

Exploration and exhaustibility in dynamic Cournot games

MICHAEL LUDKOVSKI¹ and RONNIE SIRCAR²

¹*Department of Statistics & Applied Probability, University of California, South Hall,
Santa Barbara, CA 93106-3110*

email: ludkovski@pstat.ucsb.edu

²*ORFE Department, Princeton University, Sherrerd Hall, Princeton, NJ 08544*

email: sircar@princeton.edu

*(Received 10 February 2011; revised 17 November 2011; accepted 17 November 2011;
first published online 15 December 2011)*

We study the stochastic effect of resource exploration in dynamic Cournot models of exhaustible resources, such as oil. We firstly treat the case of a monopolist who may undertake costly exploration to replenish his diminishing reserves. We then consider a stochastic game between such an exhaustible producer and a ‘green’ producer that has access to an inexhaustible but relatively expensive source, such as solar power. The effort control variable is taken to be either continuous or discrete (switching control). In both settings, we assume that new discoveries occur according to a jump process with intensity given by the exploration effort. This leads to a study of systems of non-linear first-order delay ordinary differential equations (ODEs). We derive asymptotic expansions for the case of a small-exploration success rate and present some numerical investigations.

Key words: Cournot games; Exploration control; Stochastic differential games; exhaustible resources

1 Introduction

We initiate a study of the effect of exploration in dynamic Cournot games. In the model studied in [13], players competed in a dynamic non-co-operative game as their reserves of an exhaustible resource depleted, but there was no possibility of exploration and discovery of new reserves. Here, we incorporate controllable exploration with costs associated with discovery efforts. For an introduction to games and oligopolies, we refer to [11, 23]; for an entry into differential games, we recommend [3, 7, 10]; and for stochastic control problems, see [6]. However, the combination of dynamic games and control of jump processes is novel here to our knowledge, and we attempt, as best as we can, to give a self-contained exposition.

The full two or N -player non-zero sum differential game analysed in [13] is characterised by systems of nonlinear HJB partial differential equations which are extremely difficult to resolve numerically. Here, we concentrate mainly on reduced versions of the full exploration game which allows us to deal with ordinary differential equations (ODEs) or variational inequalities. Firstly, we study the full monopoly problem and then the game between a producer of energy from an inexhaustible (or renewable) but costly resource,

such as solar power, and a producer of energy from a cheap but ultimately exhaustible resource, such as oil, whose reserves may be temporarily replenished by costly exploration. Our goal is to understand the effect of the possibility of finding new reserves on extraction strategies and market prices.

We work throughout in a Cournot framework in which producers choose *quantities* of energy to produce and receive profit based on a single market price determined through aggregate supply. For simplicity of exposition and specific calculations and numerics, we consider a linear inverse-demand curve. In contrast to the Cournot game where the strategic variable is quantity, one may also consider a Bertrand version where producers set prices and supply as much as demanded by the market. Such a dynamic price-setting stochastic Bertrand model, more suitable for describing consumer goods markets, is studied in [16]; see also [23] for a textbook treatment of various oligopoly models.

In the simplest *static* (or one-period) Cournot problem, given the aggregate production quantity $Q \in [0, 1]$, the market price is given by the (inverse-demand) function $P(Q) = 1 - Q$. When there is one player with marginal cost of production $s_1 \in [0, 1]$, he chooses his optimal quantity $q_1 \geq 0$ to maximise his monopoly profit function

$$\Pi_1 = q_1(1 - q_1) - s_1q_1.$$

The optimal quantity and profit are given by

$$q_1^* = \frac{1}{2}(1 - s_1), \quad \Pi_1^* = \frac{1}{4}(1 - s_1)^2.$$

When there are two players with costs $(s_1, s_2) \in [0, 1]^2$ and non-negative production quantities (q_1, q_2) , the aggregate quantity is $Q = q_1 + q_2$ and each player's profit function is

$$\Pi_i = q_i(1 - q_i - q_j) - s_iq_i, \quad i = 1, 2; j \neq i.$$

In a Nash equilibrium $(q_1^*, q_2^*) \in [0, 1]^2$ for the duopoly, each player maximises profit as a best response to the other player's equilibrium strategy:

$$\Pi_i^* = \max_{q_i \geq 0} q_i(1 - q_i - q_j^*) - s_iq_i, \quad i = 1, 2; j \neq i.$$

For costs $s_1, s_2 < \frac{1}{2}$, it is easy to see that both players have positive equilibrium productions

$$q_i^* = \frac{1}{3}(1 - 2s_i + s_j), \quad \Pi_i^* = \frac{1}{9}(1 - 2s_i + s_j)^2, \quad i = 1, 2; j \neq i.$$

However, if player j 's cost is too high relative to player i 's, specifically $s_j > \frac{1}{2}(1 + s_i)$, then he is *blockaded* from production, meaning his equilibrium quantity is zero. In this case, player i has a monopoly and the Nash equilibrium is given by

$$q_i^* = \frac{1}{2}(1 - s_i), \quad q_j^* = 0, \quad \Pi_i^* = \frac{1}{4}(1 - s_i)^2, \quad \Pi_j^* = 0.$$

A full characterisation of the static N -player game for a wide class of general inverse-demand functions is given in [13, Section 2], and for Bertrand games in [16, Section 2]; a

comparison between Cournot and Bertrand in terms of the number of blockaded players is in [17]. Here, to highlight the essential effects of exploration in a transparent way, we stay with the linear inverse-demand function and the one- or two-firm cases.

In the *single-agent* case, the literature on joint exploration and production decisions was started in the late 1970s. Let us especially highlight the seminal paper of Pindyck [20] that considered a deterministic model of exploring for a non-renewable resource (that is, one which is gone when used up, but of which new discoveries can occur). He assumed exploration to be incremental and represented as a deterministic reserve addition. Pindyck showed that the resulting resource shadow price, corresponding to the marginal value of additional reserves, will firstly decrease and then increase as reserves run low. This generalised the original result of Hotelling [14] for exhaustible resources without discovery that resource price grows at the inter-temporal discount rate. A series of works extended Pindyck's model for the case of stochastic discoveries. If exploration is punctuated by large discoveries, it is best represented as a point process [1, 8, 15, 22]. This is the point of view we take in this paper. In particular, the model of [1, 8] essentially matches our monopoly setup with continuous effort in Section 2.3, except that we interpret value of production q as revenue $P(q) \cdot q$, whereas the existing works simply interpret it as utility of consumption $u(q)$.

Finally, if the total size of reserves is unknown, its dynamics can be described via a stochastic differential equation with controlled volatility and drift [21]. If the global resource is finite, like in the case of hydrocarbons, over the long term total possible discoveries are bounded and exploration becomes harder as time goes on, see [20]. In this paper, again as in [8], we make the simpler assumption that exploration can go on forever, which is akin to unbounded technological R&D and knowledge accumulation. This implies that the controlled reserves process (X_t^*) in our model is recurrent on (a subset of) \mathbb{R}_+ and moreover that zero is recurrent for (X_t^*) (i.e. the resource is intermittently exhausted and then re-discovered again).

On a broader scale, there is enormous economics literature (we refer, for example, to the textbooks [4, 7]) on management of exhaustible and non-renewable resources that is partitioned according to main applications (mineral and hydrocarbon deposits on local and global scales, search for alternative 'green' technologies, R&D innovation, etc.), key sources of uncertainty (fluctuating reserve levels, random demand shocks, timing of innovations, etc.) and their probabilistic representation (unknown parameters, single event innovation, repeated discoveries, etc).

In Section 2, we start with a single player, or monopolist. A key issue is the start-up cost of exploration which, in this case, we look at in two ways: on/off or switching control (Section 2.1), and continuous exploration effort (Section 2.3). In Section 3, we consider a two-player game with continuous exploration control. As is widely appreciated, non-zero-sum stochastic games are much more involved due to the strong coupling induced by the best-response requirements of Nash equilibrium, and we are able to make some progress analytically (Section 3.2) and numerically (Section 3.3). In Section 3.4, we describe how Hotelling's classical rule for depletion of exhaustible resources is modified in the context of a Cournot game with exploration. Throughout, we maintain discrete fixed discovery sizes to focus on the essential effect of exploration. The numerical implementations can easily be extended to random jump sizes.

2 Monopoly exploration and production model

Consider a firm in charge of an exhaustible resource, such as oil, precious metal or coal. The main operations of the firm are summarised by two activities: production and exploration. Production of the resource generates revenue but lowers remaining stocks. Conversely, exploration will likely lead to discoveries and so boost reserves, but is costly. That is, we assume that exploration is *uncertain*, so that replenishment of the resource occurs stochastically. For example, in practical hydrocarbon exploration efforts, firms will spend significant exploration funds but actual new discoveries occur infrequently and unpredictably.

Let X_t denote the firm's reserves at date t . We assume that (X_t) decreases at a (controlled) production rate $q_t \geq 0$, and increases through *jumps* thanks to discrete new discoveries. Exploration successes are represented by a counting process (N_t) whose jumps mark the discovery times. Each discovery increases reserves by a *fixed* amount $\delta > 0$, and so we have the dynamics:

$$dX_t = -q_t \mathbf{1}_{\{X_t > 0\}} dt + \delta dN_t.$$

To make discoveries, exploration *effort* is needed, with its level indicated by $a_t \geq 0$. The intensity of the point process (N_t) is λa_t , and its costs are captured via $\mathcal{C}(a_t)$ for some positive non-decreasing convex function $\mathcal{C}(\cdot)$.

Below, we will consider two main cases:

- (1) Discrete effort levels. For simplicity, we focus on binary effort $a_t \in \{0, 1\}$. Thus, either no exploration is undertaken, or a fixed effort (normalised to be 1) is applied. We also assume that to start or stop exploration, a fixed cost $K > 0$ must be paid each time.
- (2) Continuous effort level, $a_t \in \mathbb{R}_+$. In that case, we assume that $\mathcal{C}(a)$ is strictly convex for a large enough, which guarantees that optimal effort levels are finite.

At each date t , the firm sets the quantity q_t , that is the rate of production (or extraction) of the resource, and the market price is a decreasing function of the supply. To simplify, we assume a linear price (or inverse-demand) function $P = P(q) = 1 - q$. The revenue is simply price times quantity $q \cdot P(q)$. Future revenue/costs are discounted in time by a constant discount factor $r > 0$ and the objective (specified precisely in equations (2.1) and (2.17)) is to maximise total expected discounted net revenues over the infinite horizon. We also assume for simplicity that the cost of extraction is zero; in Section 3 where an oil producer competes with an energy producer with a clean inexhaustible but expensive technology (for instance solar power), the latter will have a nonzero production cost.

The controlled dynamics of the reserves can be described in terms of a *piecewise deterministic process* (PDP). Indeed, between discoveries, no new information is coming in and reserves decrease continuously according to the production schedule. At discovery dates, reserves increase via an instantaneous jump; the intensity of discoveries is a function of current reserves through the optimal exploration control. Accordingly, the corresponding theory of PDPs [6] can be brought to bear in the single-agent case, and we refer to [5, 9, 18, 22] for the abstract theory, in particular proofs of the existence and smoothness of solutions to the resulting first-order integro-differential Bellman equation and its correspondence with the solution of the stochastic control problem.

2.1 Switching exploration control

In the *discrete* exploration control case, $a_t \in \{0, 1\}$, the firm sequentially alternates between expending discovery effort to replenish supplies and doing nothing. Exploration effort costs $\mathcal{C} \equiv \mathcal{C}(1)$ dollars per unit time, while a switch to start or stop exploring consumes $0 \leq K < \mathcal{C}/r$ dollars.

A natural admissibility condition is that (a_t) has (\mathbb{P} -a.s.) a finite number of switches on any finite horizon. This requirement makes sure that there is no chattering (infinitely fast switching between exploration regimes). With positive switching cost $K > 0$, any inadmissible strategy will have infinite costs and hence be clearly sub-optimal. The admissibility condition on (q_t) is that the corresponding reserves process remains non-negative, which is simply that $q_t = 0$ when all reserves are exhausted.

Given strategies (q_t, a_t) , and denoting $\Delta a_t = a_t - a_{t-}$, the total expected revenue for the firm is then

$$\mathbb{E} \left[\int_0^\infty e^{-rt} \{ (1 - q_t)q_t - \mathcal{C}(a_t) \} dt - \sum_{\{t: \Delta a_t \neq 0\}} e^{-rt} K \right]. \tag{2.1}$$

To study the stochastic control problem of maximising equation (2.1), we adopt the usual dynamic programming approach of analysing the corresponding value functions assuming that both (q_t) and (a_t) are Markov feedback controls: $q_t = q(X_t)$ and $a_t = a(X_t)$. Let (X_t) be the reserves process starting at $X_0 = x$. Since $a \in \{0, 1\}$, we can write

$$dX_t = -q(X_t)\mathbf{1}_{\{X_t > 0\}} dt + a(X_t)\delta dN_t, \tag{2.2}$$

where now (N_t) is a standard Poisson process with intensity $\lambda > 0$. The state variables in the problem are the initial reserves x and the initial exploration mode a_0 . Due to switching costs, the current mode has an effect on the subsequent strategy. When all resources are exhausted, the firm optimises between either (i) expending effort and waiting for a new discovery, while producing zero, or (ii) quitting and receiving a zero terminal payoff.

For future reference, we recall that the production model without exploration can be solved explicitly. Denote by

$$w_0(x) = \sup_q \int_0^\infty e^{-rt} (1 - q_t)q_t \mathbf{1}_{\{X_t > 0\}} dt \quad \text{with } X_0 = x, \tag{2.3}$$

the corresponding value function, where now $\lambda = 0$. Then, $w_0(x)$ solves the Hamilton–Jacobi ODE

$$\frac{1}{4} (1 - w'_0(x))^2 - r w_0(x) = 0, \tag{2.4}$$

with initial condition $w_0(0) = 0$.

The following recipe provides a closed-form expression for w_0 .

Lemma 2.1 *The solution to the ODE*

$$(\alpha - v')^2 = \kappa v \quad \text{on } \{x > 0\}, \quad v(0) = v_0,$$

where $\alpha, \kappa > 0$ and $v_0 \geq 0$, is

$$v(x) = \frac{\alpha^2}{\kappa} (1 + \mathbf{W}(\theta(x)))^2,$$

where

$$\theta(x) = \beta e^\beta e^{-\kappa x / (2\alpha)}, \quad \beta = -1 + \frac{\sqrt{\kappa v_0}}{\alpha},$$

and \mathbf{W} is the Lambert-W function defined as the inverse function of xe^x , restricted to the range $[-1, \infty)$ and the domain $[-e^{-1}, \infty)$.

Proof This follows easily from the substitutions $v = \frac{\alpha^2}{\kappa}(1 + u)^2$, $y = e^{-\kappa x / (2\alpha)}$ and that $\mathbf{W}'(y) = \mathbf{W}(y) / (y(1 + \mathbf{W}(y)))$. □

Applying Lemma 2.1 to equation (2.4) implies that

$$w_0(x) = \frac{1}{4r} [1 + \mathbf{W}(-e^{-2rx-1})]^2. \tag{2.5}$$

2.1.1 Quasi-variational inequality

Denote by $v(x)$ (resp. $w(x)$) the value function of the producer, conditional on starting in the exploration (resp. no exploration) mode and initial resource of size x . Applying standard dynamic programming arguments [19], it follows that the pair (v, w) satisfy a first-order coupled system of variational inequalities when $x > 0$, which is given in equation (2.9) below.

When $x = 0$, and the exploration regime is 1, the agent cannot produce but continues to explore. Alternatively, he has the option to quit permanently. Let τ be the next time of exploration success. Then, an argument based on Bellman principle gives $v(0) = \max(0, \mathbb{E}[-\int_0^\tau e^{-rt} \mathcal{C} dt + e^{-r\tau} v(\delta)])$ and, using the exponential inter-arrival times for (N_t) , $\tau \sim \text{Exp}(\lambda)$, we obtain that

$$v(0) = \max\left(0, \frac{-\mathcal{C} + \lambda v(\delta)}{\lambda + r}\right). \tag{2.6}$$

When $x = 0$ and the exploration regime is 0, the player can either permanently quit or begin exploring after paying K . Hence,

$$w(0) = \max(0, v(0) - K). \tag{2.7}$$

We summarise in the following proposition, where we employ the following convenient notation for the non-local (or delay) term coming from the jump process:

$$\Delta v(x) := v(x + \delta) - v(x). \tag{2.8}$$

Proposition 2.1 *The pair (v, w) is the smallest solution of*

$$\begin{aligned} \max \left(\frac{1}{4}(1 - v'(x))^2 - \mathcal{C} + \lambda \Delta v(x) - rv(x), w(x) - K - v(x) \right) &= 0, \\ \max \left(\frac{1}{4}(1 - w'(x))^2 - rw(x), v(x) - K - w(x) \right) &= 0, \end{aligned} \tag{2.9}$$

subject to boundary conditions (2.6) and (2.7). In particular, the optimal policy is given in feedback form as $q_v^* = \frac{1-v'(x)}{2}$ and $q_w^* = \frac{1-w'(x)}{2}$. Moreover, $v(\cdot), w(\cdot) \in \mathcal{C}^1(\mathbb{R}_+) \setminus D$, where

$$\begin{aligned} D = \{x \in \mathbb{R}_+ : w(x) - K - v(x) = (1 - v'(x))^2/4 - \mathcal{C} + \lambda \Delta v(x) - rv(x) \\ \text{or } v(x) - K - w(x) = (1 - w'(x))^2/4 - rw(x)\} \cup \{0\}. \end{aligned} \tag{2.10}$$

2.1.2 Analysis of switching regions

The key to analysis of equation (2.9) is the study of the corresponding *switching regions*, which keep track of the maximisers in each max-term. When reserves are low (x small), the marginal cost of additional reserves is high and therefore exploration is valuable. When reserves are large, the threat of exhaustion is weak and the value of exploration is small. Finally, when reserves are very low and exploration is expensive, it may be cheaper to quit permanently rather than to expend efforts on exploration. Therefore, we conjecture that there exist three thresholds $0 \leq x_Q \leq x_{on} \leq x_{sat}$ ('Q' for quit, 'on' for start exploration, 'sat' for saturation), such that

$$\begin{aligned} v(x) - K = w(x), \quad \text{for } x \in [x_Q, x_{on}], \quad w(x) > v(x) - K, \quad \text{for } x \in [0, x_Q) \cup (x_{on}, \infty), \\ w(x) - K = v(x), \quad \text{for } x \in [x_{sat}, \infty), \quad v(x) > w(x) - K, \quad \text{for } x < x_{sat}. \end{aligned}$$

We do not attempt a proof about the shape of the switching regions. However, standard theory implies that the variational inequality in equation (2.9) has a unique viscosity solution. Therefore, identifying one plausible solution structure, borne out by numerical computations in Section 2.2, goes a long way to giving a practical solution.

We refer to $[0, x_{sat})$ as the continuation region for v (exploring), and $[0, x_Q) \cup (x_{on}, \infty)$ as the continuation region for w (no exploration). As we discuss below, usually (that is for moderate values of K and \mathcal{C}), $x_Q = 0$ and there is only one continuation region for w .

The lower switching boundary x_{on} determines the maximum level of reserves at which it is optimal to re-start exploration. Two cases can occur. Firstly, we (usually) have $w(0) = v(0) - K$. For $z \geq 0$ such that $v(z) \geq K$, denote by

$$w(x; z) = \frac{1}{4r} [1 + \mathbf{W}(be^{b-2r(x-z)})]^2, \quad b = -1 + \sqrt{4r(v(z) - K)}, \quad x \geq z,$$

the unique solution to $w'(x; z) = 1 - 2\sqrt{rw(x; z)}$ (where $' = \frac{d}{dx}$), with $w(z; z) = v(z) - K$, see Lemma 2.1.

Proposition 2.2 *Suppose $v(0) \geq K$. Then, $x_{on} > 0$ and the switching region must contain an interval around zero. Moreover, $w'(x_{on}) = v'(x_{on})$ match and there is a smooth fit at x_{on} .*

Proof Since $w'(0; 0) = 1 - 2\sqrt{r(v(0) - K)} < 1$, and $w(x; z)$ is continuous in both of its arguments, we see that $w(\epsilon + \epsilon'; \epsilon) < v(\epsilon) - K$ for ϵ and ϵ' small enough. However, by equation (2.9), we must have $w(x) \geq v(x) - K$ for all x ; therefore, it must be the case that $w(x) = v(x) - K$ for x small enough. Therefore, $x_{on} > 0$.

As shown above, $v'(0) = 1 > w'(0; 0)$ and as x increases from zero, eventually a level x_{on} is reached where

$$w'(x_{on}; x_{on}) = 1 - 2\sqrt{r(v(x_{on}) - K)} = v'(x_{on}). \tag{2.11}$$

With this x_{on} , we then have $w(x) = w(x; x_{on})$ for $x > x_{on}$ and $w(x) = v(x) - K$ for $x \leq x_{on}$. Since $w(0) = v(0) - K > 0 = w_0(0)$ and $v'(x) > w'(x; x)$ on $[0, x_{on})$, the full interval $[0, x_{on})$ must be a switching region for w , so that $x_Q = 0$ in this case. In equation (2.11), we obtain that the derivatives $w'(x_{on}) = v'(x_{on})$ match and therefore give rise to a *smooth fit* at x_{on} . □

In the second case, if $w(0) = 0$, then for x very small it is not optimal to switch from w to v since the fixed outlay K outweighs the gains of making discoveries before exhaustion. As a result, there is an additional interval $[0, x_Q]$, where $w(x) > v(x) - K$.

Proposition 2.3 *If $v(0) < K$, then*

$$w(x) = \begin{cases} w_0(x), & \text{for } x < x_Q, \\ v(x) - K, & x_Q \leq x < x_{on}, \\ w(x; x_{on}), & x \geq x_{on}, \end{cases}$$

where $x_Q > 0$ is found by solving $v(x_Q) - K = w_0(x_Q)$.

In other words, starting exploration is only optimal on $(x_Q, x_{on}]$.

Proof We still have that $v'(0) = 1 > w'(0)$, so that the gap between v and w will be initially increasing. By equation (2.9), for x small enough, w solves the ODE defining equation (2.5) and since $w_0(0) = 0$ we have that $w(x) = w_0(x)$ in the neighbourhood of zero. Then, $x_Q > 0$ is determined from $v(x_Q) - K = w_0(x_Q)$. □

Note that possibly $x_{on} = x_Q$, in which case $w(x) = w_0(x)$ for all x and starting exploration is never optimal. This will occur if K is too large relative to the benefit of exploration. In such a case, the producer explores for a while, until $X_t \geq x_{sat}$ at which time exploration is permanently shutdown and all remaining reserves are exhausted.

Since (X_t) increases through jumps only, exploration is always terminated at jump-time, whenever (X_t) enters the region $[x_{sat}, \infty)$. In contrast, exploration starting date can be anticipated, since the hitting time of $[0, x_{on}]$ is predictable. These different behaviours at the switching boundary are the reason for the following result.

Proposition 2.4 *There is no smooth fit at x_{sat} , so that $v'(x_{sat}-) < v'(x_{sat}+) = w'(x_{sat})$.*

Proof It is clear that $v'(x_{\text{sat}}-) \leq v'(x_{\text{sat}}+)$; if this were not the case, then it would be sub-optimal to stop exploration at level x_{sat} since waiting infinitesimally would reduce the gap between w and v .

Towards a contradiction, suppose that $v'(x_{\text{sat}}) = w'(x_{\text{sat}})$. Then, v is differentiable everywhere. In fact, based on the ansatz of the shape of the switching regions and the delay ODE satisfied by $v(x)$ on $(0, x_{\text{sat}})$, the derivative $U(x) := v'(x)$ is itself continuously differentiable on $x > 0$ and satisfies

$$-U'(x)(1 - U(x))/2 + \lambda(U(x + \delta) - U(x)) - rU(x) = 0, \quad U(0) = 1. \tag{2.12}$$

Making use of the fact that $U(x) = w'(x)$ for $x \geq x_{\text{sat}}$ and that w is concave (i.e. $w'(x + \delta) - w'(x) < 0$), we find that

$$U'(x_{\text{sat}}) = \frac{-2rU(x_{\text{sat}}) + 2\lambda(U(x_{\text{sat}} + \delta) - U(x_{\text{sat}}))}{1 - U(x_{\text{sat}})} < \frac{-2rw'(x_{\text{sat}})}{1 - w'(x_{\text{sat}})} = w''(x_{\text{sat}}).$$

By continuity of U' and w'' , it follows that $U' < w''$ in a neighbourhood of x_{sat} . Let $x' = \sup\{x < x_{\text{sat}} : U(x) = w'(x)\}$; by above $x' < x_{\text{sat}}$. If $x' > 0$, then

$$\begin{aligned} U'(x') &= \frac{-2rU(x') + 2\lambda(U(x' + \delta) - U(x'))}{1 - U(x')} \\ &\leq \frac{-2rw'(x') + 2\lambda(w'(x' + \delta) - w'(x'))}{1 - w'(x')} < w''(x') \end{aligned}$$

since $w'(x) > 0$ for all x . This is a contradiction; thus $x' = 0$ and $U(x) < w'(x)$ for all $x < x_{\text{sat}}$ (the above analysis also implies that U is convex). In other words, v is always more concave than w . However, this contradicts the result $U(x_{\text{on}}) = w'(x_{\text{on}})$ shown above. □

2.1.3 Optimal strategy and controlled reserves process

To summarise, when $v(0) > K$ we look for critical thresholds $x_{\text{sat}} > x_{\text{on}}$ such that

$$\left\{ \begin{array}{ll} \frac{1}{4}(1 - v'(x))^2 - \mathcal{C} + \lambda[v(x + \delta) - v(x)] - rv(x) = 0, & x \leq x_{\text{sat}}, \\ \frac{1}{4}(1 - w'(x))^2 - rw(x) = 0, & x \geq x_{\text{on}}, \\ v(x) = w(x) - K, & x > x_{\text{sat}}, \\ \text{and } w(x_{\text{on}}) = v(x_{\text{on}}) - K. \end{array} \right. \tag{2.13}$$

The optimal exploration strategy is to start exploring whenever the supply drops below x_{on} and to stop whenever it increases beyond x_{sat} . The gap $[x_{\text{sat}}, x_{\text{on}}]$ is due to the discrete switching costs which create a hysteresis band: for $x \in [x_{\text{sat}}, x_{\text{on}}]$, it is optimal to maintain the *current* exploration strategy. In total, there are four free variables in equation (2.13): the location of x_{on} and x_{sat} , the level $v(0)$ and the level $w(x_{\text{sat}})$. The four constraints are: $w(x_{\text{on}}) = v(x_{\text{on}}) - K$, $w(x_{\text{sat}}) = v(x_{\text{sat}}) + K$, $w'(x_{\text{on}}) = v'(x_{\text{on}})$ and $v'(0) = 1$. It appears difficult to state any closed-form algebraic equation linking the free variables above with

these constraints. We will see that this structure of solution is borne out by numerical computations in Section 2.2.

Remark 1 If there is smooth fit at x_{on} , our analysis shows that $v'(x) \leq w'(x)$ for all $x \geq x_{on}$, and hence there is at most one crossing point between $v(x) + K$ and $w(x)$ on $[x_{on}, \infty)$. This confirms our ansatz about the shape of the exploration regions.

The structure of equation (2.13) implies that the optimal controlled reserves process (X_t^*) undergoes a cyclic behaviour. Suppose that $X_0^* = 0$ (and $v(0) > K$); then exploration is started, and once $X_t^* > 0$ production begins as well. After a few more discoveries (it is possible that intermittently $X_t^* = 0$ again and production is temporarily halted), we will have $X_t^* > x_{sat}$ and exploration is shut down. Existing reserves are then exhausted until $X_t^* = x_{on}$ when exploration is re-started. Eventually, a long interval between discoveries will again lead to $X_t^* = 0$. In our model, the potential future discoveries are unlimited and, therefore, this cycle perpetuates itself. Moreover, due to the Markov controls, each cycle is independent of the others and (X_t^*) forms a renewal process with renewal times corresponding to $X_t^* = 0$ and $a_t^* = 1$ (exhaustion with exploration). Note that the optimal reserves process is bounded, $X_t^* \in [0, x_{sat} + \delta]$, and the second w -continuation region $[0, x_0)$ is never reached by (X_t^*) .

2.2 Numerical solution and illustrations

To solve the system (2.9) requires dealing with a first-order non-linear delay ODE coupled with a first-order nonlinear ODE via two free boundaries. Fixing $\Delta x = \delta/N$ and upper threshold $\bar{M}\delta$, we discretise (2.9) via a uniform grid approximation on $[0, \bar{M}\delta]$ with step size Δx . The corresponding derivatives are replaced with finite-difference operators, taking $v_n \simeq v(\Delta xn), w_n \simeq w(\Delta xn)$. For instance, the first half of equation (2.9) for $1 \leq n \leq (\bar{M}-1)N$ becomes

$$\max(F_n, G_n) = 0, \quad \text{where} \tag{2.14}$$

$$\begin{cases} F_n := \frac{1}{4} \left(1 - \frac{v_n - v_{n-1}}{\Delta x}\right)^2 - \mathcal{C} + \lambda[v_{n+N} - v_n] - rv_n, \\ G_n := w_n - K - v_n. \end{cases} \tag{2.15}$$

At the boundaries, we directly apply $v_0 = \frac{-c + \lambda v_N}{\lambda + r}$ and for $n > (\bar{M} - \delta)N, v_n = w_n - K$ which implicitly assumes that $x_{sat} < \bar{M} - \delta$. Note that to avoid oscillations in computing $v'(x)$ (needed to find d_v^*), we apply one-sided finite-difference approximation in equation (2.15).

Repeating the same for the second half of equation (2.9), we obtain a system of $2(\bar{M}N + 1)$ (quadratic) equations in the $2(\bar{M}N + 1)$ unknowns $v_0, \dots, v_{\bar{M}N}, w_0, \dots, w_{\bar{M}N}$. This system can now be solved using standard methods to obtain the unique solution $(v_n, w_n)_{n=0}^{\bar{M}}$. Finally, one may recover the free boundaries $x_{on} \simeq \underline{m}\Delta x$ and $x_{sat} \simeq \bar{m}\Delta x$ by finding the indices $\underline{m} = \max\{n : w_n = v_n - K\}$ and $\bar{m} = \min\{n : G_n = 0\}$.

The max-operator in $\max(F_n, G_n)$ is highly non-linear; in particular, at the optimum, where by complementarity at least one of the terms is zero, we obtain a singularity.

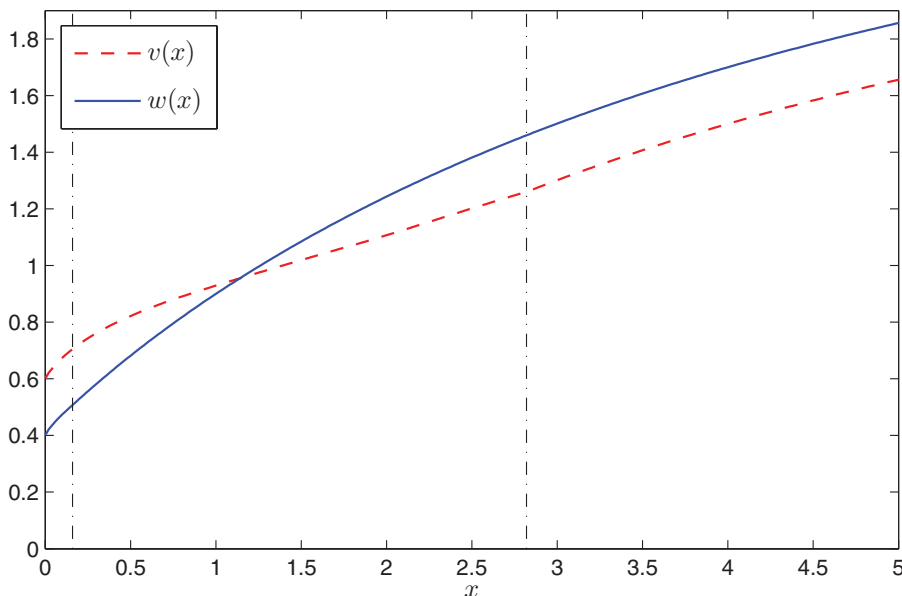


FIGURE 1. (Colour online) Solution of the single-agent problem: $v(x)$ (dash line) and $w(x)$ (solid line). The parameters are $\delta = 1, \lambda = 1, r = 0.1, \mathcal{C} = 0.25, K = 0.2$. The vertical lines indicate the switching thresholds $x_{\text{on}} = 0.17$ and $x_{\text{sat}} = 2.82$.

This presents a difficulty in applying Jacobian-based methods on the large vector equation (2.14). To regularise the problem, we apply the Fischer–Burmeister (FB) function $F_{FB}(f, g) := \sqrt{f^2 + g^2} + f + g$ that satisfies

$$\max(f, g) = 0 \iff F_{FB}(f, g) = 0,$$

and is smooth outside the origin $(0, 0)$. After experimenting with solving equation (2.14) for typical parameter ranges, we find that this approach is stable and exhibits fast (super-linear) convergence in Δx .

To illustrate our results, we consider the monopoly switching exploration control with parameters $\delta = 1, \lambda = 1, r = 0.1, \mathcal{C} = 0.25, K = 0.2$. We solve the system (2.14) using $\Delta x = 0.02$ and $\bar{M} = 6$ which leads to a 602-dim. system of quadratic equations, with the solution plotted in Figure 1. Note that the far-field limit in this case is $(4r)^{-1} = 2.5$. We also find that $x_Q = 0, x_{\text{on}} = 0.17, x_{\text{sat}} = 2.82$.

Figure 2 illustrates a counterintuitive phenomenon that occurs due to lack of smooth fit at x_{sat} . Recall that $q_v^*(x) = \frac{1}{2}(1 - v'(x))$ and we expect production to fall as reserves decrease. However, Figure 2 shows that due to the upward jump in v' at x_{sat} , for some $x < x_{\text{sat}}$, $x \mapsto q_v^*(x)$ is decreasing. This occurs because the marginal value of reserves is higher in the w -regime than in the v -regime. As (X_t) moves to the left and further away from x_{sat} , the likelihood of imminently discontinuing exploration decreases, so that the marginal value of reserves $v'(X_t)$ shrinks and production is ramped up.

As an illustration of the comparative statistics, we show in Figure 3 the effect of switching costs K and exploration cost \mathcal{C} on the optimal thresholds x_{on} and x_{sat} . We

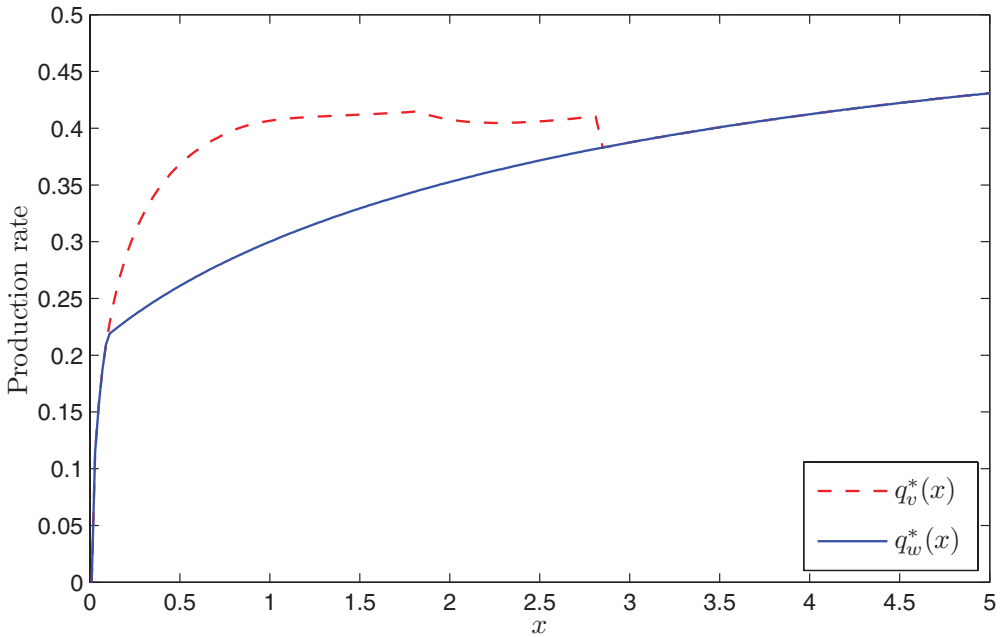


FIGURE 2. (Colour online) Optimal production rates: $q_v^*(x)$ (dash line) and $q_w^*(x)$ (solid line). Note the discontinuity of $q_v^*(x)$ at $x_{\text{sat}} = 2.82$ and the fact that $q_v^*(x)$ is decreasing around $x = 2$.

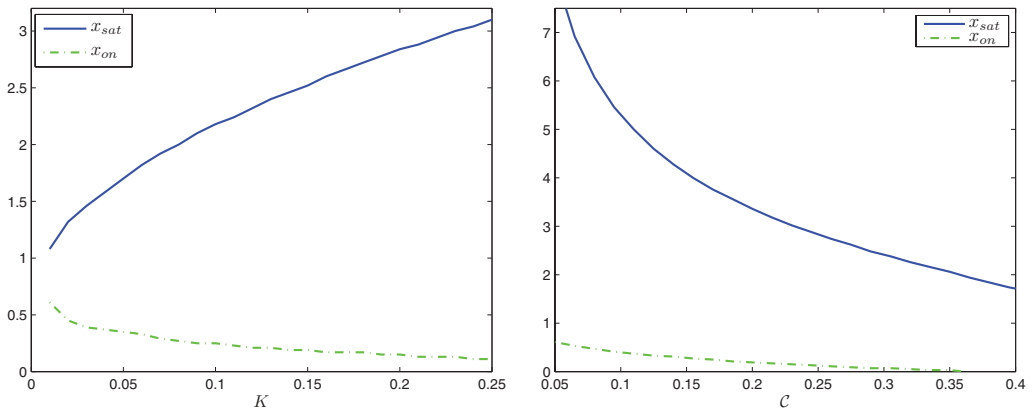


FIGURE 3. (Colour online) Switching thresholds x_{on} and x_{sat} as a function of switching costs K (left panel) and exploration costs \mathcal{C} (right panel). Other parameters are as in Figure 1.

expect that as K increases, the hysteresis interval $[x_{\text{on}}, x_{\text{sat}}]$ becomes wider, since the firm would want to minimise switching frequency. Similarly, if the exploration costs \mathcal{C} decrease, exploration is more profitable and the saturation level x_{sat} rises. The lower threshold x_{on} also rises, since it is not necessary to slow the production rate q_t^* as much before exploration resumes. Note that for $\mathcal{C} > 0.35$, $x_{\text{on}} = 0$ which means that once exploration is stopped, it is never optimal to re-start it since $v(0) < K$.

Finally, we note that the numerical scheme (2.15) can also be interpreted as a Markov chain approximation approach. Namely, let us approximate (X_t) with a controlled discrete-time Markov chain $(X_t^{(S)})$ living on $S = \{0, \Delta x, \dots\}$. Over one time-step Δt , $(X_t^{(S)})$ may move one step to the left (with probability proportional to the production rate q) or jump $N = \delta/\Delta x$ steps to the right (with probability $\lambda\Delta t$). Then, an argument based on Bellman principle implies that the resulting value function $v^{(S)}$ in the exploration regime satisfies

$$v^{(S)}(x) = \max \left(w^{(S)}(x) - K, \sup_{q \geq 0} \mathbb{E} \left[e^{-r\Delta t} v^{(S)}(X_{\Delta t}^{(S)}) + \{q(1 - q) - \mathcal{C}\} \Delta t \right] \right),$$

with $w^{(S)}$ the analogue of w . Assuming that $v^{(S)}(x) > w^{(S)}(x) - K$, and after re-arranging this leads to

$$\frac{1 - e^{-r\Delta t}}{\Delta t} v^{(S)}(x) = \sup_q q(1 - q) + e^{-r\Delta t} q \frac{v^{(S)}(x - \Delta x) - v^{(S)}(x)}{\Delta x} + e^{-r\Delta t} \lambda \Delta v^{(S)}(x) - \mathcal{C},$$

and similarly for $w^{(S)}$, which essentially becomes equation (2.15) for small-enough Δt .

2.3 Continuous exploration control

In the model complementary to equation (2.1), the exploration control a is continuous, $a_t \in \mathbb{R}_+$. Such a model has been studied in detail in [8, 22]. Relative to the analysis of the previous section, this model is both more difficult (since we now have a two-dimensional continuous control to deal with) and simpler (since there are no longer any free boundaries).

We continue to interpret a_t as the effort level; higher effort increases the arrival rate of new discoveries. The point process of discoveries (N_t) is such that

$$\bar{N}_t \triangleq N_t - \int_0^t a_s \lambda ds \tag{2.16}$$

is a martingale. Higher effort also leads to higher costs $\mathcal{C}(a_t)$.

Assumption 1 *The cost function \mathcal{C} is increasing, differentiable and strictly convex with*

$$\mathcal{C}(0) = 0, \quad \text{and} \quad \lim_{a \rightarrow \infty} \mathcal{C}'(a) = +\infty.$$

One economic interpretation of the varying exploration effort in such models is through the spatial distribution of resources in the ground, first described by [1]. Suppose that the non-renewable resource deposits are spatially distributed according to a Poisson random measure ν with constant rate λ . Recall that this means that any region of area $|A|$ contains deposits in the amount $\nu(A) \sim \text{Poisson}(\lambda|A|)$, and that deposits in non-overlapping areas are independent. Exploration now corresponds to *sweeping* through areas searching for deposits, mimicking actual exploration of, e.g., hydrocarbons in different geographical regions. Then, exploration effort is the *speed* of the sweep. Indeed, if the speed is constant at a units of area per unit of time, then new deposits are discovered at rate $a\lambda$. If the

speed is varying over time, the resulting discovery process (N_t) marking discover times has precisely the same distribution as in equation (2.16). The convex costs of exploration then refer to the fact that sweeping areas at higher speeds leads to dis-economy of scale due to scarce exploration resources (e.g. drilling equipment, qualified workforce, etc.).

The overall objective is now written as

$$v(x) := \sup_{q,a} \mathbb{E} \left[\int_0^\infty e^{-rt} \{q_t(1 - q_t) - \mathcal{C}(a_t)\} dt \mid X_0 = x \right], \tag{2.17}$$

and the corresponding Hamilton–Jacobi–Bellman equation for the value function $v(x)$ is

$$\sup_{0 \leq q \leq 1} \{q(1 - q) - qv'\} + \sup_{a \geq 0} \{a\lambda\Delta v(x) - \mathcal{C}(a)\} - rv = 0. \tag{2.18}$$

The assumptions on $\mathcal{C}(\cdot)$ ensure that there is a unique maximiser a^* to the second supremum, which is characterised by the first-order condition

$$a^*(x) = (\mathcal{C}')^{-1}(\lambda\Delta v(x)). \tag{2.19}$$

At $x = 0$, production must be suspended but exploration continues. Since the controls are Markov, until the next discovery time σ_1 the exploration level is constant. It follows that the boundary condition at $x = 0$ is given by (cf. [22])

$$v(0) = \sup_{a \geq 0} \mathbb{E} \left[e^{-r\tau} v(\delta) - \int_0^\tau e^{-rt} \mathcal{C}(a) dt \right] = \sup_{a \geq 0} \frac{\lambda av(\delta) - \mathcal{C}(a)}{\lambda a + r}. \tag{2.20}$$

We remark that if $\mathcal{C}'(0) > 0$ and r or $\mathcal{C}(\cdot)$ are too large, it is possible that the optimiser is $a^*(0) = 0$ in which case it is optimal to not explore at $x = 0$ and $v(0) = 0$. Recall that $a^*(x)$ is decreasing, so that if $a^*(0) = 0$, then $a^*(x) = 0$ for all x and the model reduces to the classical case with no exploration: $v(x) = w_0(x)$ given in (2.5).

Since equation (2.18) no longer has any free boundaries and the two controls q and a are separable, this problem is more analytically tractable than the switching version in Section 2.1. We will use equation (2.17) as the prototype for the game model in Section 3. Since the solution of equation (2.18) will be seen as a special case of the more general model in Section 3, we keep the discussion of equation (2.18) brief and only make a few remarks here about the resulting exploration control a^* .

Because $v(x)$ is concave (this is easily verified by showing that the marginal value of additional inventory is decreasing), we have from equation (2.19) that $a^*(x)$ is decreasing in x . Moreover, the uniform bound $v(x) \leq (4r)^{-1}$ implies $a^*(x) \leq \bar{a} := (\mathcal{C}')^{-1}(\lambda(4r)^{-1})$. At the same time, if $\mathcal{C}'(0) = 0$, then $a^*(x) > 0$ for all x . Indeed, in this case, for $\epsilon > 0$ small enough, the cost of exploration $\mathcal{C}(\epsilon)$ is negligible compared to the expected gain $\simeq \epsilon r^{-1} \Delta v(x)$.

Remark 2 We observe that if $a^* > 0$ for all x , then (X_t^*) is an irreducible Markov chain on \mathbb{R}_+ and therefore is recurrent on its full state space. It follows that $\sup_t X_t^* = +\infty$ and reserves will become arbitrarily large infinitely often. Such behaviour would appear to be unrealistic for describing non-renewable resources, and suggests that we should take

$\mathcal{C}'(0) > 0$. If $\mathcal{C}'(0) > 0$, then there exists a saturation level x_{sat} such that $a^*(x) = 0$ for $x > x_{\text{sat}}$, and (X_t^*) would be positive recurrent on $[0, x_{\text{sat}} + \delta)$ only.

3 Duopoly with a green producer

We now return to a competitive situation and analyse the dynamic Nash equilibrium when the firm with a depleting non-renewable resource is competing against a second producer. The other producer has inexhaustible reserves, but higher cost of production. This corresponds to sustainable production from ‘green’ sources (e.g. solar energy, wind power). Thus, producer 1 faces exhaustion and therefore needs exploration; producer 2 always has infinite resources, but also positive fixed production costs $c > 0$. The two producers compete against each other through the Cournot game equilibrium. This setup corresponds to the ‘axis-game’ in [13, Section 5.1], but with potential for discovery by player 1. A related problem in the context of mean-field games with convex costs of extraction, but without exploration, is analysed in [12].

The remaining reserves (X_t) of Player 1 follow, similar to (2.2):

$$dX_t = -q_1(X_t) \mathbf{1}_{\{X_t > 0\}} dt + \delta dN_t, \tag{3.1}$$

where q_1 is his rate of production, and (N_t) is a controlled point process with intensity λa_t . Here, Player 1 has exhaustible oil resources which he extracts at zero cost, and which is subject to random discoveries. Player 1 stimulates discoveries via a continuous exploration effort $a_t \in \mathbb{R}_+$ that is subject to cost $\mathcal{C}(a_t)$, where \mathcal{C} satisfies Assumption 1. When he runs out, he cannot produce until there is a discovery, and so we constrain $q_1(0) = 0$. Player 2 produces from an inexhaustible resource (e.g. solar) which is expensive: his marginal cost of production is $c > 0$. His rate of production is $q_2(X_t)$.

The value functions for each player are, respectively (where g stands for ‘green’):

$$v(x) = \sup_{q_1, a} \mathbb{E} \left[\int_0^\infty e^{-rt} [q_1(X_t)(1 - q_1(X_t) - q_2(X_t)) - \mathcal{C}(a_t)] dt \mid X_0 = x \right],$$

$$g(x) = \sup_{q_2 \geq 0} \mathbb{E} \left[\int_0^\infty e^{-rt} q_2(X_t)((1 - q_1(X_t) - q_2(X_t)) - c) \mathbf{1}_{\{X_t > 0\}} dt + \int_0^\infty e^{-rt} \frac{1}{4} (1 - c)^2 \mathbf{1}_{\{X_t = 0\}} dt \mid X_0 = x \right].$$

The second term in the definition of g expresses that player 2 has a monopoly while player 1 is out of reserves. We also stress that q_2 must be non-negative: for large enough c and x , we will see that player 2 is *blockaded* in that his cost of production is so high and his competitor’s reserves of the cheaper resource are so plentiful that his equilibrium strategy is not to produce anything until player 1 has run down his reserves some more. In this situation, player 1 has a temporary monopoly while player 2 sits it out.

The ODEs for v and g are

$$\sup_{q_1, a} \{(1 - q_1 - q_2^*)q_1 - q_1 v'(x) - \mathcal{C}(a) + a\lambda \Delta v(x)\} - rv(x) = 0,$$

$$\sup_{q_2 \geq 0} \{(1 - q_1^* - q_2 - c)q_2\} - q_1^* g'(x) + a^*(x)\lambda \Delta g(x) - rg(x) = 0,$$

where $\Delta v(x)$ is the non-local or jump term given in equation (2.8), and similarly $\Delta g(x)$, and $a^*(x) = \operatorname{argsup}_{a \geq 0} \{-\mathcal{C}(a) + a\lambda\Delta v(x)\}$ is the optimal exploration effort.

Motivated by the discussion of Remark 2 in Section 2.3, we take power costs

$$\mathcal{C}(a) = \frac{1}{\beta} a^\beta + \kappa a, \quad \beta > 1, \kappa \geq 0. \tag{3.2}$$

Note that $\mathcal{C}'(0) = \kappa$, so $\kappa > 0$ corresponds to the existence of a finite saturation point x_{sat} . The optimal effort is given by

$$a^*(x) = [(\lambda\Delta v(x) - \kappa)^+]^{\gamma-1},$$

where $z^+ = \max(z, 0)$ and

$$\gamma = \frac{\beta}{\beta - 1}. \tag{3.3}$$

If at reserve level $x > 0$, player 2 is *not* blockaded, the feedback production strategies are given by

$$q_1^*(x) = \frac{1}{3}(1 - 2v'(x) + c) \quad q_2^*(x) = \frac{1}{3}(1 + v'(x) - 2c), \quad x > 0, \tag{3.4}$$

and the ODEs become

$$\begin{aligned} \frac{1}{9}(1 - 2v' + c)^2 + \frac{1}{\gamma} [(\lambda\Delta v(x) - \kappa)^+]^\gamma - rv &= 0, \\ \frac{1}{9}(1 + v' - 2c)^2 - \frac{1}{3}(1 - 2v' + c)g' + \frac{1}{\gamma} [(\lambda\Delta v(x) - \kappa)^+]^{\gamma-1} \Delta g(x) - rg &= 0. \end{aligned} \tag{3.5}$$

If player 2 is blockaded at some $x > 0$, we have $q_2^* = 0$ and $q_1^* = \frac{1}{2}(1 - v')$ and so

$$\begin{aligned} \frac{1}{4}(1 - v')^2 + \frac{1}{\gamma} [(\lambda\Delta v(x) - \kappa)^+]^\gamma - rv &= 0, \\ -\frac{1}{2}(1 - v')g' + \frac{1}{\gamma} [(\lambda\Delta v(x) - \kappa)^+]^{\gamma-1} \Delta g(x) - rg &= 0. \end{aligned} \tag{3.6}$$

When $c = 1$, player 2 is never in the game, and the above model reduces to the monopoly setting of Section 2.3.

Using the same argument as for equation (2.20), the boundary conditions are

$$v(0) = \sup_a \frac{a\lambda v(\delta) - \mathcal{C}(a)}{\lambda a + r}, \quad g(0) = \frac{(1 - c)^2/4 + \lambda a^*(0)g(\delta)}{\lambda a^*(0) + r}.$$

The far-field limits are

$$x \rightarrow \infty : \quad v \rightarrow \frac{1}{9r}(1 + c)^2, \quad g \rightarrow \frac{1}{9r}(1 - 2c)^2,$$

if $c < 1/2$ and $v \rightarrow (4r)^{-1}$, $g \rightarrow 0$ if $c > 1/2$. Observe that while as before $v(x)$ is increasing, $g(x)$ is decreasing since more reserves lower the game value for the green producer.

We observe that the above systems yield an independent equation for v similar to equation (2.18), and then a first-order linear delay-ODE for g in terms of v . However, the equilibrium production and exploration strategies depend only on v , and so we will not deal with the g equation.

3.1 Case of no exploration

We firstly consider the case of no exploration possibility, namely $\lambda = 0$. This resembles the axis game considered in [13], however there it was assumed player 1 becomes an inexhaustible ‘green’ producer upon his reserves being exhausted, whereas here he simply quits altogether. We denote player 1’s value function by v_0 , and the two firms’ equilibrium strategies by $(q_1^{(0)*}, q_2^{(0)*})$.

Proposition 3.1 *For $c \leq \frac{1}{2}$, there is no blockading and the no-exploration value function $v_0(x)$, solution of equation (3.5) with $\lambda = 0$ and boundary condition $v_0(0) = 0$, is given by*

$$v_0(x) = \frac{(1 + c)^2}{9r} (1 + \mathbf{W}(\theta(x)))^2, \tag{3.7}$$

where

$$\theta(x) = -\exp(-krx - 1) \in [-e^{-1}, 0), \quad k := \frac{9}{4(1 + c)}. \tag{3.8}$$

The equilibrium production rates are given by

$$q_1^{(0)*}(x) = \frac{(1 + c)}{3} (1 + \mathbf{W}(\theta(x))), \tag{3.9}$$

$$q_2^{(0)*}(x) = \frac{1}{3} \left(1 - 2c - \frac{1}{2}(1 + c)\mathbf{W}(\theta(x)) \right). \tag{3.10}$$

Proof The solution to the ODE follows easily from Lemma 2.1. It remains to check there is no blockading. The formulas (3.9) and (3.10) are obtained by substituting for v'_0 into equation (3.4) using

$$v'_0 = \frac{1}{2} (1 + c - 3\sqrt{rv_0}) = -\frac{1}{2}(1 + c)\mathbf{W}(\theta(x)).$$

Since $q_2^{(0)*} = \frac{1}{3}(1 - 2c + v'_0)$, the duopoly production rate becomes negative only if $v'_0 < 2c - 1 < 0$. But it is easy to see that v_0 is increasing, strictly concave with $v'_0(0) = \frac{1}{2}(1 + c)$ and $v'_0(\infty) = 0$. Therefore, $q_2^{(0)*} > 0$ for all $x \geq 0$. \square

For $c > \frac{1}{2}$, blockading does occur for large enough x .

Proposition 3.2 *When $c > \frac{1}{2}$, player 2 is blockaded for $x \geq x_b$, where*

$$x_b = \frac{4}{9r} \left((1 + c) \log \left(\frac{1 + c}{2(2c - 1)} \right) - 3(1 - c) \right). \tag{3.11}$$

For $x \in [0, x_b]$, the value function $v_0(x)$ is given by equation (3.7) and the equilibrium strategies $(q_1^{(0)*}, q_2^{(0)*})$ by equations (3.9) and (3.10), and for $x \geq x_b$ by

$$\begin{aligned}
 v_0(x) &= \frac{1}{4r} (1 + \mathbf{W}(\theta(x - x_b)))^2, & \theta(x) &= (1 - 2c)e^{1-2c-2rx}, \\
 q_1^{(0)*}(x) &= \frac{1}{2} (1 + \mathbf{W}(\theta(x - x_b))), \\
 q_2^{(0)*}(x) &= 0.
 \end{aligned}
 \tag{3.12}$$

Proof Using the no-blockade solution (3.7) in the formula $q_2^{(0)*} = \frac{1}{3}(1 - 2c + v'_0)$ shows that $q_2^{(0)*} = 0$ at $x = x_b$ defined by equation (3.11), and $q_2^{(0)*} < 0$ for $x > x_b$. So, for $x \geq x_b$, player 1 has a monopoly: $q_2^{(0)*} = 0$ and $q_1^{(0)*} = \frac{1}{2}(1 - v'_0)$. It follows from these that at x_b , $v'_0(x_b) = 2c - 1$ and $q_1^{(0)*}(x_b) = 1 - c$, so $v_0(x_b) = \frac{1}{r}(1 - c)^2$. Solving (3.6) with $\lambda = 0$ on $\{x > x_b\}$ and this boundary condition gives equation (3.12). \square

3.2 Small discovery rate expansion

We consider now the situation when λ is small, that is low success probability of discovery, and derive a correction to the no exploration case of Section 3.1 in the limit $\lambda \downarrow 0$. Recall from equation (3.2) that $\mathcal{C}'(0) = \kappa$. For analytical tractability in this section, we take $\kappa = 0$. We return to the more realistic $\kappa > 0$ in the numerical computations in Section 3.3.

We construct an expansion

$$v = v_0 + \lambda^\gamma v_1 + \dots, \tag{3.13}$$

where γ was defined in (3.3). Then, $v_0(x)$ is the value function with no discovery given in Propositions 3.1 and 3.2.

Proposition 3.3 *If there is no blockading at x , we have*

$$\begin{aligned}
 v_1(x) &= |\mathbf{W}(\theta(x))| \frac{(1+c)^2}{9r^2\gamma} (1 + \mathbf{W}(\theta(\delta)))^2 + k|\mathbf{W}(\theta(x))| \int_0^x \frac{F(\eta)}{|\mathbf{W}(\theta(\eta))|(1 + \mathbf{W}(\theta(\eta)))} d\eta, \\
 F(\eta) &= \frac{1}{\gamma} \left(\frac{(1+c)^2}{9r} ([1 + \mathbf{W}(\mu\theta(\eta))]^2 - [1 + \mathbf{W}(\theta(\eta))]^2) \right)^\gamma,
 \end{aligned}
 \tag{3.14}$$

where $\mu := e^{-\delta/(kr)}$, γ is given by equation (3.3), and $k, \theta(x)$ are defined in equation (3.8).

Proof Inserting equation (3.13) into equation (3.5) and collecting terms in λ^γ gives that the correction term $v_1(x)$ solves

$$q(x)v'_1 + rv_1 = F(x), \tag{3.15}$$

where

$$F(x) = \frac{1}{\gamma} (\Delta v_0(x))^\gamma, \quad q(x) = \frac{4}{3} q_1^{(0)*}(v'_0(x)) = k^{-1} (1 + \mathbf{W}(\theta(x))).$$

The boundary condition is $v_1(0) = \frac{v_0(\delta)^\gamma}{\gamma r}$, which follows from substituting $a^*(0) = (\lambda v_0(\delta))^{\gamma-1}$

into equation (2.20) and matching terms in λ^γ . To solve equation (3.15), let

$$Q(x) = \int_0^x \frac{du}{q(u)} = k \int_0^x \frac{du}{(1 + \mathbf{W}(\theta(u)))}.$$

Making the change of variable $z = \mathbf{W}(\theta(u))$ and using that $\mathbf{W}'(y) = \mathbf{W}(y)/(y(1 + \mathbf{W}(y)))$, gives $Q(x) = -\frac{1}{r} \log |\mathbf{W}(\theta(x))|$. Then, the transformation $v_1(x) = e^{-r\xi} u(\xi)$, where $\xi = Q(x)$, leads to $u' = e^{r\xi} F(Q^{-1}(\xi))$, with initial condition $u(0) = (r\gamma)^{-1} v_0(\delta)^\gamma$. Solving and restoring the transformation yields

$$v_1(x) = \frac{e^{-rQ(x)}}{\gamma r} v_0(\delta)^\gamma + e^{-rQ(x)} \int_0^x e^{rQ(\eta)} F(\eta) \frac{d\eta}{q(\eta)}, \tag{3.16}$$

which leads to the formulas (3.14). □

The expression in Proposition 3.3 can be used to find small λ approximations to the optimal production rates q_1^* and q_2^* , and the blockade point x_b . For example, when there is no blockading,

$$q_1^* = \frac{1}{3}(1 + c - 2v') \approx q_1^{(0)*} - \frac{2}{3} \lambda^\gamma v_1',$$

where the first term is simply the no-exploration production rate given in equation (3.9), and the second term can be written in terms of v_0 and v_1 using the ODE (3.15). Similarly, expanding the blockading point

$$x_b(\gamma) = x_b + \frac{\lambda^\gamma}{\gamma} x_1 + \dots,$$

where x_b is the no-exploration blockade point given in equation (3.11), leads to

$$x_1 = -\frac{v_1'(x_b)}{v_0''(x_b)}.$$

Finally, the exploration effort may be expanded to yield $a^*(x) \simeq \lambda^{\gamma-1} (\Delta v_0(x))^{\gamma-1} + o(\lambda^\gamma)$.

Figure 4 compares the asymptotic correction terms derived above to the ‘exact’ values (see Section 3.3 for how these are iteratively approximated) and shows that they are quite accurate even for moderate $\lambda = 0.5$.

Remark 3 A similar analysis shows that the monopoly value function v defined in equation (2.17) of Section 2.3 can be approximated by $v(x) = w_0(x) + \lambda^\gamma v_1(x) + \dots$, where w_0 is the no-exploration monopoly value function given in (2.5), and

$$v_1(x) = |\mathbf{W}(\theta(x))| \left(\frac{(w_0(\delta))^\gamma}{\gamma r} + \int_0^x \frac{2F(\eta) d\eta}{|\mathbf{W}(\theta(\eta))|(1 + \mathbf{W}(\theta(\eta)))} \right).$$

Here, F and θ are $F(\eta) = \frac{1}{\gamma(4r)^\gamma} ([1 + \mathbf{W}(e^{-2r\delta}\theta(\eta))]^2 - [1 + \mathbf{W}(\theta(\eta))]^2)^\gamma$ and $\theta(x) = -e^{-2rx-1}$. The optimal monopoly production rate is approximated by $q^*(x) = \frac{1+\mathbf{W}(\theta(x))}{2} + \lambda^\gamma q_1(x) + \dots$, where

$$q_1(x) = -\frac{1}{2} v_1'(x) = \frac{rv_1(x) - F(x)}{1 + \mathbf{W}(\theta(x))}.$$

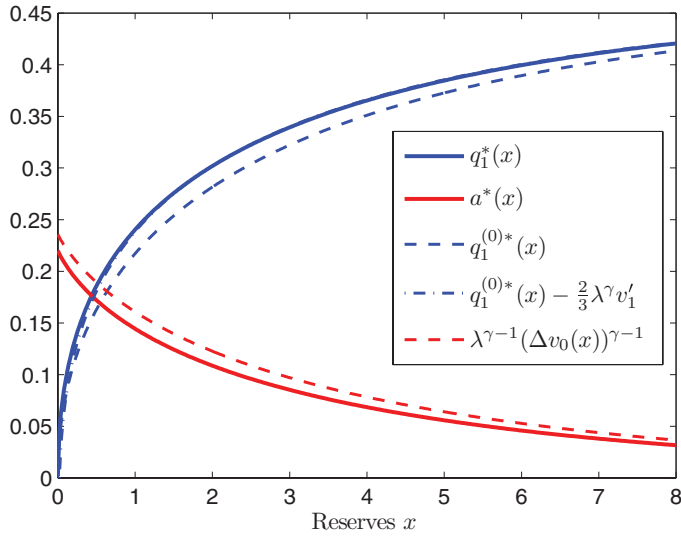


FIGURE 4. (Colour online) Optimal production and exploration controls of the exhaustible player. We take $\delta = 1, r = 0.1, \lambda = 0.5, \mathcal{C}(a) = a^2/2$ and $c = 0.4$. The dashed lines show the approximate solutions obtained from first-order asymptotic expansion with respect to λ . Note that $q_1^{(0)*} - \frac{2}{3}\lambda^\gamma v_1'$ is essentially identical to q_1^* for $x > 1$.

For the monopoly problem, a singular perturbation analysis in the regime of small discovery size at high-frequency is given in [15].

3.3 Computational analysis

Consider again the ODEs in equations (3.5) and (3.6). These may be combined in the form

$$(q_1^*)^2 - rv + \frac{1}{\gamma}((\lambda\Delta v - \kappa)^+)^{\gamma} = 0, \quad q_1^*(x) = \max\left(\frac{1}{3}(1 - 2v'(x) + c), \frac{1}{2}(1 - v'(x))\right).$$

Observe that numerically solving for $v(x)$ is challenging due to the implicit boundary condition and the presence of a ‘forward’ delay term on the semi-infinite domain \mathbb{R}_+ . We resolve this difficulty by using an iterative scheme.

Let $v^0(x) = v_0(x)$ be the no-exploration value function given in Proposition 3.2, and for $n \geq 1$, define inductively

$$(q_1^n(x))^2 - rv^n + \frac{1}{\gamma}((\lambda(v^{n-1}(x + \delta) - v^n(x)) - \kappa)^+)^{\gamma} = 0, \quad q_1^n(x) = \max\left(\frac{1}{3}(1 - 2(v^n)'(x) + c), \frac{1}{2}(1 - (v^n)'(x))\right), \tag{3.17}$$

with boundary condition

$$v^n(0) = \sup_{a \geq 0} \frac{\lambda a v^{n-1}(\delta) - \mathcal{C}(a)}{\lambda a + r}. \tag{3.18}$$

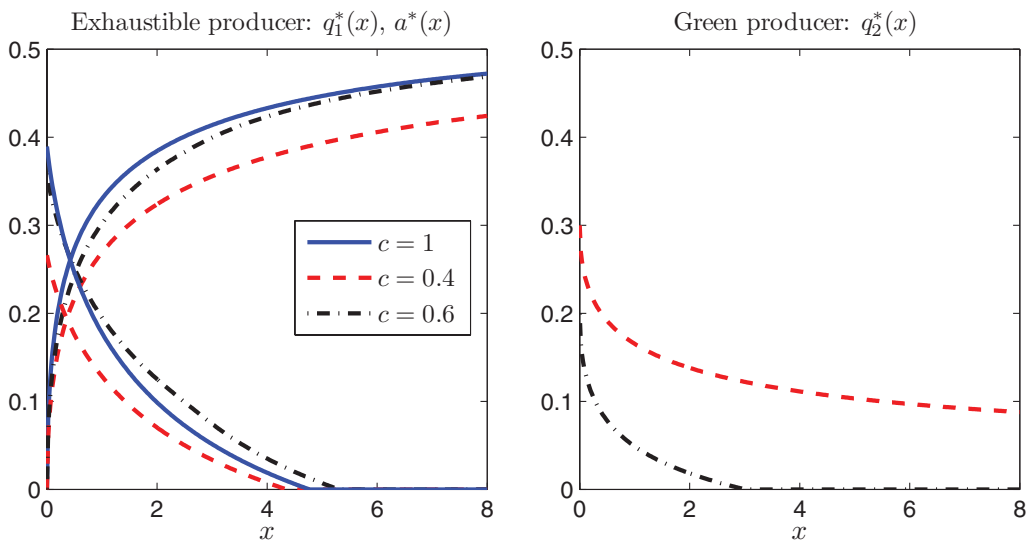


FIGURE 5. (Colour online) Left panel: optimal production rate $q_1^*(x)$ (increasing) and optimal exploration level $a^*(x)$ (decreasing) of the exhaustible producer. Right panel: production rate $q_2^*(x)$ of the green producer. The parameters are $\delta = 1, \lambda = 1, r = 0.1, \mathcal{C}(a) = 0.1a + a^2/2$ and three different levels of c . Note the blockading region for $c = 0.6$ and $x > 3.4$.

Observe that equation (3.17) is a standard first-order nonlinear ordinary differential equation with ‘source’ term $v^{n-1}(\cdot + \delta)$ and can be solved using standard tools, such as Runge–Kutta methods. Using \tilde{L} to denote the differential-difference operator on the right-hand-side of equation (3.17), we have $v^n = \tilde{L}(v^{n-1})$. We now expect that $v^n(x) \rightarrow v(x)$, so that v is a fixed point of \tilde{L} . In fact, based on control arguments, this convergence is expected to be exponential in n uniformly in x . A justification is given below for the monopoly case.

Figure 5 illustrates the numerical solution for three different values of the green production costs c . We observe that while $q_1^*(x)$ is monotone decreasing in c , the effect on $a^*(x)$ is ambiguous. Further discussion is given in Section 3.3.2.

3.3.1 Convergence of numerical method for monopoly problem

In this subsection, v refers to the value function of the monopoly problem defined in equation (2.17) with the solution given in equation (2.18). The inductive approximation is defined by

$$\frac{1}{4}(1 - (v^n)'(x))^2 - rv^n(x) + \frac{1}{\gamma} \{(\lambda(v^{n-1}(x + \delta) - v^n(x)) - \kappa)^+\}^\gamma = 0,$$

with boundary condition (3.18).

Let us recall the following characterisation for equation (3.17) in the monopoly setup of Section 2.3:

Lemma 3.1 Let σ_n be the n th discovery epoch with $\sigma_0 = 0$ and $w_0(x)$ be given by equation (2.5). Let

$$V^n(x) := \sup_{q_1, a} \mathbb{E} \left[\int_0^{\sigma_n} e^{-rt} [(1 - q_1(X_t))q_1(X_t) - \mathcal{C}(a_t)] dt + e^{-r\sigma_n} w_0(X_{\sigma_n}) \mid X_0 = x \right]. \tag{3.19}$$

Then, we have $v^n(x) = V^n(x)$ for all $n \geq 0$ and there exists $p < 1$ such that

$$\sup_{x \geq 0} |v^n(x) - v(x)| \leq p^n.$$

Sketch of Proof Usual dynamic programming arguments and the nature of Markov feedback controls imply that

$$V^n(x) = \sup_{q_1, a} \mathbb{E} \left[\int_0^{\sigma_1} e^{-rt} [(1 - q_1(X_t))q_1(X_t) - \mathcal{C}(a_t)] + e^{-r\sigma_1} V^{n-1}(X_{\sigma_1-} + \delta) \mid X_0 = x \right],$$

where σ_1 is the first arrival time of (N_t) . The form of the infinitesimal generator of the resulting (X_t^*) implies that V^n satisfies equation (3.17). Similar to equation (2.20), the boundary conditions at $x = 0$ for equation (3.19) must be

$$V^n(0) = \sup_{a \geq 0} \mathbb{E} \left[\int_0^{\sigma_1} -e^{-rt} \mathcal{C}(a) dt + e^{-r\sigma_1} V^{n-1}(\delta) \right],$$

leading to equation (3.18). It immediately follows that $V^n(x)$ is monotone increasing in x and bounded by $(4r)^{-1}$. Moreover, as before, the respective exploration controls must be bounded $a^{(n)*}(x) \leq \bar{a}$. Hence, $\mathbb{E}[\sigma_n] \geq n(\lambda\bar{a})^{-1}$ and therefore $\sigma_n \rightarrow \infty$ uniformly. Consequently, the sequence V^n converges to

$$\lim_{n \rightarrow \infty} V^n(x) = \sup_{q_1, a} \mathbb{E} \left[\int_0^{\infty} e^{-rt} [(1 - q_1(X_t))q_1(X_t) - \mathcal{C}(a_t)] dt \mid X_0 = x \right] = v(x),$$

and the error can be controlled exponentially via

$$\begin{aligned} \sup_{x \geq 0} |V^n(x) - v(x)| &\leq \mathbb{E} \left[\int_{\sigma_n}^{\infty} e^{-rs} q_1(X_s)(1 - q_1(X_s)) ds \mid X_0 = x \right] \\ &\leq \mathbb{E} \left[\frac{e^{-r\sigma_n}}{r} v(X_{\sigma_n}) \right] \leq \left(\frac{\lambda\bar{a}}{r + \lambda\bar{a}} \right)^n \frac{1}{4r^2}. \end{aligned}$$

Finally, the first term in the sequence V^n is the no-exploration solution $V^0(x) \equiv w_0(x) = v^0(x)$ and by induction we therefore obtain that $V^n = v^n \forall n$. □

As a corollary, we obtain that v is the smallest *fixed point* of the functional operator \tilde{L} . A similar result in a related model was given in [9, Theorem 2] and the spirit goes back to [6].

In the game setting, we analogously expect that v^n is a game value corresponding to Nash equilibrium on the restricted horizon $[0, \sigma_n]$. The full treatment of such a construction, including proof of the convergence to a Nash equilibrium on the infinite

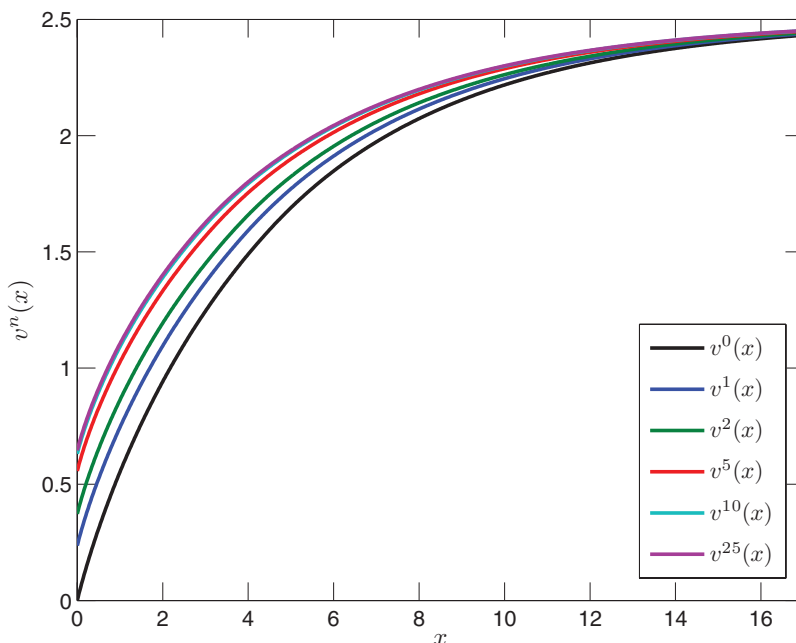


FIGURE 6. (Colour online) Convergence of v^n as $n \rightarrow \infty$ for the axis game. We take $\delta = 1$, $\lambda = 0.1$, $r = 0.1$, $\mathcal{C}(a) = 0.1a + a^2/2$ and $c = 0.6$.

horizon is beyond the scope of this paper (note that in the general game setting, the saturation and blockading regions $[x_{\text{sat}}, \infty)$ and $[x_b, \infty)$ will be also indexed by n). An example of such a computation is shown in Figure 6 where we see that the convergence is indeed very rapid (in particular, the difference $\|v^{10}(x) - v^{25}(x)\|_\infty$ is negligible).

3.3.2 Blockading and saturation levels

It is of great interest to understand how the presence of the green producer alters the behaviour of the exhaustible player. As expected, increased competition will reduce the value of the reserves and increase their marginal value, hence reducing production rates. Thus, v and q_1 are increasing in c (recall that $c = 1$ corresponds to the monopoly and $c = 0$ means the exhaustible producer is permanently undercut by the equally cheap sustainable source).

However, the impact on the exploration rate a^* is ambiguous, since while v is decreased, v' is generally increased. Numerical evidence (cf. Figure 5) shows that (i) for small values of c , increased competition reduces the exploration effort for all x ; (ii) moderate values of c increase exploration efforts for moderate x and lower them for x close to zero; (iii) for c close to 1, exploration effort is again reduced everywhere. Thus, the impact of competition has a humped shape in terms of c . The above behaviour is driven by the different market leaders in different scenarios. For small c , the green producer is the effective leader in the market, causing significant losses for the exhaustible producer who gives up and reduces efforts. For large c , the exhaustible producer is the effective

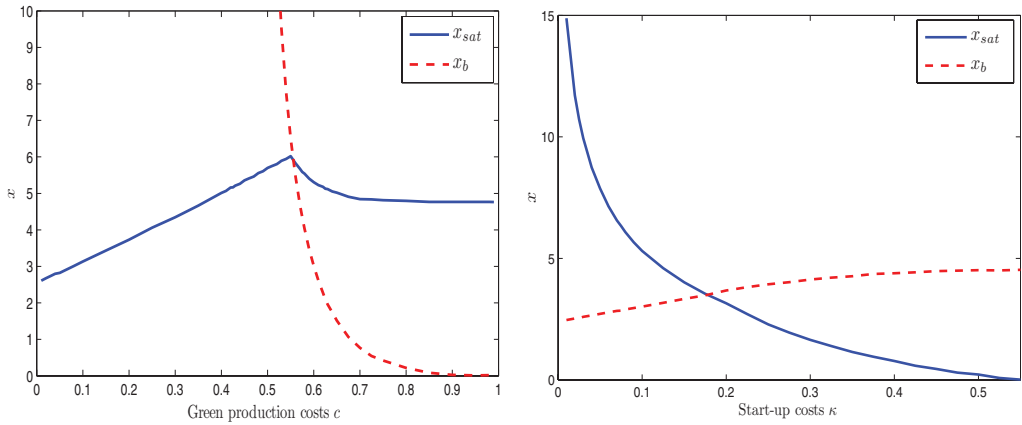


FIGURE 7. (Colour online) Left panel: x_{sat} and x_b as a function of green producer production costs c . Recall that for $c < 0.5$, $x_b = +\infty$. Right panel: x_{sat} and x_b as a function of start-up exploration costs κ . We take $\delta = 1$, $\lambda = 1$, $r = 0.1$ and $\mathcal{C}(a) = \kappa a + a^2/2$ with default values of $c = 0.6, \kappa = 0.1$.

leader and the green producer only has a small marginal negative impact. However, for moderate c , the exhaustible (*resp.* green) producer is the leader for large (*resp.* small) reserves levels. Thus, when x is close to zero, the exhaustible producer is discouraged and lowers exploration; however, when x is moderate, he is encouraged to put in extra effort to stay in front of the green producer (formally the marginal value of reserves is so high as to justify higher exploration).

Strikingly, this behaviour is also true of the saturation level x_{sat} . This effect is illustrated in the left panel of Figure 7 where we see that $c \mapsto x_{\text{sat}}(c)$ is maximised around $c = 0.55$. Note that for $c > 0.6$, $x_{\text{sat}} > x_b$, and so, compared to the monopoly case $c = 1$, the exhaustible producer will start exploration earlier to stave off competition even before the green producer enters the market. The curious shape of $x_{\text{sat}}(c)$ exhibits a ‘cusp-like’ behaviour around its maximum, though numerical experiments suggest that $x_{\text{sat}}(c)$ is in fact not quite so sharp upon zooming in.

In the right panel of Figure 7, we show the impact of the start-up exploration costs κ on x_{sat} and x_b . As expected, higher exploration costs discourage exploration, i.e. lower x_{sat} . For κ large enough, $\lambda \Delta v(0) < \kappa$ and $a^*(x) \equiv 0$, so no exploration ever takes place. Conversely, higher exploration costs benefit the green producer who can enter the game earlier, so that x_b is increasing (moderately) in κ . For κ large, $x_b \simeq 4.52$ is given from equation (3.11).

Figure 8 shows a sample path for the evolution of the game solution over time. The system state is described by (X_t^*) in the top-left panel which drives the feedback controls $q_1^*(X_t^*), a^*(X_t^*)$ and $q_2^*(X_t^*)$ in the other panels. In the given example, $c = 0.65$ and blockading occurs for $X_t^* \geq x_b \simeq 1.65$. Moreover, the saturation point is $x_{\text{sat}} = 5.02$. The recurrent behaviour of (X_t^*) is clearly visible, as the resource is repeatedly exhausted until a new discovery replenishes reserves and allows to re-start production. A string of discoveries leads to blockading of the green producer, see e.g., $t \in [42, 47]$. Higher reserves lower exploration rates and increase production. The bottom-right panel in Figure 8 shows the trajectory of the market price $P(t) = 1 - q_1^*(X_t) - q_2^*(X_t)$, where we observe first-order

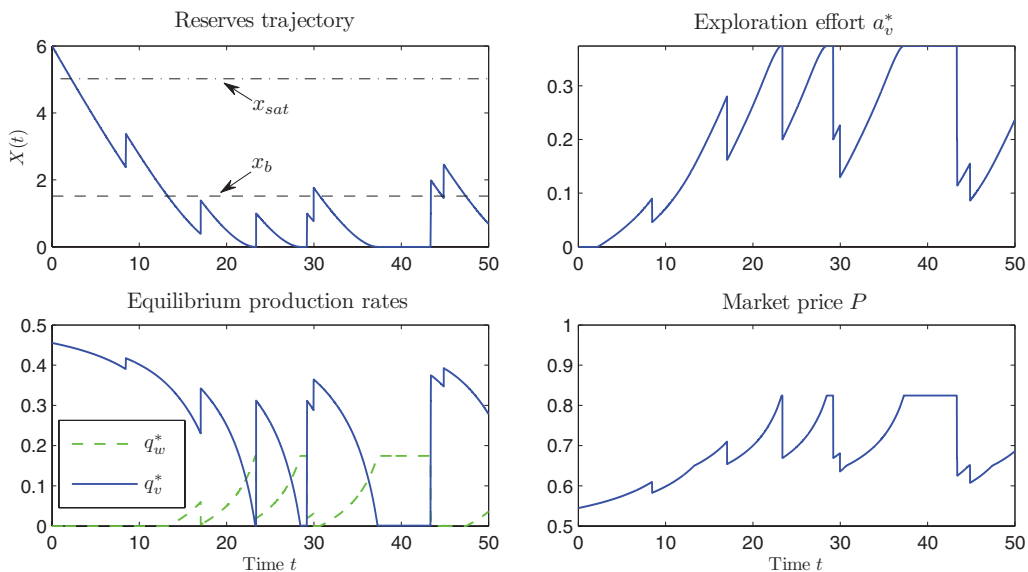


FIGURE 8. (Colour online) Trajectory of the game solution over time. Top left panel: reserves (X_t^*) of the exhaustible player; top right panel: exploration rate $a^*(X_t^*)$. Bottom left panel: production rate $q_1^*(X_t^*)$ of the exhaustible player (solid blue) and $q_2^*(X_t^*)$ of the green player (dashed green); note intermittent blockading when $X_t^* \geq x_b \simeq 1.65$. Bottom right panel: market price $P(\bar{q}) = 1 - q_1 - q_2$. We take $\delta = 1, \lambda = 1, r = 0.1, \mathcal{C}(a) = 0.1a + a^2/2$ and $c = 0.65$.

discontinuities as (X_t^*) crosses downwards the blockading point x_b around $t = 13$ and $t = 47$.

3.4 Marginal value of reserves dynamics

To highlight another key difference with the monopoly model, we now investigate the behaviour of marginal value of reserves v' along trajectories of (X_t^*) . In the monopoly setting, the marginal value of additional reserves v' has a natural interpretation as the scarcity value of the resource. We recall the stochastic Hotelling theorem of [8] which studies the evolution of this value over time. Let

$$\mathcal{D}v'(x) := \lim_{h \rightarrow 0} \frac{\mathbb{E} [v'(X_h^*) \mid X_0^* = x] - v'(x)}{h}, \tag{3.20}$$

be the time derivative of the shadow cost along trajectories of optimally managed reserves paths. [8, Theorem 4] shows that for the model of Section 2.3,

$$\mathcal{D}v'(x) = rv'(x) \quad \forall x \in \mathbb{R}_+,$$

in other words the expected growth rate of prices is equal to the discount rate r .

In the game setting, we have the following.

Proposition 3.4 *Let \mathcal{D} be as in equation (3.20). Then,*

$$\mathcal{D}v'(x) = \begin{cases} rv'(x) + q_1^*(x) \frac{\partial}{\partial x} q_2^*(x), & \text{if } x < x_b \wedge x_{\text{sat}}, \\ \frac{3}{4}rv'(x), & x_{\text{sat}} < x < x_b, \\ rv'(x), & x > x_b. \end{cases} \tag{3.21}$$

Proof Let \mathcal{S} be the infinitesimal generator of the Markov process (X_t^*) . Then, for any function $f \in \text{Dom}(\mathcal{S})$,

$$\mathcal{S}f(x) = \lambda a^*(x)\Delta f(x) - q_1^*(x)f'(x). \tag{3.22}$$

Conversely, general theory gives the analytic characterisation $\mathcal{S} \equiv \mathcal{D}$ as the time derivative along (X_t^*) in equation (3.20). Arguments similar to those in [8, Lemma 7] imply that v' is continuous and bounded. It follows that q_1^* is continuous and differentiable on \mathbb{R}_+ except at zero and x_b . Similarly, a^* is bounded, continuous and differentiable everywhere except possibly at x_{sat} . The implicit function theorem then implies that v'' exists a.e., and v' is in the domain of the operator \mathcal{S} .

Now starting with

$$rv(x) = q_1^*(x)(1 - q_1^*(x) - q_2^*(x)) - q_1^*(x)v'(x) - \mathcal{C}(a^*(x)) + \lambda a^*(x)\Delta v(x),$$

and differentiating we find

$$rv'(x) = -q_1^*(x)v''(x) + \lambda a^*(x)\Delta v'(x) - \frac{\partial}{\partial x} q_2^*(x) \cdot q_1^*(x) + \frac{\partial}{\partial x} q_1^*(x) [1 - 2q_1^*(x) - q_2^*(s) - v'(x)] + \frac{\partial}{\partial x} a^*(x) [-\mathcal{C}'(a^*(x)) + \lambda \Delta v(x)]. \tag{3.23}$$

The terms in equation (3.23) are zero either by optimality (first-order condition) of q_1^* and a^* or because $q_1^*(x)$ and/or $a^*(x)$ are constant in the neighbourhood of x . Simplifying, we end up with

$$rv'(x) = -q_1^*(x)v''(x) + \frac{\partial}{\partial x} q_2^*(x) + \lambda a^*(x)\Delta v'(x). \tag{3.24}$$

Comparing (3.22) and (3.24), we find

$$\mathcal{S}v'(x) = rv'(x) + q_1^*(x) \frac{\partial}{\partial x} q_2^*(x).$$

If player 2 is blockaded at x , $q_2^*(x) = 0$, and we recover the classical Hotelling result of [8]. Otherwise, since $q_2^* = \frac{1}{3}(1 + v' - 2c)^+$, we have $\frac{\partial}{\partial x} q_2^*(x) = v''(x)/3$ for $x < x_b$. If $x_{\text{sat}} < x < x_b$, then $a^*(x) = 0$ while $q_2^*(x) > 0$. In that case, $\mathcal{S}v'(x) = -q_1^*(x)v''(x)$, while $rv'(x) = -\frac{4}{3}q_1^*(x)v''(x)$, and we conclude $\mathcal{S}v'(x) = \frac{3}{4}rv'(x)$. \square

We conclude that in the presence of competition, shadow prices grow *slower* than at discount rate r . In particular, above the saturation level x_{sat} , marginal value of reserves grows at rate $3r/4$. Below x_{sat} , the competition effect depends on the product $q_1^*(x)v''(x)$. We note that $v'' < 0$ while $q_1^*(0) = 0$, so the overall shape of $x \mapsto q_1^*(x)v''(x)$ is humped, being zero at $x = 0$, small at x large and negative in the middle. The slowdown in the growth rate at $x = x_{\text{sat}}$ is clearly visible in the bottom-right panel of Figure 8.

The classical Hotelling rule is also modified when competition from a green producer enters in the manner of mean-field games. We refer to [12] for the calculation in that context.

4 Conclusion

We have studied the impact of exploration and discovery in models of exhaustible resources in Cournot (quantity-setting) markets. In the single-player setting with binary exploration control, we identified the critical reserves' levels x_Q , x_{on} and x_{sat} which summarise the exploration strategy. The discrete nature of the control leads to non-trivial phenomena, including discontinuity in q_v^* at x_{sat} and non-monotone $x \mapsto q_v^*(x)$. In the competition between a producer of energy from cheap non-renewable source, such as oil, and a green producer from an expensive but inexhaustible source, such as solar, our analysis shows the complex interaction between competition and strategic exploration. In particular, competition may induce earlier/higher exploration efforts; conversely exploration increases blockading effects. Moreover, Hotelling's classical result for exhaustible resources is quite different in the context of oligopolistic competition as shown in Proposition 3.4.

Tackling a genuine multi-dimensional problem in the competition between two non-renewable producers both of whom may explore remains a problem we are investigating. This involves a system of non-linear PDEs with non-local terms, and so is challenging numerically. The insight from the reduced game problem presented here may be of some help in constructing appropriate initial guesses for a numerical scheme.

Acknowledgement

The work of R. Sircar is partially supported by the NSF grant DMS-0807440. We thank Chris Costello for directing us to the economic analyses of the deterministic problem and the early results of [1, 21].

Appendix: Proof of Proposition 2.1

We first prove that the value functions are uniformly Lipschitz.

Fix x and $y > x$. Let (q, a) be an ϵ -optimal admissible strategy starting in the exploration regime and initial inventory y . Denote by $q'_t := q_t 1_{X_t^x > 0}$, $a'_t \equiv a_t$ the (admissible) strategy for initial inventory x . We use (X_t^y) (*resp.* (X_t^x)) to denote the coupled inventory processes of these two strategies with identical exploration times. Let $\sigma^0 = 0$, and define for $k \geq 1$, $v^k = \inf\{t > \sigma^{k-1} : X_t^x(q') = 0\}$ and $\sigma^k = \inf\{t > v^k : X_t^y = X_{t-}^y + \delta\}$. By construction, $q'_s = q_s$ for $s < v^1$ and $X_{v^1}^y = (y - x)$. For notational convenience, we denote for any

strategy (q, a)

$$\mathcal{H}_{t_1, t_2}(q, a) := \int_{t_1}^{t_2} e^{-r(t-t_1)} \{(1 - q_t)q_t - \mathcal{C}a_t\} dt - \sum_{t_1 \leq t \leq t_2 : \Delta a_t \neq 0} e^{-r(t-t_1)} K.$$

We now obtain

$$\begin{aligned} v(y) - v(x) &\leq v(y) - \mathcal{H}_{0, \infty}(q', a) \\ &\leq \epsilon + \mathbb{E}[\mathcal{H}_{0, v^1}(q, a) + e^{-rv^1} \mathcal{H}_{v^1, \sigma^1}(q, a) + e^{-r\sigma^1} \mathcal{H}_{\sigma^1, \infty}(q, a)] \\ &\quad - \mathbb{E}[\mathcal{H}_{0, v^1}(q, a) + e^{-rv^1} \mathcal{H}_{v^1, \sigma^1}(0, a) + e^{-r\sigma^1} \mathcal{H}_{\sigma^1, \infty}(q', a)] \\ &\leq \epsilon + \mathbb{E}[e^{-rv^1}(y - x - X_{\sigma^1-}^y) + e^{-r\sigma^1}(\mathcal{H}_{\sigma^1, \infty}(q, a) - \mathcal{H}_{\sigma^1, \infty}(q', a))], \end{aligned}$$

where the last line uses the fact that the price per unit is at most \$1, and therefore total revenue on $[v^1, \sigma^1)$ from (q_t) is at most $(X_{v^1}^y - X_{\sigma^1-}^y) = y - x - X_{\sigma^1-}^y$. Since $X_{\sigma^1}^y - X_{\sigma^1}^x = X_{\sigma^1-}^y \leq y - x$, we may repeat the argument and by induction on k find

$$\begin{aligned} v(y) - \mathbb{E}[\mathcal{H}_{0, \infty}(q', a)] - \epsilon &\leq \mathbb{E}[(y - x - X_{\sigma^1-}^y) + e^{-r\sigma^1}(\mathcal{H}_{\sigma^1, \infty}(q, a) - \mathcal{H}_{\sigma^1, \infty}(q', a))] \\ &\leq \mathbb{E}[(y - x - X_{\sigma^1-}^y) + e^{-r(v^2 - \sigma^1)}(X_{\sigma^1-}^y - X_{\sigma^2-}^y) + e^{-r\sigma^2}(\mathcal{H}_{\sigma^2, \infty}(q, a) - \mathcal{H}_{\sigma^2, \infty}(q', a))] \\ &\dots \leq \mathbb{E}[y - x + e^{-r\sigma^k}(\mathcal{H}_{\sigma^k, \infty}(q, a) - \mathcal{H}_{\sigma^k, \infty}(q', a))]. \end{aligned}$$

Now $v^k \geq \sigma_k$, the k th arrival time of the Poisson process (N_t) , and therefore $v^k \rightarrow \infty$ as $k \rightarrow \infty$. Since $\mathbb{E}[\mathcal{H}_{0, \infty}(q, a)] \leq \sup_x v(x) = (4r)^{-1}$ and using the dominated convergence theorem, we conclude that $v(y) - v(x) \leq \epsilon + (y - x)$ and the Lipschitz constant of v is ≤ 1 .

We now prove a verification argument with respect to the QVI (2.9). Let (q, a) be a fixed admissible strategy and let (\check{v}, \check{w}) be a smooth solution of equation (2.9). Denote by s^1 the first jump time of a , i.e. the first time that the exploration regime is changed. Applying Itô's formula to the a.e. differentiable function $e^{-rt}\check{v}(x)$, we have

$$\begin{aligned} e^{-rs^1}\check{v}(X_{s^1}) &= \check{v}(x) + \int_0^{s^1} \{-q_t\check{v}'(X_t) + \lambda(\check{v}(X_t + \delta) - \check{v}(X_t)) - r\check{v}(X_t)\} dt \\ &\quad + \int_0^{s^1} (\check{v}(X_t + \delta) - \check{v}(X_t)) (dN_t - \lambda dt). \end{aligned} \tag{A1}$$

Taking expectations, using the facts that the stochastic integral of the bounded process $\check{v}(X_t + \delta) - \check{v}(X_t)$ with respect to the martingale $N_t - \lambda t$ is a true martingale, $\check{v}(X_{s^1}) \geq \check{w}(X_{s^1}) - K$ and that \check{v} satisfies equation (2.9), we obtain

$$\check{v}(x) \geq \mathbb{E}\left[\int_0^{s^1} e^{-rt}\{(1 - q_t)q_t - \mathcal{C}\} dt - e^{-rs^1}K \mid X_0 = x\right] + \mathbb{E}[e^{-rs^1}\check{w}(X_{s^1}) \mid X_0 = x]. \tag{A2}$$

Repeating the same argument for $e^{-rt}\check{w}(X_t)$ on the interval $[s^1, s^2]$, we end up with

$$e^{-rs^1}\check{w}(X_{s^1}) \geq \mathbb{E}\left[\int_{s^1}^{s^2} e^{-rt}(1 - q_t)q_t dt + e^{-rs^2}(\check{v}(X_{s^2}) - K) \mid X_{s^1}\right], \tag{A3}$$

which by induction and the dominated convergence theorem (note that $e^{-rs^k} \rightarrow 0$ by admissibility of a) implies

$$\check{v}(x) \geq \mathbb{E} \left[\int_0^\infty e^{-rt} \{(1 - q_t)q_t - \mathcal{C}a_t\} dt - \sum_{t:\Delta a_t \neq 0} e^{-rt} K \mid X_0 = x \right].$$

Since (q, a) was arbitrary, $\check{v}(x) \geq v(x)$; conversely, the production strategy q^* and exploration strategy $a^* \triangleq \sum_k 1_{[s^{*,2k}, s^{*,2k+1})}$, with $s^{*,2k+1} = \inf\{t > s^{*,2k} : \check{v}(X_t) = \check{w}(X_t) - K\}$ achieve equality throughout equations (A2) and (A3). The above also shows that \check{v} and \check{w} satisfy the dynamic programming equations

$$\check{v} = L(\check{w}, 1), \quad \text{and} \quad \check{w} = L(\check{v}, 0), \tag{A4}$$

with

$$L(f, a)(x) := \sup_{q, \tau} \mathbb{E} \left[\int_0^\tau e^{-rt} \{(1 - q_t)q_t - \mathcal{C}1_{\{a=1\}}\} dt + e^{-r\tau} (f(X_\tau) - K) \mid X_0 = x \right],$$

where the supremum is over all stopping times τ adapted to the filtration of (X_t) . Inequalities (A2) and $-(A3)$ and arguments similar to those in [2] for a related piecewise-deterministic switching problem imply that the value functions (v, w) are the smallest fixed points of L (in the sense of equation (A4)) that are bigger than w_0 . We conclude that (v, w) correspond to the smallest solution of equation (2.9) which completes the proof.

References

[1] ARROW, K. J. & CHANG, S. (1982) Optimal pricing, use, and exploration of uncertain resource stocks. *J. Environ. Econ. Manag.* **9**(1), 1–10.

[2] BAYRAKTAR, E. & LUDKOVSKI, M. (2009) Optimal tracking of a hidden Markov chain under point process observations. *Stoch. Process. Appl.* **119**(6), 1792–1822.

[3] BASAR, T. & OLSDER, G. (1999) *Dynamic Noncooperative Game Theory*, Classics in Applied Mathematics, SIAM, Philadelphia.

[4] CONRAD, J. (1999) *Resource Economics*, Cambridge University Press, New York.

[5] COSTA, O. L. V. & RAYMUNDO, C. A. B. (2000) Impulse and continuous control of piecewise deterministic Markov processes. *Stoch. Stoch. Rep.* **70**(1–2), 75–107.

[6] DAVIS, M. H. A. (1993) *Markov Models and Optimization*, Chapman & Hall, London.

[7] DOCKNER, E., JØRGENSEN, S., LONG, N. V. & SORGER, G. (2000) *Differential Games in Economics and Management Science*, Cambridge University Press, Cambridge, UK.

[8] DESHMUKH, S. D. & PLISKA, S. R. (1980) Optimal consumption and exploration of nonrenewable resources under uncertainty. *Econometrica* **48**(1), 177–200.

[9] FARID, M. & DAVIS, M. H. A. (1999) Optimal consumption and exploration: A case study in piecewise-deterministic Markov modelling. *Annals Oper. Res.* **88**, 121–137.

[10] FRIEDMAN, A. (1971) *Differential Games*, John Wiley & Sons, New York, Reprinted by Dover, 2006.

[11] GIBBONS, R. (1992) *Game Theory for Applied Economists*, Princeton University Press, Princeton, NJ.

[12] GUÉANT, O., LASRY, J.-M. & LIONS, P.-L. (2010) *Mean Field Games and Oil Production*, Tech. Rep, College de France.

- [13] HARRIS, C., HOWISON, S. & SIRCAR, R. (2010) Games with exhaustible resources. *SIAM J. Appl. Math.* **70**, 2556–2581.
- [14] HOTELLING, H. (1931) The economics of exhaustible resources. *J. Polit. Econ.* **39**(2), 137–175.
- [15] HAGAN, P., WOODWARD, D., CAFLISCH, R., & KELLER, J. (1994) Optimal pricing, use and exploration of uncertain natural resources. *Appl. Math. Finance* **1**, 87–108.
- [16] LEDVINA, A. & SIRCAR, R. (2011) Dynamic Bertrand oligopoly. *Appl. Math. Optim.* **63**, 11–44.
- [17] Ledvina, A. & Sircar, R. (2011) Static and dynamic oligopoly games under asymmetric costs [online] URL: <http://www.princeton.edu/~sircar/Public/ARTICLES/static+axis-games072511.pdf>
- [18] LENHART, S. & YAMADA, N. (1992) Viscosity solutions associated with switching game for piecewise-deterministic processes. *Stoch. Stoch. Rep.* **38**(1), 27–47.
- [19] ØKSENDAL, B. & SULEM, A. (2005) *Applied Stochastic Control of Jump Diffusions*, Springer-Verlag, Berlin.
- [20] PINDYCK, R. (1978) The optimal exploration and production of nonrenewable resources. *J. Polit. Econ.* **86**, 841–862.
- [21] Pindyck, R. (1980) Uncertainty and exhaustible resource markets. *J. Polit. Econ.* **88**(6), 1203–1225.
- [22] SONER, H. M. (1985) Optimal control of a one-dimensional storage process. *Appl. Math. Optim.* **13**, 175–191.
- [23] VIVES, X. (2001) *Oligopoly Pricing: Old Ideas and New Tools*, MIT press, Cambridge, MA.