Chapter 1

Optimal Exploitation of a Mineral Resource under Stochastic Market Prices

René Caldentey, Rafael Epstein, and Denis Sauré

Abstract

In this chapter we study the operation and optimal exploitation of a mining project. We model the project as a collection of minimal extraction units or blocks, each with its own mineral composition and extraction costs. The decision maker's problem is to maximize the economic value of the project by controlling the sequence and time of extraction, as well as investing in costly capacity expansions. We use a *real options* approach based on contingent claim analysis and risk-neutral valuation to solve the problem for a fixed extraction sequence, taking as an input the stochastic process that regulates the time dynamics of futures prices. Our solution method works in two steps. First, we consider a fixed production capacity and use approximate dynamic programming to compute upper and lower bounds on the value function in terms of the spot price and mineralogical characteristics of the blocks. We use these bounds to obtain an operating policy that is asymptotically optimal as the spot price grows large. In the second step, we extend this asymptotic approximation to handle capacity expansion decisions. Our numerical computations suggest that the proposed policy is near optimal. Finally, we test our methodology in a setting based on data from a real project at Codelco (the worlds's largest copper producer).

1.1 Introduction

In this chapter, we develop a real options model for optimizing the long-term exploitation of multi-sector mining projects. This research is part of an ongoing project

Stern School of Business, New York University, e-mail: rcaldent@stern.nyu.edu · Industrial Engineering Department, University of Chile, e-mail: repstein@dii.uchile.cl, e-mail: dsaure@dii.uchile.cl

with Codelco–the World's largest copper producer–and its main theme has been the development of a long-term decision support system for production and capacity expansion plans.

Chile is the world's largest copper producer with an annual production reaching 5.7 million tons in 2013, followed by China with 1.7, and Peru with 1.2 [14]. Chile also holds 28% of the world's copper reserves, of which 17% is held by state-owned Codelco. Unsurprisingly, copper is one of Chile's most important industries, accounting for approximately 20% and 50% of the country's GDP and total exports, respectively. For countries like Chile ensuring an efficient management of their natural resources is a strategic matter. Most of the complexity associated with such a task is due to the combination of two factors: (*i*) large-scale operations involving multiple inter-temporal decisions and (*ii*) an uncertain production and market environment. In the copper industry, production uncertainty relates to factors such as heterogeneity of the mineral composition and equipment breakdowns, while marketplace uncertainty is driven by demand volatility and the stochastic evolution of market prices.

Decision-support systems based on large-scale optimization models have been successfully implemented in the copper sector (see e.g., [21] for a recent example in the Chilean copper industry), as well as in other natural resource industries (e.g., [22] in the forest industry, or [4] and [20] in the crude oil industry). By using these systems, managers evaluate alternative operational policies striving to select those that maximize the short-term and long-term profitability of their businesses. Most of these models, however, operate under deterministic input, based on average estimates of market prices and demand. As a consequence, non-adaptive operational and investment strategies, based on the so-called *discounted cashflow valuation* methodology, are predominately used by these companies, despite their failure to capture the decision maker's ability to dynamically react to a stochastically changing environment (e.g., see [37]). As a result, companies operate under suboptimal extraction and investment plans.

The real options approach overcomes these limitations of the discounted cashflow criteria by explicitly incorporating the dynamic nature of the decision-making process and the stochastic behavior of output prices and cash flows. Early research on the subject dates back to the '80s. One of the earliest example are Mc Donald and Siegel [35] who studied the optimal operation of a project under stochastic revenues and production costs when a shut-down option is available. For a comprehensive exposition on the real options approach we refer the reader to [19], [50] and [48].

In the context of natural resource management, there is an extensive real options literature that focuses on operational decisions such as determining optimal extraction policies or evaluating the options of temporarily closing-up, re-opening, or abandoning a specific project (e.g., [1], [13], [18], [34] [38], [39], [44] and [47]). An important aspect of this literature is the way in which risk over commodities' spot prices is incorporated in the valuation process. Specifically, the existence of a well-established futures market allows the use of risk-neutral (or arbitrage-free) valuation techniques similar to those used for valuing financial derivatives (e.g., [5] and [31]). An early example of this risk-neutral approach is the work of Brennan and Schwartz [7] who consider optimal extraction policies for a non-renewable natural resource. Other examples are [15], [16], [30] and [49].

Our work builds on Brennan and Schwartz [7] in the way we model the stochastic evolution of spot prices and in the resulting risk-neutral valuation approach. On the other hand, our work distinguishes itself from previous formulations in the way we model the extraction process. Most of this previous research either favors mathematical tractability by simplifying the production process, or considers realistic but intractable production models that can only be solved numerically. In this chapter, we develop a real options model that addresses some of the limitations of previous approaches, preserving tractability. Specifically, and consistent with current practice in the industry (see e.g., [21]), we model the mining project as a collection of minimal extraction blocks with different mineral composition and extraction costs. As a result, a production plan must specify an extraction sequence (that is, the order in which blocks will be extracted and processed) as well as the timing of such an extraction. We also model the option to invest in capacity expansions over time. Finally, we use real data from a Chilean mining project to illustrate the application of our methodology.

It is important to note that our model does not capture some relevant features of the operations of a mining project. Most notoriously, we do not incorporate the switching costs of temporarily idling the project or of shutting it down permanently. However, these omissions are not particularly severe for large mining companies such as Codelco that can reallocate resources when a particular project is suspended. Another limitation of our model is its focus on calculating the economic value of a prespecified sequence of blocks as opposed to determining such an optimal sequence. Two reasons support this modeling feature. First, the problem of determining an optimal dynamic (adaptive) extraction sequence is a stochastic sequencing problem in a directed network, which is difficult to solve exactly due to the curse of dimensionality. In this regard, we view this research as a necessary building block for tackling the more general stochastic network formulation. Second, our experience at Codelco suggests that in practice mining companies are interested in a relatively small set of predetermined extraction sequences. Hence, managers of these firms evaluate each of these extraction sequences and determine the one that maximizes the economic value of the mining projects.

Because of our characterization of the production process and the incorporation of stochastic market prices, our research contributes to narrow the gap between the academic literature and current practice in the copper industry. For instance, most market-leading mine planning optimization softwares (e.g., Whittle, Chronos, MineMax, and NPV Scheduler; see e.g., [32]) operate with deterministic price paths that cannot capture market risk. Also, it is worth highlighting that, although we focus on copper mining operations, our model and results can be extended to the production of other non-renewable natural resources, such as crude oil, natural gas, and other types of mineral deposits.

The rest of this chapter is organized as follows. Section 1.2 provides a description of the exact model. In Section 1.3, we assume that production capacity is fixed and we derive general properties of an optimal operating policy. We then derive a family of approximations to the optimal value of the project that include valid lower and upper bounds as special cases. We used these bounds to propose two simple extraction policies and to derive an approximate operating policy that is asymptotically optimal as the spot price and/or production capacity grow large. We conclude Section 1.3 with a set of numerical experiments that show the quality of our proposed approximations. In Section 1.4 we extend the results of the previous section to include capacity expansion decisions. Section 1.5 presents an application of our methodology to identify an optimal extraction policy for *El Diablo Regimiento*, a 230 [million ton] project at Codelco. Conclusions and future research are discussed in Section 1.6. Mathematical proofs are relegated to an Appendix, unless otherwise noted.

1.2 Model Description

We begin this section with a brief description of the mining operational process and then consider the dynamics of the spot copper prices. Subsequently, we describe our mathematical formulation of the production process. Finally, we discuss the risk-neutral valuation approach that we use to formulate the optimization problem as a dynamic programming problem.

1.2.1 Mining Operations Description

Mining operations can be seen as a sequence of stages involving geological, extraction, concentration, and refining activities. Geological activities are necessary for the discovery and characterization of new deposits. They are of great importance at the early stages of exploration and design of the mine, and are continuously required through the lifespan of a mining project for updating the geological characteristics of the mineral. Extraction activities are required to feed the concentration plants, and their structure depends on whether they are performed on an open pit mine or on an underground mine.

For an open pit operation, mineral is extracted using controlled explosions on the surface of the resource. After a blast, mineral is carried out of the pit by large trucks. In underground mines, extraction is typically conducted at specific locations (extractions points) where the material is removed using a combination of controlled explosions and gravity. In both settings, as mine sectors are located at different heights, possibly overlapping each other, upper sectors must be extracted first for the extraction to be feasible and safe. Further discussion on extracting methods can be found in [2].

The grade (percentage of copper) of the extracted material is variable, but typically below 2.0%. The material with grade over a pre-determined cut-off threshold goes though a sequence of size-reducing processes (both mechanical and chemical), concentration and refinement, which output ore with 99.9% grade that is sold as a commodity in the marketplace. Material with grade below the threshold is either left *in situ* in the case of underground mines, or sent to dump deposits in the case of open pit mines. Cut-off grade strategies are typically defined in the early stages of the planning process using an approximate “opportunity” cost of the mineral (see *e.g.* [33]).

Figure 1.1 shows schematically the entire mining operations process for the underground case. Further discussion on exploration and geological activities can be found in [16], while the concentration and refining operations are discussed in detail in [9]. A description of the entire process can be found in [11].

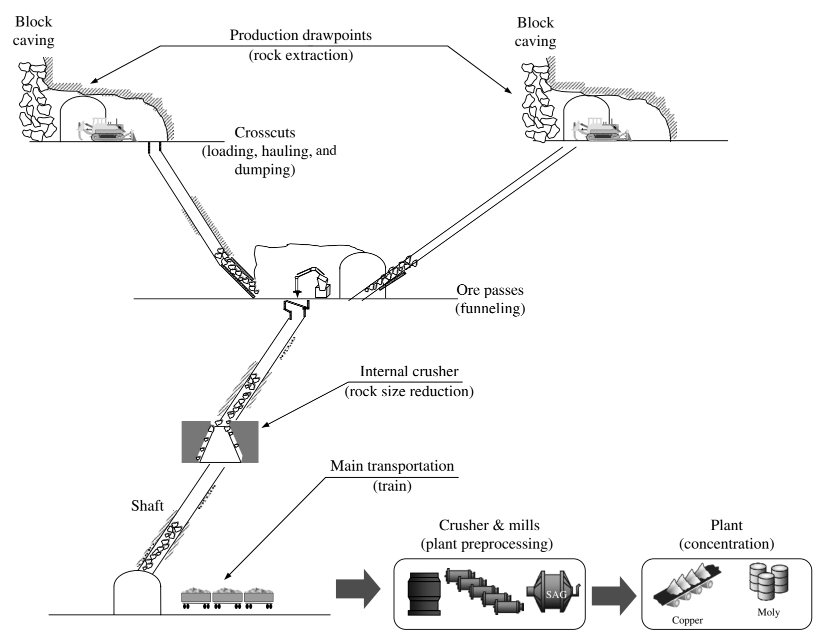

Fig. 1.1 Mining process for an underground copper mine.

1.2.2 Spot Prices

Copper's spot price is a critical ingredient in the valuation of a mining project, as it modulates project revenues, influencing extraction plans and capacity expansion, among other decisions. In addition, the inherently stochastic behavior of the spot price complicates the optimization and the evaluation of a project.

Traditionally, the early real options literature modeled commodity spot prices using Geometric Brownian Motions (GBM). Empirical evidence, analytical tractability, the unpredictability of the price path, and the early applications of GBM in mathematical finance (e.g., [41], [5], [36]) are some of the main reasons behind this choice. However, it has been recognized (e.g., [43], [49]) that many commodity prices exhibit a mean-reverting trend that captures the natural market-equilibrium tendency of these prices to revert towards a level that reflects production costs and companies' flexibility to adjust production capacity (i.e., open/close projects) to balance demand and supply in response to changing market conditions. For example, the following mean reverting process is proposed by Schwartz [42]

$$\mathrm{d}S_t = \kappa(\mu - \ln(S_t))\,S_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}B_t$$

to model the spot price S_t of a commodity, where B_t a standard Brownian Motion. (In what follows we assume that all relevant stochastic processes are embedded in a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$.) The speed of adjustment, κ, represents the degree of mean reversion to the long-run adjusted mean μ while σ measures the volatility of the price process. In a series of papers ([42], [30] and [43]), Schwartz propose three variations of a mean-reverting stochastic model driven by one, two and three factors. These models are empirically validated for copper, gold and crude oil.

Despite being economically sound, mean-reverting processes do not systematically dominate GBM models when it comes to statistically fitting the empirical data of commodity prices. Dixit and Pindyck [19] test the GBM versus a mean-reverting hypothesis using copper prices for the last 200 years and conclude that the mean reversion hypothesis should be accepted. However, they also claim that the GBM hypothesis cannot be rejected if only 30 to 40 years of data is included (see Figure 1.2). Similar conclusions are reported in [29] for ten different natural resources, including copper. More recently, [27] and [28] analyze energy commodity prices (oil and natural gas) and conclude that depending on the time period the GBM model offers a better fit than a mean-reverting one.

In what follows, for clarity of exposition and mathematical tractability, we will adopt the GBM framework, *à la* Brennan and Schwartz [7], to model the stochastic evolution of copper's spot price. We refer the interested reader to Chapter 3 in [26] and to the recent monograph [45] for further discussion on the modeling of commodity prices and their economic implications for project valuation.

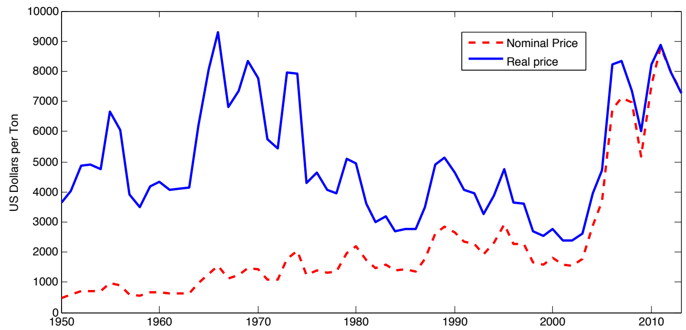

Fig. 1.2 Copper spot price evolution since 1950. (Source: London Metal Exchange)

1.2.3 Production Model

We adopt a continuous-time model to represent the operation of the mine, and assume that exploration stages have been already completed. Hence, the decision maker is mainly concerned with determining an optimal extraction and capacity expansion plan for a fixed mining project with known mineral content and quality.

Consistent with current industry practice, we represent the mine as a collection of mineral *blocks*, each having specific geological properties. These blocks represent the minimal extraction units so that production decisions are made at the block level. Executing a mining plan results in an extraction sequence for these blocks. However, not all sequences are feasible due to technical or safety reasons. For instance, different sectors of the mine are located at different elevations, usually overlapping with each other. Due to safety reasons, extraction from upper sectors must be finished before extraction from lower sectors can start. While the problem of identifying an optimal sequence is an important and challenging one, in this chapter we assume a fixed and feasible extraction sequence is given and focus exclusively on determining the timing of the extraction and how production capacity should be expanded. We note, however, that fixing the sequence of extraction is not a serious limitation as it is often the case in practice that the decision maker only wants to evaluate a few predetermined sequences. In Proposition 4 below we propose an efficient method to compare alternative extraction sequences. We also discuss a concrete example of this scenario-based valuation approach in Section 1.5.

From a mathematical standpoint, we model the mining sector as a collection of N blocks. Without loss of generality, and for notational convenience, we use a backward indexing of these blocks $\{N,\ldots,1\}$ so that block N is the first block to be extracted and block 1 is the last one. We define Q_j and L_j to be the amount of material and average grade (% of copper), respectively, available in block $j = 1,\ldots,N$.

A production policy has two basic components: (*i*) the time at which each block starts being extracted, and (*ii*) the available processing capacities at these extraction epochs. We denote by $\mathcal{T}{:=}(\tau_1,\dots,\tau_N)$ the sequence of extraction times of the blocks. That is, τ_j is time at which block j starts being extracted. (We define $\tau_{N+1}{:=}0$.) Given the extraction times $\mathcal{T}$, we denote by $\mathcal{K}{:=}(K_1,\dots,K_N)$ the vector of production capacities.[2] The j-th component of this vector, K_j, is the available production capacity at time $t=\tau_j$ when block j starts extraction. For completeness, we define K_{N+1} to be the initial level of capacity.

For simplicity, we assume that the decision maker does not increase production capacity during the extraction of a block, and expansion decisions are only made in between block extractions. In addition, we assume that all the available capacity is used during the extraction of a block. That is, the decision maker will always run the operation at 100% utilization (recall we assume that mining capacity is binding). While these assumptions provide mathematical tractability, they are not particularly restrictive from a theoretical standpoint, as one can always reduce the size of the blocks by increasing their number (N). (We can show that a "bang-bang" extraction policy is optimal when blocks are infinitesimally small.) Furthermore, in practice these decisions involve production disruptions that usually cannot occur in the middle of the extraction of a block.

Given a production policy $(\mathcal{T},\mathcal{K})$, we define $T{:=}(T_1,\dots,T_N)$ where $T_j{:=}Q_j/K_j$ is the time it takes to extract and process block j. Finally, we say a production policy $(\mathcal{T},\mathcal{K})$ is feasible if it satisfies the sequencing conditions

$$\tau_j \geq \tau_{j+1} + T_{j+1} \quad \text{and} \quad K_j \geq K_{j+1}, \quad \text{for all } j = 1,\dots,N,$$

where we define $T_{N+1}{:=}0$. Other constraints such as imposing a fixed planning horizon $\bar{T}$ or a maximum production capacity $\bar{K}$, that is,

$$\tau_1 + T_1 \leq \bar{T} \qquad \text{and} \qquad K_1 \leq \bar{K}$$

can also be included. In this chapter, we will assume an infinite horizon $\bar{T}=\infty$.

1.2.4 Project Valuation and Optimality Conditions

The long-term planning problem consists on finding an investment and operational policy that maximizes the net present value of the mineral resources. From the decision maker's perspective, this long-term value maximization amounts to selecting an optimal production policy $(\mathcal{T},\mathcal{K})$ as described in the previous section.

Determining the value of the project to be maximized is, in general, a difficult task. In practice, most mining companies consider the average discounted value of the cashflows of the project as the appropriate objective function to use. However, this

[2] For simplicity we assume that the production capacity refers to a binding *mining capacity*, rather than to treating and marketing capacity [33].

approach imposes some serious challenges in terms of selecting the appropriate discount factor, or equivalently, the correct probability measure to compute expectations. Fortunately, these complications can be circumvented by exploiting the existence of a futures market for copper, so that a no-arbitrage condition allows the use a replicating-portfolio argument to compute the market value of the mining project[3]. For more details, the reader is referred to [7] for an application of this approach in the context of a natural resource exploitation and to [46] for the general theory. In what follows, we briefly summarize the main step behind this risk-neutral valuation approach.

We can view the stream of cashflows as a derivative of the underlying copper spot price S_t, for which a futures market is available. Using a no-arbitrage argument and under a complete market assumption, it follows that the economic value of the project cashflows can be obtained using a *contingent claim* approach. To this end, let S_t be the spot price of copper which we assume evolves according to a GBM (see Section 1.2.2 for details)

$$\mathrm{d}S_t = \mu S_t \,\mathrm{d}t + \sigma S_t \,\mathrm{d}B_t, \tag{1.1}$$

where σ is the instantaneous spot price volatility, which we assume is known and constant, and μ is the drift of the spot price that could be stochastic. Following Brennan and Schwartz [7], we assume the existence of a constant *convenience yield* rate ρ on the commodity, which captures the benefit associated with physically holding the commodity instead of holding a contract for future delivery. Our constant convenience yield assumption is certainly restrictive and it is mainly imposed for mathematical tractability. For alternative and more realistic models of this convenience yield we refer the interested reader to [6], [42], [30], [12] and references therein. We also refer the reader to the recent monograph [45] for a discussion of the merit of the classical interpretation of the concept of convenience yield as a pseudo-dividend and to the chapter by Evans and Guthrie [23] in this volume for a different interpretation of this concept.

Let $\mathbb{Q}$ be a probability measure (equivalent to $\mathbb{P}$) under which the spot price, S_t, discounted at the risk-free rate, r, net of the convenience yield, ρ, is a $\mathbb{Q}$-martingale, that is,

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-(r-\rho)t} S_t \,|\, S_\tau\right] = e^{-(r-\rho)\tau} S_\tau, \qquad \text{for } \tau \le t. \tag{1.2}$$

Assuming that the spot price follows the GBM dynamics in equation (1.1), this Equivalent Martingale Measure (EMM) $\mathbb{Q}$ exists and is unique. We can compute this EMM by means of a Girsanov transformation (see Chapter 5 in [46]). Define the *market price of risk* to be

$$\lambda_t := \frac{\mu - (r-\rho)}{\sigma},$$

so that the *Radon-Nikodym derivate* of $\mathbb{Q}$ with respect to $\mathbb{P}$ is given by

[3] Copper futures and options contracts are traded on a daily basis at the London Metal Exchange (LME) with maturities ranging from 3 to 63 months. These derivatives offer buyers and sellers the opportunity to hedge their risk exposure due to spot price fluctuations.

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left(-\int_0^\infty \lambda_t \, dB_t - \frac{1}{2}\int_0^\infty \lambda_t^2 \, dt\right).$$

It is not hard to show that the spot price process S_t satisfies

$$dS_t = (r-\rho) S_t \, dt + \sigma S_t \, d\tilde{B}_t, \tag{1.3}$$

where $\tilde{B}$ is a Brownian motion under $\mathbb{Q}$ that satisfies $d\tilde{B}_t = dB_t + \lambda_t \, dt$. Solving equation (1.3) leads to

$$S_t = S_0 \exp\left(\left(r-\rho-\frac{\sigma^2}{2}\right)t + \sigma\tilde{B}_t)\right),$$

where S_0 represents the spot price at time $t = 0$.

Under the risk-neutral valuation approach, the economic value of the project for a given production policy is equal to the expected value (under $\mathbb{Q}$) of the project cashflows discounted at the risk-free rate. In our operational context, these cashflows are the difference between the revenues generated by the commercialization of the final product in the spot market minus production and capacity investment costs. Following the standard practice at Codelco, we assume that all production is immediately sold in the market, that is, the company does not hold any inventory of the final product.

Let us consider the j-th block in the extraction sequence. The extraction of this block j starts at time τ_j at a constant extraction rate K_j and finishes at time $\tau_j + T_j$. We let $W_j(S,K)$ denote the expected cashflows generated by this block discounted to time τ_j, conditional on $S_{\tau_j} = S$ and $K_j = K$. That is,

$$W_j(S,K) := \mathbb{E}^{\mathbb{Q}}\left\{\int_0^{T_j(K_j)} e^{-rt}\,[L_j K_j S_{\tau_j+t} - A_j K_j]\,dt \,\middle|\, S_{\tau_j} = S,\, K_j = K\right\},$$

where $L_j K_j$ is the rate at which copper is produced and A_j is the marginal production cost for block j. For simplicity, we assume that the marginal cost A_j is constant but depends on j. This allows us to model mining operations in which production costs tend to increase as extraction progresses, which might be attributed to the fact that the distance from the extraction points to the processing plan increases over time. (Note that in practice, however, processing costs might depend on additional factors, such as the spot price, as they should impact cut-off policies.) Using (1.2) we can show that $W_j(S,K) = R_j(K)\,S - C_j(K)$ where

$$R_j(K) := L_j \left(\frac{1-e^{-\rho T_j(K)}}{\rho}\right) K \qquad \text{and} \qquad C_j(K) := A_j \left(\frac{1-e^{-r T_j(K)}}{r}\right) K.$$

The decision maker selects a production policy that maximizes the expected cumulative discounted payoff. That is, the decision maker solves

$$F^* := \sup_{\mathcal{T},\mathcal{K}} \mathbb{E}^{\mathbb{Q}}\left[\sum_{j=1}^{N} e^{-r\tau_j}\left[R_j(K_j)S_{\tau_j} - C_j(K_j) - \gamma(K_j - K_{j+1})\right]\right]. \tag{1.4}$$

where $\gamma > 0$ is the marginal cost of capacity expansion.

A few remarks about (1.4) are in order. First, the term $e^{-r\tau_j}\gamma(K_j - K_{j+1})$ assumes that any capacity expansion takes place at the same time τ_j when the extraction of block j starts. Note, however, that while the decision maker might have the ability to build capacity at any point in time in the interval $[\tau_{j+1} + T_{j+1}, \tau_j]$, because capacity expansions are costly, it is in the decision maker's best interest to postpone this action as much as possible. Second, formulation (1.4) also assumes capacity expansions are instantaneous; otherwise, we would need to add a time lag between the time expansion begins and the time the additional capacity becomes available. Finally, note that (1.4) assumes that the cost of expansion is linear, while in practice, it might exhibit a non-linear behavior.

We can reformulate (1.4) using dynamic programming. The state space of this dynamic program represents the state of the project at the time a block has finished being extracted, and it is given by the triplet (S, j, K), where S is the spot price, j is the index of the block to be extracted next and K is the available production capacity. In this state space, we denote by $F_j(S,K)$ the expected optimal discounted profit to go.

We note that the state space description (S, j, K) is sufficient in our model because of the Markovian dynamics of the spot price and because we are assuming that capacity and production rates are fixed during the extraction of a block. Therefore, in order to derive an optimal production policy it is enough to evaluate the value function only at those time epochs when a block has finished extraction. In fact, suppose we look at the system exactly at the time block $j-1$ has finished extraction and let (S, j, K) be the state of the system at such a time. The decision maker must select the time τ_j when to start extracting block j. This time τ_j is a stopping time with respect to $\mathcal{F}_t$, the filtration generated by S_t. Finally, at this extraction time the decision maker must also decide if capacity should be expanded from the current level K to a new level K_j, with $K \le K_j \le \bar{K}$. (Recall that $\bar{K}$ is an upper bound on the maximum level of capacity that can be installed.) Putting all these pieces together, we can write the following recursion for the value function $F_j(S,K)$.

$$F_j(S,K) := \sup_{\tau_j, K_j} \mathbb{E}^{\mathbb{Q}}\Big[e^{-r\tau_j}\Big(R_j(K_j)S_{\tau_j} - C_j(K_j) - \gamma(K_j - K) + e^{-rT_j}F_{j-1}(S_{\tau_j+T_j}, K_i)\Big)\Big| S_0 = S\Big] \tag{1.5a}$$

subject to the dynamics of the spot price, S_t, in equation (1.3), (1.5b)

τ_j is an $\mathcal{F}_t$ stopping time, (1.5c)

$$T_j = \frac{Q_j}{K_j}, \qquad K \le K_j \le \bar{K}, \tag{1.5d}$$

and the border condition $F_0(S,K) = 0$ for all S, K. (1.5e)

The dynamic program (1.5) is in general difficult to solve, so in the sequel we approximate its solution in two steps. First, in Section 1.3, we study the optimal timing of extraction for a fixed capacity K. Then, in Section 1.4 we relax this assumption and derive near-optimal capacity expansion decisions.

1.2.5 Notation and Conventions

Unfortunately, notation will play an important role in our presentation. This is in part due to the fact that we approach the problem from various angles each one requiring its own set of notation. Hence, we find convenient to introduce at this point some general notation and conventions that we will use throughout the rest of the chapter.

First, we say that a function $f(S)$ is *asymptotically equal* to a function $g(S)$, which we denote by $f(S) \stackrel{S\to\infty}{\longrightarrow} g(S)$, if $\lim_{S\to\infty} |f(S) - g(S)| = 0$. Also, the first and second derivatives of a smooth function $f(S)$ with respect to S are denoted by $f'(S)$ and $f''(S)$, respectively.

We let $C^k(\mathbb{R}_+)$ be the set of real-valued continuous functions on $\mathbb{R}_+$ having derivatives of order $k \geq 0$. We also define the set C^2_+ that plays a key role in our characterization of the value function:

$$C^2_+ := \Big\{ f \in C^1(\mathbb{R}_+) \;:\; \text{there exist a scalar } \theta_f \text{ and a finite set } N_f \subseteq \mathbb{R}_+ \text{ (both possibly depend on } f\text{) such that } |f'(S)| \leq \theta_f \; \forall S \in \mathbb{R}_+ \text{ and } f''(S) \text{ exists } \forall S \in \mathbb{R}_+ \setminus N_f \Big\}.$$

Consider two arbitrary vectors $X = (X_j)$ and $\alpha = (\alpha_j)$, we define

$$X^+_{k,j} := \sum_{h=k+1}^{j} X_h, \qquad \alpha^\times_{k,j} := \prod_{h=k+1}^{j} \alpha_h \quad \text{and} \quad (\alpha^\times X)^+_{k,j} := \sum_{h=k+1}^{j} \alpha^\times_{h,j} X_h.$$

We use the specialized notation $X^+_j := X^+_{0,j}$, $\alpha^\times_j := \alpha^\times_{0,j}$ and $(\alpha^\times X)^+_j := (\alpha^\times X)^+_{0,j}$. In the usage of summations and multiplications we adopt the convention $\sum_{h=k}^{j} X_h = 0$ and $\prod_{h=k}^{j} X_h = 1$ if $j < k$.

For $j \leq N$, define the average production cost $\mathbb{C}_j(K) := C_j(K)/R_j(K)$, and

$$\mathcal{R}_{k,j}(K) := \sum_{h=k+1}^{j} e^{-\rho T^+_{h,j}(K)} R_h(K), \qquad C_{k,j}(K) := \sum_{h=k+1}^{j} e^{-r T^+_{h,j}(K)} C_h(K),$$

$\mathcal{R}_j(K) := \mathcal{R}_{0,j}(K)$ and $C_j(K) \triangleq C_{0,j}(K)$. The interpretation of these quantities is as follows. Suppose there are j blocks left, the spot price is S and the capacity is K. Then, if the decision maker decides to extract the j blocks (starting with block j and finishing with block 1) without changing capacity or stopping at any time then the discounted expected payoff of this non-idling policy would be $\mathcal{W}_j(S,K):=\mathcal{R}_j(K)S - C_j(K)$.

1.3 Optimal Production Plan with Fixed Capacity

In this section we solve formulation (1.5) under the assumption that capacity expansions are not allowed and we let K be this fixed capacity. (Thus, for convenience, we drop dependencies on K in this section.) We will relax this assumption in the following section. In this setting, we derive analytically upper and lower bounds, as well as two asymptotic approximations, for the corresponding value function. Part of the analysis in this section follows closely and extends the results in §5.2 of [19].

Let $F_j(S)$ be the maximum expected discounted profit when there are j blocks left and the spot price is S. We solve for the sequence of value functions $\{F_j(S) : j = 1,\dots,N\}$ using forward induction on j. That is, starting with the border condition $F_0(S) = 0$, we can compute sequentially $F_1(S), F_2(S),\dots,F_N(S)$ solving (1.5). To this end, suppose that we have already computed the value function $F_{j-1}(S)$ and let us solve for $F_j(S)$. For this, let us define the auxiliary function

$$G_j(S) := W_j(S) + e^{-rT_j}\,\mathbb{E}^{\mathbb{Q}}\Big[F_{j-1}(S_{T_j})\Big|S_0 = S\Big].$$

With this definition, problem (1.5) is equivalent to

$$F_j(S) = \sup_{\tau\geq 0}\mathbb{E}^{\mathbb{Q}}\Big[e^{-r\tau}G_j(S_\tau)\Big|S_0 = S\Big], \tag{1.6}$$

where the supremum is taken over the set of stopping times τ with respect to $\mathcal{F}_t$ (the filtration generated by S_t) which represents the time when block j should start being produced. To solve this optimal stopping problem, we impose optimality conditions in the form of a set of partial differential inequalities (quasi-variational inequalities or QVI) that characterize the optimal stopping time. To this end, let us define the operator $\mathcal{A}$ that applies on functions $F \in C^2_+$ as follows (see section §1.2.5 for definitions):

$$\mathcal{A}F(S) := \frac{1}{2}\sigma^2 S^2 F''(S) + (r-\rho)S\,F'(S) - rF(S), \quad \text{for all } S \in \mathbb{R}_+ \setminus N_F.$$

Definition 1 (QVI) *The function $F \in C^2_+$ satisfies the quasi-variational inequalities for problem* (1.6) *if the following three conditions are satisfied:*

$$\begin{aligned} \mathcal{A}F(S) &\leq 0 && \textit{for all } S \in \mathbb{R}_+ \setminus N_F, \\ F(S) - G_j(S) &\geq 0 && \textit{for all } S \geq 0, \textit{ and} \\ (F(S) - G_j(S))\,\mathcal{A}F(S) &= 0 && \textit{for all } S \in \mathbb{R}_+ \setminus N_F. \quad \square \end{aligned} \tag{1.7}$$

As one would expect, a solution to these QVI conditions partition the state space $S \geq 0$ into two regions: a *continuation* region in which the optimal policy is to delay production, and an *intervention* region in which production should start immediately.

$$\begin{aligned}
\text{Continuation: } & \mathcal{D}{:=}\{S \geq 0 : F(S) > G_j(S) \quad \text{and} \quad \mathcal{A}F(S) = 0\}, \\
\text{Intervention: } & \mathcal{I}{:=}\{S \geq 0 : F(S) = G_j(S) \quad \text{and} \quad \mathcal{A}F(S) \leq 0\}.
\end{aligned}$$

For every solution of the QVI we can associate a stopping time τ as follows.

Definition 2 *Let $F \in C^2_+$ be a solution of the QVI in* (1.7)*. We define the QVI-control τ as follows*

$$\tau{:=}\inf\{t \geq 0 : F(S_t) = G_j(S_t)\}.$$

We are now ready to formalize the *verification* theorem that provides the connection between the QVI conditions and the original optimization problem in (1.6).

Theorem 1 (Verification) *Let $F \in C^2_+$ be a nonnegative solution of the QVI in* (1.7)*. Then,*

$$F(S) \geq F_j(S) \quad \textit{for every } S \geq 0.$$

In addition, if the continuation region $\mathcal{D}$ is bounded and there exists a QVI-control τ associated with F, i.e.*, $\mathbb{E}^{\mathbb{Q}}[\tau] < \infty$, then it is optimal and $F(S) = F_j(S)$.*

According to this result, we can tackle the problem of determining the value function $F_j(S)$ by solving the QVI conditions. This is in general a difficult task given the *free boundary* nature of these conditions (*i.e.*, part of the problem is to determine the intervention and continuation regions). We approach this problem using an "educated guess". Intuitively, we expect that if the spot price is sufficiently large then immediate production should be an optimal decision. If this intuition is correct, then we expect that there exists a single threshold value S^*_j such that $\mathcal{D} = \{0 \leq S < S^*_j\}$ and $\mathcal{I} = \{S \geq S^*_j\}$. In what follows, we solve the QVI conditions imposing this additional condition, and then use Theorem 1 to verify that our proposed solution is indeed optimal.

Based on our previous discussion, the QVI conditions imply that

$$\begin{aligned}
0 &= \mathcal{A}F_j(S) \qquad \text{for all } S < S^*_j, \\
F_j(S) &= G_j(S) \qquad \text{for all } S \geq S^*_j.
\end{aligned} \tag{1.8}$$

The first equation above corresponds to the Hamilton-Jacobi Bellman (HJB) equation, and the second is known as a *value-matching* condition. In addition, we expect the following two conditions to hold

$$F_j(0) = 0 \qquad \text{and} \qquad F'_j(S^*_j) = G'_j(S^*_j).$$

The first condition simply states that if the price process reaches the absorbing state $S = 0$, then the value of the mining project will be zero as well. The second condition guarantees that $F_j(S)$ is differentiable at the threshold price $S = S_j^*$ (*i.e.*, a *smooth-pasting* condition). This condition is necessary to ensure that $F_j(S) \in C_+^2$.

Equation (1.8) is a second-order homogeneous ordinary differential equation. Because of its special structure, its general solution can be expressed as a linear combination of any two independent solutions. The function S^β satisfies the equation provided that β is a root of the following quadratic equation

$$\frac{1}{2}\sigma^2\beta(\beta-1)+(r-\rho)\beta-r=0.$$

The two roots are

$$\beta := \frac{1}{2}-(r-\rho)/\sigma^2+\sqrt{1/4+\left[(r-\rho)/\sigma^2\right]^2+(r+\rho)/\sigma^2}>1, \quad \text{and}$$
$$\tilde{\beta} := \frac{1}{2}-(r-\rho)/\sigma^2-\sqrt{1/4+\left[(r-\rho)/\sigma^2\right]^2+(r+\rho)/\sigma^2}<0.$$

The general solution to equation (1.8) can be written as

$$F_j(S)=M_jS^\beta+\tilde{M}_jS^{\tilde{\beta}},$$

for two constants M_j and $\tilde{M}_j$. However, the border condition at $S = 0$ implies that $\tilde{M}_j = 0$. In conclusion, our candidate solution $F_j(S)$ is given by

$$F_j(S)=\begin{cases} M_jS^\beta, \text{ if } S<S_j^*,\\ G_j(S), \text{ otherwise.}\end{cases} \tag{1.9}$$

We are now ready to state the main result of this section.

Theorem 2 *For every block j, there exists a pair of nonnegative scalars* (M_j, S_j^*) *such that the value function* $F_j(S)$ *in* (1.6) *is given by* (1.9). *The values of* M_j *and* S_j^* *are equal to*

$$M_j := \min\{M\geq 0 : MS^\beta \geq G_j(S)\,\textit{for all}\; S\geq 0\} \quad \textit{and}$$
$$S_j^* := \min\{S\geq 0 : M_j(S)^\beta = G_j(S)\}.$$

Moreover, the value function $F_j(S)$ *is increasing and convex in* $S \geq 0$.

The proof of Theorem 2, which can be found in [10], shows that the characterization above implies that M_j and S_j^* satisfy the value matching and smooth-pasting conditions

$$M_j(S_j^*)^\beta = G_j(S_j^*) \qquad \text{and} \qquad \beta M_j(S_j^*)^{\beta-1}=G_j'(S_j^*). \tag{1.10}$$

For $j = 1$, condition (1.9) reduces to

$$F_1(S) = \begin{cases} M_1 S^{\beta}, & \text{if } S < S_1^*, \\ R_1 S - C_1, & \text{if } S \geq S_1^*, \end{cases} \tag{1.11}$$

and the value-matching and smooth-pasting conditions become

$$M_1 (S_1^*)^{\beta} = R_1 S_1^* - C_1 \quad \text{and} \quad M_1 \beta (S_1^*)^{\beta-1} = R_1,$$

which lead to

$$S_1^* = \frac{\beta C_1}{(\beta-1) R_1} = \frac{\beta}{\beta-1} \mathbb{C}_1 \quad \text{and} \quad M_1 = \frac{C_1}{\beta-1} \left(S_1^*\right)^{-\beta}. \tag{1.12}$$

Recall from Section 1.2.5 that $\mathbb{C}_1$ is the average per unit extraction cost for block 1. Hence, the choice of S_1^* above guarantees a per unit net margin of $\frac{1}{\beta-1}$.

Unfortunately, extending the previous analysis to compute $F_j(S)$ for an arbitrary j is difficult because of the expectation $\mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j}) \big| S_0 = S]$ in the definition of $G_j(S)$ and there is no simple characterization of $F_j(S)$ for $j \geq 2$. Nevertheless, we have been able to establish a useful asymptotic property of $F_j(S)$ (see §1.2.5 for notation).

Proposition 1 *The value function $F_j(S)$ is asymptotically equal to $\mathcal{W}_j(S)$, that is,*

$$F_j(S) \quad \overset{S\to\infty}{\longrightarrow} \quad \mathcal{W}_j(S) = \mathcal{R}_j S - C_j.$$

Proposition 1 highlights some important properties of the value function but it does not provide tight estimates of $F_j(S)$ unless S is large. For small values of S we could use numerical methods to solve the recursion in (1.9) and get an approximation of the value function. Instead of following this numerical approach, we have chosen to derive some closed-form approximations for $F_j(S)$ that provide insight about the structure of this solution and its dependence on the model parameters. First, we develop a family of approximations for $F_j(S)$, which include valid lower and upper bounds as special cases. Then, we use asymptotic analysis to extend these bounds. We conclude this section with some numerical computations that compare the performance of these bounds.

1.3.1 Upper Bound

To obtain an upper bound on the value of $F_j(S)$ we assume that the extraction of block $j-1$ can start even if the extraction of block j is not fully completed but simply started. We will use a superscript 'U' to distinguish those quantities that are derived using this approximation. For example, $F_j^{\mathrm{U}}(S)$ denotes the value function resulting from this approximation. Because $F_j^{\mathrm{U}}(S)$ is the solution of a less restricted problem it follows that $F_j(S) \le F_j^{\mathrm{U}}(S)$.

Similar to the original optimization in (1.6), the bound $F_j^{\mathrm{U}}(S)$ satisfies the following recursion

$$F_j^{\mathrm{U}}(S) = \sup_{\tau\ge 0} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau} W_j(S_\tau) + e^{-r\tau} F_{j-1}^{\mathrm{U}}(S_\tau) \,\middle|\, S_0 = S \right],$$

with $F_0^{\mathrm{U}}(S) = 0$ for all $S \ge 0$. It is not hard to see that $F_j^{\mathrm{U}}(S)$ satisfies the HJB equation (1.8) inside the continuation region. Therefore, it follows that

$$F_j^{\mathrm{U}}(S) = \begin{cases} M_j^{\mathrm{U}} S^{\beta}, & \text{if } S \le S_j^{\mathrm{U}}, \\ R_j S - C_j + F_{j-1}^{\mathrm{U}}(S), & \text{otherwise.} \end{cases} \tag{1.13}$$

We can use backward induction to compute recursively M_j^{U} and S_j^{U}, starting at block 1. We postpone this analysis to Section 1.3.3 where we derive an algorithm that performs this task efficiently.

1.3.2 Lower Bound

We can get a lower bound for the value of $F_j(S)$ using the convexity of the value function and Jensen's inequality. We will use a superscript 'L' to denote quantities that are derived using this approximation.

Consider again the optimal stopping time problem for $F_j(S)$ in (1.6)

$$F_j(S) = \sup_{\tau\ge 0} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau} W_j(S_\tau) + e^{-r(\tau+T_j)} F_{j-1}(S_{\tau+T_j}) \,\middle|\, S_0 = S \right].$$

Suppose there exists a convex function $F_{j-1}^{\mathrm{L}}(S)$ such that $F_{j-1}^{\mathrm{L}}(S) \le F_{j-1}(S)$ for all $S \ge 0$. Then,

$$F_j(S) \ge \sup_{\tau\ge 0} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\tau} W_j(S_\tau) + e^{-r(\tau+T_j)} F_{j-1}^{\mathrm{L}}(S_{\tau+T_j}) \,\middle|\, S_0 = S \right].$$

For an arbitrary stopping time τ let $\mathfrak{F}_\tau$ be the σ-algebra generated by τ. Then, using iterated (conditional) expectation, the convexity of $F_{j-1}^{\mathrm{L}}(S)$ and condition (1.2) we get that

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\left[e^{-r(\tau+T_j)} F_{j-1}^{\mathrm{L}}(S_{\tau+T_j}) \,\middle|\, S_0 = S \right] &= \mathbb{E}^{\mathbb{Q}}\left[e^{-r(\tau+T_j)} \mathbb{E}^{\mathbb{Q}}\left[F_{j-1}^{\mathrm{L}}(S_{\tau+T_j}) \,\middle|\, \mathfrak{F}_\tau \right] \,\middle|\, S_0 = S \right] \\ &\ge \mathbb{E}^{\mathbb{Q}}\left[e^{-r(\tau+T_j)} F_{j-1}^{\mathrm{L}}\left(\mathbb{E}^{\mathbb{Q}}[S_{\tau+T_j} \,|\, \mathfrak{F}_\tau] \right) \,\middle|\, S_0 = S \right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[e^{-r(\tau+T_j)} F_{j-1}^{\mathrm{L}}\left(e^{(r-\rho)T_j} S_\tau \right) \,\middle|\, S_0 = S \right] \end{aligned}$$

and so

$$F_j(S) \geq \sup_{\tau \geq 0} \mathbb{E}^{\mathbb{Q}} \left[e^{-r\tau} W_j(S_\tau) + e^{-r(\tau+T_j)} F^{\mathrm{L}}_{j-1}\left(e^{(r-\rho)T_j} S_\tau\right) \Big| S_0 = S \right].$$

From this bound, we derive the following recursion for $F^{\mathrm{L}}_j(S)$

$$F^{\mathrm{L}}_j(S) = \sup_{\tau \geq 0} \mathbb{E}^{\mathbb{Q}} \left[e^{-r\tau} W_j(S_\tau) + e^{-r(\tau+T_j)} F^{\mathrm{L}}_{j-1}\left(e^{(r-\rho)T_j} S_\tau\right) \Big| S_0 = S \right],$$

with $F^{\mathrm{L}}_0(S) = 0$ for all $S \geq 0$. Using a line of arguments similar to the one used to derive (1.9), we can show that

$$F^{\mathrm{L}}_j(S) = \begin{cases} M^{\mathrm{L}}_j S^{\beta}, & \text{if } S \leq S^{\mathrm{L}}_j, \\ R_j S - C_j + e^{-rT_j} F^{\mathrm{L}}_{j-1}(e^{(r-\rho)T_j} S), & \text{otherwise}, \end{cases} \tag{1.14}$$

where M^L_j and S^L_j satisfy value-matching and smooth-pasting conditions, and that $F^{\mathrm{L}}_j(S) \leq F_j(S)$ for all $S \geq 0$ and j.

In the following section, we provide a general method to compute $F^{\mathrm{L}}_j(S)$ and verify that it is indeed a convex function as required (see Corollary 1 below). Our approach is based on a general family of approximations that includes $F^{\mathrm{U}}_j(S)$ and $F^{\mathrm{L}}_j(S)$ as special cases.

1.3.3 (α,η)*-Approximations*

The recursions that define the upper bound $F^{\mathrm{U}}_j(S)$ in (1.13) and the lower bound $F^{\mathrm{L}}_j(S)$ in (1.14) share a similar structure that can be exploited to derive a unified approximation method.

Definition 3 *Let $\alpha = (\alpha_j)$ and $\eta = (\eta_j)$ be two positive vectors. We say that a set of continuous and differentiable functions $\{\mathcal{F}_j(S) : j = 0,\ldots,N\}$ is an (α,η)-approximation of the value functions in (1.9) if $\mathcal{F}_0(S) = 0$ for all $S \geq 0$ and*

$$\mathcal{F}_j(S) = \begin{cases} \mathcal{M}_j S^{\beta}, & \text{if } S \leq \mathcal{S}_j, \\ R_j S - C_j + \alpha_j \mathcal{F}_{j-1}(\eta_j S), & \text{otherwise}, \end{cases} \qquad j = 1,\ldots,N. \tag{1.15}$$

Because $\mathcal{F}_j(S)$ is continuous and differentiable, the values of $\mathcal{S}_j$ and $\mathcal{M}_j$ are implicitly determined imposing value matching and smooth pasting conditions similar to those in equation (1.10).

Note that this family of approximations generalizes the upper and lower bounds. Indeed, it follows from (1.13) that the upper bound $F^{\mathrm{U}}_j(S)$ is a special case of (1.15) with $\alpha_j = \eta_j = 1$. Similarly, we can recover the lower bound $F^{\mathrm{L}}_j(S)$ if we chose $\alpha_j = \exp(-rT_j)$ and $\eta_j = \exp((r-\rho)T_j)$.

In what follows, we derive an efficient algorithm that solves (1.15) for an arbitrary (α,η)-approximation. In order to get some intuition on how the algorithm works, let us first consider the special case of two blocks, that is, $N = 2$.

Using backward induction, we first compute $\mathcal{F}_1(S)$. In this case, the solution to (1.15) is identical to the solution in (1.11) and (1.12). That is,

$$\mathcal{F}_1(S) = \begin{cases} \mathcal{M}_1 S^{\beta}, & \text{if } S \le \mathcal{S}_1, \\ R_1 S - C_1, & \text{if } S \ge \mathcal{S}_1, \end{cases} \quad \text{where } \mathcal{S}_1 = \left(\frac{\beta}{\beta-1}\right)\mathbb{C}_1 \quad \text{and} \quad \mathcal{M}_1 = \left(\frac{C_1}{\beta-1}\right)(\mathcal{S}_1)^{-\beta}. \tag{1.16}$$

Based on this solution, we can solve for $\mathcal{F}_2(S)$. As before, we compute the value of $\mathcal{M}_2$ and $\mathcal{S}_2$ using the value matching and smooth pasting conditions

$$\mathcal{M}_2(\mathcal{S}_2)^{\beta} = R_2\mathcal{S}_2 - C_2 + \alpha_2\mathcal{F}_1(\eta_2\mathcal{S}_2) \quad \text{and} \quad \beta\mathcal{M}_2(\mathcal{S}_2)^{\beta-1} = R_2 + \alpha_2\eta_2\mathcal{F}_1'(\eta_2\mathcal{S}_2).$$

We identify two possible cases depending on the value of $\mathcal{F}_1(\eta_2\mathcal{S}_2)$. Suppose first that $\mathcal{S}_1 \ge \eta_2\mathcal{S}_2$, then $F_1(\eta_2\mathcal{S}_2) = \mathcal{M}_1(\eta_2\mathcal{S}_2)^{\beta}$ and the value-matching and smooth-pasting conditions imply that

$$\mathcal{S}_2 = \left(\frac{\beta}{\beta-1}\right)\mathbb{C}_2 \qquad \text{and} \qquad \mathcal{M}_2 = \alpha_2\eta_2^{\beta}\mathcal{M}_1 + \left(\frac{C_2}{\beta-1}\right)(\mathcal{S}_2)^{-\beta}.$$

The corresponding value of $\mathcal{F}_2(S)$ has three pieces

$$\mathcal{F}_2(S) = \begin{cases} \mathcal{M}_2 S^{\beta} & \text{if } S, \le \mathcal{S}_2, \\ R_2 S - C_2 + \alpha_2\eta_2^{\beta} S^{\beta}\mathcal{M}_1, & \text{if } \mathcal{S}_2 \le S \le \mathcal{S}_1/\eta_2, \\ (R_2 + \alpha_2\eta_2 R_1)S - (C_2 + \alpha_2 C_1), & \text{if } S \ge \mathcal{S}_1/\eta_2. \end{cases}$$

Note that the requirement $\mathcal{S}_1 \ge \eta_2\mathcal{S}_2$ is equivalent to $\mathbb{C}_1 \ge \eta_2\mathbb{C}_2$.

Let us now consider the case where $\mathcal{S}_1 < \eta_2\mathcal{S}_2$. It follows that $\mathcal{F}_1(\eta_2\mathcal{S}_2) = R_1\eta_2\mathcal{S}_2 - C_1$ and the value-matching and smooth-pasting conditions lead to

$$\mathcal{S}_2 = \left(\frac{\beta}{\beta-1}\right)\frac{C_2 + \alpha_2 C_1}{R_2 + \alpha_2\eta_2 R_1} \quad \text{and} \quad \mathcal{M}_2 = \left(\frac{C_2 + \alpha_2 C_1}{\beta-1}\right)(\mathcal{S}_2)^{-\beta}$$

and

$$\mathcal{F}_2(S) = \begin{cases} \mathcal{M}_2 S^{\beta}, & \text{if } S \le \mathcal{S}_2, \\ (R_2 + \alpha_2\eta_2 R_1)S - (C_2 + \alpha_2 C_1), & \text{if } S \ge \mathcal{S}_2. \end{cases} \tag{1.17}$$

In this case, one can show that the condition $\mathcal{S}_1 < \eta_2\mathcal{S}_2$ is equivalent to $\mathbb{C}_1 < \eta_2\mathbb{C}_2$, which is consistent with the previous case.

Note that the actual value of $\mathcal{F}_2(S)$ depends on the relationship between $\mathbb{C}_1$ and $\eta_2\mathbb{C}_2$. Interestingly, for the case $\mathbb{C}_1 < \eta_2\mathbb{C}_2$ the value of $\mathcal{F}_2(S)$ in (1.17) is analogous to the value of $\mathcal{F}_1(S)$ in (1.16). Indeed, in this case we can combine the two blocks into a single one so that the solution in (1.17) is equivalent to a single-block project with modified extraction cost $C_2+\alpha_2 C_1$ and modified mineral content $R_2+\alpha_2\eta_2 R_1$. To get some intuition about why the two blocks are "pooled" together, let us consider the lower bound approximation. In this case the condition $\mathcal{S}_1 < \eta_2\mathcal{S}_2$ is equivalent to $S_1^{\text{L}} < \exp((r-\rho)T_2)S_2^{\text{L}} = \mathbb{E}^{\mathbb{Q}}[S_{T_2}|S_0 = S_2^{\text{L}}]$. In other words, blocks 1 and 2 are combined when the threshold price for block 1 is below the expected value of the spot price at the time when extraction of block 2 is completed. Hence, in expectation, the extraction of blocks 1 and 2 is performed without interruption and so we can view these two blocks as a single one.

The following proposition extends the previous two-block analysis to the case of an arbitrary number of blocks. Embedded in this proposition, there is an algorithm that takes as input a j-block project with characteristics $\{(C_k,R_k,\alpha_k,\eta_k),\ k=1,\dots,j\}$ and produces a $\tilde{j}$-block project with characteristics $\{(\widetilde{C}_k,\widetilde{R}_k,\widetilde{\alpha}_k,\widetilde{\eta}_k),\ k=1,\dots,\tilde{j}\}$ and $\tilde{j}\le j$. The algorithm aggregates blocks using the same criteria discussed above. The resulting sequence $\{(\widetilde{C}_k,\widetilde{R}_k,\widetilde{\alpha}_k,\widetilde{\eta}_k),\ k=1,\dots,\tilde{j}\}$ satisfies some properties that greatly simplify the computation of $\mathcal{F}_j(S)$. (See section 1.2.5 for notation)

Proposition 2 *Consider a project with j blocks with characteristics* $\{(C_k,R_k,\alpha_k,\eta_k),\ k=1,\dots,j\}$ *and use the following algorithm to create an artificial sequence of (possibly aggregated) blocks.*

<u>ALGORITHM:</u>

Step 0: (Initialization) *Set* $\widetilde{C}_k = C_k$, $\widetilde{R}_k = R_k$, $\widetilde{\alpha}_k = \alpha_k$ *and* $\widetilde{\eta}_k = \eta_k$, $k = 1,\dots,j$ *and* $\tilde{j} = j$.

Step 1: *Compute the auxiliary variables*

$$\widetilde{\theta}_k := \widetilde{\alpha}_k\widetilde{\eta}_k, \qquad \widetilde{\mathbb{C}}_k := \frac{\widetilde{C}_k}{\widetilde{R}_k} \qquad \text{and} \qquad \widetilde{\mathbb{C}}_{k,l} := \frac{(\widetilde{\alpha}^{\times}\widetilde{C})^{+}_{k-1,l}}{(\widetilde{\theta}^{\times}\widetilde{R})^{+}_{k-1,l}}, \quad \text{for all } k,l=1,\dots,\tilde{j},\ k\le l.$$

Step 2: *Find* $\tilde{k} = \min\{2\le k\le\tilde{j}\ :\ \widetilde{\mathbb{C}}_{k-1} < \widetilde{\eta}_k\widetilde{\mathbb{C}}_k\}$. *If such* $\tilde{k}$ *does not exist then stop.*

Step 3: *Find* $\tilde{h} = \max\{1\le h\le\tilde{k}-1\ :\ \widetilde{\eta}^{\times}_{h,\tilde{k}}\widetilde{\mathbb{C}}_{h+1,\tilde{k}} \le \widetilde{\mathbb{C}}_h\}$. *If such* $\tilde{h}$ *does not exist then set* $\tilde{h}=0$.

Step 4: *Define* $\xi = \tilde{k}-\tilde{h}-1$ *and introduce the following transformation:* $\tilde{j} = \tilde{j}-\xi$ *and*

$$(\widetilde{R}_k,\ \widetilde{C}_k,\ \widetilde{\alpha}_k,\ \widetilde{\eta}_k) = \begin{cases} (\widetilde{R}_k,\ \widetilde{C}_k,\ \widetilde{\alpha}_k,\ \widetilde{\eta}_k), & \text{if } k\le\tilde{h},\\ ((\widetilde{\theta}^{\times}\widetilde{R})^{+}_{\tilde{h},\tilde{k}},\ (\widetilde{\alpha}^{\times}\widetilde{C})^{+}_{\tilde{h},\tilde{k}},\ \widetilde{\alpha}^{\times}_{\tilde{h},\tilde{k}},\ \widetilde{\eta}^{\times}_{\tilde{h},\tilde{k}}), & \text{if } k=\tilde{h}+1,\\ (\widetilde{R}_{k+\xi},\ \widetilde{C}_{k+\xi},\ \widetilde{\alpha}_{k+\xi},\ \widetilde{\eta}_{k+\xi}), & \text{if } \tilde{h}+2\le k\le\tilde{j}. \end{cases}$$

(Note that in this step we have created a new block $\tilde{h}+1$ *by aggregating all the blocks from* $\tilde{h}+1$ *to* $\tilde{k}$*, hence the total number of blocks has decreased by* ξ*.)*

Step 5: *Goto step 1.* □

After the algorithm has stopped no further block aggregation is possible. The output of the algorithm is a modified project that has $\tilde{\jmath}$ blocks. The k-th block in this new sequence has mineral content $\widetilde{R}_k$ and extraction cost $\widetilde{C}_k$. For this modified sequence of (possibly aggregated) blocks we define

$$\widetilde{\mathcal{S}}_k = \left(\frac{\beta}{\beta-1}\right)\widetilde{\mathbb{C}}_k \qquad \text{and} \qquad \widetilde{\mathcal{M}}_k = \widetilde{\alpha}_k \widetilde{\eta}_k^{\beta}\, \widetilde{\mathcal{M}}_{k-1} + \left(\frac{\widetilde{C}_k}{\beta-1}\right)(\widetilde{\mathcal{S}}_k)^{-\beta}, \qquad k \le \tilde{\jmath}, \tag{1.18}$$

with $\widetilde{\mathcal{M}}_0 = 0$. Finally, for block j in the original configuration we have that $\mathcal{S}_j = \widetilde{\mathcal{S}}_{\tilde{\jmath}}$ and

$$\mathcal{F}_j(S) = (\widetilde{\theta}^{\times} \widetilde{R})^{+}_{h,\tilde{\jmath}}\, S - (\widetilde{\alpha}^{\times} \widetilde{C})^{+}_{h,\tilde{\jmath}} + \widetilde{\mathcal{M}}_h\, \widetilde{\alpha}^{\times}_{h,\tilde{\jmath}}\, (\widetilde{\eta}^{\times}_{h,\tilde{\jmath}})^{\beta}\, S^{\beta}, \tag{1.19}$$

where $h = \max\{0 \le k \le \tilde{\jmath} \,|\, \widetilde{\mathcal{S}}_k \ge \widetilde{\eta}^{\times}_{h,\tilde{\jmath}}\, S\}$ and $\widetilde{\mathcal{S}}_0 = \infty$.

Corollary 1 *The function $\mathcal{F}_j(S)$ in equation (1.19) is convex in S.*

Example 1: To illustrate the mechanics of the algorithm in Proposition 2, let us consider a six-block mining sector with the following characteristics.

Block	R_k	C_k	T_k	$\mathbb{C}_k$
1	0.25	14	1.2	56
2	0.3	9	1.6	30
3	0.4	16	1	40
4	0.32	10	2	31.25
5	0.35	12.25	0.7	35
6	0.4	18	0.9	45

Consider a discount factor $r = 0.12$ and a convenience yield $\rho = 0.06$. Let us specialize the result in Proposition 2 to the case of the lower bound $F^{\mathrm{L}}(S)$. For this, set $\alpha_k = e^{-rT_k}$ and $\eta_k = e^{(r-\rho)T_k}$, for $k \le 6$.

In the first iteration of the algorithm we find (step 2) that $\tilde{k} = 3$. We then compute $\widetilde{\mathbb{C}}_{2,3} \cdot e^{(r-\rho)\cdot T^{+}_{2,3}} = 41.07 < \widetilde{\mathbb{C}}_1$ and $\widetilde{\mathbb{C}}_{3,3} \cdot e^{(r-\rho)\cdot T^{+}_{3,3}} = 40 > \widetilde{\mathbb{C}}_2$ and conclude (step 3) that $\tilde{h} = 1$. From step 4, we get $\xi = 1$ and the new number of blocks is $\tilde{\jmath} = 5$ (blocks 2 and 3 are pooled together). The following four tables summarize the resulting values of $\widetilde{R}_k$ and $\widetilde{C}_k$ after the first, second, third and fourth iterations of the algorithm. Note that in order to update the values of $\widetilde{\alpha}_k$ and $\widetilde{\eta}_k$ it is sufficient to update the values of the processing time $\widetilde{T}_k$.

As we can see, the algorithm finishes after four iterations and in the final configuration the mining project consists of only two blocks. The initial block 1 and a new block 2 that aggregates the original blocks 2 to 6. From equation (1.18), we derive the threshold price for block 6 in the original block configuration which is

$$\widetilde{\mathcal{S}}_6 = \left(\frac{\beta}{\beta-1}\right)\left(\frac{53.39}{1.58}\right) = \frac{33.79\beta}{\beta-1},$$

After Iteration 1

Block Number New	Original	$\widetilde{R}_k$	$\widetilde{C}_k$	$\widetilde{T}_k$
1	1	0.25	14	1.2
2	2 and 3	0.68	24	2.6
3	4	0.32	10	2
4	5	0.35	12.25	0.7
5	6	0.4	18	0.9

After Iteration 2

Block Number New	Original	$\widetilde{R}_k$	$\widetilde{C}_k$	$\widetilde{T}_k$
1	1	0.25	14	1.2
2	2, 3 and 4	0.93	29.5	4.6
3	5	0.35	12.25	0.7
4	6	0.4	18	0.9

After Iteration 3

Block Number New	Original	$\widetilde{R}_k$	$\widetilde{C}_k$	$\widetilde{T}_k$
1	1	0.25	14	1.2
2	2, 3, 4 and 5	1.25	39.43	5.3
3	6	0.4	18	0.9

After Iteration 4

Block Number New	Original	$\widetilde{R}_k$	$\widetilde{C}_k$	$\widetilde{T}_k$
1	1	0.25	14	1.2
2	2, 3, 4, 5 and 6	1.58	53.39	6.2

The interpretation of this price is as follows: As soon as the spot price goes above $\widetilde{\mathcal{S}}_6$ we should start extracting block 6. □

The algorithm in Proposition 2 provides a simple method to reduce the size of the mining project by appropriately aggregating blocks and then computing the value function and threshold prices for the modified block configuration. In practice, blocks cannot be pooled together and must be extracted one at a time and so the modified sequence of blocks is of limited practical use for extraction purposes. Nevertheless, we can define a simple feasible extraction policy based on the solution proposed by Proposition 2 as follows.

Extraction Policy based on the Approximation $\mathcal{F}_j(S)$:

1. Consider a sector with j blocks. Using the algorithm in proposition 2, aggregate blocks to obtain a new block configuration with $\tilde{j} \leq j$ blocks.
2. For this artificial configuration compute the threshold price $\widetilde{\mathcal{S}}_{\tilde{j}}$ using equation (1.18).
3. For the original block configuration with j blocks start extracting block j as soon as the spot price exceeds $\widetilde{S}_{\tilde{j}}$. Note that $\widetilde{\mathcal{S}}_{\tilde{j}} = \mathcal{S}_j$, which the optimal threshold price for the approximation $\mathcal{F}_j(S)$.
4. Once the extraction of block j is completed, iterate this sequence of steps for the remaining $j-1$ blocks. □

The previous policy uses the artificial configuration of blocks proposed by Proposition 2 only to compute the threshold price that determines when the first block (in the original sequence) should start being processed. For instance, the results in Example 1 suggests that block 6 should start processing as soon as the spot price satisfies $S_t \geq 33.79\beta/(\beta-1)$.

We conclude this section with a brief discussion on how the sequence of extraction is chosen in practice. When a mining project is designed, its extraction sequence is implicitly built. For example, in underground mines, a common design rule is to select the initial extraction front near blocks with higher grade. The idea behind this (greedy) rule is to extract the better material first at the lowest marginal cost. This simple rule has two important consequences from the point of view of our solution. First, the quality of the ore tends to be a decreasing function of the extraction front and so the parameter R_k is usually increasing in k (remember that we have indexed blocks backward with block 1 being the last block in the sequence). Second, as the operation advances through the extraction fronts, it is executed farther away from the initial front and the extracted materials must be moved a longer distance. This additional transportation increases the marginal extraction cost, that is, C_k is generally decreasing in k. Hence, we expect $\mathbb{C}_k = \frac{C_k}{R_k}$ to be a decreasing function of k under this greedy design. According to step 2 in the algorithm in Proposition 2, there is no block aggregation if $\mathbb{C}_{k-1} \geq \eta_k \mathbb{C}_k$. Hence, we can roughly say that there should be no block aggregation under this greedy design rule if η_k does not exceed one by much. This condition holds trivially for the case of the upper bound ($\eta_k = 1$). For the lower bound, $\eta_k = e^{(r-\rho)T_k}$ and so we expect no block aggregation if the discount factor and/or the processing times are small.

1.3.4 Asymptotic Approximations

In this subsection we characterize the limiting behavior of the upper and lower bounds as the spot price goes to ∞ and use it to propose two simple approximations for the value function.

Figure 1.3 (left panel) plots the value function $F_j(S)$ (numerically computed), the upper bound $F^U_j(S)$ and the lower bound $F^L_j(S)$ as a function of S using the data in Example 1. We note that both bounds perform well for small values of S, however, as S gets large the lower bound performs substantially better. The upper bound has an optimality gap $F^U_j(S) - F_j(S)$ that increases monotonically with S. This is in part due to the fact that the upper bound assumes that it is possible to extract all blocks simultaneously; an option that is more valuable when S is large. Furthermore, we can show that for S sufficiently large the upper and lower bounds are linear functions of S. The dashed lines in Figure 1.3 (left panel) represent these linear asymptotes. Based on the results in Proposition 2 we have the following corollary whose proof follows directly from this proposition and it is omitted (see §1.2.5 for notation).

Corollary 2 *Consider a project with j blocks and let $\mathcal{F}_j(S)$ be the approximation in (1.15) for some pair (α_k, η_k), $k = 1, \ldots, j$. Let $(\widetilde{R}_k, \widetilde{C}_k, \widetilde{S}_k, \widetilde{M}_k, \widetilde{\alpha}_k, \widetilde{\eta}_k, \widetilde{\theta}_k)$, $k = 1, \ldots, \tilde{j}$ be the characteristics of the resulting mining project produced by the algorithm in Proposition 2. Then, for S sufficiently large the approximation $\mathcal{F}_j(S)$ is a linear function of S. In particular,*

$$\mathcal{F}_j(S) = (\theta^{\times} R)^{+}_{\tilde{j}}\, S - (\alpha^{\times} C)^{+}_{\tilde{j}}, \qquad \text{for all } S \geq \widetilde{S}_1 / \widetilde{\eta}^{\times}_{\tilde{j}}. \tag{1.20}$$

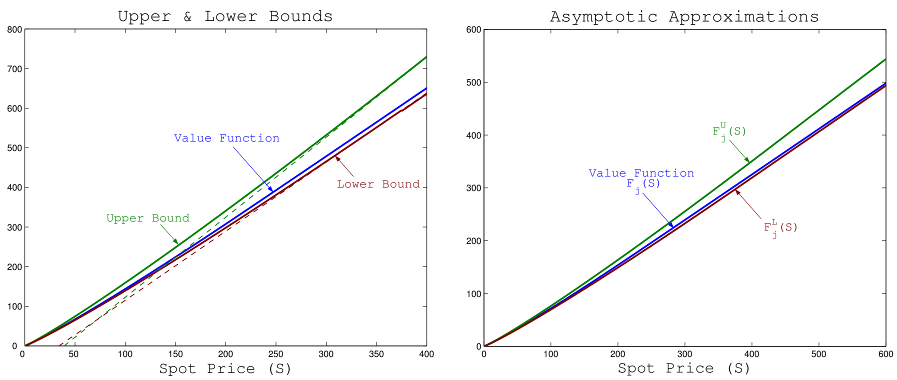

Fig. 1.3 LEFT PANEL: Value function $F_j(S)$, upper bound $F^{\mathrm{U}}_j(S)$ and lower bound $F^{\mathrm{L}}_j(S)$ as a function of the spot price S (in [cUS$/lb]) using the data in Example 1. The dashed lines correspond to linear asymptotes for the upper and lower bound approximations. RIGHT PANEL: Asymptotic approximations of the value function based on equations (1.21) and (1.22).

For the special case of the upper bound $F^{U}_j(S)$, $\alpha_k = \eta_k = 1$ and we get

$$F^{U}_j(S) = R^{+}_j S - C^{+}_j, \qquad \textit{for all } S \ge \widetilde{S}_1.$$

Similarly, if we set $\alpha_k = e^{-rT_k}$ and $\eta_k = e^{(r-\rho)T_k}$ we recover the lower bound $F^{L}_j(S)$ and equation (1.20) reduces to

$$F^{L}_j(S) = \mathcal{R}_j S - C_j, \qquad \textit{for all } S \ge \widetilde{S}_1\, e^{-(r-\rho)T^{+}_{1,j}}.$$

Note that $F^{\mathrm{L}}_j(S)$ has exactly the same linear asymptote than the one derived for $F_j(S)$ in Proposition 1. This explains the quality of the lower bound $F^{\mathrm{L}}_j(S)$ depicted in Figure 1.3 (left panel) as S grows large.

Corollary 2 also suggests a simple approximation for $F_j(S)$ based on these linear asymptotes. Recall from condition (1.9) that $F_j(S)$ satisfies

$$F_j(S) = \begin{cases} M_j S^{\beta_1}, & \text{if } S \le S^*_j, \\ W_j(S) + e^{-rT_j}\, \mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j}) \big| S_0 = S], & \text{if } S \ge S^*_j, \end{cases}$$

with $F_0(S) = 0$. As we mentioned before, the difficult part of solving this recursion is computing the expectation $\mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j}) \big| S_0 = S]$. Because this expectation is evaluated for values of S greater than the threshold S^*_j, we can get a simple (asymptotic) approximation if we replace $F_{j-1}(S)$ by one of the linear asymptotes derived in Corollary 2.

Using the upper bound asymptote $F^{\mathrm{U}}_{j-1}(S) = R^{+}_j S - C^{+}_j$ and the martingale property (1.2) we get that $\mathbb{E}^{\mathbb{Q}}[R^{+}_j S - C^{+}_j \big| S_0 = S] = R^{+}_j e^{(r-\rho)T_j} S - C^{+}_j$. Thus, we can approximate $F_j(S)$ by

$$\widehat{F}^{\mathrm{U}}_j(S) := \begin{cases} \widehat{M}^{\mathrm{U}}_j S^{\beta_1}, & \text{if } S \le \widehat{S}^{\mathrm{U}}_j, \\ \left(R_j + e^{-\rho T_j} R^+_{j-1}\right) S - \left(C_j + e^{-r T_j} C^+_{j-1}\right), & \text{if } S \ge \widehat{S}^{\mathrm{U}}_j. \end{cases} \tag{1.21}$$

(We will use a hat '∧' to denote quantities that are derived using the asymptotic approximation.) Using the value matching and smooth pasting conditions we obtain

$$\widehat{S}^{\mathrm{U}}_j = \frac{\beta\left(C_j + e^{-rT_j} C^+_{j-1}\right)}{(\beta-1)\left(R_j + e^{-\rho T_j} R^+_{j-1}\right)} \quad \text{and} \quad \widehat{M}^{\mathrm{U}}_j = \left(\frac{C_j + e^{-rT_j} C^+_{j-1}}{\beta - 1}\right)\left(\widehat{S}^{\mathrm{U}}_j\right)^{-\beta}.$$

Using exactly the same steps, we can get an alternative approximation for $F_j(S)$ based on the lower bound asymptote in Corollary 2.

$$\widehat{F}^{\mathrm{L}}_j(S) := \begin{cases} \widehat{M}^{\mathrm{L}}_j S^{\beta}, & \text{if } S \le \widehat{S}^{\mathrm{L}}_j, \\ \mathcal{R}_j S - C_j, & \text{if } S \ge \widehat{S}^{\mathrm{L}}_j. \end{cases} \tag{1.22}$$

with

$$\widehat{S}^{\mathrm{L}}_j = \frac{\beta C_j}{(\beta-1)\mathcal{R}_j} \quad \text{and} \quad \widehat{M}^{\mathrm{L}}_j = \left(\frac{C_j}{\beta-1}\right)\left(\widehat{S}^{\mathrm{L}}_j\right)^{-\beta}.$$

Figure 1.3 (right panel) plots the values of $\widehat{F}^{\mathrm{U}}_j(S)$ and $\widehat{F}^{\mathrm{L}}_j(S)$ as well as the value function $F_j(S)$ (numerically computed). As we can see, $\widehat{F}^{\mathrm{L}}_j(S)$ performs quite well over the entire range of prices. This in part due to the fact that by construction $\widehat{F}^{\mathrm{L}}_j(S)$ has exactly the same linear behavior than $F_j(S)$ and $F^{\mathrm{L}}_j(S)$ as S goes to infinity.

In an effort to support the conclusions that we have drawn from Figure 1.3, we conclude this section comparing numerically the performance of the upper bound $F^{\mathrm{U}}_j(S)$, the lower bound $F^{\mathrm{L}}_j(S)$ and the asymptotic approximations $\widehat{F}^{\mathrm{U}}_j(S)$ and $\widehat{F}^{\mathrm{L}}_j(S)$. We measure this performance as the average relative error of these approximations across a large range of initial spot prices for different values of the model parameters. More specifically, if $\mathcal{F}_j(S)$ is an arbitrary approximation for the value function $F_j(S)$ then we measure the performance of this approximation by

$$\mathcal{P}(\mathcal{F}_j) := \frac{1}{S_{\max} - S_{\min}} \int_{S_{\min}}^{S_{\max}} \frac{|\mathcal{F}_j(S) - F_j(S)|}{F_j(S)} \, \mathrm{d}S.$$

We choose the interval of spot prices $[S_{\min}, S_{\max}]$ large enough so that it includes almost the entire range of historical spot prices of copper. In particular, we chose $S_{\min} = 1.3$K [US$/Ton] and $S_{\max} = 13$K [US$/Ton].

Table 1.1 presents the value of $\mathcal{P}(\mathcal{F}_j)$ for the four approximations using the data of Example 1. In the table on the left we vary the volatility of the spot price σ^2. The middle table compares the performance of these approximations for different values of the discount factor r. Finally, in the right table we vary the extraction capacity K. In all three cases, we can see that F^{L}_j and $\widehat{F}^{\mathrm{L}}_j$ have a significantly better performance than F^{U}_j and $\widehat{F}^{\mathrm{U}}_j$. This is consistent with our previous discussion based on Figure 1.3.

In addition, we note that the asymptotic approximation $\widehat{F}^{\mathrm{L}}$ has the best performance across all instances with an average error between 1% and 3%.

σ^2	F^{U}	F^{L}	$\widehat{F}^{\mathrm{U}}$	$\widehat{F}^{\mathrm{L}}$
0.5	0.20	0.08	0.13	0.03
1.0	0.20	0.08	0.13	0.04
1.5	0.21	0.07	0.13	0.04
2.0	0.21	0.06	0.14	0.03
2.5	0.22	0.06	0.14	0.03
3.0	0.22	0.05	0.15	0.03
3.5	0.23	0.04	0.15	0.02
4,0	0.23	0.03	0.16	0.02
4.5	0.24	0.02	0.16	0.01
5.0	0.24	0.02	0.16	0.01
Av.	22.0%	5.1%	14.4%	2.6%

r	F^{U}	F^{L}	$\widehat{F}^{\mathrm{U}}$	$\widehat{F}^{\mathrm{L}}$
0.1	0.20	0.09	0.13	0.04
0.2	0.19	0.05	0.13	0.02
0.3	0.19	0.02	0.13	0.01
0.4	0.19	0.01	0.14	0.01
0.5	0.19	0.00	0.14	0.00
0.6	0.19	0.01	0.15	0.00
0.7	0.19	0.01	0.15	0.00
0.8	0.19	0.01	0.15	0.00
0.9	0.19	0.01	0.16	0.01
1.0	0.20	0.01	0.16	0.01
Av.	19.2%	2.1%	14.2%	1.0%

K	F^{U}	F^{L}	$\widehat{F}^{\mathrm{U}}$	$\widehat{F}^{\mathrm{L}}$
1.0	0.20	0.08	0.13	0.04
2.0	0.09	0.06	0.06	0.03
3.0	0.06	0.04	0.03	0.02
4.0	0.04	0.04	0.02	0.02
5.0	0.03	0.03	0.02	0.02
6.0	0.03	0.03	0.01	0.01
7.0	0.02	0.02	0.01	0.01
8.0	0.02	0.02	0.01	0.01
9.0	0.02	0.02	0.01	0.01
10.0	0.02	0.02	0.01	0.01
Av.	5.2%	3.5%	3.0%	1.7%

Table 1.1 Performance measure ($\mathcal{P}$) for the approximations F^{U}, F^{L}, $\widehat{F}^{\mathrm{U}}$ and $\widehat{F}^{\mathrm{L}}$ as a function of the spot price volatility σ^2 (left panel), discount factor r (center panel) and extraction capacity K (right panel). The data used in these computations is described in Example 1.

In terms of the sensitivity of these results, we can see that the volatility of the spot price σ^2 has a different impact on these approximations. Both $\mathcal{P}(F^{\mathrm{U}})$ and $\mathcal{P}(\widehat{F}^{\mathrm{U}})$ increase with σ^2 while the opposite is true for $\mathcal{P}(F^{\mathrm{L}})$ and $\mathcal{P}(\widehat{F}^{\mathrm{L}})$. The results in the middle panel in Table 1.1 suggest that the discount factor r does not have a significant effect on the approximations. Finally, the extraction capacity K affects these four approximations in a similar way, they are all monotonically decreasing with K. This behavior is a consequence of the following result.

Proposition 3 *Let $F_j(S,K)$, $F^{\mathrm{U}}_j(S,K)$ and $F^{\mathrm{L}}_j(S,K)$ be the value function and upper and lower bounds, respectively, for block j when the spot price is S and the extraction capacity is K. Then, in the limit as K goes to infinity the upper and lower bound approximations converge to the true value function. That is,*

$$\lim_{K\to\infty} F^{\mathrm{U}}_j(S,K) = \lim_{K\to\infty} F^{\mathrm{L}}_j(S,K) = \lim_{K\to\infty} F_j(S,K), \qquad \textit{for all } S \geq 0.$$

Hence, in the limit as K goes to infinity, $\mathcal{P}(\widetilde{F}) = \mathcal{P}(\widehat{F}) = 0$.

Proof: We only provide a sketch of the proof. As $K \to \infty$, one can show the algorithm in Proposition 2 produces exactly the same sequence of aggregated blocks for the upper and lower bound approximations. From this observation, it follows that the upper and lower bounds have the same limit: $\lim_{K\to\infty} F^{\mathrm{U}}_j(S,K) = \lim_{K\to\infty} F^{\mathrm{L}}_j(S,K)$. Finally, since $F_j(S,K)$ is bounded above and below by $F^{\mathrm{U}}_j(S,K)$ and $F^{\mathrm{L}}_j(S,K)$, respectively, the result follows. □

We conclude this section with a simple observation that is particularly useful when selecting an optimal sequence of extraction. Suppose we have a mining sector with j blocks and we want to compare two possible sequences of extraction π^1 and π^2.

Based on Proposition 1, the expected discounted value of the project under sequence π^i is asymptotically equal to $\mathcal{R}_j^{\pi^i} S - C_j^{\pi^i}$, $i = 1, 2$. Hence, for S sufficiently large the best sequence is the one that maximizes the value of $\mathcal{R}_j^{\pi^i}$. For moderate value of S, on the other hand, the comparison is not straightforward. However, we can try to extend this condition if we use the asymptotic approximation $\widehat{F}^{\mathrm{L}}(S)$ instead of the real value function $F(S)$ to perform the comparison between π^1 and π^2.

Proposition 4 *Consider two possible sequences of extraction π^1 and π^2 for a mining project with j blocks. Let $\widehat{F}_i^{\mathrm{L}}(S) = \mathcal{R}_j^{\pi^i} S - C_j^{\pi^i}$ be the (lower bound) asymptotic approximation for the value function if sequence π^i is used, $i = 1, 2$. Then $\widehat{F}_1^{\mathrm{L}}(S) \geq \widehat{F}_2^{\mathrm{L}}(S)$ for all $S \geq 0$ if and only if the following two conditions are satisfied:*

$$\mathcal{R}_j^{\pi^1} \geq \mathcal{R}_j^{\pi^2} \qquad \text{and} \qquad \left(\frac{\mathcal{R}_j^{\pi^1}}{\mathcal{R}_j^{\pi^2}} \right)^{\beta} \geq \left(\frac{C_j^{\pi^1}}{C_j^{\pi^2}} \right)^{\beta - 1}.$$

1.4 Capacity Expansions

In the previous section we derived a set of approximations for the value function assuming a fixed processing capacity K. In this section, we relax this assumption and show how to extend some of these approximations to include capacity expansion decisions. In particular, we will only discuss how to extend the lower bound asymptote $\widehat{F}_j^{\mathrm{L}}(S)$ that has shown the best numerical performance.

Using the notation in Section 1.2, we let quantities depend on K, when appropriate. For example, recall that $F_j(S,K)$ denotes the value function for a single-sector project when there are j blocks left, the spot price is S and the processing capacity is K. We find convenient to define $\bar{F}_j(S) := F_j(S, \bar{K})$, $\bar{W}_j(S) := W_j(S, \bar{K})$, $\bar{R}_j := R_j(\bar{K})$, $\bar{C}_j := C_j(\bar{K})$ and $\bar{T}_j := T_j(\bar{K})$, where $\bar{K}$ is the upper bound on the maximum level of capacity.

For a given S and K let us define the auxiliary function

$$G_j(S,K) := \sup_{K \leq \tilde{K} \leq \bar{K}} W_j(S, \tilde{K}) + e^{-r T_j(\tilde{K})} \mathbb{E}^{\mathbb{Q}} \left[F_{j-1}(S_{T_j(\tilde{K})}, \tilde{K}) \middle| S_0 = S \right] - \gamma(\tilde{K} - K), \tag{1.23}$$

and let $K_j^*(S,K)$ be the value of $\tilde{K}$ at which the maximum is attained. This function computes the optimal expected payoff if the state of the system (S,K) and the decision maker is forced to start production immediately. In this case, and similar to equation (1.6), the dynamic programming recursion takes the form

$$F_j(S,K) = \sup_{\tau \geq 0} \mathbb{E}^{\mathbb{Q}} \left[e^{-r\tau} G_j(S_\tau, K) \middle| S_0 = S \right],$$

where $F_0(S,K) = 0$. Using a similar line of arguments to the one used to derive equation (1.9) from equation (1.6), we can show that there exist two functions $M_j(K)$ and $S^*_j(K)$ such that

$$F_j(S,K) = \begin{cases} M_j(K)S^{\beta}, \text{ if } S \le S^*_j(K), \\ G_j(S,K), \text{ if } S \ge S^*_j(K). \end{cases} \tag{1.24}$$

The functions $M_j(K)$ and $S^*_j(K)$ are determined using the smooth-pasting and value-matching conditions

$$\beta G_j(S^*_j(K),K) = G'_j(S^*_j(K),K)\,S^*_j(K) \qquad \text{and} \qquad M^j(K) = G_j(S^*_j(K),K)\,(S^*_j(K))^{-\beta},$$

where G'_j denotes the derivative of G_j with respect to its first parameter. Most of the difficulty of computing the value function $F_j(S,K)$ in (1.24) boils down to determining the auxiliary function $G_j(S,K)$. Equation (1.24) suggests that we only need to compute the value of $G_j(S,K)$ for S greater than the threshold $S^*_j(K)$. As in the previous section, we will use an asymptotic approximation as a proxy for $G_j(S,K)$ in this range.

Proposition 5 *In the limit as S goes to infinity, the optimal capacity $K^*_j(S,K)$ converges to the upper bound $\bar{K}$ and the function $G_j(S,K)$ converges to a linear function of the price. In particular,*

$$G_j(S,K) \xrightarrow{S\to\infty} \bar{\mathcal{R}}_j S - \bar{\mathcal{C}}_j - \gamma(\bar{K} - K),$$

where

$$\bar{\mathcal{R}}_j := \sum_{k=1}^{j} \bar{R}_k e^{-\rho \bar{T}^+_{k,j}} \qquad \textit{and} \qquad \bar{\mathcal{C}}_j := \sum_{k=1}^{j} \bar{C}_k e^{-r \bar{T}^+_{k,j}}.$$

Proof: It follows directly from Proposition 1 and it is left to the reader. □

If we use this linear asymptotic behavior of $G_j(S,K)$ in equation (1.24), we get the following approximations of $S^*_j(K)$ and $M_j(K)$.

$$S^*_j(K) \approx \left(\frac{\beta}{\beta-1}\right)\left(\frac{\bar{\mathcal{C}}_j + \gamma(\bar{K}-K)}{\bar{\mathcal{R}}_j}\right) \qquad \text{and,}$$
$$M_j(K) \approx \left(\frac{\bar{\mathcal{R}}_j}{\beta}\right)^{\beta}\left(\frac{\beta-1}{\bar{\mathcal{C}}_j + \gamma(\bar{K}-K)}\right)^{\beta-1}. \tag{1.25}$$

According to these solutions, the threshold price $S^*_j(K)$ is a linear and decreasing function of K. That is, when capacity is large a small spot price is enough to induce the decision maker to start production.

In order to complete our approximation when capacity expansion is possible, we also need to approximate $K^*_j(S,K)$, that is, how the decision maker should expand capacity over time. The optimization in equation (1.23) is in general difficult to solve. However, since $F_j(S,K) = G_j(S,K)$ for S sufficiently large, we can exploit one more time the asymptotic approximation in Proposition 5 to get $K_j^*(S,K) \approx \max\left\{K \,;\, \mathcal{K}_j(S)\right\}$, where

$$\mathcal{K}_j(S) := \underset{0\le\tilde{K}\le\bar{K}}{\operatorname{argmax}} \left\{\mathcal{R}(\tilde{K})S - C(\tilde{K}) - \gamma\tilde{K}\right\}. \tag{1.26}$$

The function $\mathcal{K}_j(S)$ represents the optimal capacity expansion if there is no installed capacity and the spot price is S. (Note that this is an increasing function of S.) Define $\bar{S}_j := \inf\{S \ge 0 : \mathcal{K}_j(S) = \bar{K}\}$. For the decision maker's, $\bar{S}_j$ is the threshold price above which it is always optimal to expand capacity to its maximum possible level $\bar{K}$ independently of current capacity.

The left panel in Figure 1.4 plots the values of $\mathcal{K}_j(S)$ for $j = 1,\dots,6$ using the data in Example 1. As expected, we can see that $\mathcal{K}_1(S) \le \mathcal{K}_2(S) \le \cdots \le \mathcal{K}_6(S)$ for all S. This ordering reflects the fact that additional capacity is more valuable when the mining project has more blocks. This monotonicity also implies that $\bar{S}_6 \le \bar{S}_5 \le \cdots \le \bar{S}_1$.

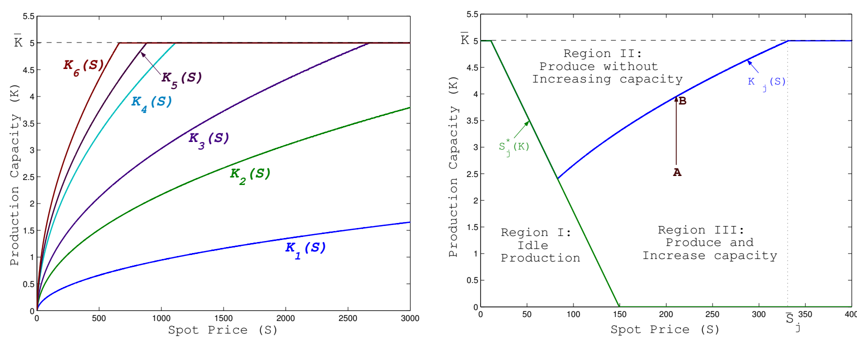

Fig. 1.4 Left Panel: Capacity expansion function $\mathcal{K}_j(S)$ for $j = 1,\dots 6$ using the data in Example 1, $\gamma = 10$ and a maximum capacity $\bar{K} = 5$. Right Panel: Optimal production and capacity expansion decisions based on the switching curves $S_j^*(K)$ and $\mathcal{K}_j(S)$ using the data of Example 1, $\gamma = 5$ and $\bar{K} = 5$. Both panels display spot prices in [cUS$/lb].

Based on the thresholds $S_j^*(K)$ and $\mathcal{K}_j(S)$ we can divide the state space $\{(S,K) : 0 \le K \le \bar{K}$ and $S \ge 0\}$ in three subregions depicted in the right panel in Figure 1.4. In Region I $:= \{(S,K) : 0 \le S \le S_j^*(K), K \le \bar{K}\}$ the spot price is very low and the decision maker is better off idling production until the price reaches the threshold $S_j^*(K)$. On the other hand, in Region II $:= \{(S,K) : S_j^*(K) \le S, \mathcal{K}_j(S) \le K \le \bar{K}\}$ the spot price and capacity are both high and production should start but no capacity expansion is required. Finally, in Region III $:= \{(S,K) \in \mathcal{S} \,:\, S_j^*(K) \le S$ and $0 \le K \le \mathcal{K}_j(S)\}$ the spot price is high but production capacity is low. In this region, the decision maker should expand capacity from K to $\mathcal{K}_j(S)$ and produce. For example, if the system is in point A in Figure 1.4 then capacity should be expanded to point B and then production should start.

We note that the negative slope of $S^*_j(K)$ implies that the opportunity cost of idling production increases as capacity increases. In other words, a large mining project will tend to operate almost independently of the price while a small project will turn production on and off as the spot price oscillates. On the other hand, the positive slope of $\mathcal{K}_j(S)$ reflects the intuitive fact that capacity becomes more valuable as the spot price increases.

1.5 Case Study

In this section, we use the methodology proposed in the previous sections to estimate the economic value of a mining project at *El Teniente*. This mine is located in the central region of Chile, 3,000 meters above the sea level, and is reported as the largest underground copper mine in the world. With a processing capacity of almost 50 [million tons/year], it produced more than 450,000 metric tons of refined copper in 2013. This mine has multiple active sectors and is in continuous expansion. One of these sectors is *El Diablo Regimiento*, which started production in 2005 and is scheduled to finish its operations circa 2020. Because of its unusual spatial distribution, several extraction sequences were considered, each one requiring a different economic evaluation. The project is currently on its third phase (out of five), and mining operators continually are pressured to evaluate large changes to production plans in limited amount of time.

In what follows, we show how we can use the methodology proposed in this chapter to tackle the sequencing problem for *El Diablo Regimiento*. Based on the original extraction sequence, we divide the almost 230 million tons of material in this sector into ten blocks. Figure 1.5 shows schematically the spatial distribution of these blocks. Table 1.2 (left panel) summarizes mineral content, grade and extraction time for the ten blocks in *El Diablo Regimiento*. Based on the spatial distribution of the blocks, we evaluate six extraction sequences (shown in the right panel), where sequence $N1$ is the one considered in the original design at *El Teniente*. The first block on each of these sequences is the first one to be extracted. We note that this set of sequences represents only a small fraction of the total number of possible extraction sequences.

Because production costs depend on the actual sequence of extraction, we do not have a fixed extraction cost for each block. For the purpose of the computational experiments reported in this section, we use a simplified method to approximate these extraction costs. If we let $\pi = (\pi_1, \pi_2, \ldots, \pi_{10})$ be any of the six sequences that we consider, then the marginal extraction cost (in [cUS$/lb]) for the j^{th} block in this

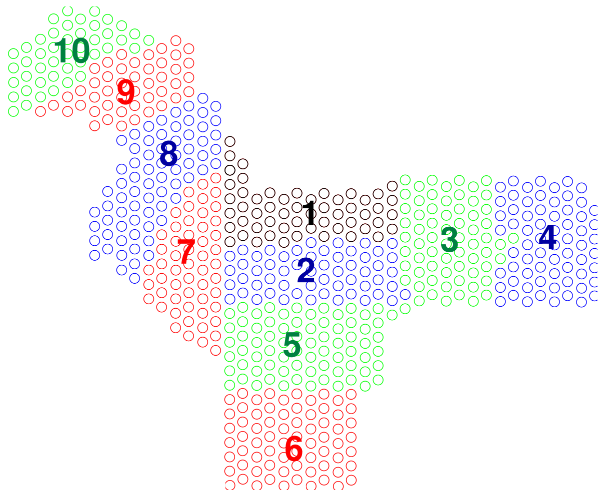

Fig. 1.5 Block spatial distribution at *El Diablo Regimiento*.

Block	Q_j [*million tons*]	L_j [%]	T_j [*years*]
1	21.415	0.827	2.93
2	21.268	0.915	2.91
3	29.526	0.823	4.04
4	28.351	0.881	3.88
5	24.854	0.845	3.40
6	23.931	0.848	3.28
7	21.476	0.768	2.94
8	26.110	0.727	3.58
9	14.913	0.694	2.04
10	13.126	0.776	1.80

Sequence	Order
N1	1-2-3-4-5-6-7-8-9-10
N2	10-9-8-7-1-2-5-3-4-6
N3	4-3-2-1-7-8-9-10-5-6
N4	6-5-2-1-3-4-7-8-9-10
N5	1-2-7-8-9-10-5-6-3-4
N6	1-2-5-3-7-8-6-4-9-10

Table 1.2 Left Panel: Mineral content Q_j, copper grade L_j and extraction time T_j for the ten blocks in *El Diablo Regimiento*. Right Panel: Six feasible extraction sequences.

sequence π is[4]

$$A_j^{\pi} = 0.4857 + 0.0162 \cdot d_{\pi_1 \pi_j}$$

where d_{ij} denotes the distance between blocks i and j (see Table 1.3). In other words, we are approximating the marginal extraction cost of a block as an affine function of the distance from the block to the initial extraction front, understanding that in practice, additional factors should be considered.

Finally, we considered a fixed production capacity of 7.3 [million tons/year] for this sector. For this exercise, we set additional parameters as follows: a recovery factor of 85% (this value further penalizes L_i and represents the loss in recovery of mineral

[4] The intercept and slope were estimated using production costs at *El Teniente* considering the extraction sequence $N1$ and using aggregate .

d_{ij}	1	2	3	4	5	6	7	8	9	10
1	0	90	190	321	191	314	159	189	275	367
2	90	0	102	234	174	295	157	240	350	438
3	190	102	0	133	233	336	182	299	424	507
4	321	234	133	0	343	425	272	399	532	607
5	191	174	233	343	0	124	327	380	454	547
6	314	295	336	425	124	0	450	503	569	661
7	159	157	182	272	327	450	0	128	263	335
8	189	240	299	399	380	503	128	0	136	209
9	275	350	424	532	454	569	263	136	0	93
10	367	438	507	607	547	661	335	209	93	0

Table 1.3 Distance (in meters) among the blocks at *El Diablo Regimiento*.

during the concentration and refinement processes), $r = 12\%$, $\sigma = 0.5$, and $\rho = 6\%$ (as a reference, [12] report an average value of the (instantaneous) convenience yield for copper of 6.3%).

For each of the six sequences we compute the value function (numerically) and the asymptotic approximation based on the lower bound in equation (1.22). Table 1.4 summarizes the results. We conclude that the best extraction sequence (as measured

Price	N1		N2		N3		N4		N5		N6		Relative
S	F	$\widehat{F}^{\mathrm{L}}$	F	$\widehat{F}^{\mathrm{L}}$	F	$\widehat{F}^{\mathrm{L}}$	F	$\widehat{F}^{\mathrm{L}}$	F	$\widehat{F}^{\mathrm{L}}$	F	$\widehat{F}^{\mathrm{L}}$	Error
50	607	552	556	499	598	543	599	541	587	535	602	549	9.1%
100	1367	1251	1255	1130	1350	1232	1352	1227	1325	1213	1356	1245	8.5%
150	2143	2007	1975	1812	2120	1975	2123	1967	2078	1945	2125	1996	6.3%
200	2926	2803	2704	2534	2896	2761	2901	2751	2837	2717	2901	2786	4.2%
250	3721	3614	3446	3295	3684	3567	3691	3562	3608	3505	3689	3590	2.9%
300	4509	4416	4183	4053	4465	4363	4474	4363	4373	4283	4469	4384	2.1%
350	5298	5219	4921	4810	5247	5159	5259	5164	5138	5062	5251	5178	1.5%
400	6097	6030	5670	5577	6040	5965	6054	5975	5914	5849	6043	5982	1.1%
450	6888	6832	6410	6335	6823	6762	6840	6776	6681	6628	6827	6777	0.8%
500	7679	7635	7152	7092	7607	7558	7627	7577	7449	7406	7611	7571	0.6%
550	8480	8446	7902	7859	8401	8364	8423	8388	8226	8194	8404	8375	0.4%
600	9272	9248	8644	8617	9186	9160	9211	9189	8995	8972	9189	9169	0.3%

Table 1.4 Value function for the six extraction sequences in *El Diablo Regimiento*. Prices in column S are expressed in [cUS$/lb]. The F columns represent the numerically computed value function (in M US$) and the $\widehat{F}^{\mathrm{L}}$ columns represent the asymptotic approximation using the lower bound in equation (1.22) (in M US$).

by the value function F) is given by N1 (the original sequence). Similarly, if we use the asymptotic approximation $\widehat{F}^{\mathrm{L}}$ to decide we also conclude that N1 is the best extraction sequence. In terms of the value of this project, the relative error between F and $\widehat{F}^{\mathrm{L}}$ is reported in the far most right column in Table 1.4. Note that the values computed using the asymptotic approximations have a relative small error that ranges from 9% to 0.3%. This error is decreasing in the spot price S, a result consistent with Proposition 1 and Corollary 2. The fact that sequence N1 is the best alternative is also consistent with Proposition 4. Indeed, analyzing the data for this instance we notice that sequence N1 has the best asymptotic behavior with the highest slope $\mathcal{R}^{N1} = 24.78$ among all six sequences.

We conclude this section discussing how to use the results of section 4 to estimate an optimal capacity expansion policy for *El Diablo Regimiento*. Supposing that initially there is no installed capacity we can use equations (1.25) and (1.26) to obtain an approximation for the optimal initial capacity. In these computations we considered $\gamma = 7.5$ [US$/(Ton/year)] and $\overline{K} = 20$ [M Tons/year].

For each of the six sequences we compute the asymptotic approximation based on the lower bound in equation (1.24) and the expansion rule in equation (1.26). Table 1.5 summarizes the results.

Price	N1		N2		N3		N4		N5		N6	
S	K^*	F	K^*	F	K^*	F	K^*	F	K^*	F	K^*	F
30	0.01	66.78	0.01	58.23	0.01	61.82	0.01	58.75	0.01	68.09	0.01	66.69
40	5.1	89.04	0.01	77.64	0.9	82.43	0.01	78.33	4.7	90.79	4.4	88.91
50	7.75	111.29	0.01	97.05	4.65	103.04	0.01	97.91	7.2	113.49	7.4	111.14
60	10.6	133.55	5.2	116.47	7.25	123.65	6.7	117.5	9.95	136.19	10.45	133.37
70	13.9	155.81	7.3	135.88	10.05	144.25	9.4	137.08	13.05	158.89	13.85	155.6
80	17.5	178.07	9.55	155.29	13.25	164.86	12.45	156.66	16.5	181.59	17.6	177.83
90	20	200.33	12.1	174.7	16.85	185.47	15.85	176.25	20	204.28	20	200.06
100	20	222.59	14.95	194.11	20	206.08	19.55	195.83	20	226.98	20	222.29
110	20	244.85	18.1	213.52	20	226.68	20	215.41	20	249.68	20	244.51
120	20	267.11	20	232.93	20	247.29	20	234.99	20	272.38	20	266.74
130	20	289.36	20	252.34	20	267.9	20	254.58	20	295.08	20	288.97
140	20	311.62	20	271.75	20	288.51	20	274.16	20	317.78	20	311.2

Table 1.5 Optimal capacity K^* (in [M Tons/year]) and expected value F (numerically computed, in [M US$]) for the six sequences of extraction at *El Diablo Regimiento* as a function of the price S in [cUS$/lb]. The value of K^* is computed using the asymptotic approximation $\widehat{F}^L$ and equation (1.26).

Interestingly, the optimal sequence in this case is N5 as opposed to N1 that is optimal when capacity is fixed at 7.3 [M Tons/year]. Because sequence N1 is the one considered in the original design of *El Diablo Regimiento*, it seems that management at *El Teniente* has not fully valued the option of increasing capacity. Of course, there are other practical considerations that are not included in our model that might explain this discrepancy. Finally, we note that it is optimal to expand capacity to its maximum level $\bar{K} = 20$ [M Tons/year] if the spot price exceeds 100 [cUS$/lb]. This is a rather small value compared to the current spot price which is around 300 [cUS$/lb], but higher than the prevalent price at the early stages of the planning process, circa 2003.

1.6 Conclusions and Future Research

In this chapter, we develop a tractable continuous-time model of a mining operation and propose a methodology to compute near-optimal production and capacity expansion policies.

On the modeling side, we represent the mining project as a finite collection of mineral blocks. These blocks are defined in terms of ore content, mineralogical composition and extraction costs. In our model, an optimal production policy defines the sequence in which blocks should be extracted as well as the timing of extraction. Our discrete block representation of the mine is consistent with current practice and deviates from previous research that commonly models mine characteristics (such as ore content, grade and production costs) as continuous variables. In this respect, we believe our model contributes to bridge the gap between the academic research and current practice in the mining industry, and represents a first step towards a methodology for optimal sequence selection.

We use a two-step approach to approximate the optimal operating and capacity expansion policy. First, in Section 1.3, we fix the sequence and production capacity and solve for the optimal timing of extraction contingent upon the evolution of the spot price. In Proposition 1, we derive general properties of an optimal policy and show that the value of the project is asymptotically equal to an affine function of the price. Unfortunately, for moderate values of the price we do not obtain a simple characterization of the value function. For this reason, we derive in §1.3.1 and §1.3.2 upper and lower bounds on the value function, respectively, and use them to propose two simple extraction policies. In addition, we use these bounds in §1.3.4 to derive a pair of asymptotic approximations to the value function. Out of these approximations, the one derived using asymptotic analysis and based on the lower bound in equation (1.22), $\widehat{F^{\mathrm{L}}}(S)$, turns out to be asymptotically equal to the true value function. Moreover, the set of numerical computations in Table 1.1 shows that $\widehat{F^{\mathrm{L}}}(S)$ performs extremely well for a wide range of prices and other parameters with an average error of 2%. We conclude Section 1.3 with Proposition 4 that provides necessary and sufficient conditions to decide when a sequence of extraction dominates another one for all values of the spot price.

In Section 1.4, we undertake the second step of our solution approach. There, we show how to extend the models of the previous section to identify efficient capacity expansion policies. Our discussion is based on the asymptotic approximation $\widehat{F^{\mathrm{L}}}(S)$ but the same methodology can be extended to other approximations of the value function. The resulting production/capacity policy is of the threshold type. Specifically, the state space (S,K) (spot price, installed capacity) is partitioned into three regions (see the right panel in Figure 1.4). In region I, the spot price is relatively small and the optimal policy is to idle production. In region II both the price and the capacity are large and so production is performed with no increase in capacity. Finally, in region III the spot price is high but capacity is relatively small. The policy in this case is to increase capacity to a level that depends on the current spot price and then produce accordingly. Our analysis provides a set of simple equations that characterize the threshold functions that separate these three regions. As a general rule, we observe that as the capacity of a project increases the option to idle production becomes less valuable. In other words, large mining projects will tend to operate almost independently of the output price while small projects will switch production on and off as the spot price oscillates.

We conclude the chapter, in Section 1.5, with an application of our methodology to a real instance of the problem at *El Teniente*. The example is based on a real project, called *El Diablo Regimiento*. Our analysis shows that the original planned sequence is optimal if production capacity is fixed at its nominal value of 7.3 (M Tons/year). However, if we allow the production capacity to be optimally chosen then it turns out a different sequence maximizes the economic value of *El Diablo Regimiento*.

There are a number of possible extensions to our model. First of all, an important component of an optimal production policy is the sequence in which blocks are extracted. In this chapter, we do not handle explicitly the question of how to dynamically choose this sequence. Instead, we take a scenario-based approach and assume that the decision maker has identified a set of potential sequences that wants to evaluate. This *open-loop* approach is indeed consistent with how mining projects are evaluated at Codelco. However, it lacks the flexibility of adjusting the sequence of extraction based on the evolution of the spot price. On the other hand, as we note in Section 1.2.4, the problem of dynamically adjusting the sequence of extraction has a combinatorial structure which makes it extremely hard to solve. Extending our methodology to explicitly include dynamic sequencing is a challenging research problem that is important from both theoretical and practical standpoints.

Another interesting direction in which our model could be extended is by looking more carefully at the relationship between spot price and production levels. In our model the spot price in equation (1.1) is independent of the output of the mining project. This is a standard assumption in the literature, which is reasonable if the producer is a small player with limited market power. However, this is arguably not the case for a company like Codelco that produces 10% of the world's copper production. In this situation, we should expect some correlation between output and spot (and futures) price trajectories (see [8]). This type of *large investor* effect has received some attention in the mathematical finance literature (e.g. [17], [25]) but it seems to have been overlooked in the operational context of real options (see [24] for an analysis of this issue for the German natural gas market).

Acknowledgements The authors would like to thank Nicola Secomandi and an anonymous referee for their helpful and constructive comments on the initial version of this chapter.

Appendix

Proof of Theorem 1

Let $F(S) \in \mathcal{C}_+^2$ be a solution to QVI conditions in equation (1.7). Given the assumptions on $F(\cdot)$, we can apply integration by parts followed by Itô's lemma (see [40]) to get

$$e^{-r\tau} F(S_\tau) = F(S) + \int_0^\tau e^{-rt} \mathcal{A}F(S_t)\,\mathrm{d}t + \int_0^\tau e^{-rt} F'(S_t)\sigma S_t\,\mathrm{d}\tilde{B}_t.$$

Using the fact that $\mathcal{A}F(S) \leq 0$ (first QVI condition) and the non negativity of F yields

$$0 \leq e^{-r\tau} F(S_\tau) \leq F(S) + \int_0^\tau e^{-rt} F'(S_t)\sigma S_t\,\mathrm{d}\tilde{B}_t.$$

Hence, the process

$$Y_t := F(S) + \int_0^t e^{-ru} F'(S_u)\sigma S_u\,\mathrm{d}\tilde{B}_u$$

is a nonnegative local $\mathbb{Q}$-martingale, and, hence, a $\mathbb{Q}$-supermatingale. Taking expectation with respect to $\mathbb{E}^{\mathbb{Q}}[\cdot]$ in the inequality above we obtain

$$\mathbb{E}^{\mathbb{Q}}[e^{-r\tau} F(S_\tau)] \leq F(S).$$

Furthermore, using the second QVI condition gives

$$\mathbb{E}^{\mathbb{Q}}[e^{-r\tau} G_j(S_\tau)] \leq F(S).$$

Because this inequality holds for any stopping time τ, we conclude that $F(S) \geq F_j(S)$. Finally, we note that all the inequalities above become equalities for the QVI-control associated to $F(\cdot)$. This follows from Dynkin's formula and the fact that the QVI-control is the first exit time from a bounded set (continuation region $\mathcal{D}$). □

Proof of Proposition 1

We will derive an upper an lower bound approximation for $F_j(S)$ from which the result will follow. First, we can get a lower bound on $F_j(S)$ if we assume that the decision maker is unable to idle production when prices are below the production thresholds $\{S_k^*\}$. Under this non-idling restriction it follows that

$$F_j(S) \geq \mathbb{E}^{\mathbb{Q}}\left[\sum_{k=1}^{j} e^{-rT_{k,j}^+} W_j(S_{\mathcal{T}_{k,j}})\right] = \sum_{k=1}^{j} e^{-rT_{k,j}^+}\left(R_k\,\mathbb{E}^{\mathbb{Q}}[S_{T_{k,j}^+} | S_0 = S] - C_k\right) = \mathcal{R}_j S - C_j. \tag{A1}$$

To get an upper bound, let us introduce a modified price process $\mathcal{S}_t$ given by

$$\mathcal{S}_t = S_t + \sum_{k:\mathcal{T}_{k,j}\le t} (S_k^* - \mathcal{S}_{T^+_{k,j}-})^+, \qquad \mathcal{S}_{0-} = S_0.$$

Recall that S_k^* is the switching price for block k, that is, the extraction of block k should start as soon as the spot price exceeds this threshold. Let us define the auxiliary value function $\mathcal{F}_j(S)$ which is the expected payoff for a project with j blocks under the modified price process $\mathcal{S}_t$ and using the switching prices $\{S_k^* : 1 \le k \le j\}$ to control production.

It is not hard to see that $\mathcal{S}_t \ge S_t$ pathwise, hence it follows that $F_j(S) \le \mathcal{F}_j(S)$ for all S. In addition, because of the specific construction of $\mathcal{S}_t$ it follows that under $\mathcal{S}_t$ the decision maker will never idle production. That is, $\mathcal{S}_{T^+_{k,j}} \ge S_k^*$ (a.s.) for all $1 \le k \le j$. Therefore, we have

$$\begin{aligned} F_j(S) \le \mathcal{F}_j(S) &= \sum_{k=1}^{j} e^{-rT^+_{k,j}} \left(R_k \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{\mathcal{T}_{k,j}} | S_0 = S] - C_k\right) \\ &= \mathcal{R}_j S - C_j + \sum_{k=1}^{j} e^{-rT^+_{k,j}} R_k \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T^+_{k,j}} - S_{T^+_{k,j}} | S_0 = S]. \qquad \text{(A2)} \end{aligned}$$

Combining (A1) and (A2), we get

$$0 \le F_j(S) - (\mathcal{R}_j S - C_j) \le \sum_{k=1}^{j} e^{-rT^+_{k,j}} R_k \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T^+_{k,j}} - S_{T^+_{k,j}} | S_0 = S].$$

To complete the proof, we need to show that the term on the right goes to 0 as S goes to infinity. In order to see this, we first note that by the definition of $\mathcal{S}_t$ we have

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T^+_{k,j}} - S_{T^+_{k,j}} | S_0 = S] &= \sum_{n=k}^{j} \mathbb{E}^{\mathbb{Q}}\left[(S_k^* - \mathcal{S}_{T^+_{k,j}-})^+ | S_0 = S\right] \\ &\le \sum_{n=k}^{j} S_k^* \mathbb{Q}(\mathcal{S}_{T^+_{k,j}-} \le S_k^* | S_0 = S) \\ &\le \sum_{n=k}^{j} S_k^* \mathbb{Q}(S_{T^+_{k,j}} \le S_k^* | S_0 = S), \end{aligned}$$

where the last inequality uses the (a.s.) facts that $\mathcal{S}_t \ge S_t$ and S_t is continuous. Under measure $\mathbb{Q}$, S_t is log-normal with drift $r - \rho - \sigma^2/2$ and diffusion coefficient σ (see equation (1.3)). So, it is not hard to show that

$$\mathbb{Q}(S_{T^+_{k,j}} \le S_k^* | S_0 = S) = \mathbb{Q}\left(\tilde{B}_{T^+_{k,j}} \ge \frac{1}{\sigma}\left[\ln\left(\frac{S}{S_k^*}\right) + (r - \rho - \frac{\sigma^2}{2}) T^+_{k,j}\right]\right),$$

where $\tilde{B}_t$ is a standard $\mathbb{Q}$-Brownian motion.

Hence, for $S \ge S_j^m \exp\left(-\min_k\left\{(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right\}\right)$, we can bound the tail probability by (e.g., [3, Theorem XIII-2.1])

$$\mathbb{Q}\left(\tilde{B}_{T_{k,j}^+} \ge \frac{1}{\sigma}\left[\ln\left(\frac{S}{S_k^*}\right) + (r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right]\right) \le \exp\left(-\frac{1}{\sigma^2}\left[\ln\left(\frac{S}{S_k^*}\right)+(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right]^2\right)$$
$$\le \exp\left(-\frac{1}{\sigma^2}\min_{1\le k\le j}\left\{\left[\ln\left(\frac{S}{S_j^m}\right)+(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right]^2\right\}\right).$$

Based on this bound, it is not hard to show that for $S \ge S_j^m \exp\left(-\min_k\left\{(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right\}\right)$ we have

$$\sum_{k=1}^{j} e^{-rT_{k,j}^+} R_k\,\mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T_{k,j}^+} - S_{T_{k,j}^+}|S_0 = S] \le S_j^m R_j^m \frac{j(j+1)}{2}\exp\left(-\frac{1}{\sigma^2}\min_{1\le k\le j}\left\{\left[\ln\left(\frac{S}{S_j^m}\right)+(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right]^2\right\}\right),$$

which goes to 0 as S goes to infinity. This completes the proof. □

Proof of Proposition 2

We use the following intermediate result.

Lemma 1 *Consider a j-block project with characteristics* $(R_k, C_k, \alpha_k, \eta_k)$, $k = 1, \ldots, j$. *Suppose that*

$$\frac{C_{k-1}}{R_{k-1}} \ge \eta_k \frac{C_k}{R_k}, \qquad k = 2, 3, \ldots, j.$$

Then, for any $k \le j-1$

$$\eta_{k,j}^{\times}\,\frac{(\alpha^{\times} C)_{k-1,j}^{+}}{(\theta^{\times} R)_{k-1,j}^{+}} \le \frac{C_k}{R_k}, \qquad \textit{where } \theta = (\theta_k) = (\alpha_k\,\eta_k).$$

Proof Lemma 1: The proof follows using backward induction on $k = j-1, j-2, \ldots$ and it is left to the reader. □

We divide the proof of Proposition 2 in two parts.

Part I: Let us first prove the correctness of equations (1.18) and (1.19). For this, we will consider the "modified" sequence of blocks produced by the algorithm and we will show that these equations do characterize $\mathcal{S}_j$, $\mathcal{M}_j$, and $\mathcal{F}_j$ for this modified sequence.

We find convenient to drop the tildes '∼' in the notation. Consider an arbitrary j-block project with characteristics $(R_k, C_k, \alpha_k, \eta_k)$, $k = 1, \ldots, j$ such that

$$\mathbb{C}_{k-1} \ge \eta_k\,\mathbb{C}_k, \qquad k = 2, 3, \ldots, j. \tag{A3}$$

Note that according to step 2 in the algorithm, condition (A3) ensures that there is no block aggregation, as required.

Let us now use induction on j to prove that for a j-block project satisfying condition (A3) the sets of threshold prices $\{\mathcal{S}_k\}_{k=1}^{j}$ and constants $\{\mathcal{M}_k\}_{k=1}^{j}$ are given by equation (1.18) and the approximation $\mathcal{F}_j(S)$ satisfies equation (1.19), that is,

$$\mathcal{F}_j(S) = (\theta^{\times} R)^{+}_{h,j} S - (\alpha^{\times} C)^{+}_{h,j} + \mathcal{M}_h \alpha^{\times}_{h,j} (\eta^{\times}_{h,j})^{\beta} S^{\beta}, \tag{A4}$$

where $h = \max\{0 \leq k \leq j \,|\, \mathcal{S}_k \geq \eta^{\times}_{k,j} S\}$ and $\mathcal{S}_0 = \infty$.

- For $j = 1$ the result follows directly from equation (1.16).
- Let us assume that the result is true for some $j-1$. That is, the values of $\{\mathcal{S}_k\}_{k=1}^{j-1}$ and $\{\mathcal{M}_k\}_{k=1}^{j-1}$ are given by equation (1.18) and $\mathcal{F}_{j-1}(S)$ is given by equation (A4). Combining condition (A3) and the value of $\mathcal{S}_k$ ($k = 1, \ldots, j-1$) in equation (1.18) we conclude that

$$\mathcal{S}_{k-1} \geq \eta_k \mathcal{S}_k, \qquad k = 2, 3, \ldots, j-1. \tag{A5}$$

- We now prove the result for j. First of all, let us show that $\mathcal{S}_{j-1} \geq \eta_j \mathcal{S}_j$. Suppose, by contradiction, that this is not the case, *i.e.*, $\mathcal{S}_{j-1} < \eta_j \mathcal{S}_j$. Then, condition (A5) and the fact that $\mathcal{S}_0 = \infty$ imply that there exists a $\hat{k} \leq j-2$ such that

$$\eta_{\hat{k}+1} \mathcal{S}_{\hat{k}+1} < \eta_{\hat{k}+1} \eta_{\hat{k}+2} \cdots \eta_j \mathcal{S}_j \leq \mathcal{S}_{\hat{k}} \quad \text{or equivalently} \quad \eta_{\hat{k}+1} \mathcal{S}_{\hat{k}+1} < \eta^{\times}_{\hat{k},j} \mathcal{S}_j \leq \mathcal{S}_{\hat{k}}. \tag{A6}$$

Now, by the definition of $\mathcal{S}_j$ and $\mathcal{M}_j$ and the value matching and smooth pasting conditions we get

$$\mathcal{M}_j \mathcal{S}_j^{\beta} = R_j \mathcal{S}_j - C_j + \alpha_j \mathcal{F}_{j-1}(\eta_j \mathcal{S}_j) \quad \text{and} \quad \beta \mathcal{M}_j \mathcal{S}_j^{\beta-1} = \mathcal{R}_j + \alpha_j \eta_j \mathcal{F}'_{j-1}(\eta_j \mathcal{S}_j).$$

Using the induction hypothesis we can replace $\mathcal{F}_{j-1}(\eta_j \mathcal{S}_j)$ using equation (A4). For this, note that the value of the index h used in (A4) to evaluate $\mathcal{F}_{j-1}(\eta_j \mathcal{S}_j)$ is exactly equal to $\hat{k}$ in (A6). In fact, at $S = \eta_j \mathcal{S}_j$, h is equal to $\max\{0 \leq k \leq j-1 \,|\, \mathcal{S}_k \geq \eta^{\times}_{k,j-1} (\eta_j \mathcal{S}_j)\}$ or equivalently $\max\{0 \leq k \leq j-1 \,|\, \mathcal{S}_k \geq \eta^{\times}_{k,j} \mathcal{S}_j\}$. This is $\hat{k}$ by definition and we obtain

$$\mathcal{F}_{j-1}(\eta_j \mathcal{S}_j) = (\theta^{\times} R)^{+}_{\hat{k},j-1} \eta_j \mathcal{S}_j - (\alpha^{\times} C)^{+}_{\hat{k},j-1} + \mathcal{M}_{\hat{k}} \alpha^{\times}_{\hat{k},j-1} (\eta^{\times}_{\hat{k},j-1})^{\beta} (\eta_j \mathcal{S}_j)^{\beta}.$$

After some algebra, the value matching and smooth pasting conditions imply

$$\mathcal{S}_j = \left(\frac{\beta}{\beta-1}\right) \frac{(\alpha^{\times} C)^{+}_{\hat{k},j}}{(\theta^{\times} R)^{+}_{\hat{k},j}} \quad \text{and} \quad \mathcal{M}_j = \mathcal{M}_{\hat{k}} \alpha^{\times}_{\hat{k},j} (\eta^{\times}_{\hat{k},j})^{\beta} + \left(\frac{(\alpha^{\times} C)^{+}_{\hat{k},j}}{\beta-1}\right) \mathcal{S}_j^{\beta}. \tag{A7}$$

However, condition (A3), the induction hypothesis $\mathcal{S}_{\hat{k}+1} = \beta C_{\hat{k}+1} / ((\beta-1) R_{\hat{k}+1})$ and Lemma 1 imply

$$\eta^{\times}_{\hat{k},j}\mathcal{S}_j = \eta_{\hat{k}+1}\,\eta^{\times}_{\hat{k}+1,j}\mathcal{S}_j = \eta_{\hat{k}+1}\,\eta^{\times}_{\hat{k}+1,j}\left(\frac{\beta}{\beta-1}\right)\frac{(\alpha^{\times}C)^{+}_{\hat{k},j}}{(\theta^{\times}R)^{+}_{\hat{k},j}} \le \eta_{\hat{k}+1}\left(\frac{\beta}{\beta-1}\right)\frac{C_{\hat{k}+1}}{R_{\hat{k}+1}} = \eta_{\hat{k}+1}\,\mathcal{S}_{\hat{k}+1}.$$

This inequality contradicts (A6) and we conclude that $\mathcal{S}_{j-1} \ge \eta_j \mathcal{S}_j$ as claimed. This conclusion implies that $\hat{k} = j-1$ and we can compute the values of $\mathcal{S}_j$ and $\mathcal{M}_j$ replacing $\hat{k}$ by $j-1$ in equation (A7) which leads to

$$\mathcal{S}_j = \left(\frac{\beta}{\beta-1}\right)\frac{C_j}{R_j} \qquad \text{and} \qquad \mathcal{M}_j = \alpha_j \eta_j^{\beta}\mathcal{M}_{j-1} + \left(\frac{C_j}{\beta-1}\right)\left(\mathcal{S}_j\right)^{-\beta},$$

proving equation (1.18) as required. Finally, from the condition $\eta_j \mathcal{S}_j \le \mathcal{S}_{j-1}$ and the induction hypothesis it is straightforward to show that $\mathcal{F}_j(S)$ is given by equation (1.19), which completes the induction.

Part II: Let us now carry out the second part of the proof. In this part, we prove that for an arbitrary j-block project with characteristics $(R_k, C_k, \alpha_k, \eta_k)$, $k = 1, \ldots, j$, the values of $\mathcal{F}_j(S)$ and $\mathcal{S}_j$ are given by equations (1.18) and (1.19). The difference with respect to Part I is that we are not assuming that condition (A3) is satisfied.

We will proceed one more time using induction on the number of blocks, j.

- For $j = 1$ the result follows directly from equation (1.16).
- Let us suppose that the result is true for $j-1$.
- Let us prove the result for j. The induction hypothesis implies that the value of $\mathcal{F}_{j-1}(S)$ is derived using a modified sequence of blocks that satisfies condition (A3). Furthermore, all that we need to know about the characteristics of blocks $\{j-1, j_2, \ldots, 1\}$ to compute $\mathcal{F}_j(S)$ is contained in $\mathcal{F}_{j-1}(S)$. Hence, we can assume without loss of generality that the sequence of blocks $\{j-1, j_2, \ldots, 1\}$ does satisfy condition (A3), that is,

$$\frac{C_{k-1}}{R_{k-1}} \ge \eta_k \frac{C_k}{R_k}, \qquad k = 2, 3, \ldots, j-1. \tag{A8}$$

If this condition is also satisfied for block j then the entire sequence satisfies condition (1.16) and the result follows from Part I. Hence, we will assume that block j does not satisfy (A3), that is,

$$\frac{C_{j-1}}{R_{j-1}} < \eta_j \frac{C_j}{R_j}. \tag{A9}$$

In the remainder of this proof, we will apply the algorithm in Proposition 2 to a sequence of blocks satisfying conditions (A8) and (A9) and we will verify that the value of $\mathcal{F}_j(S)$ and $\mathcal{S}_j$ are given by equations (1.18) and (1.19).

First, note that the inequality in (A9) and condition (A8) imply that $\tilde{k} = j$ in Step 2 of the algorithm.

We now let $\tilde{h}$ be the solution to Step 3 in the algorithm, that is,

$$\tilde{h} = \max\{1 \le h \le j-1 \,:\, \eta^{\times}_{h,j}\,\mathbb{C}_{h+1,j} \le \mathbb{C}_h\}.$$

Using these values of $\tilde{k}$ and $\tilde{h}$, Step 4 of the algorithm will pool together blocks $\tilde{h}+1, \tilde{h}+2, \ldots, j$ into a single block. Hence, after this first iteration of the algorithm, the resulting sequence of blocks has $\tilde{\jmath} = \tilde{h}+1$ blocks with characteristics

$$\widetilde{R}_k = R_k, \quad \widetilde{C}_k = C_k, \quad \widetilde{\alpha}_k = \alpha_k, \quad \widetilde{\eta}_k = \eta_k, \qquad k = 1, \ldots, \tilde{\jmath}-1$$

and

$$\widetilde{R}_{\tilde{\jmath}} = (\theta^{\times} R)^{+}_{\tilde{h},j}, \quad \widetilde{C}_{\tilde{\jmath}} = (\alpha^{\times} C)^{+}_{\tilde{h},j}, \quad \widetilde{\alpha}_{\tilde{\jmath}} = \alpha^{\times}_{\tilde{h},j} \quad \widetilde{\eta}_{\tilde{\jmath}} = \eta^{\times}_{\tilde{h},j}.$$

Note that by (A8) and the definition of $\tilde{h}$, the resulting sequence satisfies

$$\frac{\widetilde{C}_{k-1}}{\widetilde{R}_{k-1}} \ge \widetilde{\eta}_k \frac{\widetilde{C}_k}{\widetilde{R}_k}, \qquad k = 2, 3, \ldots, \tilde{\jmath}.$$

Therefore, after the first iteration the algorithm will stop. Using this modified sequence, equation (1.18) leads to

$$\mathcal{S}_j = \left(\frac{\beta}{\beta-1}\right)\frac{\widetilde{C}_{\tilde{\jmath}}}{\widetilde{R}_{\tilde{\jmath}}} = \left(\frac{\beta}{\beta-1}\right)\frac{(\alpha^{\times} C)^{+}_{\tilde{h},j}}{(\theta^{\times} R)^{+}_{\tilde{h},j}}.$$

To verify the correctness of this solution, let us compute $\mathcal{S}_j$ using its definition in equation (1.15). We have

$$\mathcal{F}_j(S) = \begin{cases} \mathcal{M}_j S^{\beta}, & \text{if } S \le \mathcal{S}_j, \\ R_j S - C_j + \alpha_j \mathcal{F}_{j-1}(\eta_j S), & \text{otherwise.} \end{cases}$$

Recall that the value of $\mathcal{F}_{j-1}(S)$ is known by the induction hypothesis and it is given by equation (1.19). This induction hypothesis and, in particular condition (A8), imply

$$\frac{\mathcal{S}_{k-1}}{\eta^{\times}_{k-1,j-1}} \ge \frac{\mathcal{S}_k}{\eta^{\times}_{k,j-1}}, \qquad k = 2, \ldots, j-1.$$

Let us suppose that the value of $\mathcal{S}_j$ satisfies

$$\frac{\mathcal{S}_{\bar{h}+1}}{\eta^{\times}_{\bar{h}+1,j-1}} < \eta_j \mathcal{S}_j \le \frac{\mathcal{S}_{\bar{h}}}{\eta^{\times}_{\bar{h},j-1}}, \tag{A10}$$

for some $\bar{h} \le j-1$. These inequalities and equation (1.19) imply

$$\mathcal{F}_{j-1}(\eta_j \mathcal{S}_j) = (\theta^{\times} R)^{+}_{\bar{h},j-1}\, \eta_j \mathcal{S}_j - (\alpha^{\times} C)^{+}_{\bar{h},j-1} + \mathcal{M}_{\bar{h}}\, \alpha^{\times}_{\bar{h},j-1}\, (\eta^{\times}_{\bar{h},j-1})^{\beta} (\eta_j \mathcal{S}_j)^{\beta}.$$

Using this condition and the value matching and smooth pasting conditions we can show, after some algebra,

$$\mathcal{S}_j = \left(\frac{\beta}{\beta-1}\right)\frac{(\alpha^\times C)^+_{\tilde{h},j}}{(\theta^\times R)^+_{\tilde{h},j}}.$$

Hence, in order for this solution to be consistent with the inequalities in (A10) we need

$$\begin{aligned}\bar{h} &= \max\left\{1 \le h \le j-1 \,:\, \eta_j\left(\frac{\beta}{\beta-1}\right)\frac{(\alpha^\times C)^+_{h,j}}{(\theta^\times R)^+_{h,j}}\frac{\mathcal{S}_h}{\eta^\times_{h,j-1}}\right\}\\ &= \max\{1 \le h \le j-1 \,:\, \eta^\times_{h,j}\,\mathbb{C}_{h+1,j} \le \mathbb{C}_h\}.\end{aligned}$$

Hence, we have that $\bar{h} = \tilde{h}$, which proves that the value of $\mathcal{S}_j$ in (1.18) is indeed correct. Now that we know the value of $\mathcal{S}_j$, it is a matter of simple (but tedious) calculations to verify that the values of $\mathcal{M}_j$ and $\mathcal{F}_j(S)$ are exactly those reported in equations (1.18) and (1.19). □

Proof of Proposition 4

The proof follows directly from equation (1.22). Indeed, suppose first that $F^{\rm L}_1(S) \ge F^{\rm L}_2(S)$ for all S. Then, for S sufficiently small (that is, $S \le \min\{S^{*\pi_1}_j, S^{*\pi_2}_j\}$, where $S^{*\pi_i}_j$ is the threshold price under sequence π_i, i=1,2) equation (1.22) implies that $M^{\pi_1}_j \ge M^{\pi_2}_j$. This inequality is equivalent to

$$\left(\frac{\mathcal{R}^{\pi^1}_j}{\mathcal{R}^{\pi^2}_j}\right)^{\beta} \ge \left(\frac{C^{\pi^1}_j}{C^{\pi^2}_j}\right)^{\beta-1}.$$

Similarly, for S sufficiently large $F^{\rm L}_1(S) \ge F^{\rm L}_2(S)$ implies that $\mathcal{R}^{\pi^1}_j \ge \mathcal{R}^{\pi^2}_j$.

Conversely, let us suppose that the following conditions are satisfied:

$$\mathcal{R}^{\pi^1}_j \ge \mathcal{R}^{\pi^2}_j \quad\text{and}\quad \left(\frac{\mathcal{R}^{\pi^1}_j}{\mathcal{R}^{\pi^2}_j}\right)^{\beta} \ge \left(\frac{C^{\pi^1}_j}{C^{\pi^2}_j}\right)^{\beta-1}.$$

Then, equation (1.22) immediately implies that $F^{\rm L}_1(S) \ge F^{\rm L}_2(S)$ for S sufficiently small or sufficiently large. Finally, the inequality $(F^{\rm L}_1(S) \ge F^{\rm L}_2(S))$ extends to all $S \ge 0$ by the convexity of $F^{\rm L}_1(S)$ and $F^{\rm L}_2(S)$. □

References

1. N. Aleksandrov, R. Espinoza, and L. Gyurkó. Optimal oil production and the world supply of oil. *Journal of Economic Dynamics and Control*, 37(7):1248–1263, 2013.

2. C. Alford, M. Brazil, and D. Lee. Optimization on underground mining. In A. Weintraub, C. Romero, T. Bjørndal, and R. Epstein, editors, *Handbook on Operations Research in Natural Resources*, pages 561–578. Springer-Verlag, New York, NY, 2007.
3. S. Asmussen. *Applied Probability and Queues*. Springer-Verlag, New York, NY, 2003.
4. T. Baker and L. Ladson. Successive linear programming at Exxon. *Management Science*, 31(3):264–274, 1985.
5. F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
6. M. Brennan. The price of convenience and the valuation of commodity contingent claims. In D. Lund and B. Øksendal, editors, *Stochastic Models and Option Values: Applications to Resources, Environment and Investment Problems*, pages 33–72, 1991.
7. M. Brennan and E. Schwartz. Evaluating natural resources investments. *Journal of Business*, 58(2):135–157, 1985.
8. P. Brown and M. Pincheira. The price-stock relationship in the copper market, a surprising approach. In *Fourth International Conference Copper 99, Phoenix, AZ*, 1999.
9. R. Caldentey and S. Mondschein. Policy model for pollution control in the copper industry, including a model for the sulfuric acid market. *Operations Research*, 51(1):1–16, 2003.
10. R. Caldentey and D. Saure. Network sequencing under market uncertainty. Working paper, New York University, 2015.
11. R. Caro, R. Epstein, P. Santibanez, and A. Weintraub. An integrated approach to the long-term planning process in the copper mining industry. In A. Weintraub, C. Romero, T. Bjørndal, and R. Epstein, editors, *Handbook on Operations Research in Natural Resources*, pages 595–609. Springer-Verlag, New York, NY, 2007.
12. J. Casassus and P. Collin-Dufresne. Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance*, 60(5):2283–2331, 2005.
13. J. Cherian, J. Patel, and Ilya Khripko. Optimal extraction of nonrenewable resources when costs cumulate. In M. Brennan and L. Trigeorgis, editors, *Project Flexibility, Agency, and Competition: New Developments in the Theory and Application of Real Options*, pages 224–253. Oxford University Press, New York, NY, 1999.
14. Codelco. Yearbook. http://www.codelco.com/memoria2013/, accessed in September 2014.
15. G. Cortazar and J. Casassus. Optimal timing of a mine expansion: Implementing a real options model. *The Quarterly Review of Economics and Finance*, 38(3):755–769, 1998.
16. G. Cortazar, E. Schwartz, and J. Casassus. Optimal exploration investments under price and geological-technical uncertainty: A real options model. *R&D Management*, 31(3):181–189, 2001.
17. D. Cuoco and J. Cvitanić. Optimal consumption choices for a "large" investor. *Journal of Economic Dynamics and Control*, 22(3):401–436, 1998.
18. R. Dimitrakopoulos and S. Sabour. Evaluating mine plans under uncertainty: Can the real options make a difference? *Resources Policy*, 32(3):116–125, 2007.
19. A. Dixit and R. Pindyck. *Investment Under Uncertainty*. Princeton University Press, Princeton, NJ, 1994.
20. J. Dyer, R. Lund, J. Larsen, and R. Leone. A decision support system for prioritizing oil and gas exploration activities. *Operations Research*, 38(3):386–396, 1990.
21. R. Epstein, M. Goic, A. Weintraub, J. Catalan, P. Santibanez, R. Urrutia, R. Cancino, S. Gaete, A. Aguayo, and F. Caro. Optimizing long-term production plans in underground and open-pit copper mines. *Operations Research*, 60(1):4–17, 2012.
22. R. Epstein, R. Morales, J. Serón, and A. Weintraub. Use of OR systems in the Chilean forest industry. *Interfaces*, 29(1):7–29, 1999.
23. L. Evans and G. Guthrie. Commodity prices and the option value of storage. In N. Secomandi, editor, *Real Options in Energy and Commodity Markets*. Springer-Verlag, New York, NY, 2016.
24. B. Felix, O. Woll, and C. Weber. Gas storage valuation under limited market liquidity: an application in Germany. *The European Journal of Finance*, 19(7):715–733, 2013.

25. R. Frey. Perfect option hedging for a large trader. *Finance and Stochastics*, 2(2):115–141, 1998.
26. H. Geman. *Commodities and Commodity Derivatives. Modeling and Pricing for Agriculturals, Metals and Energy*. John Wiley & Sons, Chichester, UK, 2005.
27. H. Geman. Energy commodity prices: Is mean-reversion dead? *The Journal of Alternative Investments*, 8(2):31–45, 2005.
28. H. Geman. Mean reversion versus random walk in oil and natural gas prices. In M. Fu, R. Jarrow, J. Yen, and R. Elliot, editors, *Advances in Mathematical Finance*, Berlin, Germany, 2007. Birkhäuser.
29. M. Gersovitz and C. Paxson. The economics of Africa and the prices of their exports. Technical Report 68, Princeton University, Princeton, NJ, 1990.
30. R. Gibson and E. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance*, 45(3):959–976, 1990.
31. J. Hull. *Options, Futures and Other Derivative Securities*. Pearson Prentice-Hall, Upper saddle River, NJ, 2006.
32. W. Hustrulid, M. Kuchta, and R. Martin. *Open Pit Mine Planning and Design*. Taylor & Francis, Bristol, PA, 2013.
33. K. Lane. *The Economic Definition of Ore: Cut-Off Grades in Theory and Practice*. Mining Journal Books Limited, London, UK, 1988.
34. R. Lumley and M. Zervos. A model for investing in the natural resource industry with switching costs. *Mathematics of Operations Research*, 26(4):637–653, 2001.
35. R. McDonald and D. Siegel. Investment and the valuation of firms when there is an option to shut down. *International Economic Review*, 26(2):331–349, 1985.
36. R. Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.
37. S. Myers. Finance theory and financial strategy. *Interfaces*, 14(1):126–137, 1984.
38. R. Pindyck. The optimal exploitation and production of nonrenewable resources. *Journal of Political Economy*, 86(5):841–861, 1978.
39. R. Pindyck. The optimal production of an exhaustible resource when price is exogenous and stochastic. *The Scandinavian Journal of Economics*, 83(2):277–288, 1981.
40. P. Protter. *Stochastic Integration and Differential Equations*, volume 21. Springer, Berlin, Germany, 2nd edition, 2004.
41. P. Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6(2):13–39, 1965.
42. E. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 52(3):923–973, 1997.
43. E. Schwartz and J. Smith. Short term variations and long-term dynamics in commodity prices. *Management Science*, 46(7):893–911, 2000.
44. N. Secomandi. Optimal commodity trading with a capacitated storage asset. *Management Science*, 56(3):449–467, 2010.
45. N. Secomandi and D.J. Seppi. Real options and merchant operations of energy and other commodities. *Foundations and Trends in Technology, Information and Operations Management*, 6(3-4):161–331, 2014.
46. S. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer-Verlag, New York, NY, 2004.
47. M. Slade. Valuing managerial flexibility: An application of real-option theory to mining investments. *Journal of Environmental Economics and Management*, 41(2):193–233, 2001.
48. H. Smit and L. Trigeorgis. *Strategic investment: Real options and games*. Princeton University Press, Princeton, NJ, 2012.
49. J.E. Smith and K.F. McCardle. Options in the real world: Lessons learned in evaluating oil and gas investments. *Management Science*, 47(1):1–15, 1999.
50. L. Trigeorgis. *Real Options: Managerial Flexibility and Strategy in Resource Allocation*. MIT Press, Cambridge, MA, 1996.