilibrium in the stage game arising from materialized advertisement decisions until then. Subsequently, Joosten [2009] established Nash equilibria regarding the advertisement strategies in the infinitely repeated game. Joosten [2016b] modeled the short run phase explicitly, assuming a rich structure connecting past advertisement decisions and the momentary strategic environment, mainly captured in the momentary demand functions.

Our first extension adds a level of explanation and motivation regarding modeling the consequences of past advertising decisions on current demand functions. We assume that a demand enhancing effect common to both firms occurs and that the division of the benefits among them depends on their past advertisement efforts relative to the total efforts, in line with standard modes of modeling in contest theory cf., e.g., Vojnović [2015]).

Another extension deals with the short term competition. We examine *standard outcomes of duopoly distinguished in industrial organization*, i.e., Cournot and Bertrand equilibrium, sequential equilibrium in quantities (Stackelberg-Cournot) and in prices (Stackelberg-Bertrand), collusion in quantities and prices. The standard Bertrand model is the most natural 'place' to introduce contest theory. We show how to transform a Bertrand model into a Cournot model and vice versa. The proposed route is to obtain a Bertrand model, if necessary by transforming a Cournot model, add the desired extension and, if necessary, transform it back into a Cournot model.

Advertising influences immediate sales directly and future sales in a cumulative fashion (cf., e.g., Friedman [1983]) and we aim to capture both. We assume regarding the immediate, or short-run, effects, that the own advertising increases the own immediate sales (potential) given the action of the other firm. Advertising may also cause immediate externalities. Friedman [1983] distinguishes predatory and cooperative advertising. An increase in advertising efforts of one firm leads to a sales decrease of the other in the former type of advertising, and to an increase in the latter. We engineer features to allow representations of both short term effects of advertising.

To capture the cumulative effect of *current* advertisement on *future* sales we introduce the notion of market potential varying in time under the past advertising efforts of the firms (see e.g., Joosten [2016b]). The *current* market potential of each firm is determined by its own but also by its opponent's *past* efforts. A higher effort of either firm leads to an increase of the mar-

[2]Shapley [1953] introduced stochastic games, the non-terminating version is due to Hoffman & Karp [1966]. We follow the latter approach using the limiting average reward criterion to evaluate the infinite stream of stage payoffs.

ket potentials (ceteris paribus). Advertising is therefore cooperative in its cumulative effects on the market potentials. How much of a firm's market potential materializes as immediate sales, depends on the current advertisement decisions and on the relative past effort in advertising.

Each period the agents engage in strategic one-shot interaction in an environment determined by their past advertising efforts and their current advertising decisions, both known for the relevant strategic one-shot decision. We assume a stage to be split up in two action phases. First, both firms independently and simultaneously advertise, or not, as determined by their long term advertisement strategies, and pay associated advertisement costs. Next, they take a (one-shot) decision to maximize their stage game profits, and for the stage game the advertisement costs are sunk. Under Bertrand competition the firms set prices, under Cournot competition quantities. A second aspect to be dealt with is timing, i.e., we distinguish simultaneous and sequential decisions, in the latter case we add the term Stackelberg. Finally, we consider that the firms may collude, which means that they try to maximize the cumulative stage profits.

The current demand functions are determined 'roughly' by past advertising, whereas present efforts fix the demand functions in detail. To be more specific, high past cumulative advertisement expenditures shift the sales potential curves upward, high individual past expenditure make the own demand less price elastic; asymmetric efforts lead to a tilt of the same curves, making the more active firm less vulnerable to a price decrease of the less active one. Current demand functions are then fixed by the advertisement actions in the first action phase moment of the stage and *unless both firms advertise each loses part of its potential.*

We assume that the firms wish to maximize average profits over an infinite time-horizon. A strategy in this framework is a prescription for the entire time horizon consisting of advertising, and one-shot decisions for which the intertemporal effects of advertising efforts should be considered. Akin to backward induction approaches, we first solve the one-shot games, taking their outcomes as one of the options mentioned, Cournot, Bertrand, Stackelberg-Cournot, Stackelberg-Bertrand and collusion. Then, we determine equilibria with respect to the long term advertising strategies employing modifications of techniques traditionally used to analyze infinitely repeated games. We find, as in the Folk Theorem for repeated games, that a continuum of rewards can be supported by equilibria involving 'threats'.

2 Positioning of our model

The games to be engineered here in the sense of Aumann [2008] are games with joint frequency dependent stage payoffs, JFD-games for short, which generalize games with frequency dependent stage payoffs, or FD-games. The latter were introduced by Brenner & Witt [2003], and fully classified and analyzed by Joosten, Brenner & Witt [2003]. (J)FD-games are stochastic games with the advantage that intertemporal externalities can be modeled conveniently cf., e.g., Joosten [2007, 2009, 2016a,b, 2019] for applications.

Dorfman & Steiner [1954] examine the effects of advertising in a static monopoly and derive necessary conditions for the optimal level of advertising. In a dynamic monopolistic model, Nerlove & Arrow [1962] treat advertisement expenditures similar to investments in a durable good. This durable good is called goodwill which is assumed to influence current sales. Investments in advertisement increase the stock of goodwill, which simultaneously depreciates over time. Nerlove and Arrow derive necessary conditions for optimal advertising, thus generalizing the Dorfman-Steiner result. Friedman [1983] in turn generalizes the Nerlove-Arrow model to allow oligopolistic competition in advertising and derives necessary conditions for the existence of a noncooperative equilibrium (Nash [1951]).

Our market potential in the Bertrand model is close to goodwill in Nerlove & Arrow [1962] and Friedman [1983]. Though the modeling of the changes in time in the former model follows Vidale & Wolfe [1957], the authors quote Waugh [1959] as a source of inspiration. Vidale & Wolfe [1957] present empirical evidence of the positive effects of past advertising on current sales. Furthermore, current sales do not collapse if advertising is stopped, but deteriorate slowly over time. The treatment of changes in goodwill over time in Friedman [1983] follows Prescott [1973].

The body of work in economics on advertising is large. One source of variety is the modeling of time-related aspects. For instance, is the model static (e.g., Dorfman & Steiner [1954]), or does the strategic environment change (e.g., Nerlove & Arrow [1962])? Another source of variety is the market under consideration, e.g., monopoly (Nerlove & Arrow [1962]), oligopoly (Friedman [1983]), leader-follower oligopoly (Kydland [1977]). A third one is possible combinations of advertising with other marketing instruments, e.g., Schmalensee [1978] combines advertising and quality. A fourth is the entity to be influenced, sales (e.g., Nerlove & Arrow [1962]) or market shares (e.g., Fershtman [1984]). A fifth is the strategic dimension in which competition on the market is chosen, for instance Cournot (e.g., Joosten [2009]) or Bertrand competition (e.g., Cellini *et al.* [2008], Joosten [2016b]). A sixth is the distinction between cooperative and predatory advertising (cf., Friedman [1983]) which concerns the type of externalities of advertising assumed (cf., e.g., Cellini & Lambertini [2003], Joosten [2009]). For cooperative (predatory) advertising an increase in advertisement expenditures increases (decreases) the other firm's demand. Unfortunately Friedman [1983] does not make timing distinctions, but Joosten [2009,2016b] distinguishes short term and long term spillovers which may or may not be of opposite character.

Nelson [1970] makes a distinction between search and experience goods. The characteristics of the former are known before purchase, whereas those of the latter can be determined only after. Advertising differs for the two types of goods because the information conveyed to the consumers differs. *Informative* advertising provides information on e.g., the price, availability or characteristics of a product; *persuasive* advertising tries to generate consumer interest for a product, often by association or through rather indirect 'channels'. Informative advertising is directed at search goods, persuasive advertising aims at experience goods.

Van Cayseele & Furth [1996], Davidson & Deneckere [1986,1990], and Benoît & Krishna [1987] regard price as a fast strategic variable, and production or capacity as slow. Scherer [1980], Brander & Harris [1983], Fershtman & Muller [1986] note a tendency to collude in the fast variables, and to compete in the slow ones.

So, we model informative and persuasive advertising in the long run with cooperative or predatory characteristics on different time scales for a dynamic duopoly in which firms additionally compete in a short run strategic variable, price or quantity, whilst allowing different timing of decisions, i.e., simultaneous, sequential, or even collusion in endogenously changing stage games, integrating elements from Contest Theory. Strategic variables range from 'slow' (advertising), over 'moderate '(quantities) to 'fast' (prices).

3 The stage games

The advertisement game is played by two firms (players) A and B at discrete moments in time called stages. Each stage the players take two decisions sequentially. First, long term strategies specify for the current stage whether to advertise or not; next, the firms interact knowing the advertisement decisions taken. The stage game decisions are assumed independent, i.e., there are no externalities in the time dimension.

We devote some space to the Bertrand model and build in our Contest Theory extension here, as this seems the most convincing option. We obtain the Cournot model from the extended Bertrand model.

3.1 The Bertrand model for the stage game

We provide the specifics of the model for Bertrand competition in relation to a given advertising history. We have two firms competing in advertising in the same market, and they are labeled A and B. We capture the history of advertising in the following matrix:

$$\rho_t = \begin{bmatrix} [\rho_t]_{1,1} & [\rho_t]_{1,2} \\ [\rho_t]_{2,1} & [\rho_t]_{2,2} \end{bmatrix},$$

where $[\rho_t]_{1,1}$ ($[\rho_t]_{2,2}$) is the proportion of stages both firms advertised (did not advertise), and $[\rho_t]_{1,2}$ ($[\rho_t]_{2,1}$) is the proportion of stages Firm A (B) advertised and the opponent did not. Entries of the matrix add up to one.

The market potentials for both firms are given by

$$\begin{aligned} MP_A(\rho_t) &= D_A(\rho_t) - z_1(\rho_t)p_A + z_2(\rho_t)p_B \\ &= D_{A,0} + q(\rho_t)D(\rho_t) - z_1(\rho_t)p_A + z_2(\rho_t)p_B \\ MP_B(\rho_t) &= D_B(\rho_t) - z_3(\rho_t)p_B + z_4(\rho_t)p_A \\ &= D_{B,0} + (1 - q(\rho_t))\, D(\rho_t) - z_3(\rho_t)p_B + z_4(\rho_t)p_A. \end{aligned}$$

$D_{A,0}$ and $D_{B,0}$ are positive constants, $D(\rho_t)$ represents the current expansion of the market given the advertisement history. $D(\rho) = 0$ if $\rho_{2,2} = 1$, i.e., if neither firm advertises the increase in market potentials is zero. We assume

$$\frac{\partial D(\rho)}{\partial\left(\rho_{1,1} + \rho_{1,2}\right)}, \frac{\partial D(\rho)}{\partial\left(\rho_{1,1} + \rho_{2,1}\right)} \geq 0,$$

i.e., market potentials (weakly) increase if either firm increases its advertisement efforts. Now, $q(\rho_t)$ is the proportion that Firm A obtains of this market growth, the complement goes to the competitor.

We impose $z_1(\rho_t), z_3(\rho_t) > z_2(\rho_t), z_4(\rho_t) > 0$ for all $\rho_t \in \Delta^{2\times 2} = \{x \in \mathbb{R}^{2\times 2} |\ x_{i,j} \geq 0$ for all $i, j \in \{1, 2\}$ and $\sum_{i,j\in\{1,2\}} x_{i,j} = 1\}$, allowing to interpret the goods as imperfect substitutes as an increase of the price of A, ceteris paribus, causes the market potential of A (B) to go down (up).

3.1.1 Market demand for the second phase of the stage game

Demand materializes from market potential as follows. If both firms advertise in the current stage the resulting demand will be equal to the market potentials. In all other cases, market potential partly evaporates. Most market potential evaporates if neither firm advertises, in case only one firm advertises, less market potential evaporates, and the firm advertising losses less of its market potential than the firm not advertising. Consider

$$1 = \theta_A^{1,1} = \theta_A^{1,1} > \theta_A^{1,2} = \theta_B^{2,1} \geq \theta_A^{2,1} = \theta_B^{1,2} \geq \theta_A^{2,2} = \theta_B^{2,2} > 0,$$

then the following system notation formalizes these ideas

$$x_A^{i,j}(\rho_t) = \theta_A^{i,j} MP_A(\rho_t) \text{ and } x_B^{i,j}(\rho_t) = \theta_B^{i,j} MP_B(\rho_t),$$

where superscript $1,1$ (resp. $2,2$) indicates that both firms advertise (resp. do not) in the same stage; $1,2$ (resp. $2,1$) indicates that Firm A (B) advertises and the other does not in the same stage. So, $x_A^{1,1}(\rho_t)$ is Firm A's demand given ρ_t and given that both firms advertise in the first phase.

Profits of the firms for the second phase given $i, j \in \{1, 2\}$ are

$$\begin{aligned} \kappa_A^{i,j}(\rho_t, p_A) &= x_A^{i,j}(\rho_t)\,(p_A - c_A) - c_{A,0} = \theta_A^{i,j} MP_A(\rho_t)\,(p_A - c_A) - c_{A,0}, \\ \kappa_B^{i,j}(\rho_t, p_B) &= x_B^{i,j}(\rho_t)\,(p_B - c_B) - c_{B,0} = \theta_B^{i,j} MP_B(\rho_t)\,(p_B - c_B) - c_{B,0}. \end{aligned}$$

Here, c_A, c_B are variable unit costs, and $c_{A,0}, c_{B,0}$ are fixed costs.

3.2 Contest Theory and market potential growth

The formulation of the market potentials is comparable to standard formulations of demand functions, except for the parts

$$q(\rho_t)D(\rho_t) \text{ and } (1 - q(\rho_t))\, D(\rho_t).$$

The part depending on $D(\rho_t)$ is easily explained. Suppose there is a large population of consumers that can be influenced by advertising. If advertising occurs over some time, the potentials grow proportional to $D(\rho_t)$, whereas how much each firm obtains from this depends on $q(\rho_t)$. To motivate the specific functional form of $q(\rho_t)$, we reason as follows. Given past advertisement summarized by ρ_t, each consumer chooses product A with probability $q(\rho_t)$, and the other product with the complementary probability. Applying a law-of-large-numbers argument, we may approximate the stochastic demand increases on the aggregate level by the deterministic functions presented.

To connect the advertisement history to the stochastic choice model yielding probability $q(\rho_t)$, we employ Contest Theory for which we require a measure of effort. We assume $c_i^{ADV} = \beta_0 \cdot e_{i,s}^{\beta_1}$, with $\beta_0, \beta_1 > 0$, i.e., the advertisement costs of Firm $i = A, B$ in each period s, increase monotonically in the advertisement costs. To formalize the effects of past efforts in advertisement we take the cumulative past advertisement efforts, i.e.,

$$E_{A,t} = \sum_{s=1}^{t} e_{A,t} \text{ resp. } E_{B,t} = \sum_{s=1}^{t} e_{B,t}.$$

We assume that the advertisement efforts in each period are dyadic. Moreover, the lowest level of effort is 0, i.e., the agent does nothing. This means that $E_{A,s}, E_{B,s} \in \{0, e^H\}$ for all $s = 1, 2, ...$ As a result the firms pay a fixed fee of c, *each period in which they advertise*, i.e.

$$c_{A,t}^{ADV} = c_{B,t}^{ADV} = \beta_0 \cdot \left(e^H\right)^{\beta_1} \text{ for all } t \geq 2.$$

Next, we rewrite for given ρ_t the expression above as follows

$$E_{A,t} = \sum_{s=1}^{t} e_{A,t} = \sum_{s=1}^{t} e^H \cdot i_{A,s} = e^H \cdot t \left([\rho_t]_{1,1} + [\rho_t]_{1,2}\right),$$

where $i_{A,s}$ is an indicator function, such that $i_{A,s} = 1$ if Firm A advertised at stage $s \leq t$, and $i_{A,s} = 0$ otherwise. We take the same definitions and interpretations mutatis mutandis for Firm B. Note then that for given ρ_t

$$\frac{E_{A,t}}{E_{A,t} + E_{B,t}} = \frac{\left([\rho_t]_{1,1} + [\rho_t]_{1,2}\right)}{\left([\rho_t]_{1,1} + [\rho_t]_{1,2}\right) + \left([\rho_t]_{1,1} + [\rho_t]_{2,1}\right)}.$$

Adopting a proportionally allocation Contest Success Function from Vojnović [2015, Chapter 4] for Firm A, we establish a link between the efforts and the variables comprising the advertisement past

$$q(\rho_t) = \frac{[\rho_t]_{1,1} + [\rho_t]_{1,2}}{2[\rho_t]_{1,1} + [\rho_t]_{1,2} + [\rho_t]_{2,1}}.$$

Vojnović [2015, Chapter 4] provides several interesting alternatives which for the present setting are easily connected to the matrix ρ_t, too.

3.3 The Cournot model for the stage game

For our extended Bertrand model, the following result establishes the inverse demand functions of the corresponding Cournot model. For conciseness, define $j(1) = 2$, $j(2) = 0 = j(4)$, $j(3) = -2$, and notations:

$$\begin{aligned} g_k(\rho_t) &= \frac{z_{k+j(k)}(\rho_t)}{z_1(\rho_t) z_3(\rho_t) - z_2(\rho_t) z_4(\rho_t)} \text{ for } k = 1, ..., 4, \\ Y_A(\rho_t) &= \frac{z_3(\rho_t) D_A(\rho_t) + z_2(\rho_t) D_B(\rho_t)}{z_1(\rho_t) z_3(\rho_t) - z_2(\rho_t) z_4(\rho_t)}, \\ Y_B(\rho_t) &= \frac{z_4(\rho_t) D_A(\rho_t) + z_1(\rho_t) D_B(\rho_t)}{z_1(\rho_t) z_3(\rho_t) - z_2(\rho_t) z_4(\rho_t)}. \end{aligned}$$

Lemma 1 *For $\theta_A^{i,j}, \theta_B^{i,j}$, $i, j \in \{1, 2\}$, the inverse demand functions of the Cournot model corresponding with the extended Bertrand model are given by*

$$\begin{aligned} p_A^{i,j}(\rho_t) &= Y_A(\rho_t) - g_1(\rho_t) \frac{x_A \rho_t}{\theta_A^{i,j}} - g_2(\rho_t) \frac{x_B(\rho_t)}{\theta_B^{i,j}}, \\ p_B^{i,j}(\rho_t) &= Y_B(\rho_t) - g_3(\rho_t) \frac{x_B(\rho_t)}{\theta_B^{i,j}} - g_4(\rho_t) \frac{x_A(\rho_t)}{\theta_A^{i,j}}. \end{aligned}$$

Profits of the firms for the second phase given $i, j \in \{1, 2\}$ are

$$\kappa_A^{i,j}(\rho_t, x_A) = x_A \left(p_A^{i,j} - c_A\right) - c_{A,0}, \text{ and } \kappa_B^{i,j}(\rho_t, x_B) = x_B \left(p_B^{i,j} - c_B\right) - c_{B,0}.$$

Example 1 Let the following parameters, functions and matrix be given

$$\begin{aligned} 1 &= \theta_A^{1,1} = \theta_B^{1,1},\ \theta_A^{1,2} = \theta_B^{2,1} = \frac{7}{8},\ \theta_A^{2,1} = \theta_B^{1,2} = \frac{5}{8},\ \theta_A^{2,2} = \theta_B^{2,2} = \frac{1}{2}, \\ q(\rho) &= \frac{\rho_{1,1} + \rho_{1,2}}{2\rho_{1,1} + \rho_{1,2} + \rho_{2,1}},\ D_A = D_B = 100, \\ z_1(\rho) &= 24 - 6\left(\rho_{1,1} + \rho_{1,2}\right),\ z_2(\rho) = 8 - 4 \cdot \left(\rho_{1,1} + \rho_{2,1}\right), \\ z_3(\rho) &= 24 - 6\left(\rho_{1,1} + \rho_{2,1}\right),\ z_4(\rho) = 8 - 4 \cdot \left(\rho_{1,1} + \rho_{1,2}\right), \\ D(\rho) &= 100\left(3\rho_{1,1} + \rho_{1,2} + \rho_{2,1}\right),\ \rho = \begin{bmatrix} \frac{1}{2} & \frac{1}{8} \\ 0 & \frac{3}{8} \end{bmatrix}. \end{aligned}$$

Then, for $\rho_t = \rho$, we have

$$\begin{aligned} D(\rho) &= 162.5,\ q(\rho) = \frac{5}{9}, \\ z_1(\rho) &= 20.25,\ z_2(\rho) = 6,\ z_3(\rho) = 21,\ z_4(\rho) = 5.5, \\ MP_A(\rho) &= 190.28 - 20.25 p_A + 6 p_B,\ MP_B(\rho) = 172.22 - 21 p_B + 5.5 p_A \end{aligned}$$

By the proof of Lemma 1, the inverse demand functions for the Cournot system are determined by $p^{i,j} = Z^{-1}(\rho) \cdot \widetilde{D}(\rho_t) - Z^{-1}(\rho) \left[\Theta^{i,j}\right]^{-1} \cdot x$. Then making all items concrete for $\rho_t = \rho$, we obtain for given $i, j \in \{1, 2\}$

$$\widetilde{D}(\rho) = \begin{bmatrix} 190.28 \\ 172.22 \end{bmatrix},\ \Theta^{i,j} = \begin{bmatrix} \theta_A^{i,j} & 0 \\ 0 & \theta_B^{i,j} \end{bmatrix},\ Z^{-1}(\rho) = \tfrac{4}{1569} \begin{bmatrix} 21 & 6 \\ 5.5 & 20.25 \end{bmatrix}.$$

This leads to the following inverse demand functions given $i, j \in \{1, 2\}$

$$\begin{aligned} p_A^{i,j}(\rho) = 12.821 - \tfrac{84}{1569} \tfrac{x_A(\rho)}{\theta_A^{i,j}} - \tfrac{24}{1569} \tfrac{x_B(\rho)}{\theta_B^{i,j}}, \\ p_B^{i,j}(\rho) = 11.559 - \tfrac{81}{1569} \tfrac{x_B(\rho)}{\theta_B^{i,j}} - \tfrac{22}{1569} \tfrac{x_A(\rho)}{\theta_A^{i,j}}. \end{aligned}$$

4 Stage game rationality (known realizations i, j)

In order to reach their objectives, the firms set quantities the Cournot setting, and prices in the Bertrand setting. For each, we distinguish three further variants, one is that the firms maximize their profits simultaneously and independently, the second is that they take decisions sequentially, in the third they jointly maximize total profits.

4.1 Bertrand: simultaneous, independent decisions

In this framework the maximization problem for Firm k, $k = A, B$ is

$$\max_{p_k} \kappa_k^{i,j}(\rho_t, p_k) = \max_{p_k} \theta_k^{i,j} MP_A(\rho_t)(p_k - c_k) - c_{k,0},\ k = A, B.$$

From the first-order conditions of the maximization problems above, we obtain the following useful functions.[3]

Definition 2 *In the Bertrand stage game the best reply functions are*

$$\begin{aligned} p_A(p_B, \rho_t) &= \frac{D_A(\rho_t) + z_2(\rho_t) p_B + z_1(\rho_t) c_A}{2 z_1(\rho_t)}, \\ p_B(p_A, \rho_t) &= \frac{D_B(\rho_t) + z_4(\rho_t) p_A + z_3(\rho_t) c_B}{2 z_3(\rho_t)}. \end{aligned}$$

[3] For notational convenience and efficiency we omit ρ_t from the notations in the textual parts. **Only** for formal parts we reinsert ρ_t.

So, if, for instance, Firm B sets a certain price p_B, then it is in Firm A's best interest to set price at $p_A(p_B, \rho_t)$.

Definition 3 *A Bertrand equilibrium is a pair of prices* $\left(p_A^B(\rho_t), p_B^A(\rho_t)\right)$ *satisfying*

$$\left(p_A^B(\rho_t), p_B^A(\rho_t)\right) = \left(p_A(p_B^A(\rho_t), \rho_t), p_B(p_A^B(\rho_t), \rho_t)\right).$$

So, both strategic variables are mutual best replies. The following reformulates this equilibrium as a function of the primitive functions of the model.

Lemma 4 *The Bertrand equilibrium* $\left(p_A^B(\rho_t), p_B^B(\rho_t)\right)$ *is given by*

$$\begin{aligned} p_A^B(\rho_t) &= \frac{2z_3(\rho_t)\left[D_A(\rho_t) + z_1(\rho_t)c_A\right] + z_2(\rho_t)\left[D_B(\rho_t) + z_3(\rho_t)c_B\right]}{4z_1(\rho_t)\, z_3(\rho_t) - z_2(\rho_t)\, z_4(\rho_t)}, \\ p_B^B(\rho_t) &= \frac{2z_1(\rho_t)\left[D_B(\rho_t) + z_3(\rho_t)c_B\right] + z_4(\rho_t)\left[D_A(\rho_t) + z_1(\rho_t)c_A\right]}{4z_1(\rho_t)\, z_3(\rho_t) - z_2(\rho_t)\, z_4(\rho_t)}. \end{aligned}$$

4.2 Stackelberg Bertrand: sequential decisions

In this framework, one firm announces its price first, and the other reacts in an optimal fashion. Assume without loss of generality that Firm A announces its price first, i.e., A becomes the leader, B is the follower. Knowing p_A, Firm B sets it price according to the best reply function. Anticipating this, Firm A's maximization problem becomes

$$\max_{p_A} \kappa_A^{i,j}(\rho_t, p_A)|_{p_B = p_B(p_A)}.$$

Let $p_A^*(\rho_t) = \operatorname{argmax}_{p_A} \theta_A^{i,j}\left(D_A(\rho_t) - z_1(\rho_t)p_A + z_2(\rho_t)p_B(p_A, \rho_t)\right)(p_A - c_A)$. The next concept formulates the solution to both maximization problems.

Definition 5 *A Stackelberg Bertrand equilibrium with Firm A as the leader and Firm B as the follower is a pair of prices* $\left(p_A^{LB}(\rho_t), p_B^{FB}(\rho_t)\right)$ *satisfying*

$$\left(p_A^{LB}(\rho_t), p_B^{FB}(\rho_t)\right) = \left(p_A^*(\rho_t), p_B(p_A^{LB}, \rho_t), \rho_t\right).$$

This in turn leads to the following statement.

Lemma 6 *The Stackelberg Bertrand equilibrium* $\left(p_A^{LB}(\rho_t), p_B^{FB}(\rho_t)\right)$ *with Firm A as the leader and Firm B as the follower is given by*

$$\begin{aligned} p_A^{LB}(\rho_t) &= \frac{1}{2}c_A + \frac{2z_3(\rho_t)\, D_A(\rho_t) + z_2(\rho_t)\left[D_B(\rho_t) + z_3(\rho_t)\, c_B\right]}{4z_1(\rho_t)\, z_3(\rho_t) - 2z_2(\rho_t)\, z_4(\rho_t)}, \\ p_B^{FB}(\rho_t) &= \frac{D_B(\rho_t) + z_3(\rho_t)\, c_B + \frac{1}{2}z_4(\rho_t)\, c_A}{2z_3(\rho_t)} + \\ &\quad \frac{z_4(\rho_t)}{2z_3(\rho_t)} \frac{2z_3(\rho_t)\, D_A(\rho_t) + z_2(\rho_t)\left[D_B(\rho_t) + z_3(\rho_t)\, c_B\right]}{4z_1(\rho_t)\, z_3(\rho_t) - 2z_2(\rho_t)\, z_4(\rho_t)}. \end{aligned}$$

4.3 Collusion Bertrand: joint profit maximization

In the Collusion Bertrand framework the maximization problem becomes

$$\max_{p_A, p_B} [\kappa_A^{i,j}(\rho_t, p_A) + \kappa_B^{i,j}(\rho_t, p_B)].$$

The following prices solve this maximization problem.

Lemma 7 *The Collusion Bertrand outcome* $\left(p_A^{CB}(\rho_t), p_B^{CB}(\rho_t)\right)$ *is*

$$\begin{aligned}
p_A^{CB}(\rho_t) &= \frac{2\theta_B^{i,j} z_3(\rho_t)\left(\theta_A^{i,j}\left(D_A(\rho_t) + z_1(\rho_t) c_A\right) - \theta_B^{i,j} z_4(\rho_t) c_B\right)}{4\theta_A^{i,j}\theta_B^{i,j} z_1(\rho_t) z_3(\rho_t) - \left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)^2} + \\
&\quad \frac{\left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)\left(\theta_B^{i,j}\left(D_B(\rho_t) + z_3(\rho_t) c_B\right)\right)}{4\theta_A^{i,j}\theta_B^{i,j} z_1(\rho_t) z_3(\rho_t) - \left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)^2} \\
&\quad - \frac{\left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)\theta_A^{i,j} z_2(\rho_t) c_A}{4\theta_A^{i,j}\theta_B^{i,j} z_1(\rho_t) z_3(\rho_t) - \left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)^2}, \\
p_B^{CB}(\rho_t) &= \frac{2\theta_A^{i,j} z_1(\rho_t)\left(\theta_B^{i,j}\left(D_B(\rho_t) + z_3(\rho_t) c_B\right) - \theta_A^{i,j} z_2(\rho_t) c_A\right)}{4\theta_A^{i,j}\theta_B^{i,j} z_1(\rho_t) z_3(\rho_t) - \left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)^2} + \\
&\quad \frac{\left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)\left(\theta_A^{i,j}\left(D_A(\rho_t) + z_1(\rho_t) c_A\right)\right)}{4\theta_A^{i,j}\theta_B^{i,j} z_1(\rho_t) z_3(\rho_t) - \left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)^2} \\
&\quad - \frac{\left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)\theta_B^{i,j} z_4(\rho_t) c_B}{4\theta_A^{i,j}\theta_B^{i,j} z_1(\rho_t) z_3(\rho_t) - \left(\theta_A^{i,j} z_2(\rho_t) + \theta_B^{i,j} z_4(\rho_t)\right)^2}.
\end{aligned}$$

4.4 Cournot: simultaneous, independent decisions

In the Cournot framework the maximization problems for the two firms are

$$\begin{aligned}
\max\nolimits_{p_A} \kappa_A^{i,j}(\rho_t, x_A) &= \max\nolimits_{x_A} x_A \left(YC_A - g_1 \tfrac{x_A}{\theta_A^{i,j}} - g_2 \tfrac{x_B}{\theta_B^{i,j}}\right) - c_{A,0}, \\
\max\nolimits_{pB} \kappa_B^{i,j}(\rho_t, x_B) &= \max\nolimits_{x_B} x_B \left(YC_B - g_3 \tfrac{x_B}{\theta_B^{i,j}} - g_4 \tfrac{x_A}{\theta_A^{i,j}}\right) - c_{B,0},
\end{aligned}$$

where for notational convenience and enhanced efficiency, we wrote YC_k instead of $Y_k - c_k$ for $k = A, B$. Rearranging the first-order conditions of these maximization problems induces the following pair of functions.

Definition 8 *In the Cournot stage game the best reply functions are*

$$
\begin{aligned}
x_A(x_B, \rho_t) &= \frac{\theta_A^{i,j}}{\theta_B^{i,j}} \frac{\theta_B^{i,j} YC_A(\rho_t) - g_2(\rho_t) x_B}{2 g_1(\rho_t)}, \\
x_B(x_A, \rho_t) &= \frac{\theta_B^{i,j}}{\theta_A^{i,j}} \frac{\theta_A^{i,j} YC_B(\rho_t) - g_4(\rho_t) x_A}{2 g_3(\rho_t)}.
\end{aligned}
$$

So, if, for instance, Firm B sets quantity x_B, then it is in Firm A's best interest to set quantity $x_A(x_B, \rho_t)$.

Definition 9 *A Cournot equilibrium is a pair of quantities* $\left(x_A^C(\rho_t), x_B^C(\rho_t)\right)$ *satisfying*

$$
\left(x_A^C(\rho_t), x_B^C(\rho_t)\right) = \left(x_A(x_B^C(\rho_t), \rho_t), x_B(x_A^C(\rho_t), \rho_t)\right).
$$

So, the strategic variables are mutual best replies. If Firm A sets quantity $x_A^C(\rho_t)$ the profit maximizing quantity of Firm B is $x_B^C(\rho_t)$, and vice versa.

Lemma 10 *The Cournot equilibrium* $\left(x_A^C(\rho_t), x_B^C(\rho_t)\right)$ *is given by*

$$
\begin{aligned}
x_A^C(\rho_t) &= \theta_A^{i,j} \frac{2 g_3(\rho_t) YC_A(\rho_t) - g_2(\rho_t) YC_B(\rho_t)}{4 g_1(\rho_t) g_3(\rho_t) - g_2(\rho_t) g_4(\rho_t)}, \\
x_B^C(\rho_t) &= \theta_B^{i,j} \frac{2 g_1(\rho_t) YC_B(\rho_t) - g_4(\rho_t) YC_A(\rho_t)}{4 g_1(\rho_t) g_3(\rho_t) - g_2(\rho_t) g_4(\rho_t)}.
\end{aligned}
$$

4.5 Stackelberg Cournot: sequential decisions

In the Stackelberg-Cournot framework, one firm announces its quantity first, and the other reacts in an optimal fashion. Assume without loss of generality that Firm A announces its quantity first, i.e., Firm A becomes the leader, Firm B is the follower. Knowing the quantity x_A set, Firm B uses its best reply. Anticipating Firm B's move, Firm A's maximization problem becomes

$$
\max_{x_A} \kappa_A^{i,j}(\rho_t, x_A)|_{x_B = x_B(x_A, \rho_t)}.
$$

Let $x_A^*(\rho_t) = \operatorname{argmax}_{x_A} x_A \left(YC_A - g_1 \frac{x_A}{\theta_A^{i,j}} - g_2 \frac{x_B(x_A, \rho_t)}{\theta_B^{i,j}} \right)$. Then, the following concept is a solution to both firms' maximization problems.

Definition 11 *A Stackelberg Cournot equilibrium with Firm A as the leader and Firm B as the follower is a pair* $\left(x_A^{LC}(\rho_t), x_B^{FC}(\rho_t)\right)$ *satisfying*

$$
\left(x_A^{LC}(\rho_t), x_B^{FC}(\rho_t)\right) = \left(x_A^*(\rho_t), x_B(x_A^{LC}(\rho_t))\right) = \left(x_A^*(\rho_t), x_B(x_A^*(\rho_t))\right).
$$

This give rise to the following.

Lemma 12 *The Stackelberg Cournot equilibrium* $\left(x_A^{LC}(\rho_t), x_B^{FC}(\rho_t)\right)$ *with Firm A as the leader and Firm B as the follower is given by*

$$\begin{aligned} x_A^{LC}(\rho_t) &= \theta_A^{i,j}\left(\frac{2g_3(\rho_t)YC_A(\rho_t)-g_2(\rho_t)YC_B(\rho_t)}{4g_1(\rho_t)g_3(\rho_t)-2g_2(\rho_t)g_4(\rho_t)}\right), \\ x_B^{FC}(\rho_t) &= \frac{\theta_B^{i,j}}{2g_3(\rho_t)}\frac{YC_B(\rho_t)(4g_1(\rho_t)g_3(\rho_t)-g_2(\rho_t)g_4(\rho_t))}{4g_1(\rho_t)g_3(\rho_t)-2g_2(\rho_t)g_4(\rho_t)} \\ &\quad -\frac{\theta_B^{i,j}g_4(\rho_t)YC_A(\rho_t)}{4g_1(\rho_t)g_3(\rho_t)-2g_2(\rho_t)g_4(\rho_t)}. \end{aligned}$$

4.6 Collusion Cournot: joint profit maximization

In the Collusion Cournot framework the maximization problem becomes

$$\max_{x_A,x_B}[\kappa_A^{i,j}(\rho_t,x_A)+\kappa_B^{i,j}(\rho_t,x_B)].$$

The solutions lead to the following.

Lemma 13 *The Collusion Cournot outcome* $\left(x_A^{CC}(\rho_t), x_B^{CC}(\rho_t)\right)$ *is given by*

$$\begin{aligned} & x_A^{CC}(\rho_t) \\ = \;& \theta_A^{i,j}\theta_B^{i,j}\frac{2g_3(\rho_t)\theta_A^{i,j}\theta_B^{i,j}YC_A(\rho_t)-\left(\theta_A^{i,j}g_2(\rho_t)+\theta_B^{i,j}g_4(\rho_t)\right)YC_B(\rho_t)}{4g_1(\rho_t)g_3(\rho_t)\theta_A^{i,j}\left(\theta_B^{i,j}\right)^2-\left(\theta_A^{i,j}g_2(\rho_t)+\theta_B^{i,j}g_4(\rho_t)\right)^2}, \\ & x_B^{CC}(\rho_t) \\ = \;& \theta_A^{i,j}\theta_B^{i,j}\frac{2g_1(\rho_t)\theta_A^{i,j}\theta_B^{i,j}YC_B(\rho_t)-\left(\theta_A^{i,j}g_2(\rho_t)+\theta_B^{i,j}g_4(\rho_t)\right)YC_A(\rho_t)}{4g_1(\rho_t)g_3(\rho_t)\left(\theta_A^{i,j}\right)^2\theta_B^{i,j}-\left(\theta_A^{i,j}g_2(\rho_t)+\theta_B^{i,j}g_4(\rho_t)\right)^2}. \end{aligned}$$

5 Strategies and rewards

At stage t, the firms (or players) know the current state and the history of play, i.e., the state visited and actions chosen at stage $u<t$ denoted by $\left(\rho_u, j_u^A, j_u^B, l_u^A, l_u^B\right)$. Here, $j_u^A, j_u^B \in J=\{1,2\}$ and $l_u^A, l_u^B \in \mathbb{R}^+ \cup \{0\}$. A **strategy** prescribes at all stages, for any state and history, a mixed action to be used by a player. The sets of all strategies for A respectively B will be denoted by $\mathcal{X}^A$ respectively $\mathcal{X}^B$, and $\mathcal{X} \equiv \mathcal{X}^A \times \mathcal{X}^B$. The payoff to player k, $k=A,B$, at stage t, is stochastic and depends on the strategy-pair $(\pi,\sigma)\in\mathcal{X}$; the **expected stage payoff** is denoted by $R_t^k(\pi,\sigma)$.

It seems convenient to rewrite the strategy π if Firm A as

$$\pi = \left(\left(\pi_1^{ADV}, \pi_1^{SGV}\right), \left(\pi_2^{ADV}, \pi_2^{SGV}\right), \ldots\right),$$

where the first component relates to the long term advertising strategy and the second to the state game variable (price or quantity). We denote the strategy σ of Firm B similarly. Furthermore, we write

$$\pi^{ADV} = \left(\pi_1^{ADV}, \pi_2^{ADV}, ...\right) \text{ and } \pi^{ADV} = \left(\pi_1^{SGV}, \pi_2^{SGV}, ...\right).$$

The realizations of the former are then $j_u^A, j_u^B \in J = \{1, 2\}$ and $l_u^A, l_u^B \in \mathbb{R}^+ \cup \{0\}$ of the latter, $u = 1, 2, ...$ Firms interact strategically either in a Cournot setting for all stages, or in a Bertrand setting for all stages.

As mentioned in the previous sections, past advertising determines the setting (or state) for each stage game summarized in the matrix ρ_t. For this, we need the realizations of the advertisement strategies themselves, i.e.,

$$\left(\left(j_1^A, j_1^B\right), \left(j_2^A, j_2^B\right), ..., \left(j_t^A, j_t^B\right)\right).$$

The matrix ρ_t, informally introduced in Section 3.1, can be made to stick formally by the following

$$\rho_t = \begin{bmatrix} \frac{\#\{j_u^A=1 \text{ and } j_u^B=1 |\ 1 \leq u \leq t\}}{t} & \frac{\#\{j_u^A=1 \text{ and } j_u^B=2 |\ 1 \leq u \leq t\}}{t} \\ \frac{\#\{j_u^A=2 \text{ and } j_u^B=1 |\ 1 \leq u \leq t\}}{t} & \frac{\#\{j_u^A=2 \text{ and } j_u^B=2 |\ 1 \leq u \leq t\}}{t} \end{bmatrix}.$$

Note that the advertising strategies may be stochastic, but for the matrix above, the realizations are taken, i.e., action pairs materialized already.

The firms receive an infinite stream of stage payoffs during the play, and they are assumed to wish to maximize their average rewards. For a given pair of strategies (π, σ), Firm k's **average reward,** $k = A, B$, is given by $\gamma^k(\pi, \sigma) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^T R_t^k(\pi, \sigma)$; $\gamma(\pi, \sigma) \equiv \left(\gamma^A(\pi, \sigma), \gamma^B(\pi, \sigma)\right)$.

It may be quite hard to determine the **set of feasible (average) rewards** F, directly. It is common in the analysis of stochastic games to limit the scope of strategies on the one hand, and to focus on rewards on the other. First, we focus on rewards from strategies which are pure and jointly convergent with respect to the advertising framework.

A strategy is **pure**, if at *each* stage a **pure action** is chosen, i.e., the action is chosen with probability 1. The set of pure strategies for player k is $\mathcal{P}^k$, and $\mathcal{P} \equiv \mathcal{P}^A \times \mathcal{P}^B$. The strategy pair $(\pi, \sigma) \in \mathcal{X}$ is **jointly convergent** if and only if a matrix $\rho \in \Delta^{2\times 2}$ exists such that for all $\varepsilon > 0$:

$$\limsup_{t \to \infty} \Pr_{\pi, \sigma} \left[\left| \frac{\#\{j_u^A = i \text{ and } j_u^B = j |\ 1 \leq u \leq t\}}{t} - \rho_{i,j} \right| \geq \varepsilon \right] = 0.$$

where $\Pr_{\pi,\sigma}$ denotes the probability under (π, σ), and $i, j \in J$. Let $\mathcal{JC}$ denote the set of jointly-convergent strategy pairs. Under a pair of jointly-convergent strategies, the relative frequency of each action pair $(i, j) \in J^2$ converges with probability 1 to ρ_{ij} in the terminology of Billingsley [1986, p. 274]. Then, we write '$\{\rho_t(\pi, \sigma)\}_{t=1}^\infty$ converges with probability 1 to ρ'.

The **set of jointly-convergent pure-strategy rewards** is given by

$$P^{\mathcal{JC}} \equiv cl\left\{\left(x^1, x^2\right) \in \mathbb{R}^2 | \ \exists_{(\pi,\sigma)\in\mathcal{P}\cap\mathcal{JC}} : \left(\gamma^k\left(\pi,\sigma\right), \gamma^k\left(\pi,\sigma\right)\right) = \left(x^1, x^2\right)\right\},$$

where *cl* S is the closure of the set S. For any pair of rewards in this set, we can find a pair of jointly-convergent pure strategies that yield rewards arbitrarily close to the original pair of rewards.

We call the strategy π **outcome stationary** if $\pi_s^{SGV}\left(\rho_{s-1}, j_s^A, j_s^B\right) = \pi_t^{SGV}\left(\rho_{t-1}, j_t^A, j_t^B\right)$ whenever $\left(\rho_{s-1}, j_s^A, j_s^B\right) = \left(\rho_{t-1}, j_t^A, j_t^B\right)$, and $\pi_t^{SGV}(\cdot)$ is continuous in the first argument for all $t > 1$. For the other firm we use the same definition, mutatis mutandis. The set of outcome-stationary strategies for Firm $k = A, B$ is $\mathcal{OS}^k$, and $\mathcal{OS} = \mathcal{OS}^A \times \mathcal{OS}^B$. For outcome-stationary strategies the following results immediately from Billingsley [1986].

Lemma 14 *Let* $(\pi, \sigma) \in \mathcal{P} \cap \mathcal{JC} \cap \mathcal{OS}$*, i.e.,*

$$(\pi,\sigma) = \left\{\left(\left(j_u^A, l_u^A\left(\rho_{u-1}, j_u^A, j_u^B\right)\right), \left(j_u^B, l_u^B\left(\rho_{u-1}, j_u^A, j_u^B\right)\right)\right)\right\}_{u\in\mathbb{N}},$$

then numbers $l^A(\rho, i, j), l^B(\rho, i, j)$ *exist such that*

$$\lim_{t\to\infty} l_t^k\left(\rho_{t-1}, i, j\right) = l^k\left(\rho, i, j\right), \ k = A, B.$$

The set of **jointly-convergent outcome-stationary pure-strategy rewards** dependent on $\widetilde{\pi}, \widetilde{\sigma}$ is given by

$$\begin{aligned} P^{\mathcal{JC}}\left(\widetilde{\pi}, \widetilde{\sigma}\right) \equiv \ & cl\{\left(x^1, x^2\right) \in \mathbb{R}^2 | \ \exists_{(\pi,\sigma)\in\mathcal{P}\cap\mathcal{JC}\cap\mathcal{OS}} : \left(\gamma^k\left(\pi,\sigma\right), \gamma^k\left(\pi,\sigma\right)\right) \\ & = \left(x^1, x^2\right) \text{ and } \pi^{SGV} = \widetilde{\pi}, \sigma^{SGV} = \widetilde{\sigma}\}. \end{aligned}$$

For any $\left(x^1, x^2\right) \in P^{\mathcal{JC}}\left(\widetilde{\pi}, \widetilde{\sigma}\right)$, we can find a pair of jointly-convergent outcome-stationary pure strategies that yield rewards arbitrarily nearby.

Remark 1 *Note that the strategies* $l_t^k\left(\rho_{t-1}, j_t^A, j_t^B\right), \ k = A, B$ *constituting rational outcomes for the stage games at any point t in time presented in Section 4, depend on* $\left(\rho_{t-1}, j_t^A, j_t^B\right)$*. If combined with jointly-convergent strategies, by continuity* $\lim_{t\to\infty} l_t^k\left(\rho_{t-1}, i, j\right) = l^k\left(\rho_{t-1}, i, j\right), \ k = A, B.$

5.1 Payoffs to average rewards

Let $(\pi, \sigma) \in \mathcal{P} \cap \mathcal{JC} \cap \mathcal{OS}$, i.e.,

$$(\pi,\sigma) = \left\{\left(\left(j_u^A, l_u^A\left(\rho_{u-1}, j_u^A, j_u^B\right)\right), \left(j_u^B, l_u^B\left(\rho_{u-1}, j_u^A, j_u^B\right)\right)\right)\right\}_{u\in\mathbb{N}},$$

then ρ exists such that $\{\rho_t(\pi, \sigma)\}_{t=1}^{\infty}$ converges with probability 1 to ρ. Moreover, by Lemma 14 numbers $l^A(\rho, i, j), l^B(\rho, i, j)$ exist such that

$$\lim\nolimits_{t\to\infty} l_t^k\left(\rho_{t-1}, i, j\right) = l^k\left(\rho, i, j\right), \ k = A, B.$$

Let $\varphi(\rho) = (\varphi_A(\rho), \varphi_B(\rho))$ with

$$\begin{aligned}\varphi_A(\rho) &= \textstyle\sum_{(i,j)\in J^2} \rho_{i,j}\kappa_A^{i,j}\left(\rho, l^A(\rho,i,j)\right) - \left(\rho_{1,1}+\rho_{1,2}\right) c_A^{ADV},\\ \varphi_B(\rho) &= \textstyle\sum_{(i,j)\in J^2} \rho_{i,j}\kappa_B^{i,j}\left(\rho, l^B(\rho,i,j)\right) - \left(\rho_{1,1}+\rho_{2,1}\right) c_B^{ADV},\end{aligned}$$

where $\kappa_A^{i,j}(\cdot)$ is the second phase profit function as given in Section 3.1 for prices and in Section 3.2 for quantities. We may interpret $\varphi_A(\rho), \varphi_B(\rho)$ as the average profits of both firms under this pair of jointly-convergent outcome-stationary pure strategies, as the profits converge because

$$\lim_{t\to\infty} \kappa_k^{i,j}\left(\rho_t, l^k\left(\rho_{t-1}, i, j\right)\right) = \kappa_k^{i,j}\left(\rho, l^k(\rho,i,j)\right),\ k = A, B.$$

So, for fixed (i,j), the associated second phase profits converge as time goes to infinity. To obtain the average rewards we recall that $\rho_{i,j}$ is precisely the proportion of the stages in which during the first phase action pair (i,j) was used. So, (i,j) materializes at proportion $\rho_{i,j}$ of the stages, hence summing up over all possible entries yields the average associated second phase profits. From the latter, the average advertising expenditures are subtracted. Taken together, this yields the average profits of both firms.

The following result is very similar to one in Joosten *et al.* [2003] for FD-games, we omit its proof as it would be a copy of the one given there. Let $CP^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma})$ denote the convex hull of $P^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma})$.

Theorem 15 *The set of all jointly-convergent outcome-stationary pure-strategy rewards dependent on $(\widetilde{\pi},\widetilde{\sigma})$ is given by $P^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma}) = \bigcup_{\rho\in\Delta^{2\times 2}} \varphi(\rho)$. Moreover, each pair of rewards in $CP^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma})$ is feasible.*

Remark 2 *Theorem 15 implies that $P^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma})$ can be visualized elegantly, see e.g., Figures 1, 4 and 5. Furthermore, take $\left(\pi^{SGV},\sigma^{SGV}\right) = (\widetilde{\pi},\widetilde{\sigma})$, then $P^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma}) \subseteq P^{\mathcal{JC}} \subseteq cl\ F$. Figure 1 illustrates that already for a fixed pair $(\widetilde{\pi},\widetilde{\sigma})$ the set $P^{\mathcal{JC}}(\widetilde{\pi},\widetilde{\sigma})$ may contain an infinite number of elements. Let $P^{\mathcal{JC}\cap\mathcal{OS}} \equiv cl\{\left(x^1,x^2\right)\in\mathbb{R}^2 |\ \exists_{(\pi,\sigma)\in\mathcal{P}\cap\mathcal{JC}\cap\mathcal{OS}}\exists_{(\widetilde{\pi},\widetilde{\sigma})} : \left(\gamma^k(\pi,\sigma), \gamma^k(\pi,\sigma)\right) = \left(x^1,x^2\right)$ and $\pi^{SGV} = \widetilde{\pi}, \sigma^{SGV} = \widetilde{\sigma}\}$. Then, $P^{\mathcal{JC}\cap\mathcal{OS}} \subseteq P^{\mathcal{JC}} \subseteq cl\ F$. Furthermore, an infinite number of myopic strategies exist. This cries out for equilibrium refinements according to some e.g., Gintis [2001].*

6 Equilibria and 'refinements'

The strategy pair (π^*,σ^*) is an **equilibrium**, if no player can improve by unilateral deviation, i.e., $\gamma^A(\pi^*,\sigma^*) \geq \gamma^A(\pi,\sigma^*)$, $\gamma^B(\pi^*,\sigma^*) \geq \gamma^B(\pi^*,\sigma)$ for all $\pi\in\mathcal{X}^A, \sigma\in\mathcal{X}^B$. An equilibrium is called **subgame perfect** if for each possible state and possible history (even unreached states and histories) the subsequent play corresponds to an equilibrium, i.e., no player can improve by deviating unilaterally from then on.

In the construction of equilibria for repeated games, 'threats' play an important role. A threat specifies the conditions under which one player will punish the other, as well as the subsequent measures. We call $v = \left(v^A, v^B\right)$ the **threat point**, where $v^A = \min_{\sigma \in \mathcal{X}^B} \max_{\pi \in \mathcal{X}^A} \gamma^A(\pi, \sigma)$, and $v^B = \min_{\pi \in \mathcal{X}^A} \max_{\sigma \in \mathcal{X}^B} \gamma^B(\pi, \sigma)$. So, v^A is the highest amount A can get if B tries to minimize his average payoffs. Under a pair of **individually rational** rewards each player receives at least his threat-point reward.

To present the general idea of the result of Joosten *et al.* [2003] to come, we adopt terms from Hart [1985], Forges [1986] and Aumann & Maschler [1995]. First, there is a 'master plan' to obtain certain outcomes which is followed by each player as long as the other does so too; then there are 'punishments' which come into effect if a deviation from the master plan occurs. The master plan is a sequence of 'intra-play communications' between the players, the purpose of which is to decide by which equilibrium the play is to continue. The outcome of the communication period is determined by a 'jointly controlled lottery', i.e., at each stage of the communication period the players randomize with equal probability on both actions; at the end of the communication period one sequence of pairs of action choices materializes determining the equilibrium to be played.

Detection of deviation from the master plan *after* the communication period is easy as both players use pure actions on the equilibrium path from then on. Deviation *during* the communication period by using an *alternative randomization* on the actions is impossible to detect. However, it can be shown that no alternative unilateral randomization yields a higher reward. So, the outcome of the procedure is an equilibrium. We restate here the major result from Joosten *et al.* [2003] which applies to general games with frequency-dependent stage payoffs as well as to JFD-games.

Theorem 16 *(Joosten, Brenner & Witt [2003]) Each pair of rewards in the convex hull of all individually-rational pure-strategy rewards can be supported by an equilibrium. Moreover, each pair of rewards in the convex hull of all pure-strategy rewards giving each player strictly more than the threat-point reward, can be supported by a subgame-perfect equilibrium.*

In the next two subsections we tackle the subjects of 'sustainable' punishing, and stage-game rationality from Section 4. The first requires that punishing does not yield negative rewards to the punisher. Sustainability of punishment has not been conceived as an equilibrium refinement criterion but it may reduce the set of equilibria and equilibrium rewards, nevertheless.

6.1 Sustainable punishing and sustainable threats

We aim to find sustainable threats, i.e., threats based on sustainable punishments, which give nonnegative rewards to the punisher. A standard Folk

Theorem approach relies on a threat points induced by punishing the opponent in the harshest manner possible for departing from an equilibrium path. In trying to do so, the punisher may incur eternal losses,which should be regarded as not sustainable, as the punisher might go bankrupt. The surviving firm then becomes a monopolist and may set advertising, and prices or quantities at will to maximize average profits. That obviously takes out the sting of punishing. The following may serve as an example.

Example 2 (Example 1 cont'd) The problem of finding a threat point is complex. Here, we 'merely' establish an upper bound for the threat point rewards. To simplify computations significantly we let the punisher pick a fixed price for all punishment stages in the Bertrand set-up. The game with pricing decisions of the punisher fixed is a unichain stochastic game, hence with respect to the long-run strategic variable stationary optimal strategies exist, cf., Hoffman & Karp [1966], Thuijsman [1992]. We also fix the long-term advertising decisions for the punisher, hence the punishee faces a Markov decision problem (cf., Hordijk *et al.* [1983], Blackwell [1962]).

If for instance Firm A were to punish Firm B by never advertising at all and by setting its price equal to variable costs, then Firm A's average profits would be -200. The optimal strategy of the other firm is pure and stationary. If Firm B advertises with a relative frequency of y in the long run, the play induces the following matrix of long run relative frequencies

$$\rho = \begin{bmatrix} 0 & 0 \\ y & 1-y \end{bmatrix}.$$

Recall then that the maximization problem for Firm B simplifies to

$$\max_{y,p_B} \left(\frac{7}{8}y + \frac{1}{2}(1-y)\right)(124 + 100y - (24-6y)p_B)(p_B - 3) - 200y - 200.$$

We visualized the function to be maximized in Figure 5 which is maximized for $y = 1$ (**Claim 1**). Thus, this problem for Firm B simplifies further to

$$\max_{p_B} \tfrac{7}{8}(224 - 18p_B)(p_B - 3) - 400.$$

The profit maximizing prize is therefore $p_B = \frac{278}{36} = 7.722\,2$. So, this means

$$\max_{y,p_B} \Pi_B^{p_A=3,x=0} = -48.785.$$

This punishment strategy of Firm A can keep the opponent's long term average profits at at most -48.785, but we saw already that it is not sustainable.

Now, suppose Firm A sets a price of 7.55 and advertises each stage with a probability of 0.69. Then, this price and probability imply (**Claim 2**)

$$MP_B = 139.562 + \tfrac{y(169y+69)}{y+0.69} - (24-6y)p_B.$$

The maximization problem for Firm B becomes

$$\max_{y,p_B} (0.375y + 0.58625) \left(139.562 + \frac{y(169y+69)}{y+0.69} - (24-6y)p_B\right)(p_B - 3)$$
$$-200y - 200$$

A visual inspection of Firm B's profit manifold (cf., Figure 7) reveals that its optimal advertising effort is $y = 1$ again. This implies (**Claim 3**)

$$MP_B = 280.39 - 18p_B.$$

For the computation of the optimal price for Firm B we take the best response function, which yields $p_B = \frac{280.39+3\cdot 18}{36} = 9.288\,6$. So, the long term average profits for Firm B are at most

$$\max_{y,p_B} \Pi_B^{p_A=7.55,x=0.69} = 284.26.$$

Hence, Firm A has this myopic strategy to keep Firm B's average profits at at most 284.26. Meanwhile, Firm A's own profits are

$$\Pi_A^{p_A=7.55,x=0.69,y=1,p_B=9.288\,6} = 1.3109 > 0.$$

So, Firm A's current combination of strategic variables to punish is sustainable. Then, clearly

$$\begin{aligned} v^B &\leq \min_{x,p_A} \max_{y,p_B} \Pi_B(x,p_A,y,p_B) \leq \max_{y,p_B} \Pi_B^{p_A=3,x=0} = -48.785 \\ &\leq \max_{y,p_B} \Pi_B^{p_A=7.55,x=0.69} = 284.26. \end{aligned}$$

Here, $\min_{x,p_A} \max_{y,p_B} \Pi_B(x,p_A,y,p_B)$ is the reward for Firm B obtained by the Firm A minimizing his opponents maximum reward while both are restricted to myopic strategies. As both firms are symmetric, we have

$$\left(\max_{x,p_A} \Pi_A^{p_B=7.55,y=0.69}, \max_{y,p_B} \Pi_B^{p_A=7.55,x=0.69}\right) \leq (284.26, 284.26).$$

All reward pairs giving each firm at least a reward of 284.26 can be supported by a Nash equilibrium involving sustainable threats (cf., Figure 1). ■

Our approach is not haphazard. First we noted that for most of the relevant combinations for Firm A to punish in Example 2, the optimal advertising effort of Firm B is $y = 1$. In the general case we would inspect the average reward function for Firm B, to find its optimal advertising effort. Under the assumption that the latter is known, the problem for Firm A becomes easier and we now show how to obtain a set of myopic sustainable punishing strategies. Note that $y = 1$ implies

$$MP_A(\rho) = 100\frac{1+2x+2x^2}{1+x} - (24-6x)p_A + 4p_B,$$
$$MP_B(\rho) = 100\frac{2+3x}{1+x} - 18p_B + (8-4x)p_A.$$

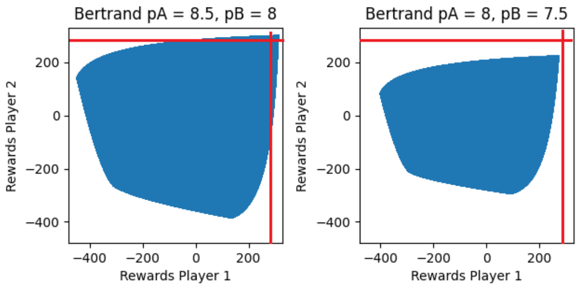

Figure 1: The set $P^{\mathcal{JC}}(\widetilde{\pi}, \widetilde{\sigma})$ with $(\widetilde{\pi}_t, \widetilde{\sigma}_t) = (8.5, 8)$ and $(\widetilde{\pi}_t, \widetilde{\sigma}_t) = (8, 7.5)$ for Examples 1 & 2. Rewards to the North-East of the sustainable threat point $(284.26, 284.26)$ can be supported by an equilibrium involving sustainable threats. On the right however, no such points exist.

The maximization problem for Firm B is therefore reduced to

$$\max_{p_B} \left(\tfrac{7}{8} + \tfrac{1}{8}x\right) MP_B(\rho)(p_B - 3) - 400.$$

The optimal price follows from Firm B's best response function

$$p_B(p_A) = \frac{100\frac{2+3x}{1+x} + 54 + (8-4x)p_A}{36}.$$

Now, we return to the rewards of Firm A. These are given by

$$\left(\tfrac{5}{8} + \tfrac{3}{8}x\right) MP_A(\rho)|_{p_B = p_B(p_A)}(p_B - 3) - 200x - 200.$$

Figure 2 depicts sets of combinations of (x, p_A) giving Firm A average rewards close to zero if $y = 1$, $p_B = p_B(p_A)$.
In Example 2, we **chose** the pair $x = 0.69$ and $p_A = 7.55$ using Figure 2. A **full analysis** of all combinations in the set is pending, but we found

$$\max_{y, p_B} \Pi_B^{p_A=5.15, x=0.96} = 247.17, \ \Pi_A^{p_A=5.15, x=0.96, y=1, p_B=p_B(5.15)} = 0.057473.$$

Hence, **all reward pairs giving each firm at least a reward of** 247.17 **can be supported by an equilibrium involving sustainable threats.** We now turn to an illustration for the Cournot framework.

Example 3 For the Cournot model we have the following. Suppose that Firm A tries to punish his opponent by never advertising and setting $x_A = 50$. Then, we establish (**Claim 4**)

$$p^{2,1} = \begin{bmatrix} p_A^{2,1} \\ p_B^{2,1} \end{bmatrix} = \frac{1}{256-56y} \begin{bmatrix} 1600 - 100y - 200y^2 \\ 1600 + 1200y \end{bmatrix} - \frac{1}{256-56y} \begin{bmatrix} \frac{8}{5}(12-3y)x_A + \frac{8}{7}(4-2y)x_B \\ \frac{96}{7}x_B + \frac{32}{5}x_A \end{bmatrix},$$

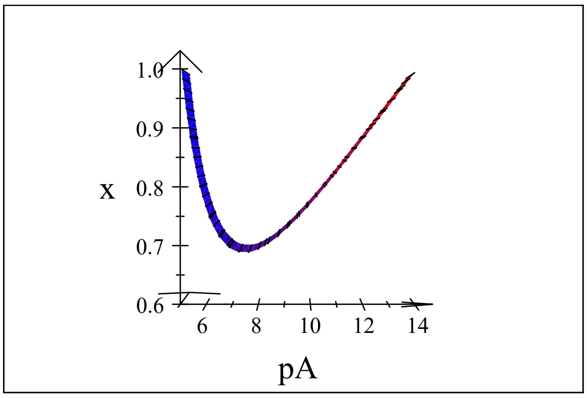

Figure 2: Combinations (x, p_A) for which Firm A has a reward close to zero, resulting from Firm B trying to maximize its own rewards.

$$p^{2,2} = \begin{bmatrix} p_A^{2,2} \\ p_B^{2,2} \end{bmatrix} = \frac{1}{256-56y}\left(\begin{bmatrix} 1600-100y-200y^2 \\ 1600+1200y \end{bmatrix}\right) - \frac{1}{256-56y}\begin{bmatrix} (24-6y)\,x_A + (8-4y)\,x_B \\ 24x_B + 8x_A \end{bmatrix}.$$

Recall that the second phase profit functions are given by

$$\kappa_A^{i,j}(\rho_t, x_A) = x_A\left(p_A^{i,j} - c_A\right) - c_{A,0},\ \kappa_B^{i,j}(\rho_t, x_B) = x_B\left(p_B^{i,j} - c_B\right) - c_{B,0}.$$

Therefore,

$$\begin{aligned}
\kappa_A^{2,1}(\rho_t, x_A) &= x_A\left(\tfrac{1600-100y-200y^2-\frac{8}{5}(12-3y)x_A-\frac{8}{7}(4-2y)x_B}{256-56y} - 3\right) - 200,\\
\kappa_B^{2,1}(\rho_t, x_B) &= x_B\left(\tfrac{1600+1200y-\frac{96}{7}x_B-\frac{32}{5}x_A}{256-56y} - 3\right) - 200y - 200,\\
\kappa_A^{2,2}(\rho_t, x_A) &= x_A\left(\tfrac{1600-100y-200y^2-(24-6y)x_A-(8-4y)x_B}{256-56y} - 3\right) - 200,\\
\kappa_B^{2,2}(\rho_t, x_B) &= x_B\left(\tfrac{1600+1200y-24x_B-8x_A}{256-56y} - 3\right) - 200.
\end{aligned}$$

Maximizing Firm B's profits for $(i,j) = (2,1)$ and $(i,j) = (2,2)$ yields

$$\begin{aligned}
x_B(x_A)|_{(i,j)=(2,1)} &= \tfrac{7}{192}\left(832 + 1368y - \tfrac{32}{5}x_A\right),\\
x_B(x_A)|_{(i,j)=(2,2)} &= \tfrac{832+1368y-8x_A}{48}.
\end{aligned}$$

Figure 8 in the Appendix depicts Firm B's average profits assuming that in the two situations Firm B uses a best response to Firm A's quantity set $(x_A = 50)$, and the latter not advertising at all which is given by

$$y \cdot \kappa_B^{2,1}(\rho, x_B(x_A)|_{(i,j)=(2,1)}) + (1-y) \cdot \kappa_B^{2,2}(\rho, x_B(x_A)|_{(i,j)=(2,2)}).$$

We obtain the insight that Firm B maximizes its average profits using $y = 1$. (**Claim 5**) So, Firm B's best reply simplifies to

$$x_B = \tfrac{7}{192}\left(832 + 1368 - \tfrac{32}{5}x_A\right)|_{x_A=50} = \tfrac{1645}{24} = 68.542.$$

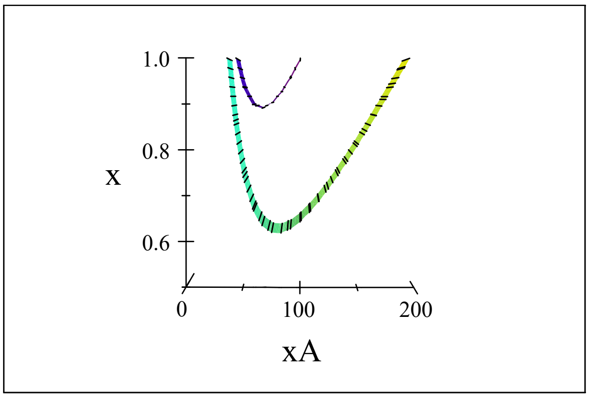

Figure 3: Combinations (x, x_A) for which Firm A has near zero profits if $y = 1$, $x_B = x_B(x_A)$. One set belongs to $(i, j) = (1, 1)$, the other to $(i, j) = (2, 1)$. The latter is located above the former.

and therefore

$$\left[x_B\left(\frac{1600+1200-\frac{96}{7}x_B-\frac{32}{5}x_A}{200}-3\right)\right]_{x_A=50,x_B=68.542}-400=-77.854.$$

So, Firm A possesses a myopic strategy to keep Firm B's rewards at at most -77.854, however we find

$$\kappa_A^{2,1}(\rho_t, x_A)|_{x_A=50,y=1,x_B=68.542} = \left[x_A\left(\frac{1600-100y-200y^2-\frac{8}{5}(12-3y)x_A-\frac{8}{7}(4-2y)x_B}{256-56y}-3\right)\right]_{x_A=50,y=1,x_B=68.542}-200 = -244.17.$$

Hence, in order to inflict a long term average **loss** of 77.854 on Firm B, Firm A must suffer a loss of 244.17 itself.

Having found again that Firm B maximizes profits by using $y = 1$ and having the same observation for multiple cases examined, we set $y^* = y = 1$. This leads to best response functions

$$x_B^{1,1}(x_A) = \frac{1}{6}\frac{2200+3138x-362x^2-400x^3}{(4-x)(x+1)} - \frac{1}{6}\frac{(4-2x)x_A}{4-x}$$
$$x_B^{2,1}(x_A) = \frac{7}{48}\frac{2200+3138x-362x^2-400x^3}{(4-x)(x+1)} - \frac{7}{48}\frac{(4-2x)x_A}{4-x}$$

We fill these functions into the profit functions of Firm A and evaluate them. We search combinations (x, x_A) giving near zero profits to Firm A, provided that Firm B advertises forever and uses best responses in each of the two situations. Figure 3 visualizes such sets of combinations (x, x_A) .
Using Figure 3 to obtain candidates for sustainable punishing we establish

$$\kappa_A^{1,1}(\rho_t, x_A)\big|_{y^*,x_B^{1,1}(x_A),x_A=37.75,x=0.95} = 248.87$$
$$\kappa_A^{2,1}(\rho_t, x_A)\big|_{y^*,x_B^{2,1}(x_A),x_A=48.25,x=0.95} = 218.5.$$

Thus, $\max_{y,x_B^{1,1},x_B^{2,1}} \Pi_B^{x_A^{1,1}=48.25,x_A^{2,1}=37.75,x=0.95} = 247.35$. Clearly,

$$
\begin{aligned}
v^B &\leq \min_{x,x_A}\max_{y,x_B} \Pi_B\left(x,x_A,y,x_B\right) \leq \max_{y,x_B} \Pi^{x_A^{2,1}=50=x_A^{2,2},x=0} = -78.438 \\
&\leq \max_{y,x_B^{1,1},x_B^{2,1}} \Pi_B^{x_A^{1,1}=48.25,x_A^{2,1}=37.75,x=0.95} = 247.35.
\end{aligned}
$$

Here, $\min_{x,x_A}\max_{y,x_B} \Pi_B\left(x,x_A,y,x_B\right)$ is the reward for Firm B obtained by the Firm A minimizing his opponents maximum reward while both are restricted to myopic strategies. The firms are symmetric, therefore

$$\max_{x,x_A^{1,1},x_A^{1,2}} \Pi_A^{x_B^{1,1}=48.25,x_B^{1,2}=37.75,,y=0.95} = 247.35.$$

6.2 Stage-game rationality: an equilibrium refinement

In principle, equilibria may exist even for rather arbitrarily chosen prices, see Figure 1. For stage-game rationality, we impose the stage game optimal strategies derived in Section 4, that are outcome stationary if taken for all stages and all possible histories. Figure 4 visualizes the jointly-convergent outcome-stationary pure-strategy rewards for the Bertrand and the Cournot equilibrium (Lemma 10). Comparing the areas of equilibrium rewards under strategies involving simple punishment strategies or sustainable ones, we note significant reductions in size.

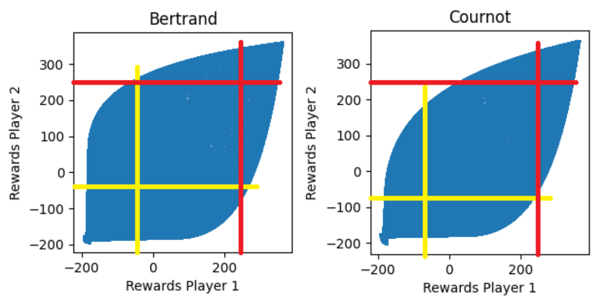

Figure 4: The set $P^{\mathcal{JC}}\left(\widetilde{\pi},\widetilde{\sigma}\right)$ with $\left(\widetilde{\pi}_t,\widetilde{\sigma}_t\right) = \left(p_A^B(\rho_t), p_B^B(\rho_t)\right)$ and $\left(\widetilde{\pi}_t,\widetilde{\sigma}_t\right) = \left(x_A^C(\rho_t), x_B^C(\rho_t)\right)$ Each reward North East of $(-48.785, -48.785)$ resp. $(-77.854, -77.854)$ can be supported by an equilibrium using threats. Each reward North East of $(247.17, 247.17)$ resp. $(247.35, 247.35)$ can be supported by an equilibrium using sustainable threats.

6.3 Applications to the Stackelberg and collusion variants

In Section 4 we treated six variants of short run equilibria in order to guarantee stage game-rationality. We already dealt with the cases in which the firms take simultaneous and independent decisions, Figure 5 visualizes the remaining variants. The sustainable threats remain identical, as they do not rely on any form of stage-game rationality. What changes is exclusively due to stage-game rationality, and it involves the sets of jointly-convergent outcome-stationary pure-strategy rewards for each of the four cases.

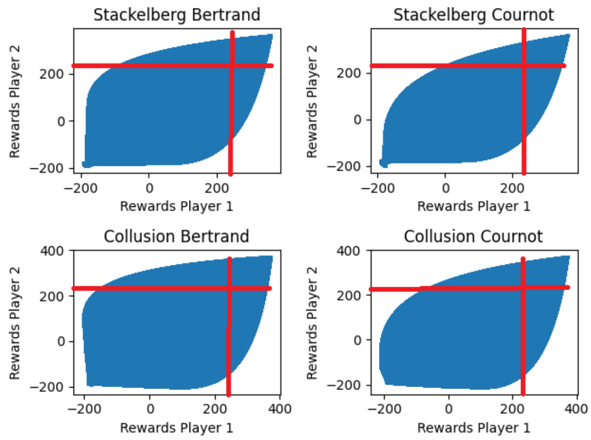

Figure 5: The jointly-convergent outcome-stationary pure-strategy rewards for the four remaining cases with respect to stage game rationality. As before the rewards to the North East of the intersection of the red lines can be supported by an equilibrium involving sustainable threats.

7 Conclusion

We presented a dynamic model of long term advertising, building in ideas from Contest Theory. Long run advertising in a Bertrand setting changes the height and shape of each firm's market potential in a gradual fashion. Contest Theory motivates the changes in height of the market potentials in a Bertrand setting, which are taken proportional to the historic advertisement efforts of the firms. As in Joosten [2016b], we split up each stage game into two decision phases. In the first one, the long term advertisement strategies materialize determining the strategic environment for the second phase in detail. If both firms advertised at that stage, market potentials turn into immediate demand functions. Otherwise, a fraction of the market potential 'evaporates'. So, advertising has both gradual long term effects through the height and shape of the market potential functions, but also sharp immediate effects on demand, in line with e.g., Friedman [1983].

A non-terminating series of one-shot stage games is played. The stage games are mutually independent, i.e., the actions chosen in the second decision phase do not influence the transition probabilities to another state, nor the future payoffs. Any such transition is exclusively determined by the realizations of the long term advertising strategies. The strategic environment for every single stage game depends on the advertising history, though.

We derive a Cournot model from a Bertrand model extended with the Contest Theory module under rather weak assumptions, and the relationship between them is one-to-one. The Bertrand model seems the most natural one to introduce Contest Theory, so our motivations focus on this model.

We use standard solutions from microeconomics for the stage games. We treat the timing of decisions, in one variant we assume that the firms take decisions in the stage game simultaneously, in the other sequentially. To the sequential-decisions variant the term Stackelberg is added. So, we formulate for each stage game the simple Cournot or Bertrand equilibria for the simultaneous decisions case, and the Stackelberg-Cournot and Stackelberg-Bertrand equilibria for the sequential decisions case. In all four cases the firms try to maximize their individual stage payoffs.

An interesting alternative[4] is collusion during the stage games, i.e., the firms maximize their joint stage-game payoffs. Since here too, a distinction along the strategic variable is required, we have a solution in the collusion-Cournot case and one in the collusion-Bertrand case.

A strategy here is a plan for the entire duration of the play, prescribing for each stage whether to advertise or not, and prescribing which level the relevant strategic variable should have in the stage game. We assume that both firms interact in either the Cournot setting for all stages, or the Bertrand setting for all stages. Beyond this, the setup dictates nothing beforehand to the firms. The advertisement part of the strategies determine what the range of possibilities is that may be obtained in a stage game, as well as the transitions between the states, the stage game strategic variables determine only the immediate profits and as such have only effect on the rewards in an aggregate. The concept of a strategy consisting of two substrategies, so to speak, causes some challenges

We analyze the games presented with a method similar to the Folk Theorem used in the study of repeated games. We regard this a major strength, large sets of equilibria are relatively easily found and the variety of mathematical equations allowed is almost limitless as we only need a weak continuity assumption to validate our method of analysis. A drawback is that there may exist a multitude of equilibria and equilibrium rewards. Additional criteria, such as Pareto-efficiency or subgame perfectness, might reduce the number of equilibria and associated rewards. Each example treated here has a unique Pareto-efficient outcome which can be supported by an equilibrium involving threats. However, subgame perfectness in the context of infinitely repeated or stochastic games hardly ever provides a significant reduction of

[4]Scherer [1980], Brander & Harris [1983], Fershtman & Muller [1986] revealed a tendency to collude in fast variables, so treating collusion seems appropriate.

the set of equilibrium rewards (cf., e.g., Van Damme [1991]).

Equilibrium selection was not a goal of our explorative investigation. However, the six well-established solutions from microeconomics for the stage games, combined with advertisement strategies to form jointly-convergent pure-strategies constitute restrictions on the scope of possibilities as illustrated. Furthermore, if the latter are used for the construction of equilibria with sustainable threats, the resulting equilibria form a subset of the set of equilibria. So, ex-post we seem to have introduced equilibrium refinements.

This paper's direct predecessors, Joosten [2009, 2016b], are *entirely* in the class of *ETP-ESP* games, as the stage game strategy parts was assumed solved already by a particular solution from microeconomics. Here however, we do not restrict the stage game strategies **a priori**. Hence, the games presented are extensions of this earlier work, because the strategy of each player consists of two fully developed independent parts, the advertisement part and the stage game part. In working towards an economically sound solution though, we restrict the range of stage-game strategies to proceed analysis along the lines of earlier *ETP-ESP* games.

The field of *ETP-ESP* games is expanding rapidly. Tools establishing jointly convergent pure-strategy rewards for environmental problems (e.g., Joosten [2019]) and Small Fish Wars (Joosten [2007, 2016a], Joosten & Meijboom [2018], Joosten & Samuel [2017, 2020] and Joosten & Harmelink [2023]), exist, yet there is no general tool to find threat points in FD-games.

8 Appendix

Proof of Lemma 1. We rewrite the Bertrand system given $i, j = 1, 2$ as

$$x(\rho_t) = \Theta^{i,j}\left(\widetilde{D}\left(\rho_t\right) - Z\left(\rho_t\right)p\right) \text{ with}$$
$$x(\rho_t) = \begin{bmatrix} x_A^{i,j}(\rho_t) \\ x_B^{i,j}(\rho_t) \end{bmatrix},\ Z\left(\rho_t\right) = \begin{bmatrix} z_1(\rho_t) & -z_2(\rho_t) \\ -z_4(\rho_t) & z_3(\rho_t) \end{bmatrix},$$
$$\Theta^{i,j} = \begin{bmatrix} \theta_A^{i,j} & 0 \\ 0, & \theta_B^{i,j} \end{bmatrix},\ \widetilde{D}\left(\rho_t\right) = \begin{bmatrix} D_A \\ D_B \end{bmatrix} = \begin{bmatrix} D_{A,0} + q(\rho_t)D(\rho_t) \\ D_{B,0} + \left(1 - q(\rho_t)\right)D(\rho_t) \end{bmatrix}$$

By our assumptions $Z^{-1}\left(\rho_t\right)$ exists, inducing the following implications

$$\begin{gathered} x(\rho_t) = \Theta^{i,j}\left(\widetilde{D}\left(\rho_t\right) - Z\left(\rho_t\right)p\right) \Longrightarrow \\ \left[\Theta^{i,j}\right]^{-1} x(\rho_t) = \widetilde{D}\left(\rho_t\right) - Z\left(\rho_t\right)p \Longrightarrow \\ Z^{-1}\left(\rho_t\right)\left[\Theta^{i,j}\right]^{-1} x(\rho_t) = Z^{-1}\left(\rho_t\right)\widetilde{D}\left(\rho_t\right) - p \Longrightarrow \\ p = Z^{-1}\left(\rho_t\right)\widetilde{D}\left(\rho_t\right) - Z^{-1}\left(\rho_t\right)\left[\Theta^{i,j}\right]^{-1} x(\rho_t). \end{gathered}$$

The inverse demand functions presented are obtained by setting

$$G(\rho_t) = Z^{-1}\left(\rho_t\right),\quad \begin{bmatrix} Y_A\left(\rho_t\right) \\ Y_B\left(\rho_t\right) \end{bmatrix} = Z^{-1}\left(\rho_t\right)\widetilde{D}\left(\rho_t\right). \qquad \blacksquare$$

Proof of Lemma 4. We solve first-order conditions for maximization, i.e.,

$$0 = \theta_A^{i,j}\left(D_A - 2z_1 p_A + z_2 p_B + z_1 c_A\right),$$
$$0 = \theta_B^{i,j}\left(D_B - 2z_3 p_B + z_4 p_A + z_3 c_B\right).$$

Deleting thetas, multiplying by $2z_3$ respectively z_2, adding them up, yields

$$0 = 2z_3 D_A - 4z_1 z_3 p_A + 2z_1 z_3 c_A + z_2 D_B + z_2 z_4 p_A + z_2 z_3 c_B \Longrightarrow$$
$$p_A = \frac{2z_3(D_A + z_1 c_A) + z_2(D_B + z_3 c_B)}{4z_1 z_3 - z_2 z_4}.$$

Mutatis mutandis, we obtain $p_B = \frac{2z_1(D_B + z_3 c_B) + z_4(D_A + z_1 c_B)}{4z_1 z_3 - z_2 z_4}$. ■

Proof of Lemma 6. The first-order condition for Firm A's maximization problem is equivalent to the following (we dispense with $\theta_A^{i,j}$)

$$0 = \left(-z_1 + \left[\frac{z_2 z_4}{2z_3}\right]\right)(p_A - c_A) + D_A - z_1 p_A + z_2\left[\frac{D_B + z_4 p_A + z_3 c_B}{2z_3}\right]$$
$$= -\left(\frac{2z_1 z_3 - z_2 z_4}{2z_3}\right)(2p_A - c_A) + \left[\frac{2z_3 D_A + z_2[D_B + z_3 c_B]}{2z_3}\right].$$

The latter implies

$$p_A = \frac{1}{2}c_A + \left(\frac{2z_3}{4z_1 z_3 - 2z_2 z_4}\right)\left[\frac{2z_3 D_A + z_2[D_B + z_3 c_B]}{2z_3}\right].$$

Reformulating the solution as p_A^L we obtain

$$p_A^L = \frac{1}{2}c_A + \left[\frac{2z_3 D_A + z_2[D_B + z_3 c_B]}{4z_1 z_3 - 2z_2 z_4}\right].$$

The follower uses his best-reply against the leader's price

$$p_B^F = p_B\left(p_A^L\right) = \frac{D_B + z_4 p_A^L + z_3 c_B}{2z_3}$$
$$= \frac{D_B + z_3 c_B}{2z_3} + \frac{z_4}{2z_3}\left[\frac{1}{2}c_A + \left[\frac{2z_3 D_A + z_2[D_B + z_3 c_B]}{4z_1 z_3 - 2z_2 z_4}\right]\right]$$
$$= \frac{D_B + z_3 c_B + \frac{1}{2}z_4 c_A}{2z_3} + \frac{z_4}{2z_3}\left[\frac{2z_3 D_A + z_2[D_B + z_3 c_B]}{4z_1 z_3 - 2z_2 z_4}\right]. \quad ■$$

Proof of Lemma 7. The first-order conditions for the maximization problem are

$$0 = \theta_A^{i,j}\left(D_A - 2z_1 p_A + z_2 p_B + z_1 c_A\right) + \theta_B^{i,j} z_4\left(p_B - c_B\right),$$
$$0 = \theta_B^{i,j}\left(D_B - 2z_3 p_B + z_4 p_A + z_3 c_B\right) + \theta_A^{i,j} z_2\left(p_A - c_A\right).$$

Multiplying the first equation by $2\theta_B^{i,j} z_3$ and the second by $\theta_A^{i,j} z_2 +\ \theta_B^{i,j} z_4$, then adding them up to eliminate p_B yields

$$0 = 2\theta_B^{i,j} z_3\left(\theta_A^{i,j}\left(D_A - 2z_1 p_A + z_1 c_A\right) - \theta_B^{i,j} z_4 c_B\right) +$$
$$\left(\theta_A^{i,j} z_2 + \theta_B^{i,j} z_4\right)\left(\theta_B^{i,j}\left(D_B + z_4 p_A + z_3 c_B\right) + \theta_A^{i,j} z_2\left(p_A - c_A\right)\right).$$

Isolating the terms containing p_A and bringing them to the other side of the equality sign yields

$$\begin{aligned}&\left[4\theta_A^{i,j}\theta_B^{i,j}z_1z_3-\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)\right]p_A\\ &=2\theta_B^{i,j}z_3\left(\theta_A^{i,j}\left(D_A+z_1c_A\right)-\theta_B^{i,j}z_4c_B\right)+\\ &\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)\left(\theta_B^{i,j}\left(D_B+z_3c_B\right)-\theta_A^{i,j}z_2c_A\right).\end{aligned}$$

Then, renaming the solution p_A^{CB} which is mnemonic for collusion in the Bertrand framework, we obtain

$$p_A^{CB}=\frac{2\theta_B^{i,j}z_3\left(\theta_A^{i,j}(D_A+z_1c_A)-\theta_B^{i,j}z_4c_B\right)}{4\theta_A^{i,j}\theta_B^{i,j}z_1z_3-\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)^2}+\frac{\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)\left(\theta_B^{i,j}(D_B+z_3c_B)-\theta_A^{i,j}z_2c_A\right)}{4\theta_A^{i,j}\theta_B^{i,j}z_1z_3-\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)^2}.$$

Hence, mutatis mutandis we establish

$$p_B^{CB}=\frac{2\theta_A^{i,j}z_1\left(\theta_B^{i,j}(D_B+z_3c_B)-\theta_A^{i,j}z_2c_A\right)}{4\theta_A^{i,j}\theta_B^{i,j}z_1z_3-\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)^2}+\frac{\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)\left(\theta_A^{i,j}(D_A+z_1c_A)-\theta_B^{i,j}z_4c_B\right)}{4\theta_A^{i,j}\theta_B^{i,j}z_1z_3-\left(\theta_A^{i,j}z_2+\theta_B^{i,j}z_4\right)^2}.\ \blacksquare$$

Proof of Lemma 10. The first-order conditions for the maximization problems are

$$\begin{aligned}0&=YC_A-2g_1\frac{x_A}{\theta_A^{i,j}}-g_2\frac{x_B}{\theta_B^{i,j}},\\ 0&=YC_B-2g_3\frac{x_B}{\theta_B^{i,j}}-g_4\frac{x_A}{\theta_A^{i,j}}.\end{aligned}$$

Multiplying the first by $2g_3$, the second by $-g_2$, then adding them up to eliminate x_B yields the following implications

$$\begin{aligned}&0=2g_3YC_A-g_2YC_B-\frac{4g_1g_3-g_2g_4}{\theta_A^{i,j}}x_A\implies\\ &x_A=\theta_A^{i,j}\frac{2g_3YC_A-g_2YC_B}{4g_1g_3-g_2g_4}.\end{aligned}$$

Mutatis mutandis, $x_B=\theta_B^{i,j}\frac{2g_1YC_B-g_4YC_A}{4g_1g_3-g_2g_4}$. $\blacksquare$

Proof of Lemma 12. Rewriting Firm A's maximization problem yields

$$\begin{aligned}&\max_{x_A}x_A\left(YC_A-g_1\frac{x_A}{\theta_A^{i,j}}-\frac{g_2}{\theta_B^{i,j}}\frac{\theta_B^{i,j}}{\theta_A^{i,j}}\frac{\theta_A^{i,j}YC_B-g_4x_A}{2g_3}\right)-c_{A,0}\\ &=\max_{x_A}x_A\left(YC_A-g_1\frac{x_A}{\theta_A^{i,j}}-\frac{g_2}{\theta_A^{i,j}}\frac{\theta_A^{i,j}YC_B-g_4x_A}{2g_3}\right)-c_{A,0}.\end{aligned}$$

The first-order condition yields the following implications

$$\begin{aligned}&0=YC_A-2\frac{g_1-\frac{g_2g_4}{2g_3}}{\theta_A^{i,j}}x_A-\frac{g_2}{\theta_A^{i,j}}\frac{\theta_A^{i,j}YC_B}{2g_3}\\ &=\frac{2g_3YC_A-g_2YC_B}{2g_3}-\frac{4g_1g_3-2g_2g_4}{2\theta_A^{i,j}g_3}x_A\implies\\ &x_A=\theta_A^{i,j}\left(\frac{2g_3YC_A-g_2YC_B}{4g_1g_3-2g_2g_4}\right)=x_A^{LC}.\end{aligned}$$

Next, we obtain

$$x_B\left(x_A^{LC}\right)=\frac{\theta_B^{i,j}}{\theta_A^{i,j}}\frac{\theta_A^{i,j}YC_B-g_4\theta_A^{i,j}\left(\frac{2g_3YC_A-g_2YC_B}{4g_1g_3-2g_2g_4}\right)}{2g_3}$$
$$=\frac{\theta_B^{i,j}}{2g_3}\left(YC_B-\frac{2g_3g_4YC_A-g_2g_4YC_B}{4g_1g_3-2g_2g_4}\right)$$
$$=\frac{\theta_B^{i,j}}{2g_3}\left(\frac{YC_B(4g_1g_3-g_2g_4)-2g_3g_4YC_A}{4g_1g_3-2g_2g_4}\right)=x_B^{FC}. \blacksquare$$

Proof of Lemma 13. The first-order conditions are equivalent to

$$0=YC_A-2g_1\frac{x_A}{\theta_A^{i,j}}-\left(\frac{g_2}{\theta_B^{i,j}}+\frac{g_4}{\theta_A^{i,j}}\right)x_B,$$
$$0=YC_B-2g_3\frac{x_B}{\theta_B^{i,j}}-\left(\frac{g_2}{\theta_B^{i,j}}+\frac{g_4}{\theta_A^{i,j}}\right)x_A.$$

Then, multiplying the second equation by $-\left(\frac{g_2}{\theta_B^{i,j}}+\frac{g_4}{\theta_A^{i,j}}\right)$ and the first by $2g_3$, and adding them up in order to eliminate x_B, we obtain

$$0=2g_3YC_A-4g_1g_3\frac{x_A}{\theta_A^{i,j}}-\frac{\theta_A^{i,j}g_2+\theta_B^{i,j}g_4}{\theta_A^{i,j}\theta_B^{i,j}}YC_B+\left[\frac{\theta_A^{i,j}g_2+\theta_B^{i,j}g_4}{\theta_A^{i,j}\theta_B^{i,j}}\right]^2x_A\Longrightarrow$$
$$\frac{4g_1g_3\theta_A^{i,j}\left(\theta_B^{i,j}\right)^2-\left(\theta_A^{i,j}g_2+\theta_B^{i,j}g_4\right)^2}{\left(\theta_A^{i,j}\theta_B^{i,j}\right)^2}x_A=2g_3YC_A-\frac{\theta_A^{i,j}g_2+\theta_B^{i,j}g_4}{\theta_A^{i,j}\theta_B^{i,j}}YC_B.$$

Then we find

$$x_A=\theta_A^{i,j}\theta_B^{i,j}\frac{2\theta_A^{i,j}\theta_B^{i,j}g_3YC_A-\left(\theta_A^{i,j}g_2+\theta_B^{i,j}g_4\right)YC_B}{4g_1g_3\theta_A^{i,j}\left(\theta_B^{i,j}\right)^2-\left(\theta_A^{i,j}g_2+\theta_B^{i,j}g_4\right)^2}=x_A^{CC}.$$

Mutatis mutandis, $x_B=\theta_A^{i,j}\theta_B^{i,j}\frac{2\theta_A^{i,j}\theta_B^{i,j}g_1YC_B-\left(\theta_A^{i,j}g_2+\theta_B^{i,j}g_4\right)YC_A}{4g_1g_3\left(\theta_A^{i,j}\right)^2\theta_B^{i,j}-\left(\theta_A^{i,j}g_2+\theta_B^{i,j}g_4\right)^2}=x_B^{CC}$. ■

Proof of Lemma 14. Let S be closed. It is well established that if $\{x_t\}_{t=1}^{\infty}\subset S$ converges with probability 1 to $x\in S$, then $\lim_{t\to\infty}=f(x)$ provided the function f is continuous (Billingsley [1984]). Then, the statement follows immediately. ■

Proof of claims in Example 2 (Claim 1) If Firm A sets $p_A=3$ and never advertises, then the maximization problem for Firm B simplifies to

$$\max_{y,p_B}\left(\frac{7}{8}y+\frac{1}{2}(1-y)\right)(124+100y-(24-6y)p_B)(p_B-3)-200y-200.$$

Figure 6 visualizes the function to be maximized and clearly, $y=1$ maximizes the function on the relevant domain. (Claim 2) Suppose Firm A sets a price of 7.55 and advertises each stage with a probability of 0.69.

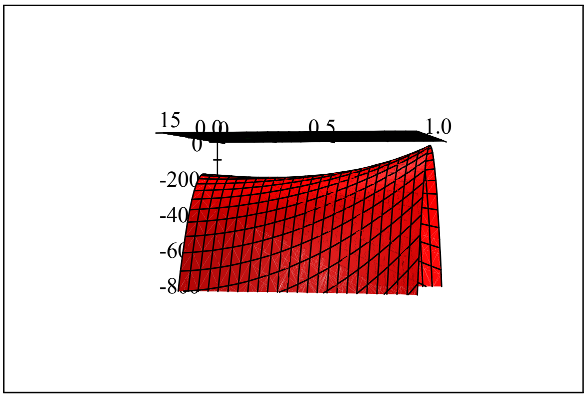

Figure 6: View from below the zero-profit plane indicated in green. On the vertical axis Firm B's profit, on the axis to the right y and to the left p_B. The highest profits are located near $y = 1$ and they are negative.

Then, we know by the results mentioned that the optimal strategy of the other firm is pure and stationary. This price and probability imply that

$$\rho = \begin{bmatrix} 0.69y & 0.69\,(1-y) \\ (1-0.69)\,y & (1-0.69)\,(1-y) \end{bmatrix}$$
$$D(\rho) = 169y + 69,\ q(\rho) = \tfrac{0.69}{y+0.69},$$
$$z_1(\rho) = 19.86,\ z_2(\rho) = 8 - 4y,\ z_3(\rho) = 24 - 6y,\ z_4(\rho) = 5.24,$$
$$\theta_A = 0.125y + 0.75875,\ \theta_B = 0.375y + 0.58625,$$
$$MP_A = 100 + \tfrac{0.69(169y+69)}{y+0.69} - 19.86 \cdot 7.55 + (8-4y)\,p_B$$
$$= \tfrac{0.69(169y+69)}{y+0.69} - 49.943 + (8-4y)\,p_B,$$
$$MP_B = 100 + \tfrac{y(169y+69)}{y+0.69} - (24-6y)p_B + 5.24 \cdot 7.55$$
$$= 139.562 + \tfrac{y(169y+69)}{y+0.69} - (24-6y)p_B.$$

So, the maximization problem for Firm B becomes

$$\max_{y,p_B} (0.375y + 0.58625)\left(139.562 + \tfrac{y(169y+69)}{y+0.69} - (24-6y)p_B\right)(p_B - 3)$$
$$-200y - 200$$

The profit function (or rather) manifold is visualized in Figure 7.

A visual inspection of Figure 7 reveals that its optimal advertising effort is $y = 1$ again. Then, the above simplifies to

$$\rho = \begin{bmatrix} 0.69 & 0 \\ 1-0.69 & 0 \end{bmatrix},\ D(\rho) = 238,\ q(\rho) = \tfrac{0.69}{1.69} = 0.408\,28,$$
$$z_1(\rho) = 19.86,\ z_2(\rho) = 4,\ z_3(\rho) = 18,\ z_4(\rho) = 5.24,$$
$$\theta_A = 0.125 + 0.75875 = 0.883\,75,\ \theta_B = 0.375 + 0.58625 = 0.961\,25,$$

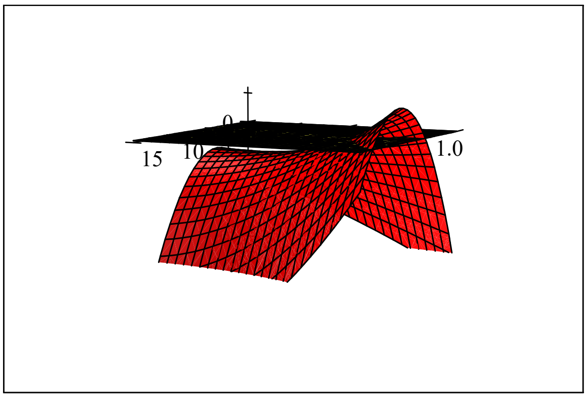

Figure 7: View from above the 0-profit plane. The manifold represents long term average profits for Firm B ($x = 0.69$, $p_A = 7.55$) for all combinations of y (axis to the right) and p_B (axis to the left).

$$MP_A = \left[\frac{0.69(169y+69)}{y+0.69} - 49.943\right]_{y=1} + 4p_B = 47.229 + 4p_B,$$
$$MP_B = \left[139.562 + \frac{y(169y+69)}{y+0.69} - (24-6y)p_B\right]_{y=1} = 280.39 - 18p_B.$$

Proof of claims in Example 3 (**Claim 4**) Recall that

$$p^{i,j} = Z^{-1}(\rho) \cdot \widetilde{D}(\rho) - Z^{-1}(\rho)\left[\Theta^{i,j}\right]^{-1} x \text{ with}$$
$$\rho = \begin{bmatrix} 0 & 0 \\ y & 1-y \end{bmatrix}, \text{ and therefore } q = 0,$$
$$z_1(\rho) = 24, \ z_2(\rho) = 8 - 4y, \ z_3(\rho) = 24 - 6y, \ z_4(\rho) = 8,$$
$$Z(\rho) = \begin{bmatrix} 24 & 4y-8 \\ -8 & 24-6y \end{bmatrix}, \ Z^{-1}(\rho) = \frac{1}{256-56y}\begin{bmatrix} 12-3y & 4-2y \\ 4 & 12 \end{bmatrix},$$
$$\widetilde{D}(\rho) = \begin{bmatrix} 100 \\ 100+100y \end{bmatrix}, \ \theta_A^{2,1} = \tfrac{5}{8}, \ \theta_B^{2,1} = \tfrac{7}{8}, \ \theta_A^{2,2} = \tfrac{1}{2} = \theta_B^{2,2}.$$

So, we find

$$p^{2,1} = \begin{bmatrix} p_A^{2,1} \\ p_B^{2,1} \end{bmatrix} = \frac{1}{256-56y}\begin{bmatrix} 1600 - 100y - 200y^2 \\ 1600 + 1200y \end{bmatrix}$$
$$-\frac{1}{256-56y}\begin{bmatrix} \frac{8}{5}(12-3y)x_A + \frac{8}{7}(4-2y)x_B \\ \frac{96}{7}x_B + \frac{32}{5}x_A \end{bmatrix},$$
$$p^{2,2} = \begin{bmatrix} p_A^{2,2} \\ p_B^{2,2} \end{bmatrix} = \frac{1}{256-56y}\begin{bmatrix} 1600 - 100y - 200y^2 \\ 1600 + 1200y \end{bmatrix}$$
$$-\frac{1}{256-56y}\begin{bmatrix} (24-6y)x_A + (8-4y)x_B \\ 24x_B + 8x_A \end{bmatrix}.$$

(**Claim 5**) Firm B uses a best response to Firm A's quantity set ($x_A = 50$), and the latter not advertising at all which is given by

$$y \cdot \kappa_B^{2,1}(\rho, x_B) + (1-y) \cdot \kappa_B^{2,2}(\rho, x_B)$$

This function is visualized in Figure 8.

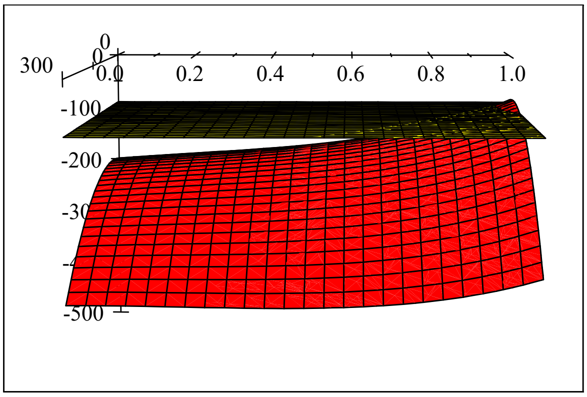

Figure 8: View from above the 0-profit plane, the plane indicated is at level -90. The manifold represents long term average profits for Firm B ($x = 0$ and $x_A = 50$) for all combinations of y (axis to the right) and x_B (axis to the left). The vertical axis is the average profit.

The restriction $y = 1$ implies

$$\rho = \begin{bmatrix} x & 0 \\ 1-x & 0 \end{bmatrix}, \text{ and } q = \tfrac{x}{2x+1-x} = \tfrac{x}{1+x},$$
$$D(\rho) = 100(3x+1-x) = 100(1+2x),$$
$$z_1(\rho) = 24-6x,\ z_2(\rho) = 4,\ z_3(\rho) = 18,\ z_4(\rho) = 8-4x,$$
$$Z(\rho) = \begin{bmatrix} 24-6x & -4 \\ 4x-8 & 18 \end{bmatrix},\ Z^{-1}(\rho) = \tfrac{1}{200-46x}\begin{bmatrix} 9 & 2 \\ 4-2x & 12-3x \end{bmatrix},$$
$$\widetilde{D}(\rho) = \begin{bmatrix} 100x\frac{2x+1}{x+1}+100 \\ 100\frac{2x+1}{x+1}+100 \end{bmatrix},\ \theta_A^{1,1} = 1 = \theta_B^{1,1},\ \theta_A^{2,1} = \tfrac{5}{8},\ \theta_B^{2,1} = \tfrac{7}{8}.$$
$$p^{1,1} = \begin{bmatrix} p_A^{1,1} \\ p_B^{1,1} \end{bmatrix} = \tfrac{1}{200-46x}\begin{bmatrix} \frac{1300+2400x+1800x^2}{x+1} \\ \frac{2800+3600x-500x^2-400x^3}{x+1} \end{bmatrix}$$
$$-\tfrac{1}{200-46x}\begin{bmatrix} 9x_A+2x_B \\ (12-3x)x_B+(4-2x)x_A \end{bmatrix},$$
$$p^{2,1} = \begin{bmatrix} p_A^{2,1} \\ p_B^{2,1} \end{bmatrix} = \tfrac{1}{200-46x}\begin{bmatrix} \frac{1300+2400x+1800x^2}{x+1} \\ \frac{2800+3600x-500x^2-400x^3}{x+1} \end{bmatrix}$$
$$-\tfrac{1}{200-46x}\begin{bmatrix} \frac{72}{5}x_A+\frac{16}{7}x_B \\ \left(\frac{96}{7}-\frac{24}{7}x\right)x_B+\left(\frac{32}{5}-\frac{16}{5}x\right)x_A \end{bmatrix}.$$

The maximization problems for Firm B are

$$\max_{x_B}\left[x_B\left(p_B^{1,1}-3\right)\right]_{\substack{p_B^{1,1}=\frac{1}{200-46x}\times \\ \left(\frac{2800+3600x-500x^2-400x^3}{x+1}-(12-3x)x_B-(4-2x)x_A\right)}} - 400$$

$$\max_{x_B} \left[x_B \left(p_B^{2,1} - 3 \right) \right]_{p_B^{2,1} = \frac{1}{200-46x} \times \left(\frac{2800+3600x-500x^2-400x^3}{x+1} - \left(\frac{32}{5} - \frac{16}{5}x\right)x_A - \left(\frac{96}{7} - \frac{24}{7}x\right)x_B \right)} - 400$$

These lead to best response functions

$$x_B = \frac{1}{6}\frac{2200+3138x-362x^2-400x^3}{(4-x)(x+1)} - \frac{1}{6}\frac{(4-2x)x_A}{4-x}$$
$$x_B = \frac{7}{48}\frac{2200+3138x-362x^2-400x^3}{(4-x)(x+1)} - \frac{7}{48}\frac{(4-2x)x_A}{4-x}$$

9 References

Aumann RJ, 2008, Game engineering in: "Mathematical Programming and Game Theory for Decision Making" (SK Neogy, RB Bapat, AK Das & T Parthasarathy, eds.) World Scientific, Singapore, *pp.* 279-285.

Aumann RJ & M Maschler, 1995, "Repeated Games of Incomplete Information", MIT Press, Cambridge, MA.

Benoît J-P & V Krishna, 1987, Dynamic duopoly: prices and quantities, *Rev Econ Studies* **54**, 23-36.

Billingsley P, 1986, "Probability and Measure", John Wiley & Sons, New York, NY.

Blackwell D, 1962, Discrete dynamic programming, *Annals Math Statistics* **33**, 719-726.

Brander H & R Harris, 1983, Anticipated collusion and excess capacity, Working Paper 530, Queen's University (Ont.), Dept. Economics.

Brenner T, & U Witt, 2003, Melioration learning in games with constant and frequency-dependent payoffs, *J Econ Behav Organ* **50**, 429-448.

Cellini R & L Lambertini, 2003, Advertising with spillover effects in a differential oligopoly game with differentiated goods, *Eur J Oper Res* **11**, 409-423.

Cellini R, L Lambertini & A Mantovani, 2008, Persuasive advertising under Bertrand competition: A differential game, *OR Letters* **36**, 381-384.

Davidson C & R Deneckere, 1986, Long-run competition in capacity, short-run competition in price, and the Cournot model, *RAND J Econ* **17**, 404-415.

Davidson C & R Deneckere, 1990, Excess capacity and collusion, *Intern Econ Rev* **31**, 521-541.

Dirven CAJM & OJ Vrieze, 1986, Technical Note-Advertising models, stochastic games and myopic strategies, *Oper Res* **34**, 645-649.

Dorfman R, & PO Steiner, 1954, Optimal advertising and optimal quality, *Amer Econ Rev* **44**, 826-836.

Fershtman C, 1984, Goodwill and market shares in oligopoly, *Economica* **51**, 271-281.

Fershtman C & E Muller, 1986, Capital investments and price agreements in semicollusive markets, *RAND J Econ* **7**, 214-226.

Forges F, 1986, An approach to communication equilibria, *Econometrica* **54**, 1375-1385.
Friedman JW, 1983, Advertising and oligopolistic equilibrium, *Bell J Econ* **14**, 464-473.
Gintis H, 2001, "Game Theory Evolving", Princeton Univ Press, Princeton, NJ.
Hart S, 1985, Nonzero-sum two-person repeated games with incomplete information, *Math Oper Res* **10**, 117-153.
Hoffman AJ & RM Karp, 1966, On nonterminating stochastic games, *Management Science* **12**, 359-370.
Hordijk A, OJ Vrieze & GL Wanrooij, 1983, Semi-Markov strategies in stochastic games, *Int J Game Theory* **12**, 81-89.
Joosten R, 2007, Small Fish Wars: A new class of dynamic fishery-management games, *ICFAI J Managerial Economics* **5**, 17-30.
Joosten R, 2009, Strategic advertisement with externalities: a new dynamic approach, in: "Modeling, Computation and Optimization" (SK Neogy, AK Das & RB Bapat, eds.), World Scientific, Singapore, *pp.* 21-43.
Joosten R, 2016a, Strong and weak rarity value: resource games with complex price-scarcity relationships, *Dyn Games & Applic* **16**, 97-111.
Joosten R, 2016b, Long-run strategic advertisement and short-run Bertrand competition, *Int Game Theory Rev* **17**, DOI: 10.1142/S0219189-15400149.
Joosten R, 2019, Strategic interaction and externalities: FD-games and pollution, in: "Understanding Economic Change: Advances in Evolutionary Economics" (U Witt, A Chai, eds.), Cambridge Univ. Press, *pp.* 288-308.
Joosten R, T Brenner & U Witt, 2003, Games with frequency-dependent stage payoffs, *Int J Game Theory* **31**, 609-620.
Joosten R, & R Harmelink, 2023, Tipping and strong rarity value in a stochastic fishery game, *Int Game Theory Rev*, (In press).
Joosten R & R Meijboom, 2018, Stochastic games with endogenous transition probabilities, in: "Mathematical Programming and Game Theory" (SK Neogy, RB Bapat & D Dubey, eds.), Springer, *pp.* 205-226.
Joosten R, & L Samuel, 2017, On stochastic fishery games with endogenous stage payoffs and transition probabilities, in: "Game Theory & Applications" (D.F. Li, X.G. Yang, M. Uetz, G.J. Xu, eds.), CCIS 758, Springer, Heidelberg, *pp.* 115-133.
Joosten R, & L Samuel, 2020, On finding large sets of rewards in two player *ETP-ESP* games, *Int Game Theory Rev* **22**, DOI: 10.1142/S021198920 400022.
Kydland F, 1977, Equilibrium solutions in dynamic dominant-player models, *J Econ Theory* **15**, 307-324.
Nash J, 1951, Noncooperative games, *Ann Math* **54**, 284-295.
Nelson P, 1970, Information and consumer behavior, *J Pol Econ* **78**, 311-329.

Nerlove M & KJ Arrow, 1962, Optimal advertising policy under dynamic conditions, *Economica* **29**, 129-142.
Prescott E, 1973, Market structure and monopoly profits: a dynamic theory, *J Econ Theory* **6**, 546-557.
Scherer F, 1980, "Industrial Market Structure and Economic Performance", Rand McNally, Chicago, IL.
Schmalensee R, 1978, A model of advertising and product quality, *J Pol Economy* **86**, 485-503.
Shapley L, 1953, Stochastic games, *PNAS U.S.A.* **39**, 1095-1100.
Sobel MJ, 1981, Myopic solutions of Markov Decision Processes and stochastic games, *Oper Res* **29**, 995-1009.
Thuijsman F, 1992, "Optimality and Equilibria in Stochastic Games", *CWI-Tract* **82**, Centre for Math & Comp Science, Amsterdam.
Van Cayseele P & D Furth, 1996, Bertrand-Edgeworth duopoly with buyouts or first refusal contracts, *Games Econ Behav* **16**, 391-417.
Van Damme EEC, 1991, "Stability and Perfection of Nash Equilibria", Springer-Verlag, Berlin.
Vidale ML, & HB Wolfe, 1957, An operations-research study of sales response to advertising, *Oper Res* **5**, 370-381.
Vojnović M, "Contest theory: incentive mechanisms and ranking methods", Microsoft Research, Cambridge Univ, UK.
Waugh FV, 1959, Needed research on the effectiveness of farm products promotions, *J Farm Econ* **41**, 364-376.