A class of short-term models for the oil industry that accounts for speculative oil storage

Yves Achdou1 · Charles Bertucci2 · Jean-Michel Lasry3 · Pierre-Louis Lions4 · Antoine Rostand5 · José A. Scheinkman6,7,8

Received: 27 October 2020 / Accepted: 15 April 2022 / Published online: 28 June 2022 © The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

We propose a plausible mechanism for the short-term dynamics of the oil market based on the interaction of a cartel, a fringe of competitive producers, and a crowd of capacity-constrained physical arbitrageurs that store the resource. The model leads to a system of two coupled nonlinear partial differential equations, with a new type of boundary conditions that play a key role and translate the fact that when storage is either full or empty, the cartel has enhanced strategic power. We propose a finite difference scheme and report numerical simulations. The latter result in apparently surprising facts: 1) the optimal control of the cartel (i.e., its level of production) is a discontinuous function of the state variables; 2) the optimal trajectories (in the state variables) are cycles which take place around the discontinuity line. These patterns help explain remarkable price swings in oil prices in 2015 and 2020.

C. Bertucci
charles.bertucci@polytechnique.edu

Y. Achdou
achdou@ljll-univ-paris-diderot.fr

A. Rostand
a.rostand@kayrros.com

J.A. Scheinkman
js3317@columbia.edu

1 Laboratoire Jacques-Louis Lions (LJLL), Université de Paris Cité and Sorbonne Université, CNRS, 75006, Paris, France
2 École Polytechnique, Paris, France
3 Université Paris-Dauphine, Paris, France
4 Collège de France, Paris, France
5 Kayrros, Paris, France
6 Columbia University, New York, USA
7 Princeton University, Princeton, NJ 08544, USA
8 NBER, Cambridge, MA 02138, USA
Keywords Economics of oil industry · Major agent facing a crowd of physical arbitrageurs · Boundary conditions linked to state constraints

Mathematics Subject Classification (2020) 49N80 · 91A16 · 91B69 · 91B52

JEL Classification C61 · C73 · D43 · L13 · Q02 · Q41

1 Introduction

1.1 Objective

We propose a model for short- and mid-term dynamics of the oil industry. Since this industry plans its development on a time scale of decades, short- and mid-term means quarters or even several years. At this time scale, the leading mechanism consists of the interactions between a monopolistic cartel (OPEC), a competitive fringe of producers, and a crowd of competitive physical arbitrageurs who store and sell oil, but are subject to capacity constraints.

As is often the case in economics, a competitive set of physical arbitrageurs limits the freedom of the cartel to implement strategies that depend on price changes. But in the oil industry, physical arbitrageurs have their own limits. Indeed, a key element in the present model is the existence of capacity constraints, minimum and maximum capacities for storage. Note that a more realistic model would consider the possibility of building storage capacity at increasing costs; we think of fixed capacity as a limit case that provides a clearer picture.

These constraints allow the cartel more strategic power when capacity limits are reached. In fact, numerical simulations in Sect. 7 show that equilibria exhibit a stable cycle, which in the deterministic limit has a phase of increasing prices, followed by a sharp price drop. The phase of price increases starts from a low price and storage at its minimum and OPEC producing less than its capacity. Initially prices increase at the rate of interest, exactly as in Hotelling’s model; thus futures exhibit normal backwardation, at least for short maturities. Physical arbitrageurs are willing to accumulate oil holdings. While prices are low, fringe producers cannot replace the capacity they lose by producing, but after prices increase sufficiently, the fringe can finance projects to gain capacity. At some point, spare storage capacity is exhausted, and OPEC chooses to cut production fast and prices rise faster than the rate of interest. This is a period in which all oil producers make large profits, the fringe can afford to increase capacity relatively fast and the production share of OPEC declines. The futures curve exhibits contango for short maturities. After a period, prices grow at the rate of interest, and in equilibrium, speculators decrease the amount they hold and storage reaches its minimum. Eventually, OPEC chooses to cut prices in a discontinuous manner and oil price go back to the beginning of the price-increasing phase. Since this drastic price cut is expected, even while the short futures exhibit contango, in the period where prices are increasing rapidly, longer futures would exhibit backwardation. The presence of noise blurs this clear picture, but as we argue below, our model qualitatively reproduces two major events in the recent history of the oil market, in 2015 and 2020.
In addition to proposing an original model for the short and middle run of the oil industry that explicitly accounts for the presence of physical arbitrageurs subject to capacity constraints, we introduce new mathematical tools: the system of PDEs (1.1), (1.2) below and the boundary conditions accounting for physical state constraints. We also report on numerical simulations that are useful for understanding the mathematical aspect of the problem and, once parameters of the model are chosen, give insights on the behaviour of the different actors in the oil industry, regarding production, prices, investments and market shares. In this introduction, we present key facts about the oil industry that motivate our model, describe the mathematical content of the paper, and explain how the paper is organised.

1.2 Some stylised facts about the oil industry

Since the 1960s, a cartel of producers has coexisted with a substantial competitive fringe in the oil industry. This survival of the cartel is partly explained by the fact that international law does not prohibit worldwide cartels, while in most countries, laws forbid lasting and non-regulated nationwide monopolies. However, regardless of the legal framework, the existence of a long-lasting equilibrium between a cartel and a competitive fringe of producers is not obvious, especially for an “exhaustible” resource.

In separate work in progress, we have proposed an economic model that is motivated by the following observations.

A first observation is that oil reserves and production capacity have been driven for several decades by production and investment. While production makes reserves decrease, new reserves can be found and new production facilities are built. Moreover, for most oil fields, the necessary investments from prospecting to construction of production facilities (capital expenditure, CAPEX) have been rather stable for several decades, of the order of 10 to 20 dollars per barrel eventually produced. Due to inefficiencies that we discuss below, the flow of CAPEX, i.e., CAPEX per year, has a decreasing yield. Indeed, it is observed that the yearly increase in production capacity is proportional to the square root of CAPEX flow. This means that if during a given year, a flow of CAPEX of $300 billion (bn) increases production capacity by 30 bn barrels, then during the same year, a flow of CAPEX of $600 bn would have increased capacity by \(\sqrt{\frac{2}{30}} \times 30 = 1.41 \times 30\) bn barrels. Notice that this observation leads us to a radical departure from Hotelling’s framework; see Hotelling [11], Dasgupta and Heal [9]. Indeed, on the time scale considered here (i.e., up to a few years), new reserves are the results of research effort and investment that are endogenously determined; therefore, exhaustibility of the resource is irrelevant. The observation that the Hotelling model is of little relevance for the study of the oil market also appears in e.g. Hamilton [10] or Covert et al. [8]. Earlier papers on mining industries that depart from Hotelling’s framework include e.g. Achdou et al. [1], Barnett [3], Barnett et al. [2], Bornstein et al. [5].

A second observation that helps explain the equilibrium observed in the oil market is that investments on capacity by fringe producers are limited by credit constraints. Indeed, most fringe producers can finance investment only if the price per barrel is sufficiently high, i.e., above a given threshold, and fringe investments increase linearly with the price of a barrel when it is above the threshold. This constraint on
investments is probably due to risk factors (operational risks, duration of projects, sovereign risk, etc.) associated to financing capacity expansion projects. The constraint on investments gives an important strategic lever to the monopolistic cartel, all the more important given that the elasticity of demand is extremely low. Indeed, in the short run, oil price elasticity is $\approx -0.04\%$ (Caldara et al. [6] estimate an oil price elasticity between $-0.02$ and $-0.08$) so that a small increase in the cartel’s production drives down prices by a substantial amount. This stops investment by the fringe and drives down their capacity, allowing the monopolist to increase market share. Conversely, when the cartel reduces its production by a relatively small amount, price increases substantially and fringe producers are able to invest. However, since all fringe producers are investing at the same time, the CAPEX required to obtain an additional unit of capacity is high. In other words, these simultaneous investments by fringe producers are inefficient. This gives the cartel more strategic power and allows the cartel to make additional profits for a rather long period.

A third observation is that the typical fringe producer faces high costs when adjusting production relative to capacity. If production is below capacity, some capacity is lost unless substantial investments are made. In contrast, OPEC producers, especially the most efficient such as Saudi Arabia, have a much lower cost for production adjustments.

The first two observations are key ingredients for the equilibrium described in long-term models such as Bornstein et al. [5] and our separate work in progress, which exhibit fluctuations between high prices that lead to immediate benefits for the cartel, but also a decrease in the cartel’s future market share, and low prices which have the opposite effect. The model studied in our separate work in progress is stationary, except for a single multiplicative factor on the demand curve. Numerical calibrations show that models that account for these two observations are capable of matching key moments of observed time series in the oil market. The third observation motivates our choice here to assume that fringe producers can only change output by changing capacity, whereas OPEC uses the quantity produced as a control variable.

1.3 Introduction to the mathematical model

We consider two state variables: $k$, the level of speculative storage, and $z$, the aggregate output of the competitive fringe. Let $U$ be the value function of the optimal control problem solved by the cartel, and $p$ the price of oil. The model we develop leads to a system of partial differential equations of the form

$$0 = -rU + H(z, p, \partial_k U) + b(k, z, p)\partial_z U + \frac{\sigma^2(k)}{2}\partial_{kk} U,$$  
(1.1)

$$0 = -rp + \partial_k H(z, p, \partial_k U)\partial_k p + b(k, z, p)\partial_z p + \frac{\sigma^2(k)}{2}\partial_{kk} p - g(k),$$  
(1.2)

for $k_{\text{min}} < k < k_{\text{max}}$ and $z_{\text{min}} < z < z_{\text{max}}$, where $g$ stands for the unit cost of storage and where the Hamiltonian $H$ and the drift $b$ are respectively given in (2.8) and (2.2) below. While equation (1.1) is a Hamilton–Jacobi–Bellmann equation, (1.2) has
the form of the master equations discussed in Bertucci et al. [4]. Hence the system (1.1), (1.2) is of the same type as those of PDEs studied in Lions [14], Cardaliaguet et al. [7], Bertucci et al. [4] for mean field games with a major agent, although (1.2) does not model a mean field game. Even without considering the boundary conditions that complement (1.1), (1.2), this system of PDEs presents major mathematical difficulties, because a sound theory of weak solutions leading to uniqueness and stability still remains to be found. Unfortunately, the present paper does not make significant progress on this aspect of mathematical analysis; however, numerical simulations reported below (the first of this kind) give insights on the solutions to (1.1), (1.2), and the observed singularities show that a notion of weak solutions is indeed necessary.

As suggested by the introduction, the boundary conditions linked to physical constraints on storage play a key role in the model. Here, we focus on the boundary \( \{k = k_{\min}\} \) (i.e., when storage is empty), because the boundary conditions at \( k = k_{\max} \) are of the same nature, mutatis mutandis. We show that the boundary conditions are

\[
-rU + \max(A, B) = 0 \quad \text{for } k = k_{\min},
\]

where the Hamiltonian \( A \) (respectively \( B \)) corresponds to strategies resulting in increasing storage (respectively to strategies which keep the storage at the minimum value). These two Hamiltonians have the form

\[
A = H_{\uparrow}(z, p(k_{\min}+, z), \partial_z U) + b(k_{\min}, z, p(k_{\min}+, z))\partial_z U,
\]

\[
B = \max_{\pi : r_{\pi} \geq b(k_{\min}, z, \pi)} F(\pi, \partial_z U),
\]

where \( H_{\uparrow} \) is the non-decreasing envelope of \( H \) (with respect to its third argument), \( p(k_{\min}+) = \lim_{k \to k_{\min}, k > k_{\min}} p(k) \) and \( F \), given by (4.4) below, corresponds to the second type of strategies. When the maximum in (1.3) is achieved by \( A \), no special condition is needed for \( p \), which means that (1.1), (1.2) will determine \( p \). On the contrary, when the maximum in (1.3) is achieved by \( B \), a Dirichlet condition should be imposed on \( p \), namely that \( p \) coincides with the optimiser in (1.4), i.e., \( p \) is given as a function of \( \partial_z U \). This means that in this regime, the price \( p \) is actually determined by the cartel and not by the physical arbitrageurs. Numerical simulations reported below will show that the system may spend a significant amount of time in this second regime in which the cartel fixes the price of oil.

To the best of our knowledge, these boundary conditions for the system (1.1), (1.2) are completely new. Since no theory is available for systems like (1.1), (1.2) in the whole space, the complete analysis of the system supplemented by the latter boundary conditions seems a fortiori out of reach at the present time.

1.4 Organisation of the paper

The remainder of the paper is organised as follows. We present the formal model in Sect. 2; mathematically, it leads to a system of partial differential equations. Section 2 also includes the description of simpler variants of the model. In Sect. 3, we
discuss the economic interpretation of the model. The issue of the boundary conditions corresponding to situations when the storage facilities are either empty or full is particularly delicate and left for Sect. 4. To the best of our knowledge, such boundary conditions are completely new. Their general mathematical analysis seems quite difficult, but we provide partial theoretical results in Sect. 5. Section 6 is devoted to a finite difference method for solving the system of PDEs supplemented with the boundary conditions. Finally, we report numerical simulations in Sect. 7. In particular, we show that equilibria may involve large and abrupt changes in oil prices. We also comment on the simulations, and stress that output from our model can qualitatively reproduce two major events that occurred in the recent history of the oil market, in 2015 and 2020.

2 Mathematical models

We consider a cartel producing a natural resource and facing both a competitive fringe of small producers and a competitive set of risk-neutral physical arbitrageurs. Even though our motivation is to understand aspects of the oil market in the short or middle term and we often use terminology linked to the oil industry (for example, oil for the resource, OPEC for the cartel), the models discussed below may be applied to other settings.

There are four types of agents: consumers, a cartel or major agent (OPEC), minor producers forming a competitive fringe, and the physical arbitrageurs that buy, store and then sell the resource. The physical arbitrageurs will most often limit price changes, but when storage capacity limits are reached, the strategic power of the cartel increases dramatically. When no resource is stored, the cartel has the power to drive prices up by cutting production; conversely, when storage is at its maximal level, the cartel can drive prices down by increasing production.

For simplicity, we do not consider the decision-making process of the competitive fringe of small producers, but instead assume that the dynamics of their global production rate is given as a function of the current state of the world; see the previous discussion on credit constraints.

We start by a mathematical presentation in which we mainly write the partial differential equations arising from the model, make some assumptions and explain what the mathematical difficulties are; the discussion of the boundary conditions, which are an important novelty, is postponed to Sect. 4.

The models described below involve two state variables: the level $k$ of speculative storage, and the level $z$ of aggregate output of the competitive fringe.

While $k$ takes values in a given interval, say $[k_{\text{min}}, k_{\text{max}}]$, the second state variable $z$ may be either discrete or continuous. More precisely, in the first variant, the global production rate $z_t$ of the competitive fringe takes its values in the interval $[z_{\text{min}}, z_{\text{max}}]$, while in the second variant, mathematically simpler but more limited from the modelling viewpoint, $z_t$ can take a finite number of values $z_j$, $j = 0, \ldots, J - 1$, for a positive integer $J$ (we only discuss the cases $J = 2$ and $J = 1$).
2.1 A model with two continuous state variables

2.1.1 The dynamics of \((k_t)\) and \((z_t)\)

The global production rate \(z_t\) of the competitive fringe is assumed to follow the dynamics

\[
dz_t = b(k_t, z_t, p_t) dt,
\]

(2.1)

where \(p_t\) stands for the price of a unit of the resource and \(b\) is a given real-valued smooth function defined on \([k_{\text{min}}, k_{\text{max}}] \times [z_{\text{min}}, z_{\text{max}}] \times \mathbb{R}_+\) of the form

\[
b(k, z, p) = \lambda p - \mu + f(k) + \tilde{b}(z).
\]

(2.2)

In (2.2), \(\lambda p - \mu\) expresses the direct impact of prices on investments, hence on production capacity, as discussed in the introduction. A typical possible choice for the second term \(f(k)\) is

\[
f(k) = a_1 \left( \frac{k_{\text{max}} - k}{k_{\text{max}} - k_{\text{min}}} \right)^2 - a_2 \left( \frac{k - k_{\text{min}}}{k_{\text{max}} - k_{\text{min}}} \right)^2,
\]

(2.3)

with suitable positive constants \(a_1\) and \(a_2\), in such a way that \(f\) has a significant effect on \(b\) only for values of \(k\) close to \(k_{\text{min}}\) or \(k_{\text{max}}\). The choice of \(f(k)\) made in (2.3) is discussed in detail and justified in Sect. 3 below.

Since we aim at simulating numerically the partial differential equation system that arises from the model, it is easier to work in a bounded domain. This is why we impose that \(z_t\) takes its values in \([z_{\text{min}}, z_{\text{max}}]\). Hence, the drift term must force \((z_t)\) to remain in \([z_{\text{min}}, z_{\text{max}}]\). Therefore we assume that \(\tilde{b}(z)\) is negative near \(z = z_{\text{max}}\), positive near \(z = z_{\text{min}}\) and vanishes away from \(z = z_{\text{min}}\) and \(z = z_{\text{max}}\). So the role of this term is purely technical, and it does not affect the numerical results far enough from \(z = z_{\text{min}}\) and \(z = z_{\text{max}}\).

Remark 2.1 We may add some noise in the dynamics of \((z_t)\) and replace (2.1) with

\[
dz_t = b(k_t, z_t, p_t) dt + \sqrt{2\nu_z} dB_t,
\]

(2.4)

where \(B\) is a Brownian motion. This randomness brings the model closer to reality, since typically a large number of shocks with relatively small amplitude arise from production or from demand. In the numerical simulations reported in Sect. 7, we take \(\nu_z > 0\).

The demand of final consumers is a decreasing function of the price of the resource; after a suitable choice of units, the simplest demand function is

\[
D(p) = 1 - \epsilon p,
\]

(2.5)

where the parameter \(\epsilon\) stands for the elasticity of demand. Note that it would be more appropriate to set \(D(p) = \max(0, 1 - \epsilon p)\), but in the regime that will be considered,
the price $p$ will never exceed $1/\epsilon$. This linear demand function fits well the observed data in the usual range of oil prices.

The control variable of the monopoly is its production rate $q_t$. Matching demand and supply yields $dk_t = (q_t + z_t - D(p_t))dt$. However, we consider instead a slightly more general dynamics of $(k_t)$, possibly including some small noise in storage capacities, namely

$$dk_t = (q_t + z_t - D(p_t))dt + \sigma(k_t)dW_t \quad \text{for } k_{\text{min}} \leq k_t \leq k_{\text{max}},$$

(2.6)

where $W$ is a Brownian motion. We suppose that the volatility $k \mapsto \sigma(k)$ is a smooth nonnegative function that vanishes at $k = k_{\text{min}}$ and $k = k_{\text{max}}$ and that the quantities $\sigma(k)k_{\text{min}}$ and $\sigma(k)k_{\text{max}}$ are bounded. This assumption plays an important role in the discussion of the boundary conditions in Sect. 4.

2.1.2 Equilibrium

We look for a stationary equilibrium. Given the unit price of the resource, the cartel solves an optimal control problem. Let $(k, z) \mapsto U(k, z)$ be the associated value function. We show below that $p_t = p(k_t, z_t)$, and that the functions $U$ and $p$ satisfy a system of two coupled partial differential equations.

The optimal control problem solved by the cartel knowing the trajectory of $(p_t)$ is

$$U(k, z) = \sup_{(q_t)} \mathbb{E} \left[ \int_0^\infty e^{-rt} \left( (p_t - c)q_t - \alpha \frac{|q_t - q_o|^2}{2} \right) dt \right] \bigg| (k_0, z_0) = (k, z),$$

where $r$ is a positive discount factor and $c$ is the cost related to the production of a unit of resource. The cost $-\alpha \frac{|q_t - q_o|^2}{2}$ is discussed and justified in Sect. 3 below.

The dynamic programming principle yields that the value function is a solution of the Hamilton–Jacobi–Bellman equation

$$0 = -rU + b(k, z, p)\partial_z U + \frac{\sigma^2(k)}{2} \partial_{kk} U$$

$$+ \sup_{q \geq 0} \left( -\alpha \frac{|q - q_o|^2}{2} + (p - c)q + (q + z - D(p))\partial_k U \right).$$

(2.7)

Introducing the Hamiltonian

$$H(z, p, \xi) = \sup_{q \geq 0} \left( -\alpha \frac{|q - q_o|^2}{2} + (p - c)q + \xi (q + z - D(p)) \right),$$

in which the maximum is reached by $q^* = \max(0, q_o + \frac{1}{\alpha}(p - c + \xi))$, (2.7) can be written as

$$-rU + H(z, p, \partial_k U) + b(k, z, p)\partial_z U + \frac{\sigma^2(k)}{2} \partial_{kk} U = 0.$$
Models for the oil industry including speculative oil storage

Since in the regime that will be considered, \( q_0 + \frac{1}{2}(p - c + \partial_k U) \) will always be non-negative, we omit for simplicity the constraint \( q \geq 0 \) in the definition of the Hamiltonian; hereafter, we set

\[
H(z, p, \xi) = \sup_{q \in \mathbb{R}} \left( -\alpha \frac{|q - q_0|^2}{2} + (p - c)q + \xi (q + z - D(p)) \right) = \frac{1}{2\alpha} (p - c + \xi)^2 + \xi (z - D(p)) + q_0 (p - c - \xi),
\]

(2.8)

and the optimal production rate at \( k_t = k \) and \( z_t = z \) is given by the feedback law

\[
q^*(k, z) = \partial_\xi H(z, p(k, z), \partial_k U(z, p)) - z + D(p(k, z)) = q_0 + \frac{1}{\alpha} (p(k, z) - c + \partial_k U(k, z)).
\]

Since storers are risk-neutral and face a cost of storage that depends on the average amount stored, a symmetric equilibrium would require that if aggregate storage satisfies \( k_{\min} < k_t < k_{\max} \), then

\[
p_t = \mathbb{E} \left[ e^{-r\delta t} p_{t+\delta t} - \int_t^{t+\delta t} e^{-r(s-t)} g(k_s) ds \right] (k_t, z_t),
\]

where \( g(k) \) is the cost of storing a unit of resource per unit of time when the level of storage is \( k \). This condition for risk-neutral competitive physical arbitrageurs facing storage costs is standard in economics. For an example of an explicit derivation, see e.g. Nutz and Scheinkman [15].

Since \( p_t = p(kt, z_t) \), Itô’s formula yields for \( k_{\min} < k < k_{\max} \) that

\[
-rp + \partial_\xi H(z, p, \partial_k U) \partial_k p + b(k, z, p) \partial_z p + \frac{\sigma^2(k)}{2} \partial_{kk} U - g(k) = 0.
\]

To summarise, the system of PDEs satisfied by \( U, p \) is

\[
0 = -rU + H(z, p, \partial_k U) + b(k, z, p) \partial_z U + \frac{\sigma^2(k)}{2} \partial_{kk} U, \tag{2.9}
\]

\[
0 = -rp + \partial_\xi H(z, p, \partial_k U) \partial_k p + b(k, z, p) \partial_z p + \frac{\sigma^2(k)}{2} \partial_{kk} p - g(k), \tag{2.10}
\]

for \( k_{\min} < k < k_{\max} \) and \( z_{\min} < z < z_{\max} \) and with \( H \) given by (2.8).

Note that (2.10) is nonlinear with respect to \( p \). It is reminiscent of the master equations discussed in Bertucci et al. [4].

2.1.3 Mathematical difficulties

It is possible to connect the problem above with the theory of mean field games involving a major agent. The system of partial differential equations (2.9), (2.10) couples a Hamilton–Jacobi–Bellman equation for the cartel (major agent) and an
equation of the type master equation for the price of the resource; see Lions [14], Lasry and Lions [12], Cardaliaguet et al. [7] and Bertucci et al. [4].

Note that in the present case, the master equation does not model a crowd of players as in mean field games, but rather an equilibrium reached by the crowd of physical arbitrageurs. It seems to be the first example in which the master equation does not involve the value of a game between competitive agents, but rather a price fixed by a (physical) arbitrage relationship. Other examples will be supplied in forthcoming papers.

Note also that it seems possible to refine the present model by considering that the physical arbitrageurs are rational agents playing a mean field game. This would lead to a more involved model of a mean field game with a major agent; see Lasry and Lions [13]. However, the resulting system of partial differential equations would have the same structure as (2.9), (2.10).

Mathematically, little is known on (2.9), (2.10). First, it needs to be supplemented by boundary conditions that are discussed below. However, even if (2.9), (2.10) are posed in the whole space or supplemented with periodic boundary conditions (which are meaningless in the present context), the main mathematical difficulty which has yet to be solved is to propose a relevant notion of weak solutions. Indeed, in Sect. 7 below, we shall see that (2.9), (2.10) may have singular solutions with discontinuities in the dynamics and the price; the latter are actually observed in the historical data (see Sect. 7). Hence it would not be reasonable to focus on classical solutions of (2.9), (2.10), and a notion of weak solutions is therefore necessary. Such weak solutions should be unique and stable with respect to small variations of the data, as is the case for entropic solutions of hyperbolic systems of conservation laws. For systems such as (2.9), (2.10), a sound way of selecting weak solutions is not available.

The boundary conditions at  \( k = k_{\text{min}} \) and  \( k = k_{\text{max}} \) arise from the physical constraints on the storage capacity. These constraints play a key role. Indeed, we show below that in some situations and when the storage level is either minimal or maximal, the cartel directly controls the price of the resource.

The resulting boundary conditions at  \( k = k_{\text{min}} \) and  \( k = k_{\text{max}} \) are extremely unusual from the mathematical point of view, have important economic implications, and are completely new to the best of our knowledge. Therefore, they will be dealt with in a separate section, Sect. 4.

2.2 A variant in which the production of the fringe is a two-state Poisson process

We now introduce a variation aiming at keeping essential features of the previous model, while simplifying the mathematics. We consider a situation in which the production rate  \( z_t \) can take only two values  \( 0 \leq z_0 < z_1 \) and is described by a stochastic Poisson process with intensities that may depend on  \( k_t \) and  \( p_t \),

\[
\begin{align*}
\mathbb{P}[z_t + \Delta t = z_0 | z_t = z_0] &= 1 - \lambda_0(k_t, p_t) \Delta t + o(\Delta t), \\
\mathbb{P}[z_t + \Delta t = z_1 | z_t = z_0] &= \lambda_0(k_t, p_t) \Delta t + o(\Delta t), \\
\mathbb{P}[z_t + \Delta t = z_1 | z_t = z_1] &= 1 - \lambda_1(k_t, p_t) \Delta t + o(\Delta t), \\
\mathbb{P}[z_t + \Delta t = z_0 | z_t = z_1] &= \lambda_1(k_t, p_t) \Delta t + o(\Delta t).
\end{align*}
\]
All the other features of the model are the same as in Sect. 2.1; in particular, the
dynamics of \((k_t)\) is still given by (2.6). The optimal value of the cartel and the price
are described by \(U(k, z_j) = U_j(k)\) and \(p(k, z_j) = p_j(k)\), where for \(j = 0, 1\), the
real-valued functions \(U_j, p_j\) are defined on \([k_{\text{min}}, k_{\text{max}}]\) and satisfy a system of four
coupled differential equations.

Introducing the Hamiltonians

\[ H_j(p, \xi) = \frac{1}{2\alpha} \left( (p - c + \xi)^2 + \xi(z_j - D(p)) \right) + q_o(p - c - \xi) \]

(we still omit the constraint that the production rate is nonnegative), and repeating
the arguments in Sect 2.1.2, we get the system of differential equations

\[ 0 = -r U_j + H_j(p_j, U'_j) + \lambda_j(k, p_j)(U_{\ell} - U_j) + \frac{\sigma^2(k)}{2} U''_j, \quad (2.11) \]

\[ 0 = -r p_j + \partial_\xi H_j(p_j, U'_j)p'_j + \lambda_j(k, p_j)(p_{\ell} - p_j) - g(k) + \frac{\sigma^2(k)}{2} p''_j, \quad (2.12) \]

for \(j = 0, 1, \ell = 1 - j\) and \(k \in (k_{\text{min}}, k_{\text{max}})\). The optimal drift of \(k_t\) in (2.6) is then
given for \(z_t = z_j\) by

\[ \partial_\xi H_j(p_j(k_t), U'_j(k_t)) = \frac{1}{\alpha} \left( (U'_j(k_t) - c + p_j(k_t)) + z_j - D(p_j(k_t)) \right) + q_o. \]

**Remark 2.2** Here again, the boundary conditions are important and non-standard.

### 2.3 An even simpler model

It is possible to simplify further the model by assuming that the production rate of
the competitive fringe is a constant \(z\). Introducing the Hamiltonian

\[ H(p, \xi) = \frac{1}{2\alpha} \left( (p - c + \xi)^2 + \xi(z - D(p)) \right) + q_o(p - c - \xi) \]

and repeating the arguments in Sect 2.1.2, we get the system of two differential equations

\[ 0 = -r U + H(p, U') + \frac{\sigma^2(k)}{2} U'', \quad (2.13) \]

\[ 0 = -r p + \partial_\xi H(p, U')p' - g(k) + \frac{\sigma^2(k)}{2} p'', \quad (2.14) \]

for \(k \in (k_{\text{min}}, k_{\text{max}})\).

### 3 Remarks on the economics of the model

This section contains remarks on the demand function, on the dynamics of fringe
production, and on the optimal control problem solved by the cartel.
Since our aim is to model the dynamics for horizons of the order of a few years, we take the demand function by final consumers as constant. This assumption is a good approximation for oil markets since except for the occurrence of major unexpected shocks – such as the arrival of the pandemic in 2020 –, there is little variation in the demand function on a yearly horizon (see Sect. 7.5 for a discussion). Nonetheless, we comment on the effect of rare disasters on equilibria; see Sect. 7.5.

The term $f(k)$ in (2.2) and (2.3) is a proxy for the time delays between the investment decisions of producers in the competitive fringe and the actual arrival of new capacities of production. Indeed, it would be a strong simplification to assume that the investment has an instantaneous effect on capacity. This simplification may be acceptable in a very-long-horizon model, i.e., at the scale of decades; it permits to keep the theoretical complexity at a reasonable level. The situation is much different when one deals with the short term, of the order of a few years: it seems necessary to model inertia effects, memory effects and anticipations of delays between investment and the creation of capacity. Nevertheless, since we wish to keep the model as simple as possible, we limit ourselves to a proxy when addressing the delay effects. The function $f(k)$ accounts for a small increase (respectively decrease) in production capacities when the storage facilities are close to empty (respectively full). Indeed, an almost empty storage must be the result of a period when the price is high, so that the investments of the producers in the competitive fringe are at a high level; the latter result in an increase of production capacity, i.e., an increase in $z$, even if the instantaneous price has decreased. The mechanism has to be reversed when the storage facilities are close to full. Hence, we choose $f$ which takes a value of the order of $1\%$ for $k \approx k_{\text{min}}$ and $-1\%$ for $k \approx k_{\text{max}}$. An accurate model with more state variables is possible, but it would be more difficult to understand and to simulate numerically (in particular, because it would increase the dimensionality of the problem).

Observe that in the oil industry, the range $k_{\text{max}} - k_{\text{min}}$ of the storage available for physical arbitrageurs is relatively small (of the order of 5–7% of annual production, see Sect. 7). Given this small range and the small elasticity of demand ($\epsilon$ in (2.5) is of the order of 0.04%), the strategy of the cartel can be implemented by tuning its production $q_t$ within a small range of values.

Indeed, it has been observed for decades that the spare production capacity of the cartel mostly varies between 3% and 5% and that the market share $q_t$ of the cartel stays close to 42%. This aspect is discussed in the last paragraph. But to keep the model focused on interactions between the cartel and the physical arbitrageurs, we describe the range of possible production strategies of the cartel by a proxy, namely, we assume that the cartel incurs an “adjustment cost” of $\alpha(q_t - q_\circ)^2 / 2$, for deviations of the cartel’s production $q_t$ from a target production $q_\circ \approx 42\%$.

As a consequence, the equilibrium $z_t$ will stay close to $z_\circ \approx 58\%$. Therefore, in the present model, since the time scale is of the order of a few years, we neglect the effect of variations in $z_t$ on the drift of $z$ in the region of interest (these small variations would induce only a small correction to the optimal strategy of the cartel).

The penalty term $\alpha(q_t - q_\circ)^2 / 2$ is motivated by the fact that for decades, the interactions between the cartel and the competitive fringe have resulted in a cartel share that averaged $q_\circ \approx 42\%$ with a standard deviation of a few %. These interactions are modelled in a separate work in progress, see also Bornstein et al. [5], and are briefly described in the introduction.
4 Boundary conditions

In this section, we discuss the boundary conditions for the full model, and more briefly for the simplified variants proposed in Sects. 2.2 and 2.3.

These boundary conditions translate mathematically the change in strategic power created by the constraints: while the price dynamics is determined by the physical arbitrageurs when storage is neither full nor empty, the price is driven by the cartel when storage is full or empty. As always in partial differential equations, these boundary conditions are key for the determination of the solution, and therefore for the behaviour of the cartel and of the physical arbitrageurs. For example, if the range \(k_{\text{max}} - k_{\text{min}}\) is very small, then the solution is mostly determined by the boundary conditions, which translates the fact that the cartel could neglect the impact of the arbitrageurs.

4.1 Boundary conditions associated with the model discussed in Sect. 2.1

No boundary conditions are needed at \(z = z_{\text{min}}\) and \(z = z_{\text{max}}\) because of the assumptions on \(b\).

For describing the boundary conditions which are linked to the state constraints \(k_{\text{min}} \leq k_t \leq k_{\text{max}}\), it is useful to introduce the nonincreasing and nondecreasing envelopes of the function \(\xi \mapsto H(z, p, \xi)\). We set

\[
H_{\downarrow}(z, p, \xi) := \max_{q \leq D(p) - z} \left( -\frac{\alpha}{2} (q - q_o)^2 + (p - c)q + \xi (q + z - D(p)) \right),
\]

\[
H_{\uparrow}(z, p, \xi) := \max_{q \geq D(p) - z} \left( -\frac{\alpha}{2} (q - q_o)^2 + (p - c)q + \xi (q + z - D(p)) \right).
\]

The Hamiltonian \(H_{\downarrow}(z, p, \xi)\) (resp. \(H_{\uparrow}(z, p, \xi)\)) corresponds to controls \(q\) such that the drift of \((k_t)\) in (2.6) is nonpositive (resp. nonnegative). It is also convenient to set

\[
H_{\text{min}}(z, p) = \min_{\xi} H(z, p, \xi) = -\frac{\alpha}{2} (D(p) - z - q_o)^2 + (p - c)(D(p) - z),
\]

(4.1)

which corresponds to the control \(q = D(p) - z\) for which the drift of \(k_t\) in (2.6) vanishes. Note that \(p \mapsto H_{\text{min}}(z, p)\) is strongly concave with respect to \(p\). It is easy to check that

\[
H(z, p, \xi) = H_{\downarrow}(z, p, \xi) + H_{\uparrow}(z, p, \xi) - H_{\text{min}}(z, p).
\]

The optimal values of \(q\) in the definition of \(H_{\downarrow}(z, p, \xi)\) and \(H_{\uparrow}(z, p, \xi)\) are

\[
q_{\downarrow}^*(z, p, \xi) = \min \left( D(p) - z, q_o + \frac{p - c + \xi}{\alpha} \right),
\]

\[
q_{\uparrow}^*(z, p, \xi) = \max \left( D(p) - z, q_o + \frac{p - c + \xi}{\alpha} \right).
\]
Hence,

\[ H_1(z, p, \xi) = \frac{1}{2} \left( \left( \sqrt{\alpha} (z - D(p) + q_0) + \frac{1}{\sqrt{\alpha}} (p - c + \xi) \right)^- \right)^2 + H_{\text{min}}(z, p), \]  

\[ (4.2) \]

\[ H_1(z, p, \xi) = \frac{1}{2} \left( \left( \sqrt{\alpha} (z - D(p) + q_0) + \frac{1}{\sqrt{\alpha}} (p - c + \xi) \right)^+ \right)^2 + H_{\text{min}}(z, p). \]  

\[ (4.3) \]

**Remark 4.1** Note that the function \( p \mapsto H_{\text{min}}(z, p) + \xi b(k, z, p) \) is strongly concave.

We start with the boundary conditions at \( k = k_{\text{min}} \). In view of the assumption made that \( \sigma \) vanishes at \( k_{\text{min}} \), it is not restrictive to focus on the deterministic case, and we take \( \sigma = 0 \) for simplicity.

The state constraint \( k_t \geq k_{\text{min}} \) implies that \( q^*(k_{\text{min}}, z) + z - D(p(k_{\text{min}}, z)) \geq 0 \). Two situations may occur.

If

\[ \frac{\partial_k U(k, z) - c + p(k, z)}{\alpha} + z - D(p) + q_0 > 0 \quad \text{for } k \text{ near } k_{\text{min}}, \]

then the optimal strategy results in increasing the level of storage. This means that in (2.10), the drift \( \partial_k H(z, p, \partial_k U) \) is positive for \( k \) near \( k_{\text{min}} \), and no boundary condition is needed for \( p \).

On the other hand, if

\[ \frac{\partial_k U(k, z) - c + p(k, z)}{\alpha} + z - D(p) + q_0 \leq 0 \quad \text{for } k \text{ near } k_{\text{min}}, \]

then the optimal drift of \( (k_t) \) in (2.6) vanishes at \( k = k_{\text{min}} \), which implies that we obtain \( q + z - D(p(k_{\text{min}}, z)) = 0 \). This relation and the strict monotonicity of \( D \) imply that \( p \) can be considered as the control variable at \( k = k_{\text{min}} \). In other words, the cartel directly controls the price in this situation.

Ruling out arbitrage opportunities but taking into account the state constraints, it is immediate that the price process satisfies

\[ p_t \geq \mathbb{E} \left[ e^{-r\delta t} p_{t+\delta t} - \int_t^{t+\delta t} e^{-r(s-t)} g(k_s) ds \right | (k_t = k_{\text{min}}, z_t)]. \]

Since the optimal drift of \( (k_t) \) is 0, we obtain

\[ r p(k_{\text{min}}, z) - b(k_{\text{min}}, z, p(k_{\text{min}}, z)) \partial_z p(k_{\text{min}}, z) + g(k_{\text{min}}) \geq 0. \]

Returning to the cartel, we deduce from the considerations above that among the strategies consisting of keeping \( k_t \) fixed at \( k_{\text{min}} \) for \( z_t = z \), the optimal one is
\[ q^* = D(p^*) - z, \]
\[ p^* = \arg \max_{\pi : r \pi \geq b(k_{\min}, z, \pi) \partial_z p - g(k_{\min})} F(\pi, \partial_z U), \]

where

\[ F(\pi, \partial_z U) = H_{\min}(z, \pi) + b(k_{\min}, z, \pi) \partial_z U \quad (4.4) \]

and \( H_{\min}(z, \pi) \) is defined in (4.1). Note that \( \pi^* \) is unique from Remark 4.1 and depends on \( z, \partial_z U, \partial_z p \). In this situation, the nonlinear boundary condition

\[ p = p^*(z, \partial_z U, \partial_z p) \quad (4.5) \]

must be imposed at \((k_{\min}, z)\).

To summarise, setting \( p(k_{\min}+, z) = \lim_{k \to k_{\min}, k > k_{\min}} p(k, z) \) another way of formulating the boundary conditions at \( k = k_{\min} \) is as follows.

– The nonlinear condition (4.5), i.e.,

\[ p = p^*(z, \partial_z U, \partial_z p), \]

understood in a weak sense, i.e., it holds only if \( \partial_k U(k, z) - c + p(k, z) \), the optimal drift, is \( \leq 0 \) near \( k = k_{\min} \), and where \( p^*(z, \partial_z U, \partial_z p) \) achieves the maximum in (4.6) below.

– The equation for \( U \) can be written as

\[ -r U + \max(A, B) = 0 \]

with

\[ A = H_{1}(z, p(k_{\min}+, z), \partial_k U) + b(k_{\min}, z, p(k_{\min}+, z)) \partial_z U, \]
\[ B = \max_{\pi : r \pi \geq b(k_{\min}, z, \pi) \partial_z p - g(k_{\min})} F(\pi, \partial_z U), \quad (4.6) \]

and \( F \) is given by (4.4). Note that to the best of our knowledge, this set of boundary conditions, associated to the system (2.9), (2.10) and to the state constraint \( k \geq k_{\min} \), has never been proposed.

Arguing as above and setting \( p(k_{\max}-, z) = \lim_{k \to k_{\max}, k < k_{\max}} p(k, z) \), the boundary conditions at \( k = k_{\max} \) can be written as follows.

– A nonlinear condition for \( p \) of the form

\[ p = p^{**}(z, \partial_z U, \partial_z p), \]

understood in a weak sense, i.e., it holds only if \( \partial_k U(k, z) - c + p(k, z) \), the optimal drift, is \( \geq 0 \) near \( k = k_{\max} \), where \( p^{**}(z, \partial_z U, \partial_z p) \) achieves the maximum in (4.7) below (it is unique from Remark 4.1).
– An equation for $U$, namely

$$-rU + \max(C, D) = 0,$$

with

$$C = H_\downarrow(z, p(k_{\max} -, z), \partial_k U) + b(k_{\max}, z, p)\partial_z U,$$

$$D = \max_{\pi : r\pi \leq b(k_{\max}, z, \pi)\partial_z p - g(k_{\max})} G(\pi, \partial_z U)$$

and

$$G(\pi, \partial_z U) = H_{\min}(z, \pi) + b(k_{\max}, z, \pi)\partial_z U.$$

### 4.2 Boundary conditions associated with the model in Sect. 2.2

The boundary conditions associated with the system (2.11), (2.12) are obtained in the same manner as in the previous case. To avoid repetitions, we focus on the boundary at $k = k_{\min}$, because the needed modifications with respect to Sect. 4.1 are similar for $k = k_{\max}$ and $k = k_{\min}$. The interested reader can easily find the boundary conditions at $k = k_{\max}$ from Sect. 4.1 and what follows.

As above, we set

$$H_{j,\min}(p) = \min_\xi H_j(p, \xi) = -\frac{\alpha}{2}(D(p) - z_j - q_o)^2 + (p - c)(D(p) - z_j),$$

$$H_{j,\downarrow}(p, \xi) = \max_{q \leq D(p) - z_j} \left(-\frac{\alpha}{2}(q - q_o)^2 + (p - c)q + \xi(q + z_j - D(p))\right)$$

$$= \frac{1}{2}\left(\left(\sqrt{\alpha}(z_j - D(p) + q_o) + \frac{1}{\sqrt{\alpha}}(p - c + \xi)\right)^{-}\right)^2 + H_{j,\min}(p),$$

$$H_{j,\uparrow}(p, \xi) = \max_{q \geq D(p) - z_j} \left(-\frac{\alpha}{2}(q - q_o)^2 + (p - c)q + \xi(q + z_j - D(p))\right)$$

$$= \frac{1}{2}\left(\left(\sqrt{\alpha}(z_j - D(p) + q_o) + \frac{1}{\sqrt{\alpha}}(p - c + \xi)\right)^{+}\right)^2 + H_{j,\min}(p).$$

The optimal values of $q$ in the definition of $H_{j,\downarrow}(p, \xi)$ and $H_{j,\uparrow}(p, \xi)$ are

$$q_{j,\downarrow}^*(p, \xi) = \min\left(D(p) - z_j, q_o + \frac{p - c + \xi}{\alpha}\right),$$

$$q_{j,\uparrow}^*(p, \xi) = \max\left(D(p) - z_j, q_o + \frac{p - c + \xi}{\alpha}\right).$$

Setting $p_{\ell+} = \lim_{k \to k_{\min}, k > k_{\min}} p_\ell(k)$, $\ell = 0, 1$, the boundary conditions at $k = k_{\min}$ are as follows: for $i = 0, 1$ and $j = 1 - i$,

– a condition of the form

$$p_j = p_j^*(U_j, U_\ell, p_{\ell+}) \quad \text{with } \ell = 1 - j.$$
understood in a weak sense, i.e., it holds only if \( U_j'(k) - c + p_j(k) \leq 0 \) for \( k \) near \( k_{\text{min}} \), where \( p_j^*(U_j, U_\ell, p_\ell+) \) achieves the maximum in (4.8) below (it is supposed to be unique);

– the equation for \( U_j \) can be written

\[-r U_j + \max(A, B) = 0\]

with

\[
A = H_{j, \uparrow}(p_j +, U_j') + \lambda_j(k_{\text{min}}, p_j+)(U_\ell - U_j),
\]

\[
B = \max_{p : (r + \lambda_j(k_{\text{min}}, p)) p - \lambda_j(k_{\text{min}}, p) p_\ell + g(k_{\text{min}}) \geq 0} F_j(p, U_j, U_\ell),
\] (4.8)

where \( \ell = 1 - j \), and

\[
F_j(p, U_j, U_\ell) = H_{j, \min}(p) + \lambda_j(k_{\text{min}}, p)(U_\ell - U_j).
\]

### 4.3 Boundary conditions associated with the model discussed in Sect. 2.3

Here also, we focus on \( k = k_{\text{min}} \) to avoid repetitions. Let us set

\[
H_{\text{min}}(p) = \min_\xi H(p, \xi) = -\frac{\alpha}{2}(D(p) - z - q_0)^2 + (p - c)(D(p) - z),
\]

\[
H_{\downarrow}(p, \xi) = \max_{q \leq D(p) - z} \left(-\frac{\alpha}{2}(q - q_0)^2 + (p - c)q + \xi(q + z - D(p))\right) = \frac{1}{2}\left(\left(\sqrt{\alpha}(z - D(p) + q_0) + \frac{1}{\sqrt{\alpha}}(p - c + \xi)\right)^2 + H_{\text{min}}(p)\right),
\]

\[
H_{\uparrow}(p, \xi) = \max_{q \geq D(p) - z} \left(-\frac{\alpha}{2}(q - q_0)^2 + (p - c)q + \xi(q + z - D(p))\right) = \frac{1}{2}\left(\left(\sqrt{\alpha}(z - D(p) + q_0) + \frac{1}{\sqrt{\alpha}}(p - c + \xi)\right)^2 + H_{\text{min}}(p)\right).
\]

Setting \( p^+ = \lim_{k \to k_{\text{min}}, k > k_{\text{min}}} p(k) \), the boundary conditions at \( k = k_{\text{min}} \) are:

– A condition of the form

\[ p = p^* \]

understood in a weak sense, i.e., it holds only if \( U'(k) - c + p(k) \leq 0 \) for \( k \) near \( k_{\text{min}} \), where \( p^* \) achieves the maximum in (4.10) below (\( p^* \) is unique).

– The equation for \( U \) can be written as

\[-r U + \max(A, B) = 0, \quad (4.9)\]

with

\[
A = H_{\uparrow}(p+, U'),
\]

\[
B = \max_{p : r p + g(k_{\text{min}}) \geq 0} H_{\text{min}}(p). \quad (4.10)
\]
5 Mathematical analysis of the boundary conditions in the one-dimensional model

In this section, we discuss the boundary conditions introduced in Sect. 4.3. Because the system (2.13), (2.14) presents major mathematical difficulties even in the absence of boundary conditions, we focus here on a local characterisation of the solutions near the boundaries. Since the roles played by the boundaries \( \{ k = k_{\text{min}} \} \) and \( \{ k = k_{\text{max}} \} \) are symmetrical, we restrict ourselves to values of \( k \) close to \( k_{\text{min}} \). To simplify the notation, we set \( k_{\text{min}} = 0 \).

Note first that in the case in which the drift \( \partial_k H(p, \partial_k U) \) is positive near \( k = 0 \), the functions \( p \) and \( U \) are expected to be smooth near the boundary. In this case, if we were to impose a Dirichlet boundary condition, it would not be satisfied by the weak solution of the problem.

Hence, we focus on the case in which the drift \( \partial_k H(p, V) \) points towards the boundary (i.e., \( \partial_k H(p, V) \leq 0 \)). If \( (U, p) \) is a classical solution of (2.13), (2.14) in an interval \( (0, k^*) \), writing \( V = \partial_k U \), the system of differential equations satisfied by \( p \) and \( V \) in \( (0, k^*) \) is

\[
0 = -r V + (\sigma(k)\sigma'(k) + \partial_k H(p, V))V' + \partial_p H(p, U')p' + \frac{\sigma^2(k)}{2} V'',
\]

\[
0 = -rp + \partial_k H(p, V)p' - g(k) + \frac{\sigma^2(k)}{2} p''.
\]

Since \( \sigma \) vanishes at \( k = 0 \), it plays no role in what follows and we can suppose that \( \sigma = 0 \) in \( (0, k^*) \). Hence, we are interested in the system of differential equations

\[
0 = -r V + \partial_k H(p, V)V' + \partial_p H(p, V)p',
\]

\[
0 = -rp + \partial_k H(p, V)p' - g(k).
\]

First, we are going to explain why a singular behaviour should be expected in general at the boundary \( \{ k = 0 \} \). Second, we give a result of existence and uniqueness provided that \( k^* \) is small enough.

5.1 A singularity is expected

Let us explain why a singular behaviour should be expected near the boundary \( \{ k = 0 \} \). We make the ansatz

\[
V(k) = V(0) + \gamma k^n + o(k^n),
\]

\[
p(k) = p(0) - \beta k^m + o(k^m),
\]

with \( n, m \leq 1 \). For shortening the notation, we define the pair

\[
(V_0, p_0) = (V(0), p(0)).
\]

Assume that there is no singularity and that \( m = n = 1 \); in this situation, from the assumption made on the sign of the drift near the boundary and the constraint \( k_\gamma \geq 0 \),
we deduce that

$$0 = \partial_\xi H(p_0, V_0) = \left(\frac{1}{\alpha} + \epsilon\right)p_0 + \frac{1}{\alpha} V_0 - \frac{c}{\alpha} + z - 1 - q_0. \quad (5.5)$$

Then, plugging the ansatz for $V$ and $p$ into (5.1), (5.2) and focusing on the zeroth-order terms, we obtain that

$$r V_0 = \beta \left(\frac{1}{\alpha} (p_0 - c + V_0) + \epsilon V_0 + q_0\right),$$
$$rp_0 = -g(0). \quad (5.6)$$

The equations in (5.6) and (5.5) form a linear system which is over-determined except for a single value of $\beta$. Thus the values of $V_0$, $p_0$ and $\beta$ are determined. Passing to the first-order terms in the expansion of the system, we obtain two second-order polynomial equations in $\gamma$ and $\beta$, while $\beta$ is already known. It is then easy to observe that for a generic choice of the parameters, this system of second-order equations is not consistent with the already obtained values of $V_0$, $p_0$ and $\beta$.

**Remark 5.1** Recall that if the drift is positive near $k_{\text{min}}$, no singularity is expected.

**5.2 Characterisation of the singularity**

We first characterise formally the singularity. Then we cast the system of differential equations in a new set of variables coming from the ansatz, and we state a local existence and uniqueness result.

**Proposition 5.2** If $V$ and $p$ satisfy (5.3) and (5.4), then $n = m = 1/2$ and the pair $(V_0, p_0)$ is completely determined by the values of $z$, $\epsilon$, $\alpha$, $q_0$. Moreover, if we have $(\alpha \epsilon)^2 + \alpha \epsilon > 1$, there is at most one pair $(\gamma, \beta)$ such that (5.3) and (5.4) hold.

**Remark 5.3** The condition on $\alpha \epsilon$ in Proposition 5.2 is fulfilled in the numerical simulations in Sect. 7 below.

**Remark 5.4** The value of $p_0$ is obviously $p^*$ which was defined in Sect. 4.3.

**Proof of Proposition 5.2** Plugging the ansatz (5.3), (5.4) into (5.1), (5.2), and using the boundedness of $g$, we deduce that

$$(1 + \alpha \epsilon)V_0 + p_0 = c - \alpha q_0 \quad (5.7)$$

by identifying the higher-order terms in the expansion. From the state constraint and the sign assumption on the drift, we also obtain

$$(1 + \alpha \epsilon)p_0 + V_0 = c + \alpha (1 - z) - \alpha q_0. \quad (5.8)$$
Since $\alpha \varepsilon \notin \{ -2, 0 \}$, we deduce that

$$V_0 = \frac{z - 1 + \varepsilon (c - \alpha q_0)}{\varepsilon (2 + \alpha \varepsilon)},$$

$$p_0 = \frac{\varepsilon (c - \alpha q_0) + (1 + \alpha \varepsilon)(1 - z)}{\varepsilon (2 + \alpha \varepsilon)}. \quad (5.9)$$

Identifying the higher-order terms in the expansion, we see that if $n \neq m$, then $\beta = \gamma = 0$. Therefore, $n = m$. Now, if $2n - 1 \notin \{0, 1\}$, identifying the terms of order $2n - 1$ leads to

$$-(1 + \alpha \varepsilon)\beta + \gamma = 0,$$

$$(1 + \alpha \varepsilon)\gamma - \beta = 0.$$  

The latter system yields that $\gamma = \beta = 0$. Thus $n \in \{ 1/2, 1 \}$. The only possible value of $n$ is $1/2$ since the case $n = 1$ has already been ruled out.

Considering the zeroth-order terms, we conclude that

$$0 = -r V_0 + \frac{1}{2\alpha} (\gamma - \beta)^2 - \varepsilon \gamma \beta,$$

$$0 = -r p_0 - \frac{\beta}{2\alpha} (\gamma - \beta) + \frac{\varepsilon}{2} \beta^2 - g(0). \quad (5.10)$$

Let us introduce the parameter

$$\lambda = \frac{r V_0}{r p_0 + g(0)},$$

which is well defined since $g \geq 0$, $p_0 > 0$. Observe that $\lambda \geq 0$. We deduce that

$$\frac{1}{2\alpha} (\gamma - \beta)^2 - \varepsilon \gamma \beta = -\lambda \frac{\beta}{2\alpha} (\gamma - \beta) + \lambda \frac{\varepsilon}{2} \beta^2,$$

then that

$$\gamma^2 + (1 - \lambda (1 + \alpha \varepsilon)) \beta^2 - (2 + 2\alpha \varepsilon - \lambda) \gamma \beta = 0.$$  

Defining the numbers $x_\pm$ by

$$x_\pm = \frac{2 + 2\alpha \varepsilon - \lambda \pm \sqrt{(2 + 2\alpha \varepsilon - \lambda)^2 - 4(1 - \lambda (1 + \alpha \varepsilon))}}{2},$$

we finally obtain

$$(\gamma - \beta x_+)(\gamma - \beta x_-) = 0.$$  

Rewriting the second equation in (5.10), we obtain

$$2\alpha g(k_{\min}) + 2\alpha r p_0 = \beta ((\alpha \varepsilon + 1)\beta - \gamma).$$
Thus $\gamma = x_+\beta$ is impossible if $x_+ > 1 + \alpha\epsilon$. An easy calculation leads to the fact that this last condition is satisfied (i.e., $x_+$ is large enough) if
\[(\alpha\epsilon)^2 + \alpha\epsilon - 1 > 0.\]

Proposition 5.2 suggests to make the change of variables
\[V(k) = V_0 + \sqrt{k}\tilde{V}(\sqrt{k}),\]
\[p(k) = p_0 + \sqrt{k}\tilde{p}(\sqrt{k})\] for $k$ in $(0,k^*)$. We notice that $\partial_\xi H$ and $\partial_p H$ are affine functions of $(p,\xi)$; indeed,
\[\partial_\xi H(p,\xi) = A(p,\xi) + z - 1 + q_0 - \frac{c}{\alpha},\]
\[\partial_p H(p,\xi) = B(p,\xi) + q_0 - \frac{c}{\alpha},\]
where $A$ and $B$ are the linear functions
\[A(p,\xi) = (\epsilon + \frac{1}{\alpha})p + \frac{1}{\alpha}\xi,\]
\[B(p,\xi) = \frac{1}{\alpha}p + \left(\epsilon + \frac{1}{\alpha}\right)\xi.\]

We deduce from (5.1), (5.2) and (5.7), (5.8) that $\tilde{V}$ and $\tilde{p}$ solve
\[0 = -rV_0 - r\sqrt{k}\tilde{V}(\sqrt{k}) + \frac{\sqrt{k}}{2}A(\tilde{p}(\sqrt{k}),\tilde{V}(\sqrt{k}))(\tilde{V}'(\sqrt{k}) + \frac{\tilde{V}(\sqrt{k})}{\sqrt{k}})\]
\[+ \frac{\sqrt{k}}{2}B(\tilde{p}(\sqrt{k}),\tilde{V}(\sqrt{k})),\] (5.13)
\[0 = -rp_0 - r\sqrt{k}\tilde{p}(\sqrt{k}) + \frac{\sqrt{k}}{2}A(\tilde{p}(\sqrt{k}),\tilde{V}(\sqrt{k}))(\tilde{p}'(\sqrt{k}) + \frac{\tilde{p}(\sqrt{k})}{\sqrt{k}})\]
\[+ g(k).\] (5.14)

Let us now remark that identifying the zeroth-order terms in the previous equations, we are naturally brought back to the arguments in the proof of Proposition 5.2. Hence, using the notation in Proposition 5.2 and setting
\[(\tilde{V}(0),\tilde{p}(0)) = (\gamma,-\beta),\] (5.15)
we can rewrite (5.13) and (5.14) as
\[A(\tilde{p}(t),\tilde{V}(t))\tilde{V}'(t) + B(\tilde{p}(t),\tilde{V}(t))\tilde{p}'(t)\]
\[= 2r\tilde{V}(t) + \frac{1}{t}(A(-\beta,\gamma)\gamma - A(\tilde{p}(t),\tilde{V}(t))\tilde{V}(t)\]
\[B(-\beta,\gamma)\beta - B(\tilde{p}(t),\tilde{V}(t))\tilde{p}(t)),\] (5.16)
\[ A(\tilde{p}(t), \tilde{V}(t)) \tilde{p}'(t) = 2r \tilde{p}(t) + \frac{1}{t} \left( -A(-\beta, \gamma) \beta - A(\tilde{p}(t), \tilde{V}(t)) \tilde{p}(t) \right) + 2 \frac{g(t^2) - g(0)}{t} \]  
\quad \text{for } t \in (0, \sqrt{k^*}). \]

Assuming that \( A(\tilde{p}(t), \tilde{V}(t)) \neq 0 \) for all \( t \in [0, \sqrt{k^*}] \), (5.16), (5.17) become

\[ \tilde{V}'(t) = 2A^{-1}(\tilde{p}(t), \tilde{V}(t)) \times \left( r\tilde{V}(t) - B(\tilde{p}(t), \tilde{V}(t))A^{-1}(\tilde{p}(t), \tilde{V}(t))(r\tilde{p}(t) + \frac{g(t^2) - g(0)}{t}) \right) \]
\[ + \frac{1}{t} \left( A^{-1}(\tilde{p}(t), \tilde{V}(t))A(-\beta, \gamma) \beta - \tilde{V}(t) \right) - \frac{\beta}{t} A^{-1}(\tilde{p}(t), \tilde{V}(t))B(-\beta, \gamma) \]
\[ + \frac{\beta}{t} B(\tilde{p}(t), \tilde{V}(t))A^{-2}(\tilde{p}(t), \tilde{V}(t))A(-\beta, \gamma), \]
\[ \tilde{p}'(t) = 2A^{-1}(\tilde{p}(t), \tilde{V}(t)) \left( r\tilde{p}(t) + \frac{g(t^2) - g(0)}{t} \right) \]
\[ - \frac{1}{t} \left( \beta \frac{A(-\beta, \gamma)}{A(\tilde{p}(t), \tilde{V}(t))} + \tilde{p}(t) \right). \]

Setting \( X(t) = (\tilde{V}(t), \tilde{p}(t)) \), we can rewrite (5.18), (5.19) as

\[ \dot{X}(t) = f(t, X(t)) + \frac{F(X(t))}{t}, \]

where \( f \) is a function from \( \mathbb{R}_+ \times \mathbb{R}^2 \) to \( \mathbb{R}^2 \) which is smooth near \((0, X_0)\), where \( X_0 = (\alpha, -\beta) \), and \( F \) is a function from \( \mathbb{R}^2 \) to \( \mathbb{R}^2 \) which is smooth near \( X_0 \) and has \( F(X_0) = 0 \).

**Lemma 5.5** Let \( f \) be a function from \( \mathbb{R}_+ \times \mathbb{R}^2 \) to \( \mathbb{R}^2 \) which is smooth near \((0, X_0)\), where \( X_0 = (\alpha, -\beta) \), and \( F \) a function from \( \mathbb{R}^2 \) to \( \mathbb{R}^2 \) which is smooth near \( X_0 \) and such that \( X_0 \) is the unique zero of \( F \). If the spectrum of \( DF(X_0) \) does not contain any integer, then there is one and only one smooth solution (of class \( C^\infty \)) of (5.20).

**Proof** For brevity, we only sketch the proof. The main idea consists in noticing that the ordinary differential equation (5.20) allows one to identify all the derivatives of a given smooth solution \( X \) at \( t = 0 \). On the other hand, away from \( t = 0 \) where (5.20) is singular, standard arguments may be used.

First, observe that if \( X \) is a Lipschitz solution of (5.20), then it must satisfy \( X(0) = X_0 = F^{-1}(0) \). Then (5.20) can be written as

\[ \dot{X}(t) = f(t, X(t)) + \frac{DF(X_0)(X(t) - X_0)}{t} + \frac{F(X(t)) - DF(X_0)(X(t) - X_0)}{t}. \]

Passing to the limit as \( t \to 0 \) and using the assumption that \( F \) is smooth and that \( 0 \) is not an eigenvalue of \( DF(X_0) \), this implies that \( X \) is \( C^1 \)-regular near \( t = 0 \) and that
\[ \dot{X}(0) \text{ is characterised by } (\text{Id} - DF(X_0)) \dot{X}(0) = f(0, X(0)). \]

Taking the derivative of (5.20) yields
\[ \ddot{X}(t) = \partial_t f(t, X(t)) + Df(t, X(t)) \dot{X}(t) + \frac{DF(X(t)) \dot{X}(t)}{t} - \frac{F(X(t)) - F(X(0))}{t^2}. \]

Passing to the limit as \( t \to 0 \) is possible and yields that \( \ddot{X}(0) \) is characterised by
\[ \ddot{X}(0) = \partial_t f(0, X(0)) + Df(0, X(0)) \dot{X}(0) \]
\[ - \frac{1}{2} DF(X_0) \ddot{X}(0) + \frac{1}{2} \dot{X}(0) D^2 F(X_0) \dot{X}(0), \]
which has a unique solution since \(-2\) is not an eigenvalue of \( DF(X_0) \) and by the fact that \( \dot{X}(0) \) and \( X_0 \) are already known. Repeating this argument, we deduce that for all \( n \geq 0, X^{(n)}(0) \) is characterised by the equation.

To establish uniqueness of smooth solutions, consider two smooth solutions \( X \) and \( Y \) of (5.20). Since \( X \) and \( Y \) have the same derivatives at 0, we deduce that for any \( \alpha > 0, \)
\[ Z(t) = t^{-\alpha} (X(t) - Y(t)) \to 0 \quad \text{as} \quad t \to 0. \]

Using the regularity of \( f \) and \( F \), we obtain that there exists a constant \( C \) which depends on \( f, F, X \), and \( Y \), but not on \( \alpha \), such that
\[ \dot{Z}_i(t) \leq C \|Z(t)\|_\infty + \frac{(DF(X_0)Z(t))_i}{t} - \frac{\alpha Z_i(t)}{t} \]
for \( i = 1, 2. \) Then choosing \( \alpha \) sufficiently large, we see that
\[ \|\dot{Z}(t)\|_\infty \leq C \|Z(t)\|_\infty. \quad (5.21) \]

Since \( Z(0) = 0, (5.21) \) implies that \( Z(t) = 0, \) hence \( X(t) = Y(t) \) for all \( t \) in the interval in which both \( X \) and \( Y \) are defined. This concludes the proof of uniqueness.

Existence can be obtained by using the fact that all the derivatives of a solution are known at \( t = 0. \)

\[ \square \]

**Theorem 5.6** For a generic choice of parameters such that the condition in Proposition 5.2 is satisfied, there exists \( k^* > 0 \) (depending on the parameters) such that there exists a unique pair \((V, p)\) of functions in \( C^\gamma([0, k^*)) \cap C^1((0, k^*))\) which solve (5.1), (5.2) in \((0, k^*)\) and satisfy the initial conditions \( \partial_\xi H(p(0), V(0)) = 0 \) and \((V(0), p(0)) = (V_0, p_0)\), where \( V_0 \) and \( p_0 \) are given by (5.9).

**Proof** If such a solution \((V, p)\) exists, then it must be of the form (5.11), (5.12) for some \( \tilde{V} \) and \( \tilde{p} \). Let us note that for a generic choice of the parameters, the values of \( V(0), p(0), \tilde{V}(0) \) and \( \tilde{p}(0) \) are necessarily given by Proposition 5.2 and (5.15). The existence and uniqueness of \( \tilde{p} \) and \( \tilde{V} \) follows from Lemma 5.5, the assumptions of which are satisfied for a generic choice of the parameters.

\[ \square \]
Remark 5.7 The value $k^*$ is the bound of the interval in which the maximal solution $(V, p)$ of (5.18), (5.19) is defined. It may happen that $k^* = k_{\text{max}}$.

Remark 5.8 Note that because of (4.9), $U$ near $k_{\text{min}}$ is completely characterised from the knowledge of $V$ and $p$ near $k_{\text{min}}$. Indeed, $U(k_{\text{min}})$ only depends on $V(k_{\text{min}})$ and $p(k_{\text{min}})$; see (4.9). Then $U$ near $k_{\text{min}}$ is obtained by integrating $V$.

Theorem 5.6 suggests that the solution $(V, p)$ of (5.1), (5.2) is completely determined near the boundary $\{k = k_{\text{min}} = 0\}$. At first sight, this could raise some issues, considering that the original problem is stated on, say, $(0, k_{\text{max}})$ with boundary conditions at the two extremities, namely on how the solutions coming from both sides can match. Our interpretation is the following. First, the solution $(V, p)$ is only characterised in the interval in which (5.18), (5.19) has a unique solution. If the system becomes unsolvable at some $k^* < k_{\text{max}}$, this allows a connection phenomenon with the rest of the system. Second, in systems such as (2.13), (2.14), there may be shocks, and the shocks may arise when two maximal solutions coming from both sides of $(k_{\text{min}}, k_{\text{max}})$ connect. Hence the shocks allow a general behaviour for the solution far from the boundaries.

Remark 5.9 Let us comment on the extension of Theorem 5.6 to the model discussed in Sects. 2.2 and 4.2. If the drifts related to both $j = 0$ and $j = 1$ point in the same direction near the boundary $\{k = k_{\text{min}}\}$, then the previous analysis can be easily adapted. If this is not the case, then the arguments above cannot be applied except in very small intervals of definition in $k$.

6 Approximation by a finite difference method

As already mentioned, the solutions of the system (2.9), (2.10) of PDEs may be discontinuous. The numerical scheme must be designed in order to handle these discontinuities.

Let us focus on the case when

$$b(k, z, p) \frac{\partial p}{\partial z} = (\phi(k) + (\lambda p - \mu)) \frac{\partial p}{\partial z} = \frac{\partial}{\partial z} \left( \phi(k) p + \frac{1}{2\lambda} (\lambda p - \mu)^2 \right).$$

We are going to use the latter conservative form in the numerical scheme for (2.10). Note that

$$\phi(k) p + \frac{1}{2\lambda} (\lambda p - \mu)^2 = \frac{\lambda}{2} \left( p - \mu + \frac{\phi(k)}{\lambda} \right)^2 - \frac{\phi^2(k)}{2\lambda} + \frac{\mu}{\lambda} \phi(k).$$

It is useful to introduce the numerical flux function

$$\Psi(k, p_\ell, p_r) = \begin{cases} \max_{p_\ell \leq p \leq p_r} (\phi(k) p + \frac{1}{2\lambda} (\lambda p - \mu)^2), & p_\ell \leq p_r, \\ \min_{p_r \leq p \leq p_\ell} (\phi(k) p + \frac{1}{2\lambda} (\lambda p - \mu)^2), & p_\ell \geq p_r, \end{cases}$$

Springer
and straightforward calculus leads to

\[
\Psi(k, p_\ell, p_r) = -\frac{\phi^2(k)}{2\lambda} + \frac{\mu}{\lambda} \phi(k) + \frac{\lambda}{2} \max \left( \left( p_\ell - \frac{\mu}{\lambda} + \frac{\phi(k)}{\lambda} \right)^2, \left( p_r - \frac{\mu}{\lambda} + \frac{\phi(k)}{\lambda} \right)^2 \right).
\]

Take a uniform grid on \([k_{\min}, k_{\max}] \times [z_{\min}, z_{\max}]\) and set \(k_i = k_{\min} + i \Delta k\) for \(0 \leq i \leq N\), \(z_j = z_{\min} + j \Delta z\) for \(0 \leq j \leq M\), with \(\Delta k = \frac{k_{\max} - k_{\min}}{N}\) and \(\Delta z = \frac{k_{\max} - k_{\min}}{M}\). The discrete approximations of \(U(k_i, z_j)\) and \(p(k_i, z_j)\) are respectively named \(U_{i,j}\) and \(p_{i,j}\).

### 6.1 The discrete version of the system (2.9), (2.10)

We use for the three-nodes centred finite difference approximation of \(\partial_{kk}U\) the notation

\[
(D^2\!_k U)_{i,j} = \frac{U_{i+1,j} - 2U_{i,j} + U_{i-1,j}}{\Delta k^2}.
\]

Consider also the first-order one-sided finite difference approximations of \(\partial_k U\) and \(\partial_z U\), namely

\[
(D_{k,\ell} U)_{i,j} = \frac{U_{i,j} - U_{i-1,j}}{\Delta k}, \quad (D_{k,r} U)_{i,j} = \frac{U_{i+1,j} - U_{i,j}}{\Delta k},
\]

\[
(D_{z,\ell} U)_{i,j} = \frac{U_{i,j} - U_{i,j-1}}{\Delta z}, \quad (D_{z,r} U)_{i,j} = \frac{U_{i,j+1} - U_{i,j}}{\Delta z}.
\]

The advection term (sometimes named drift term) with respect to \(z\) in (2.9), namely \(b(k, z, p)\partial_z U\), will be discretised with a first-order upwind scheme. The discrete version of the Hamiltonian \(H\) involves the function \(\mathcal{H} : \mathbb{R}^4 \to \mathbb{R}\),

\[
\mathcal{H}(z, p, \xi_\ell, \xi_r) = H_\downarrow(z, p, \xi_\ell) + H_\uparrow(z, p, \xi_r) - H_{\min}(z, p),
\]

where \(H_\downarrow, H_\uparrow\) and \(H_{\min}\) are respectively defined in (4.2), (4.3) and (4.1). Note that \(\mathcal{H}\) is nonincreasing with respect to \(\xi_\ell\) and nondecreasing with respect to \(\xi_r\).

The discrete version of (2.9) (monotone and first-order scheme) is

\[
0 = -r U_{i,j} + \frac{\sigma^2(k_i)}{2}(D^2_k U)_{i,j} + \mathcal{H}(z_j, p_{i,j}, (D_{k,\ell} U)_{i,j}, (D_{k,r} U)_{i,j}) + \max\left(0, b(k_i, z_j, p_{i,j})\right)(D_{z,r} U)_{i,j} + \min\left(0, b(k_i, z_j, p_{i,j})\right)(D_{z,\ell} U)_{i,j}
\]

(6.1)
for \(i = 1, \ldots, N - 1\) and \(j = 0, \ldots, M\). Note that the scheme is actually well defined at \(j = 0\) (with a slight abuse of notation), because it does not involve \(U_{i-1}\) since \(b(k_i, z_{\min}, p_i, j) \geq 0\). A similar remark can be made for \(j = M\).

For the discrete version of (2.10), we choose

\[
g(k_i) = -rp_{i,j} + \frac{\sigma^2(k_i)}{2} (D_k^* p)_{i,j} + \frac{\partial H}{\partial \xi}(z_j, p_{i,j}, (D_k^* U)_{i,j}) (D_k^* p)_{i,j} + \frac{\partial H}{\partial \xi}(z_j, p_{i,j}, (D_k^* U)_{i,j}) (D_k^* p)_{i,j} + \frac{1}{\Delta z} \left( \Psi(k_i, p_{i,j}, p_{i,j+1}) - \Psi(k_i, p_{i,j-1}, p_{i,j}) \right) \tag{6.2}
\]

for \(i = 1, \ldots, N - 1\) and \(j = 0, \ldots, M\).

The scheme for (2.9), (2.10) enjoys monotonicity properties: there is a comparison principle for (6.1) and (6.2). These monotonicity properties allow us to capture the already mentioned discontinuities.

**Remark 6.1** Neglecting the viscous terms \(\frac{\sigma^2}{2} \partial_{kk} U\) and \(\frac{\sigma^2}{2} \partial_{kk} P\) in (2.9), (2.10) and taking the derivative of (2.9) with respect to \(k\) leads to a weakly hyperbolic system (only weakly because the Jacobian matrix has repeated eigenvalues and an incomplete set of eigenvectors, i.e., it is not diagonalisable). The scheme proposed above may be modified in order to handle what physicists and mathematicians call **sonic points** and **rarefaction waves**, by following the ideas in Smith et al. [16]. In the application presently discussed, it turns out that such an improvement is not necessary because we did not observe any rarefaction wave. In a forthcoming paper, we consider an economic model described by a weakly hyperbolic system leading to sonic points, and we discuss a numerical scheme which copes with rarefaction waves.

**6.2 The discrete scheme at \(i = 0\)**

In order to write the discrete version of the boundary conditions at \(k = k_{\min}\), we set

\[
A_j = H^\uparrow(z_j, p_{0,j}, (D_k^* U)_{0,j}) + \max \left( 0, b(k_{\min}, z_j, p_{0,j}) \right) (D_{z,k} U)_{0,j} + \min \left( 0, b(k_{\min}, z_j, p_{0,j}) \right) (D_{z,k} U)_{0,j},
\]

\[
B_j = \max_{p : K_j(p_{0,j-1}, p, p_{0,j+1}) \leq g(k_{\min})} F_j(p, U), \tag{6.3}
\]

where

\[
K_j(p_{0,j-1}, p, p_{0,j+1}) = -rp + \frac{1}{\Delta z} \left( \Psi(k_{\min}, p, p_{0,j+1}) - \Psi(k_{\min}, p_{0,j-1}, p) \right)
\]
and

\[ F_j (p, U) = H_{\min}(z_j, p) + \max (0, b(k_{\min}, z_j, p))(D_z U)_{0,j} + \min (0, b(k_{\min}, z_j, p))(D_z U)_{0,j} . \]

The numerical scheme corresponding to the boundary condition at \( i = 0 \) consists of two equations for each \( 0 \leq j \leq M \); the first equation is

\[ -r U_{0,j} + \max (A_j, B_j) = 0, \tag{6.4} \]

and the second equation is either (6.5) or (6.6) below: if the maximum in (6.4) is achieved by \( A_j \), then

\[ g(k_{\min}) = -rp_{0,j} + \frac{\partial H}{\partial \xi}(z_j, p_{0,j}, (D_k, r U)_{0,j}) + \frac{1}{\Delta z} \left( \Psi(k_{\min}, p_{0,j+1}) - \Psi(k_{\min}, p_{0,j-1}, p_{0,j}) \right), \tag{6.5} \]

otherwise (if the maximum in (6.4) is achieved by \( B_j \))

\[ p_{0,j} = p_j^*(U, P), \tag{6.6} \]

where \( p_j^*(U, P) \) achieves the maximum in (6.3).

**Remark 6.2** There is a unique solution to

\[ rp = \frac{1}{\Delta z} \left( \Psi(k_{\min}, p, p_{0,j+1}) - \Psi(k_{\min}, p_{0,j-1}, p) \right) - g(k_{\min}). \tag{6.7} \]

Indeed, (6.7) can be written as \( \chi(q) = -g(k_{\min}) \) with

\[ \chi(q) = r \left( \frac{\mu}{\lambda} - \frac{\phi(k_{\min})}{\lambda} \right) + rq - \frac{\lambda}{2\Delta z} \max ((q_{0,j+1}^+)^2, (q^-)^2) + \frac{\lambda}{2\Delta z} \max ((q^+)^2, (q_{0,j-1}^-)^2). \]

The function \( \chi \) is increasing and \( \lim_{q \to \pm \infty} \chi(q) = \pm \infty \). Note also that

\[ \chi'(q) = r + \frac{\lambda}{\Delta z} q(\mathbf{1}_{q \leq -q_{0,j+1}^+} + 1_{q \geq q_{0,j-1}^-}). \]

Setting

\[ Q = -\left( \frac{\mu}{\lambda} - \frac{\phi(k_{\min})}{\lambda} \right) + \frac{\lambda}{2r \Delta z} (q_{0,j+1}^+)^2 - \frac{\lambda}{2r \Delta z} (q_{0,j-1}^-)^2 - \frac{g(k_{\min})}{r}. \]
we see that

\[
q = \begin{cases} 
   r - \sqrt{r^2 + \frac{2\lambda}{\Delta t} \left( r \left( \mu - \frac{\varphi(k_{\text{min}})}{x} \right) + \frac{\lambda}{2\Delta t} (q_{0,j-1})^2 + g(k_{\text{min}}) \right)} , & \text{if } Q < -q_{0,j+1}, \\
   Q, & \text{if } Q \in [-q_{0,j+1}, q_{0,j-1}], \\
   -r + \sqrt{r^2 - \frac{2\lambda}{\Delta t} \left( r \left( \mu - \frac{\varphi(k_{\text{min}})}{x} \right) - \frac{\lambda}{2\Delta t} (q_{0,j+1})^2 + g(k_{\text{min}}) \right)} , & \text{if } Q > q_{0,j-1}.
\end{cases}
\]

Then \( p \) satisfies the constraint in (6.3) if and only if \( p \geq q + \frac{\mu}{x} - \frac{\phi(k_{\text{min}})}{x} \), and \( B_j \) is computed by maximising a concave and quadratic function on \([q + \frac{\mu}{x} - \frac{\phi(k_{\text{min}})}{x}, +\infty)\).

### 6.3 The discrete scheme at \( i = N \)

For brevity, we do not write the numerical scheme corresponding to the boundary condition at \( k = k_{\text{max}} \), because the equations (two equations for each value of \( j \), \( 0 \leq j \leq M \)) may be obtained in exactly the same way as in the previous section.

### 6.4 Solving the system of nonlinear equations: a long-time approximation

The system of equations including (6.1), (6.2) for \( 0 < i < N \) and \( 0 \leq j \leq M \) and the discrete versions of the boundary conditions at \( k = k_{\text{min}} \) and \( k = k_{\text{max}} \) described above can be written in the abstract form

\[
F(U, P) = 0, \tag{6.8}
\]

where \( F \) is a nonlinear map from \( \mathbb{R}^{2(N+1)(M+1)} \) to \( \mathbb{R}^{2(N+1)(M+1)} \) such that the Jacobian matrix of \( F(U, P) \) has negative diagonal entries.

We aim at solving the discrete system (6.8) by a long-time approximation involving an explicit scheme. The reason for choosing an explicit scheme lies in the complexity of the boundary conditions. Finding an implicit or semi-implicit scheme consistent with the nonlinear boundary conditions seems challenging.

We first fix a time step \( \Delta t > 0 \). Setting then \( U^\ell = (U^\ell_{i,j})_{0 \leq i \leq N, 0 \leq j \leq M} \) and \( P^\ell = (p^\ell_{i,j})_{0 \leq i \leq N, 0 \leq j \leq M} \), we compute the sequence \((U^\ell, P^\ell)\) by the induction

\[
(U^{\ell+1}, P^{\ell+1}) = (U^\ell, P^\ell) - \Delta t F(U^\ell, P^\ell),
\]

and we expect that the sequence converges as \( \ell \to \infty \). If that case, the limit is a solution of (6.8).

### 7 Numerical simulations

#### 7.1 Parameters

The numerical simulations reported below aim at describing some aspects of the short-term dynamics of the oil market. Some of the parameters \((r, \lambda, \mu, \epsilon, q_0, c)\) come...
from the calibration of our long-time model (performed in a work in progress) to prices, CAPEX, OPEX, capacities and production observed in the last three decades. The value of $k_{max} - k_{min}$ is a qualitatively reasonable guess, since there is no direct observation of storage by arbitrageurs. Indeed, the real storage includes strategic, operational and arbitrage storage. The other parameters and functions are qualitatively reasonable guesses to obtain proxies for the investment delay effect and for the cost of storage.

We take $g(k) = 0$,

\[ b(k, z, p) = a \left( \frac{k_{max} - k}{k_{max} - k_{min}} \right)^2 - a \left( \frac{k - k_{min}}{k_{max} - k_{min}} \right)^2 + \lambda p - \mu \]

far enough from $z = z_{min}$ and $z = z_{max}$, with

\[ r = 0.1, \quad \epsilon = 4 \times 10^{-4}, \quad a = 0.01, \quad \lambda = 8 \times 10^{-4}, \quad \mu = 0.05, \]
\[ q_o = 0.42, \quad \alpha = 10^4, \quad c = 10, \quad v_z = 10^{-4}, \quad \sigma(k) = 0 \quad \text{for all } k. \]

We set $k_{min} = 0, k_{max} = 0.05, z_{min} = 0.35$ and $z_{max} = 0.75$. We refer to Appendix A for a simulation in which the cost of storage $g(k)$ is not zero and penalises situations in which the storage capacities are close to full.

In order to keep the expression of $b$ simple, we have decided not to write explicitly the perturbations of $b$ near $z = z_{min}$ and $z = z_{max}$, which do not impact the solution in the region of interest. In the same vein, recall that $z_{min}$ and $z_{max}$ are technical bounds on $z$ which are only useful for numerical purposes, i.e., in order to work in a bounded domain; another sensible choice of $z_{min}$ and $z_{max}$ would not change the solution in the region of interest.

The mesh parameters are $N = M = 200$, and the time step is $\Delta t = 0.00001$.

### 7.2 Results

![Fig. 1 Optimal Production Level $q^*$ of the Cartel as Function of $k$ and $z$. Note the Discontinuity, Whose Amplitude is Maximal Near $k = k_{min}$. The Discontinuity Disappears as $k$ Increases to $k_{max}$](image)
Fig. 2 Drift of the equilibrium storage $k$, i.e., $q^* + z - D(p^*)$, as function of $k$ and $z$. Note that this drift behaves as $-\sqrt{k - k_{\text{min}}}$ when it is nonpositive near $k = k_{\text{min}}$, and as $\sqrt{k_{\text{max}} - k}$ when it is nonnegative near $k = k_{\text{max}}$.

Fig. 3 Equilibrium drift of $z$, i.e., $b(k, z, p^*)$ as a function of $k$ and $z$. Recall that $z$ is the production of the fringe.
Models for the oil industry including speculative oil storage

**Fig. 4** Equilibrium drift of $(k, z)$ vs. $(k, z)$

**Fig. 5** Price of oil as a function of $k$ and $z$; there are also discontinuities in price

**Fig. 6** Contours of the invariant measure of $(k_t, z_t)$ (in logarithmic scale); it is concentrated around a stable cycle. Within the cycle, the density of the measure is much higher in the region close to the lines $\{k = k_{\min}\}$ and $\{k = k_{\max}\}$, because the drift is small there.
A stable cycle in the absence of noise: We simulate a trajectory by (i) computing numerically the functions $u$ and $p$, using the finite difference method, (ii) computing a trajectory starting at the point $(k, z) = (0, 0.5)$ by using a standard Euler scheme applied to (2.1)–(2.6), thus neglecting the Gaussian noise in $z$. The trajectory is rapidly attracted to a stable periodic cycle, with a period of the order of 8 years. We plot two periods of the limit cycle. Top-left: the production of the cartel vs. time. Top-right: the price vs. time. Bottom-left: the level of storage vs. time. Bottom-right: the production of the fringe vs. time. Note that these figures are not equilibrium trajectories. In the model, there is noise, hence the cycles would not be so evident or regular. In particular, the discontinuity in the cartel production would not occur at a fixed time.

7.3 Comments

The numerical results are displayed in Figures 1–7. Figure 1 contains the graph and the contour lines of the optimal production level $q^*$ of the cartel as a function of $k$ and $z$. We see that $q^*$ is discontinuous across a line in the $(k, z)$-plane (the discontinuity is smeared a little due to the small diffusion). The amplitude of the discontinuity is maximal at $k = k_{\text{min}}$ and vanishes at $k = k_{\text{max}}$, i.e., at $k = k_{\text{max}}$, the optimal production $q^*$ depends continuously on $z$.

Figures 2–4 show the dynamics of $k$ and $z$ in equilibrium. Figure 2 contains the graph and the contour lines of the drift of the storage level $k$, i.e., $q^* + z - D(p^*)$, as a function of $k$ and $z$. It is of course discontinuous across the same line as $q^*$. Roughly speaking, the region located above (respectively below) the discontinuity in the $(k, z)$ plane corresponds to a regime when storage increases (respectively decreases). Note the characteristic behaviour of the optimal drift of $k$ for small values of $z$ near $k = k_{\text{min}}$: it is negative and vanishes at $k = k_{\text{min}}$ as $-\sqrt{k - k_{\text{min}}}$, in agreement with the theoretical results in Sect. 5. Similar behaviour is observed for large values of $z$ near $k = k_{\text{max}}$: the optimal drift is positive and vanishes at $k = k_{\text{max}}$ as
The $z$-component of the drift, i.e., $b(k, z, p^*)$, discontinuous across the same line, is displayed in Figure 3. The optimal drift vector in the $(k, z)$-plane is plotted in Figure 4.

Figure 5 contains the graph and the contour lines of $p^*$ as a function of $k$ and $z$; $p^*$ also displays a discontinuity. In Figure 6, we display the contour lines of the invariant measure of the process $(k_t, z_t)$; we see that it is concentrated around a cycle that takes place around the discontinuity line. Finally, in Figure 7, with the functions $U$ and $p^*$ computed numerically by the finite difference method, we neglect the noise and simulate a trajectory starting from the point $(k, z) = (0, 0.5)$ by means of a standard Euler scheme; after a small delay, the trajectory becomes periodic with respect to time (the transition to the periodic regime is not shown in Figure 7), with a period of the order of 8 years. This is commented on and interpreted below. Note that we purposely initialised the trajectory at $(k, z) = (0, 0.5)$ which does not belong to the cycle, in order to check the fact that the cycle is attractive.

It is remarkable that the system of PDEs (2.9), (2.10), supplemented with the boundary conditions linked with the constraints on the state variable $k$ and discussed in Sect. 4.1, leads to a discontinuous optimal strategy as in Figure 1. The discontinuous solutions obtained in the simulations may seem surprising, but we argue below that the simulated optimal policy matches qualitatively well some episodes that have been observed in the past few years. We are first going to see that a cyclic strategy of the cartel is naturally linked to the singularity displayed in Figure 1. After having described the cycle, we explain how these results shed light on what has been observed in 2015 and 2020.

Note that the level of noise in the present simulation is smaller than the actual level of noise. We have deliberately made the noise smaller in order to better see the mechanism resulting from the model, which would appear less clearly in the presence of noise. Note also that, knowing $u$ and $p$ from the simulation of the system of PDEs (2.9), (2.10), we purposely simulated (2.1) and (2.6) instead of (2.4) and (2.6) (recall that $\sigma = 0$), in order to exhibit the stable cycle and its periodicity. As we can see in Figures 4, 6 and 7, the present model, in the absence of noise, leads to a cyclic behaviour on the time scale of few years (more details are given below). Since there are important noise and risk factors in the oil market, such a cyclic behaviour is not always observed, but it has actually occurred twice in the recent past, in 2015 and 2020.

### 7.4 Interpretation of the observed cycle

The cycle that is observed in the numerical simulations, see Figures 6 and 7, is drawn schematically in Figure 8. In the absence of substantial randomness, we can see four phases. Below, we describe the four phases and relate them to Figure 7.

- Phase ($\alpha$): In this phase, which corresponds to the time interval approximately from 0.1 to 0.4 (or 8.2 to 8.5) in Figure 7, storage is close to minimal, i.e., $k = k_{\text{min}}$ (recall that we only deal with the storage managed by the physical arbitrageurs). The monopolistic cartel has therefore the power to drive the price up by maintaining a low level of production. When the price goes up, the fringe producers invest in new production capacities and gradually increase their market share. At some point, the
monopolistic cartel would like to drive the price down, in order to prevent its competitors from investing. However, the monopolistic cartel is aware that when it does so by increasing its own production, arbitrageurs start storing the resource, which diminishes the intended impact of the policy. In order to prevent the period of low prices from lasting too long, it is therefore optimal for the monopolistic cartel to increase production fast, thereby initiating the phases $\beta, \gamma, \delta$ that bring the monopolistic cartel back to its zone of profit. This is what the numerical simulations show. The strategic production discontinuity appears clearly in Figure 1; indeed, it can be seen that when storage is minimal, the production of the monopolistic cartel increases brutally when $z$, the production of the fringe, crosses a critical value. Then in Figure 5, we see that this large increase in production produces an oil-price fall. This phenomenon was observed in 2015 and 2020.

– Phase ($\beta$): the time interval approximately from 0.4 to 2.4 in Figure 7. When the price has fallen, the stored quantity of resource increases rapidly, and the state $(k_t, z_t)$ drifts to the right with a velocity nearly parallel to the $k$-axis. After having dramatically increased production, the cartel may let it decrease smoothly until storage gets full. In this regime, the physical arbitrageurs fix the price, and the price increases almost linearly with respect to time, at a rate close to the interest rate $r$.

– Phase ($\gamma$): the time interval approximately from 2.4 to 4 in Figure 7. Storage is full. The monopolistic cartel can now increase its production again and maintain the low level of price as long as necessary in order to deter the fringe from investing or even to diminish existing fringe capacity (see below). The production $z$ of the fringe decreases to a value that benefits the cartel.

– Phase ($\delta$): the time interval approximately from 4 to 8.2 in Figure 7. Since the value of $z$ is low enough, the monopolistic cartel may reduce production, increasing the price. Eventually, the stored quantity in the hands of arbitrageurs decreases. When storage is empty, the cartel can start the phase $\alpha$ of the cycle and raise the price to the optimal value.

Even if some aspects of the equilibrium may vary with the choice of parameters, see for example Appendix A, the main features discussed above are robust with re-
spect to the choice of the parameters: the discontinuity and the cycle comprising a low price period in order for the cartel to recover market shares and a high price period leading to large profits. One can also see that the present model gives rise to backwardisation and contango periods; indeed, taking $p_{t+1} - p_t$ as an approximation for the slope of the futures curve, Figure 7 shows that backwardisation (resp. contango) occurs when storage is empty (resp. full). Note that given the noise, backwardisation and contango may also occur when the constraints on storage are still not binding.

### 7.5 Discussion of 2015 and 2020 events

In 2015, OPEC decided to reconquer market share that had been lost due to the fast development of the US shale industry from 2009 to 2015. The observed price drop was strong and sudden: prices dropped from $100 to $40 per barrel in a few months. We should argue that this price drop corresponds to the discontinuous transition from the phase $\alpha$ to the phase $\beta$ of the cycle; see Figure 8. At that time, the price drop was thought to be an attack against US shale. In the spirit of the present model, it was rather an attack against all competitors, aimed at recovering market share. Indeed, the OPEC strategy had a strong impact not only on the US shale industry, which proved to have more resiliency than other fringe producers, but on all fringe producers.

In 2020, Covid-19 was a “rare disaster” that caused a demand shock seven to ten times the standard deviation of the usual shocks to demand. Although it was not designed to handle such situations, our model suggests a reasonable explanation of what happened. Recall that in the present model, $z$ (resp. $q$) is the ratio of the non-OPEC capacity of production to the global level of demand (resp. the ratio of the OPEC production to the global level of demand). Before 2020, the level of demand increased regularly from year to year at an annual growth rate of the order of $2\% \pm 1\%$. In the first semester of 2020, the Covid crisis resulted in a sudden, unexpected drop of the global demand, of the order of $10\% - 15\%$. Therefore, the variable $z$ suddenly increased by $10\% - 15\%$, and the monopolistic cartel got carried to the upper side of the singularity locus displayed in Figure 8. In our model, the optimal response is to increase production. This is precisely what happened. Many observers thought then that a suicidal conflict between OPEC and Russia led to an increase in production, while demand collapsed. However, from the viewpoint of our model, this strategy was simply intended at entering the phases $\beta$, $\gamma$, $\delta$ of the cycle, which drive $z$ to the value desired by the cartel. In other words, the aim was to reduce the capacity of production of competitors as fast as possible. Indeed, after OPEC had increased its production, storage became rapidly full (stage $\beta$). The maximal level $k_{\text{max}}$ was reached in a few weeks, and the cartel could then drive prices to a very low level (future prices even went negative during a very short period). Many planned investments were scrapped, and production units were closed. After this period, OPEC started to strongly reduce its production. The production drop was strong in absolute value, but not so strong relatively to the global demand, hence in reasonable agreement with the predictions of our model (recall that all the quantities in the present model are reduced by taking the ratio over the global demand). Then, in the second semester of 2020 and in 2021, demand increased, which implied a decrease of $z$.

Hence, despite the fact that the collapse of the global demand in 2020 was a rare event (seven to ten times the standard deviation of the historical shocks to global...
demand), our model gives a satisfactory qualitative explanation of the cycle $\alpha$, $\beta$, $\gamma$, $\delta$ observed (in a very accelerated version) in 2020.

**Appendix A: A simulation with a cost of storage**

Here, we keep the parameters as in Sect. 7, except that we take $k_{\text{max}} = 0.07$ and we suppose that there is a cost of storage, which has the form

$$g(k) = 10 \left( \frac{k - k_{\text{min}}}{k_{\text{max}} - k_{\text{min}}} \right)^3.$$  

Such a cost heavily penalises the situations in which storage is close to full. The results have been obtained using exactly the same method as in Sect. 7, and are represented in Figures 9–14. The results look rather similar, but since $k_{\text{max}}$ is

---

**Fig. 9** Case 2: optimal production level of the cartel as a function of $k$ and $z$

**Fig. 10** Case 2: $k$-component of the equilibrium drift, which gives the dynamics of the storage level, as a function of $k$ and $z$
larger than in Sect. 7, and since that near to full storage is penalised, these is no phase \( \gamma \) in the cycle, as can be seen in Figure 14.
Fig. 14 Case 2: a stable cycle. Top-left: the production of the cartel. Top-right: the price. Bottom-left: the level of storage. Bottom-right: the production of the fringe

Acknowledgement This research was supported by Kayrros. Y. Achdou, C. Bertucci, J.-M. Lasry and P.-L. Lions acknowledge partial support from the Chair Finance and Sustainable Development and the FiME Lab (Institut Europlace de Finance). Y. Achdou and C. Bertucci acknowledge partial support from the ANR (Agence Nationale de la Recherche) through MFG project ANR-16-CE40-0015-01. The paper was completed when Y. Achdou spent a semester at INRIA MATHERIALS.

References

1. Achdou, Y., Giraud, P.-N., Lasry, J.-M., Lions, P.-L.: A long-term mathematical model for mining industries. Appl. Math. Optim. 74, 579–618 (2016)
2. Barnett, M., Brock, W., Hansen, L.P.: Pricing uncertainty induced by climate change. Rev. Financ. Stud. 33, 1024–1066 (2020)
3. Barnett, M.: A Run on Oil: Climate Policy, Stranded Assets, and Asset Prices. PhD thesis, The University of Chicago (2019). https://www.frbsf.org/economic-research/wp-content/uploads/sites/4/Paper-7-2019-11-8-Barnett-240PM-1st-paper.pdf
4. Bertucci, C., Lasry, J.-M., Lions, P.-L.: Some remarks on mean field games. Commun. Partial Differ. Equ. 44, 205–227 (2019)
5. Bornstein, G., Krusell, P., Rebelo, S.: A world equilibrium model of the oil market. Technical report, Northwestern University (2021). https://ideas.repec.org/p/nbr/aberwo/23423.html
6. Caldara, D., Cavallo, M., Iacoviello, M.: Oil price elasticities and oil price fluctuations. J. Monet. Econ. 103, 1–20 (2019)
7. Cardaliaguet, P., Delarue, F., Lasry, J.-M., Lions, P.-L.: The Master Equation and the Convergence Problem in Mean Field Games. Annals of Mathematics Studies, vol. 201. Princeton University Press, Princeton (2019)
8. Covert, T., Greenstone, M., Knittel, C.R.: Will we ever stop using fossil fuels? J. Econ. Perspect. 30, 117–138 (2016)
9. Dasgupta, P., Heal, G. : The optimal depletion of exhaustible resources. Rev. Econ. Stud. 41, 3–28 (1974)
10. Hamilton, J.D. : Understanding crude oil prices. Energy J. 30(2), 179–206 (2009)
11. Hotelling, H. : The economics of exhaustible resources. J. Polit. Econ. 39, 137–175 (1931)
12. Lasry, J.-M., Lions, P.-L. : Mean field games. Jpn. J. Math. 2, 229–260 (2007)
13. Lasry, J.-M., Lions, P.-L. : Mean-field games with a major player. C. R. Math. Acad. Sci. Paris 356, 886–890 (2018)
14. Lions, P.-L. : pp. 2011–2014. http://www.college-de-france.fr/default/EN/all/equ_der/
15. Nutz, M., Scheinkman, J.A. : Shorting in speculative markets. J. Finance 75, 995–1036 (2020)
16. Smith, T.A., Petty, D.J., Pantano, C. : A Roe-like numerical method for weakly hyperbolic systems of equations in conservation and non-conservation form. J. Comput. Phys. 316, 117–138 (2016)

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