Optimal Exploitation of a Nonrenewable Resource

René Caldentey
Stern School of Business, New York University,
44 West Fourth Street, Suite 8-77, New York, NY 10012, rcaldent@stern.nyu.edu.

Rafael Epstein
Industrial Engineering Department, University of Chile,
Republica 701, Santiago, Chile. repstein@dii.uchile.cl.

Denis Sauré
Industrial Engineering Department, University of Chile,
Republica 701, Santiago, Chile. dsaure@dii.uchile.cl.

Abstract

In this paper we study the long-term operation of an underground copper mining project. We model the project as a collection of *blocks* (minimal extraction units) 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 rate of extraction as well as investing on costly capacity expansions. We use standard contingent claim analysis and risk-neutral valuation to solve this problem taking as an input the stochastic process that regulates the dynamics of the copper spot and futures prices. Our solution method works in two steps. First, we consider a fixed production capacity and use approximated dynamic programming to compute upper and lower bounds for the value function in terms of the spot price and mineralogical characteristics of the blocks. We use these bounds to obtain an approximation that is asymptotically optimal as the spot price grows large. In the second step, we extend this asymptotic approximation to handle capacity expansion decisions. Numerical computations are used to evaluate the performance of our proposed policy. We conclude with an application of our methodology to a real instance of the problem at CODELCO (Chile's largest copper producer).

1 Introduction

In this paper, we develop a real options model for optimizing the long-term exploitation of an underground multi-sector copper mining project. This research is part of an ongoing project with CODELCO –Chile's largest copper producer– an 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 5.4 metric million tons of fine copper (TMF) and holding 38% of the world's copper reserves. Unsurprisingly, copper is one of Chile's most important industries, accounting for approximately 7% and 38% 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 of this problem comes from the combination of two factors: (*i*) a large scale operation involving multiple inter-temporal decisions and (*ii*) an uncertain production and market environment. In the copper industry, production uncertainty relates to the heterogeneity of the mineral composition, while marketplace uncertainty is due to the stochastic evolution of spot prices.

Decision-support systems based on large-scale optimization models have been successfully implemented in the copper sector (*e.g.*, Mondschein and Schilkrut 1997, Santibañez 2000, Caldentey and Mondschein 2003), as well as in other natural resource industries (*e.g.*, Epstein et al. 1999 in the forrest industry, or Baker and Ladson 1985, Dyer et al. 1990 in the crude oil industry). Through the use of these systems, managers are able to evaluate alternative operational policies selecting those that maximize the short-term and long-term profitability of the business. Most of these models, however, operate under deterministic inputs based on mean estimates of market prices and demand. Furthermore, in this deterministic world the so-called discounted cash flow (DCF) methodology with a non-adapted operational and investment strategy is predominately used by companies, despite the fact that it fails capturing decision makers' ability to dynamically react to a stochastically changing environment (*e.g.*, Myers 1987). As a result, companies operate under suboptimal extraction and investment plans, highly exposed to operational and financial risks.

The real options approach overcomes these limitations of the DCF 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 is McDonald and Siegel 1985 that investigated the optimal operation of a project under stochastic revenues and production costs if a shut-down option is considered. For a comprehensive exposition on real options we refer the reader to Dixit and Pindyck (1994) and Trigeoris (1996).

In the context of natural resource management, there is an extensive real options literature that has focused 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.*, Pindyck 1978 and Pindyck 1981, Schwartz et al. 2001, Brennan and Trigeorgis 2000 and Lumley and Zervos 2001). An important aspect of this real options literature for commodities (such as copper or oil) has to do with the way spot price risk is incorporated in the valuation process. Specifically, for these commodities the existence of a futures markets allows the use of risk-neutral (or arbitrage-free) valuation techniques similar to those used for valuing financial derivatives (*e.g.*, Black and Scholes 1973 and Hull 1993). A notable example of this risk-neutral approach is Brennan and Schwartz (1985) who consider optimal extraction policies for a non-renewable natural resource. Other examples are Gibson and Schwartz (1990), Samis et al. (2001) and Smith and McCardle (1999).

Our paper differs from previous formulation in the way we model the extraction process. Based on our experience working with CODELCO, most of the previous research favors mathematical tractability by oversimplifying the production process. In this work we develop a real options model that addresses some of the limitations of previous approaches. Specifically, and consistent with current practices in the industry, we model the mining project as a collection of blocks (minimal extraction units) with different mineral composition and extraction costs. As a result, a production plan must specifies a sequence of extraction (that is, the order in which blocks will be processed) as well as the rate of production. We also model the options of investing in costly capacity expansions over time. Finally, we use real data from a Chilean copper mining project to calibrate our production and cost parameters and show how to apply our methodology in a real instance at CODELCO.

Because of our detailed characterization of the production process, we believe our paper contributes to narrow the gap between the academic literature and current practices in the industry. As a side remark, and before jumping into the details of the model, we would like to highlight the fact that, although we focus on a copper mining operations, our model and results can be extended to other non-renewal natural resources such as crude oil, natural gas, and other type of mineral deposits.

The rest of this paper is organized as follows. In section 2, we provide a description of the model. This section is subdivided in five subsections that include (*i*) a brief summary of the operations of an underground mine, (*ii*) a review of spot and futures price models for commodities, (*iii*) a detailed description of the production process, (*iv*) a discussion of the risk-neutral approach that we use to value the mining project including a dynamic programming formulation of the problem and (*v*) a summary of notation and conventions that we use throughout the paper. In section 3, we fix the sequence of extraction and the production capacity and derive general properties of an optimal solution. We also compute upper and lower bounds for the optimal value of the project and use these bounds to propose two simple methods to control extraction. These bounds are also used to derive an approximation that is asymptotically optimal as the spot price and/or production capacity grow large. We conclude section 3 with a set of numerical experiments that show the quality of our proposed asymptotic approximation with an average error ranging between 1% and 3%. In section 4 we extend the results of the previous section to include capacity expansion decisions. Section 5 presents an application of our methodology to identify an optimal extraction policy for *El Diablo*, a 230 [million ton] project at *El Teniente*. Conclusions and future research are discussed in section 6. Finally, mathematical proofs are relegated to an Appendix at the end.

2 Model Description

We now present the model in detail. We begin with a brief description of the mining operational process and then consider the dynamics of the spot and futures prices of copper. Then, we describe the mathematical formulation that we use to model the production process at an underground deposit. Finally, we discuss the risk-neutral valuation approach that we use to formulate the optimization problem as a dynamic programming problem.

2.1 Mining Operations Description

Mining operations can be regarded 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 they are continuously required through the lifespan of a mining project for updating the geological characteristics of the mineral. Extraction activities are required to obtain the ore that feeds the processing 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 the blast, mineral is carried out of the pit by huge trucks. Mine sectors are also located at different heights overlapping each other. Upper sectors must be extracted first for the extraction to be feasible and safe. Since only exposed mineral can be extracted from the surface, extraction sequence follows naturally.

In underground mines, extraction is conducted by choosing specific extractions points where the mineral flows by a combination of controlled explosions and gravity. Charges of explosives are set at the base of the mineral columns. After the blast, gravity makes the mineral to fall down. This extracting method is called “block caving” and it is used by almost all Chilean underground mines. Further discussion on extracting methods can be found on Alford et al. (2005). The ore is then transported outside the mine and new explosives are set in order to continue the so-called *precipitation* of the mineral.

The material coming out of the mine has poor grade (percentage of copper) which varies from 0.6% to 2.0%. For this material to have an economic value concentration and refining must be performed. However, not all the resource is sent to refining plants. For underground mines, material for which net cash flows are negative is left in-situ. For open pit mines this material is sent to dump deposits. Lane (1988) proposes a cut-off grade strategy for distinguishing between waste and ore based on approximations to the “opportunity” cost of the material.

The concentration and refining processes are basically the same for underground and open pit mines. These processes take place at the concentration and smelting plants. Here, the mineral goes through a sequence of processes (mechanical and chemical) of size reduction, concentration and refinement. The resulting ore has 99.9% grade and it is sold as a commodity in the local and international markets as refined copper. Not all the copper coming from the mines is transformed into high quality copper and some fraction is lost in the refinement processes. This fraction depends on the quality of the mineral, the extraction rate, and the relationship between the market price and the processing cost. This fraction is the so called *recovery* rate of the process.

Figure 1 shows schematically the entire mining operations process for the underground case. Further discussion on exploration and geological activities can be found in Schwartz *et* al. Schwartz et al. (2001) while the concentration and refining operations are discussed in detail in Caldentey and Mondschein (2003). A more complete description of the entire process can be found in *The Copper Manual* Cochilco (1976).

2.2 Spot and Futures Prices

Copper spot price is a critical ingredient when evaluating mining projects, as it modulates the project revenues influencing extraction plans, cutoff grades, and capacity expansion, among other decisions. In addition, the inherently stochastic behavior of these spot prices complicates the optimization and the evaluation of our project.

Most of the contingent claim literature models on stocks spot prices using the traditional Geometric Brownian Motion (GBM). Empirical evidence, analytical tractability, the unpredictability of the path,

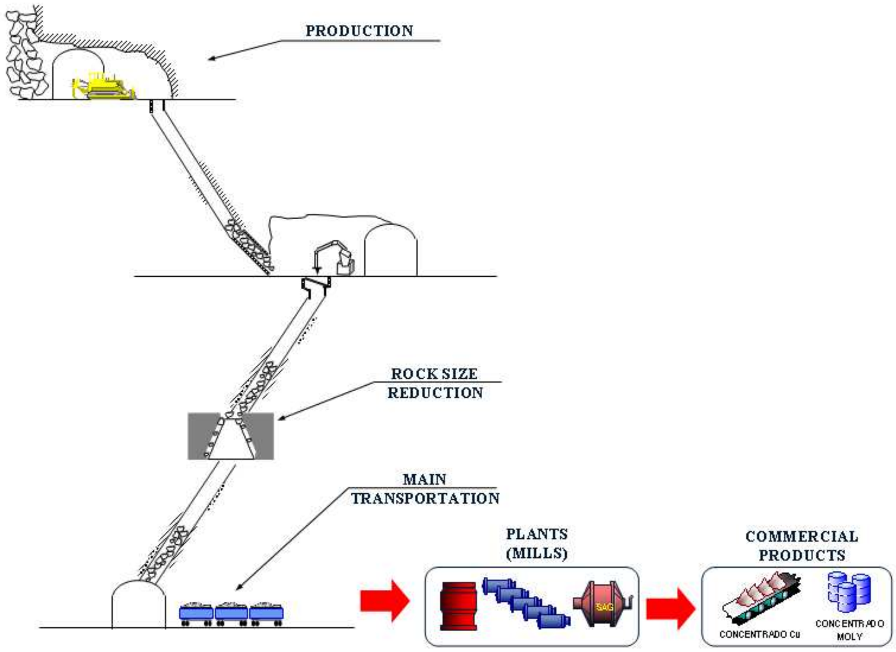


Figure 1: Mining process for an underground copper mine.

and the early applications of GBM in mathematical finance (*e.g.*, Samuelson 2001, Black and Scholes 1973, Merton 1973) are some of the main reasons behind this choice. However, it has been recognized (*e.g.*, Schwartz and Smith 2000, Smith and McCardle 1999) that commodity prices have a mean-reverting tendency that reflects companies' flexibility to open/close projects or increase/decrease production capacity in response to new market prices. Dixit and Pindyck (1994) test the GBM versus mean-reverting hypothesis using the copper spot price series 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 2). Similar conclusions are reported in Gersovitz and Paxson (1990) for ten different natural resource including copper.

In a series of papers (Schwartz 1997, Gibson and Schwartz 1990 and Schwartz and Smith 2000), Schwartz proposed three variations of a mean-reverting stochastic model driven by one, two or three factors. These models were empirically validated for copper, gold and crude oil. Alternative discrete-time models for the copper spot price are discussed in Engel and Valdés (1990). The authors make use of chronological series and ARIMA models (among others) to forecast copper prices within a five-year horizon. They conclude that the two models with the best predictability are the auto-regressive of first order and the discrete-time version of the GBM.

We assume that the copper spot price, S_t, follows a single factor mean reverting process, as described in Schwartz (1997) by the following stochastic differential equation

$$\mathrm{d}S_t = \kappa(\mu - \ln(S_t))\, S_t\, \mathrm{d}t + \sigma S_t\, \mathrm{d}B_t, \tag{1}$$

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 mean μ. Using the Kalman Filter procedure and data from

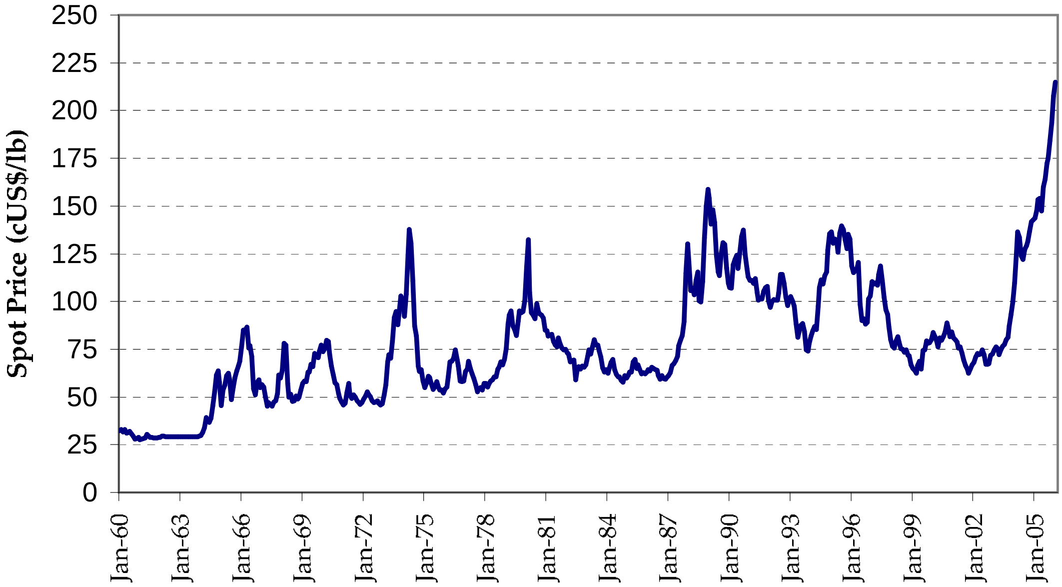


Figure 2: Copper spot price evolution since January 1960. (Source: London Metal Exchange)

1988 to 1995 the author obtained the following estimates: $\kappa = 0.369$, $\mu = 4.854$ and $\sigma = 0.233$.

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. The existence of this future market will allow us to use a replicating (or risk-neutral) argument optimize the value of the mining project under consideration.

The relationship between spot and futures price that we consider is based on the model proposed by Brennan and Schwartz (1985) in which the convenience yield[†] is proportional to the spot price with constant of proportionality ρ. For alternative models of this convenience yield we refer the reader to Casassus and Collin-Dufresne (2005) and references therein. It is well known that the futures price for the purchase of one unit of the commodity to be delivered at time τ is equal to the expected value of the spot price S_τ under the risk-neutral measure (e.g., Shreve 2004, chapter 5). Hence, if we let $F(S,\tau)$ be this futures price when the current spot price is $S_0 = S$ then

$$F(S,\tau) = S\,\exp((r-\rho)\,\tau), \tag{2}$$

where r is the risk-free discount rate. For the purpose of numerical computations, we will assume $r = 12\%$ and $\rho = 6.3\%$ which is the average value of the (instantaneous) convenience yield for copper reported in Casassus and Collin-Dufresne (2005, Table V).

2.3 Production Model

We adopt a continuous-time model to represent the operation of an underground mine. We shall also assume that the exploration stage have been already completed. Hence, the decision maker is mainly

[†]The convenience yield can be interpreted as the flow of services accruing to the holder of the spot commodity but not to the owner of a futures contract.

concerned with determining an optimal extraction and capacity expansion plan for a fixed mining project with known mineral content and quality.

As a matter of representation, and consistent with current practices at CODELCO, we divide the mine into a collection of BCUs (Basic cubication units) or *blocks* each one having specific geological properties. These blocks represent the minimal extraction units so that production decisions are made at the block level. Usually, extraction begins at the block with the best grade. Then, it continues by extracting the surrounding blocks giving priority to those with higher grade. This determines an extraction sequence. However, there are some restrictions that prevent this greedy sequence to be actually implemented. 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.

From a mathematical standpoint, we model the mine as an undirected graph $G = (\mathcal{N}, \mathcal{A})$ where $\mathcal{N} = \{1, 2 \ldots, N\}$ is the collection of BCUs or blocks. We define Q_i and L_i to be the amount of ore and average grade (% of copper), respectively, available in block $i \in \mathcal{N}$. We also denote by $\mathcal{N}_0 \subseteq \mathcal{N}$ the subset of blocks that can be chosen as initial block in a feasible production sequence.

The set of arcs $\mathcal{A}$ connecting the blocks in $\mathcal{N}$ represents the alternative sequences (paths) of production in which blocks can be extracted. For instance, if there is an arc connecting nodes i and j (*i.e.*, $(i, j) \in \mathcal{A}$) then after completing the extraction of block i it is possible to start the extraction of block j, and viceversa. Hence, the topology induced by $\mathcal{A}$ in the graph G is intended to capture the actual spatial topology of the mining project. It also captures additional production adjacency constraints among blocks such as those described above.

For a given subset of blocks $\widetilde{\mathcal{N}} \subseteq \mathcal{N}$, we define $\mathcal{A}(\widetilde{\mathcal{N}}) := \{j \in \mathcal{N} - \widetilde{\mathcal{N}} \ : \ \exists\, i \in \widetilde{\mathcal{N}} \text{ such that } (i, j) \in \mathcal{A}\}$, that is, $\mathcal{A}(\widetilde{\mathcal{N}})$ is the set of blocks that are adjacent to $\widetilde{\mathcal{N}}$. We also define Π to be the set of all possible permutations $\pi = (\pi_1, \ldots, \pi_N)$ of $\mathcal{N}$ such that $\pi_1 \in \mathcal{N}_0$ and $\pi_i \in \mathcal{A}(\{\pi_1, \ldots, \pi_{i-1}\})$ for all $i = 2, \ldots, N$. In other words, Π is the set of all possible sequences in which the blocks in $\mathcal{N}$ can be extracted.

A production strategy in our model consists of three basic components: (i) the sequence in which blocks are extracted, (ii) the time at which each block in this sequence starts processing, and (iii) the available production capacity at these production epochs. We will denote by θ a generic production strategy and by Θ the set of all feasible production strategies. In what follows, we describe in detail the set Θ and the three components of a policy $\theta \in \Theta$.

First, we let $\pi^\theta = (\pi_1^\theta, \ldots, \pi_N^\theta) \in \Pi$ be the (feasible) production schedule associated to strategy $\theta \in \Theta$. We denote by $\tau^\theta = (\tau_1^\theta, \ldots, \tau_N^\theta)$ the sequence of extraction times of the sequence of blocks in π^θ. That is, τ_i^θ is time at which the i^{th} block in the sequence π^θ starts production. Given π^θ and corresponding extraction time τ^θ, we denote by $K^\theta = (K_1^\theta, \ldots, K_N^\theta)$ the vector of production capacity. The i^{th} component of this vector, K_i^θ, is the available production capacity at time τ_i^θ when block π_i^θ starts processing. For completeness, we define $K_0^\theta = K_0$ for all $\theta \in \Theta$, where K_0 is the initial level of capacity.

Based on our experience working with CODELCO, we will assume that the decision maker does not increase production capacity during the extraction of a block. These capacity expansion decisions are only made in between block extractions. In addition, we will also 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. We make these assumptions for mathematical tractability but we note that they are not far from reality since these decisions involve production and scheduling disruptions that usually cannot occur in the middle of the extraction of a block. Furthermore, we can always reduce the size of the blocks (by increasing the cardinality of $\mathcal{N}$) and so the assumptions are not particularly restrictive from a theoretical standpoint. In fact, we can show that if we make the size of the blocks infinitesimally small then a “bang-bang” extraction policy is optimal which is consistent with our previous set of assumptions.

Given a policy $\theta \in \Theta$, we will denote by $Q^\theta = (Q_1^\theta, \ldots, Q_N^\theta)$ and $L^\theta = (L_1^\theta, \ldots, L_N^\theta)$ the vectors of ore content and average grade, respectively, for the production sequence π^θ. We also define $T^\theta = (T_1^\theta, \ldots, T_N^\theta)$ where $T_i^\theta = Q_i^\theta / K_i^\theta$ is the time it takes to process block π_i^θ under policy θ.

Finally, we define the set Θ of feasible production policies $\theta = (\pi^\theta, \tau^\theta, K^\theta)$ as those satisfying the conditions $\pi^\theta \in \Pi$ and

$$\tau_i^\theta \geq \tau_{i-1}^\theta + T_{i-1}^\theta \quad \text{and} \quad K_i^\theta \geq K_{i-1}^\theta, \quad \text{for all } i = 1, \ldots, N,$$

with border conditions $K_0^\theta = K_0$ and $\tau_0^\theta = T_0^\theta = 0$. Other constraints such as imposing a fixed planning horizon $\bar{T}$ or a maximum production capacity $\bar{K}$, that is,

$$\tau_N^\theta + T_N^\theta \leq \bar{T} \qquad \text{and} \qquad K_N^\theta \leq \bar{K}$$

can also be included. In this paper, we will assume that $\bar{T} = \infty$ which is a reasonable assumption for a company like CODELCO with more than 100 years of copper reserves.

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 resource. The investment policy in expansion projects must allow mineral extraction plan to be feasible in the short and long-term, ensuring the necessary capacities and technologies. From the decision maker’s perspective this long-term value maximization amounts to selecting an optimal production strategy $\theta \in \Theta$ 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 operators, like CODELCO, consider the average discounted cumulative cashflows of the project as the appropriate objective function to use. However, this NPV approach imposes some serious challenges in terms of selecting the appropriate discount factor, or equivalently the correct probability measure to compute expectations. Fortunately, we can avoid all these complications using the so called contingent claim valuation approach. Specifically, the existence of a futures market for copper together with a non-arbitrage condition allow us to use a replicating argument to compute the market value of our project. For more details, the readers is referred to Brennan and Schwartz (1985) for an application in the context of a natural resource exploitation and to Shreve (2004) 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 the so called *contingent* *claim* approach. Let $\mathbb{Q}$ be a probability measure (equivalent to $\mathbb{P}$) under which the spot price, S_t, discounted at the risk-free rate net of the convenience yield, $(r-\rho)$, is a $\mathbb{Q}$-martingale, that is,

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-(r-\rho)\,t}\,S_t \,|\, S_\tau\right] = e^{-(r-\rho)\,\tau}\,S_\tau, \qquad \text{for } \tau \leq t. \tag{3}$$

It follows from the single-factor dynamics of the spot price in equation (1) that 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 Shreve (2004)). For this, we define the *market price of risk* to be

$$\lambda_t \triangleq \frac{\kappa(\mu - \ln(S_t)) - (r-\rho)}{\sigma},$$

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

$$Z \triangleq \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \exp\left(-\int_0^T \lambda_t\,\mathrm{d}B_t - \frac{1}{2}\int_0^T \lambda_t^2\,\mathrm{d}t\right).$$

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

$$\mathrm{d}S_t = (r-\rho)\,S_t\,\mathrm{d}t + \sigma\,S_t\,\mathrm{d}\tilde{B}_t, \tag{4}$$

where $\tilde{B}$ is a Brownian motion under $\mathbb{Q}$ that satisfies $\mathrm{d}\tilde{B}_t = \mathrm{d}B_t + \lambda_t\,\mathrm{d}t$.

Under the risk-neutral valuation approach, the economic value of the project for a given policy $\theta \in \Theta$ 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. We will 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 i^{th} block in a feasible production strategy $\theta \in \Theta$. The extraction of this block π_i^θ starts at time τ_i^θ at a constant extraction rate K_i^θ and finishes at time $\tau_i^\theta + T_i^\theta$. We will denote by W_i^θ the time τ_i^θ discounted cashflows generated by this block.

$$W_i^\theta = \int_0^{T_i^\theta} e^{-r\,t}\,[L_i^\theta\,K_i^\theta\,S_{\tau_i^\theta+t} - A_i^\theta\,K_i^\theta]\,\mathrm{d}t,$$

where $L_i^\theta\,K_i^\theta$ is the rate at which copper is produced and A_i^θ is the marginal production cost for block π_i^θ under policy θ. From a modeling perspective, we would like the marginal cost A_i^θ to depend not only on the particular block π_i^θ but more generally on the actual sequence of extraction. This flexibility in our formulation is particularly useful for modeling underground mining operations in which production costs tend to increase as extraction progresses. This in part due to the fact that the distance from the extraction points to the processing plan increases over time. In order to capture this sequence-depend cost structure in a tractable way, we will assume that A_i^θ depends on θ only through the set $\mathfrak{N}_{i-1}^\theta := \{\pi_1^\theta,\ldots,\pi_{i-1}^\theta\}$, that is, $A_i^\theta = A_i(\mathfrak{N}_{i-1}^\theta)$. This is not a very restrictive assumption from a practical standpoint that will allows us to formulate the decision maker's optimization problem using standard dynamic programming techniques.

For every $\theta \in \Pi$, we define the decision maker's cumulative discounted payoff by

$$U^\theta = \sum_{i=1}^{N} e^{-r\,\tau_i^\theta}\left[W_i^\theta - \gamma\,(K_i^\theta - K_{i-1}^\theta)\right], \tag{5}$$

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

A few remarks about equation (5) are in order. First, the term $e^{-r\,\tau_i^\theta}\,\gamma\,(K_i^\theta - K_{i-1}^\theta)$ assumes that any capacity expansion takes place at the same time τ_i^θ when the extraction of block π_i^θ starts. One could argue, however, that the decision maker has the ability to build capacity at any point in time in the interval $[\tau_{i-1}^\theta + T_i^\theta, \tau_i^\theta]$, before the actual extraction of block i. However, because capacity expansion are costly, it is in the decision maker's best interest to postpone as much as possible this outlay which leads to equation (5). Condition (5) 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.

The optimization problem that we are interested in solving is

$$U^* = \sup_{\theta\in\Theta} \mathbb{E}^{\mathbb{Q}}[U^\theta].$$

Using equation (3) we can show that

$$\mathbb{E}^{\mathbb{Q}}[W_i^\theta | \mathcal{F}_{\tau_i^\theta}] = R_i^\theta\, S_{\tau_i^\theta} - C_i^\theta$$

where

$$R_i^\theta := L_i^\theta \left(\frac{1 - e^{-\rho\, T_i^\theta}}{\rho}\right) K_i^\theta \qquad \text{and} \qquad C_i^\theta := A_i(\mathfrak{N}_{i-1}^\theta) \left(\frac{1 - e^{-r\, T_i^\theta}}{r}\right) K_i^\theta.$$

The optimization problem can be rewritten as

$$U^* = \sup_{\theta\in\Theta} \mathbb{E}^{\mathbb{Q}}\left[\sum_{i=1}^{N} e^{-r\,\tau_i^\theta}\,\left[R_i^\theta\, S_{\tau_i^\theta} - C_i^\theta - \gamma\,(K_i^\theta - K_{i-1}^\theta)\right]\right]. \tag{6}$$

We can search for a solution to this problem using dynamic programming. The state space of this dynamic program is the triplet $(S, K, \mathfrak{N})$, where S is the spot price, K is the available production capacity and $\mathfrak{N} \subseteq \mathcal{N}$ is the subset of blocks already extracted. In this state space, we denote by $V(S, K, \mathfrak{N})$ the expected optimal discounted profit to go.

We note that the state space description $(S, K, \mathfrak{N})$ 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 strategy it is enough to evaluate the value function only at those times when a block has finished extraction. In fact, suppose we look at the system exactly at the time a block has finished extraction and we let $(S, K, \mathfrak{N})$ be the state of the system at this time. At this point in time, the decision maker must decide which is the next block to extract. This next block i must belong to the set of blocks that are adjacent to $\mathfrak{N}$, that is, $i \in \mathcal{A}(\mathfrak{N})$. In addition, the decision maker must select the time τ when to start processing this block i. This time τ is a stopping with respect to $\mathcal{F}_t$, the filtration generated by S_t. Finally, at this extraction time τ the decision maker must also decide if capacity should be expanded from the current level K to a new level $\hat{K}$, with $K \le \hat{K} \le \bar{K}$. (Recall that $\bar{K}$ is an upper bound on the maximum level of capacity that can be installed.) Putting all the pieces together, we can write the following recursion for the value function $V(S,K,\mathfrak{N})$.

$$V(S,K,\mathfrak{N}) = \sup_{\tau,i,\hat{K}} \mathbb{E}^{\mathbb{Q}}\Big[e^{-r\,\tau}\,\Big(R_i(\hat{K})\,S_\tau - C_i(\hat{K},\mathfrak{N}) - \gamma\,(\hat{K}-K) + e^{-r\,T_i}\,V(S_{\tau+T_i},\hat{K},\mathfrak{N}\cup\{i\})\Big)\,\Big|S_0 = S\Big] \quad (7)$$

subject to the dynamics of the spot price, S_t, in equation (4), (8)

τ is an $\mathcal{F}_t$ stopping time, (9)

$$i \in \mathcal{A}(\mathfrak{N}), \qquad T_i = \frac{Q_i}{\hat{K}}, \qquad K \le \hat{K} \le \bar{K}, \quad (10)$$

and the border condition $V(S,K,\mathcal{N}) = 0$ for all S, K. (11)

We use the notation $R_i(\hat{K})$ and $C_i(\hat{K},\mathfrak{N})$ to emphasize the dependence of the revenue R_i and cost C_i on $\hat{K}$ and $\mathfrak{N}$ the rate and the sequence of extraction, respectively.

The optimization problem (7)-(11) is not the prototypical dynamic program that we usually encounter in this type of manufacturing problems. What makes the problem special is the state variable $\mathfrak{N}$. Because of its discrete nature, it is not possible to derive a simple (Hamilton-Jacobi-Bellman) optimality equation for the value function in terms of its partial derivatives. Instead, we have to set an optimization problem for each possible value of $\mathfrak{N}$ and solve a sequence of optimal stopping problems using backward induction on the cardinality of $\mathfrak{N}$.

The idea is as follows. For any set $\mathfrak{N}$ and $i \in \mathcal{A}(\mathfrak{N})$ define the auxiliary value function $V(S,K,\mathfrak{N},i)$ which is the optimal discounted profit to go if the system is currently in state $(S,K,\mathfrak{N})$ and the decision maker has already decided to extract block $i \in \mathcal{A}(\mathfrak{N})$ next, although not necessarily immediately.

Suppose that we have been able to compute the value of $V(S,K,\mathfrak{N})$ for all S, K and $\mathfrak{N}$ such that $\|\mathfrak{N}\| = k$. The border condition makes this assumption trivial for $k = N$. Then, for any set $\mathfrak{N}$ such that $\|\mathfrak{N}\| = k-1$ and $i \in \mathcal{A}(\mathfrak{N})$ we can compute $V(S,K,\mathfrak{N},i)$ reformulating (7)-(11) as an optimal stopping time problem with S and K as the only state variables. We can tackle the solution to this subproblem using standard optimal control techniques (*e.g.*, chapters 10 and 11 in Øksendal (2003)). The specific details of how we solved it are relegated to sections 3 and 4. Once we have computed the auxiliary value function $V(S,K,\mathfrak{N},i)$ for every $\mathfrak{N}$ such that $\|\mathfrak{N}\| = k-1$ and $i \in \mathcal{A}(\mathfrak{N})$, we can recover the original value function solving $V(S,K,\mathfrak{N}) = \max_{i\in\mathcal{A}(\mathfrak{N})}\{V(S,K,\mathfrak{N},i)\}$.

It seems that the only difficulty of the algorithm described in the previous paragraph is solving the optimal control problem for $V(S,K,\mathfrak{N},i)$. Indeed, this is not an easy task. However, there is another important obstacle in the implementation of this algorithm. For every k there are potentially "N choose k", or $N!/(k!\,(N-k)!)$, subsets of $\mathcal{N}$ with cardinality k. Hence, the number of possible subproblems that we need to solve can be extremely large even for moderate values of N.

We will not deal with this *curse of dimensionality* directly in this paper. Instead, we will concentrate on solving the subproblem of determining optimal production and capacity investment decisions for a given sequence of extraction π. Using some standard terminology of dynamic programming (*e.g.*, Bertsekas 1995), we will solve the *open-loop* version of the problem in which the sequence of extraction has been already defined. This is an important simplification, however, we view it as the first step towards solving the general problem. Furthermore, it is often the case in practice that the decision maker wants to evaluate only a few predetermined extraction sequences. In Proposition 4 below we propose an efficient method to compare alternative sequences of extraction. We also discuss a concrete example of this scenario-based valuation in section 5.

For a fixed extraction sequence, our solution approach works in two steps. First, in section 3, we study the optimal timing of extraction given a fixed capacity K. Then, in section 4 we relax this assumption and derive near-optimal capacity expansion decisions.

2.5 Notation and Conventions

Unfortunately, notation will play an important role in our analysis. 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 paper. First, we say that a function $f(S)$ is *asymptotically equal* to a function $g(S)$, which we denote by $f(S) \overset{S\to\infty}{\longrightarrow} g(S)$, if $\lim_{S\to\infty} |f(S) - g(S)| = 0$. Also, the first and second derivatives of $f(S)$ with respect to S are denoted by $f'(S)$ and $f''(S)$, respectively.

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

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

We also use the specialize notation $X^+_j \triangleq X^+_{0,j}$, $\alpha^\times_j \triangleq \alpha^\times_{0,j}$ and $(\alpha^\times X)^+_j \triangleq (\alpha^\times X)^+_{0,j}$. 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 $j < k$.

Consider an arbitrary block j with mineral content Q_j, average grade L_j and marginal production cost A_j. Suppose the spot price is S and the production capacity is K. Then, the time it takes to extract block j is $T_j(K) \triangleq Q_j/K$ and the discounted expected profit generated by this block is

$$\begin{aligned} W_j(S,K) \quad &\triangleq \quad \mathbb{E}^{\mathbb{Q}}\left[\int_0^{T_j(K)} e^{-r\,t}\,[L^j \cdot S_t - A^j]\,K\,\mathrm{d}t \,\middle|\, S_0 = S\right] \triangleq S\,R_j(K) - C_j(K) \\ \text{where} \qquad & R_j(K) \triangleq \frac{L^j(1-e^{-\rho\,T_j(K)})\,K}{\rho} \quad \text{and} \quad C_j(K) \triangleq \frac{A^j\,(1-e^{-r\,T_j(K)})\,K}{r}. \end{aligned} \tag{12}$$

We interpret $R_j(K)$ as a modified stock of copper in block j and $C_j(K)$ as the discounted extraction cost for this block. We will denote by

$$\mathbb{C}_j(K) \triangleq \frac{C_j(K)}{R_j(K)}$$

the resulting per unit average production cost. We also define

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

$\mathcal{R}_j(K) \triangleq \mathcal{R}_{0,j}(K)$ and $\mathcal{C}_j(K) \triangleq \mathcal{C}_{0,j}(K)$. 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 strategy would be $\mathcal{W}_j(S,K) \triangleq \mathcal{R}_j(K)\,S - \mathcal{C}_j(K)$.

3 Optimal Production Plan with Fixed Capacity

In this section we solve problem (7)-(11) under the following two assumptions. First, we assume that the sequence of extraction π has been defined in advance. Without lost of generality, and for notational convenience, we index the blocks in this sequence by $\pi = \{N, N-1, \ldots, 2, 1\}$ so that block N is the first block to be processed and block 1 is the last one. The second condition that we impose is that capacity expansions are not allowed and we let K be this fixed capacity. We will relax this second assumption in the following section. Under these two conditions, 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 §5.2 in Dixit and Pindyck (1994).

We define $F_j(S, K)$ to be the maximum expected discounted profit when there are j blocks left, the spot price is S and the capacity is K. The corresponding Bellman equation for this value function is given by (see §2.5 for notation)

$$F_j(S,K) = \sup_{\tau>0} \mathbb{E}^{\mathbb{Q}} \left[e^{-r\,\tau} W_j(S_\tau, K) + e^{-r(\tau+T_j(K))}\, F_{j-1}(S_{\tau+T_j}, K) \,\middle|\, S_0 = S \right] \tag{13}$$

with border condition $F_0(S, K) = 0$ for all $S \geq 0$. The sup is taken over the set of stopping times τ (with respect to S_t) representing the time when block j should start being produced. Because K is kept constant in this section, we find convenient to drop the dependence of $F_j(S, K)$, $W_j(S, K)$, $T_j(K)$, $R_j(K)$ and $C_j(K)$ on K.

Problem (13) is an optimal stopping time problem for which the optimal policy is of the threshold type. This means that there is a threshold price S_j^* such that production of block j should start as soon as the spot price exceeds this threshold (for more details see sections 4.1.D and 5.2 in Dixit and Pindyck (1994)).

Using some standard results from optimal stopping time theory (*e.g.*, chapter 10 in Øksendal (2003)) together with the dynamics of the spot price in (4), we can derive the following HJB equation for $F_j(S)$ inside the *continuation region*, *i.e.*, the region of spot prices in which it is optimal to postpone the production of block j.

$$0 = \frac{1}{2}\sigma^2 S^2 \cdot F_j''(S) + (r-\rho)\cdot S \cdot F_j'(S) - rF_j(S), \qquad \text{for all } S \leq S_j^* \tag{14}$$

The border conditions for this HJB are

$$\begin{aligned} &F_j(0) = 0, \\ &F_j(S_j^*) = W_j(S_j^*) + e^{-r\,T_j}\, \mathbb{E}^{\mathbb{Q}} \left[F_{j-1}(S_{T_j}) \,\middle|\, S_0 = S_j^* \right] \quad \text{and} \\ &F_j'(S_j^*) = W_j'(S_j^*) + e^{-r\,T_j}\, \frac{\mathrm{d}}{\mathrm{d}S} \mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j}) | S_0 = S] \Big|_{S=S_j^*}. \end{aligned} \tag{15}$$

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 and third conditions guarantee that $F_j(S)$ is continuous and differentiable at the threshold price $S = S_j^*$. These are the so called *value matching* and *smooth pasting* conditions, respectively.

Equation (14) 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

$$\begin{aligned}\beta_1 &\triangleq \frac{1}{2}-(r-\rho)/\sigma^2+\sqrt{1/4+[(r-\rho)/\sigma^2]^2+(r+\rho)/\sigma^2}>1, \text{ and}\\ \beta_2 &\triangleq \frac{1}{2}-(r-\rho)/\sigma^2-\sqrt{1/4+[(r-\rho)/\sigma^2]^2+(r+\rho)/\sigma^2}<0.\end{aligned}$$

The general solution to equation (14) can be written as

$$F_j(S)=M_j\,S^{\beta_1}+\tilde{M}_j\,S^{\beta_2},$$

for two constants M_j and $\tilde{M}_j$. However, the border condition at $S=0$ implies that $\tilde{M}_j=0$. In conclusion, we have the following recursive solution for $F_j(S)$.

$$F_j(S)=\begin{cases} M_j\,S^{\beta_1} & \text{if } S\le S_j^*\\ R_j\,S-C_j+e^{-r\,T_j}\,\mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j})\big|S_0=S] & \text{otherwise.}\end{cases} \tag{16}$$

In what follows we redefine $\beta=\beta_1$.

To compute the values of S_j^* and M_j it is necessary to use the value matching and smooth pasting conditions. Except for the case of $j=1$, this is not easy to do analytically since we cannot get a tractable representation of $F_{j-1}(S)$ recursively from condition (16).

For $j=1$, condition (16) reduces to

$$F_1(S)=\begin{cases} M_1\,S^{\beta} & \text{if } S\le S_1^*\\ R_1\,S-C_1 & S\ge S_1^*\end{cases} \tag{17}$$

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

$$S_1^*=\frac{\beta\,C_1}{(\beta-1)\,R_1}=\frac{\beta}{\beta-1}\,\mathbb{C}_1 \quad\text{ and }\quad M_1=\frac{C_1}{\beta-1}\,(S_1^*)^{-\beta}. \tag{18}$$

Recall from section 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 $\frac{\beta}{\beta-1}-1$.

In general, extending the previous analysis to an arbitrary j is difficult because of the expectation $\mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j})\big|S_0=S]$ required in (16) and we have not been able to derive a simple characterization of $F_j(S)$ for $j\ge 2$. Nevertheless, we have been able to derive some useful properties of $F_j(S)$ that we present in the following result (see §2.5 for notation).

Proposition 1 *The value function $F_j(S)$ is increasing and convex in S. In addition, let us define $S_j^m\triangleq\max_{1\le k\le j}\{S_k^*\}$, $R_j^m\triangleq\max_{1\le k\le j}\{R_k\}$ and*

$$G_j(S)\triangleq S_j^m\,R_j^m\,\frac{j\,(j+1)}{2}\,\exp\left(-\frac{1}{\sigma^2}\min_{1\le k\le j}\left\{\left[\ln\left(\frac{S}{S_j^m}\right)+(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\right]^2\right\}\right).$$

Then, for $S \geq S_j^m \exp\Big(-\min_k \Big\{(r-\rho-\frac{\sigma^2}{2})\,T_{k,j}^+\Big\}\Big)$ *we have that*

$$0 \leq F_j(S) - \big(\mathcal{R}_j\, S - \mathcal{C}_j\big) \leq G_j(S). \tag{19}$$

Since $G_j(S)$ *converges to 0 as* S *goes to infinity, it follows that* $F_j(S)$ *is asymptotically equal to* $\mathcal{W}_j(S)$*, that is,*

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

Proof: See the appendix at the end. □

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 (16) 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 to the different parameters of the model. First, we obtain upper and lower bounds for $F_j(S)$ based on a generic type of approximation. Then, we use asymptotic analysis to extend these bounds. We conclude this section with some numerical computations that compare the performance of these bounds.

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 $F_j(S) \leq F_j^{\mathrm{U}}(S)$.

Similar to the original optimization in (13), the bound $F_j^{\mathrm{U}}(S)$ satisfies the following recursion

$$F_j^{\mathrm{U}}(S) = \sup_{\tau>0}\ \mathbb{E}^{\mathbb{Q}}\left[e^{-r\,\tau}W_j(S_\tau) + e^{-r\,\tau}\,F_{j-1}^{\mathrm{U}}(S_\tau)\,\Big|\,S_0 = S\right], \tag{20}$$

with $F_0^{\mathrm{U}}(S) = 0$ for all $S \geq 0$. It is not hard to see that $F_j^{\mathrm{U}}(S)$ satisfies the HJB equation (14) inside the continuation region. Therefore, it follows that

$$F_j^{\mathrm{U}}(S) = \begin{cases} M_j^{\mathrm{U}}\, S^{\beta} & \text{if } S \leq S_j^{\mathrm{U}} \\ R_j\, S - C_j + \widetilde{F}_{j-1}(S) & \text{otherwise.} \end{cases} \tag{21}$$

We can use backward induction to compute recursively the values of M_j^{U} and S_j^{U}, starting at block 1. We postpone this analysis to section 3.3 where we derive an algorithm that performs this task efficiently.

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.

The basic idea goes as follows. Consider again the optimal stopping time problem for $F_j(S)$ in (13)

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

Suppose we are able to find a convex function $F^{\mathrm{L}}_{j-1}(S)$ such that $F^{\mathrm{L}}_{j-1}(S) \leq F_{j-1}(S)$ for all $S \geq 0$. Then,

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

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 $F^{\mathrm{L}}_{j-1}(S)$ and condition (3) we get that

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

and so

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

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

$$F^{\mathrm{L}}_j(S) = \sup_{\tau>0} \, \mathbb{E}^{\mathbb{Q}} \left[e^{-r\,\tau} W_j(S_\tau) + e^{-r(\tau+T_j)} \, F^{\mathrm{L}}_{j-1} \left(e^{(r-\rho)\,T_j} \, S_\tau \right) \,\Big|\, S_0 = S \right], \tag{22}$$

with $F^{\mathrm{L}}_0(S) = 0$ for all $S \geq 0$. Using a similar line of arguments to the one used to derive (16), we can show that

$$F^{\mathrm{L}}_j(S) = \begin{cases} M^{\mathrm{L}}_j \, S^\beta & \text{if } S \leq S^{\mathrm{L}}_j \\ R_j \, S - C_j + e^{-r\,T_j} \, F^{\mathrm{L}}_{j-1}\big(e^{(r-\rho)\,T_j} \, S\big) & \text{otherwise.} \end{cases} \tag{23}$$

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. Our approach is based on a general family of approximations that includes $F^{\mathrm{L}}_j(S)$.

3.3 (α, η)-Approximations

The recursions that define the upper bound $F^{\mathrm{U}}_j(S)$ in (21) and the lower bound $F^{\mathrm{L}}_j(S)$ in (23) share a similar structure that we will exploit to derive a unified solution method.

Definition 1 *Let $\alpha = (\alpha_k)$ and $\eta = (\eta_k)$ be two positive vectors. We say that a set of continuous and differentiable functions $\{\mathcal{F}_k(S) \,:\, k = 0, \ldots, j\}$ is an (α, η)-approximation of the value functions in (16) if $\mathcal{F}_0(S) = 0$ for all $S \geq 0$ and*

$$\mathcal{F}_k(S) = \begin{cases} \mathcal{M}_k \, S^\beta & \text{if } S \leq \mathcal{S}_k \\ R_k \, S - C_k + \alpha_k \, \mathcal{F}_{k-1}\big(\eta_k \, S\big) & \text{otherwise,} \end{cases} \qquad k = 1, \ldots, j. \tag{24}$$

Because $\mathcal{F}_k(S)$ is continuous and differentiable, the values of $\mathcal{S}_k$ and $\mathcal{M}_k$ are implicitly determined imposing value matching and smooth pasting conditions similar to those in equation (15).

Our main motivation to consider this abstract family of approximations is because they generalize the upper and lower bounds. In fact, it follows from (21) that the upper bound $F_j^{\mathrm{U}}(S)$ is a special case of (24) with $\alpha_j = \eta_j = 1$. Similarly, we can recover the lower bound $F_j^{\mathrm{L}}(S)$ if we chose $\alpha_j = \exp(-r\,T_j)$ and $\eta_j = \exp((r-\rho)\,T_j)$. Hence, equation (24) defines a family of approximations that includes the upper and lower bounds.

In what follows, we derive an efficient algorithm that solves (24) 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, $j = 2$.

Using backward induction, we first compute $\mathcal{F}_1(S)$. In this case, the solution to (24) is identical to the solution in (17) and (18). That is,

$$\mathcal{F}_1(S) = \begin{cases} \mathcal{M}_1\,S^{\beta} & \text{if } S \le \mathcal{S}_1 \\ R_1\,S - C_1 & S \ge \mathcal{S}_1 \end{cases} \quad \text{where} \quad \mathcal{S}_1 = \left(\frac{\beta}{\beta-1}\right)\mathbb{C}_1 \quad \text{and} \quad \mathcal{M}_1 = \left(\frac{C_1}{\beta-1}\right)(\mathcal{S}_1)^{-\beta}. \tag{25}$$

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 recognize 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

$$\mathcal{F}_2(S) = \begin{cases} \mathcal{M}_2\,S^{\beta} & \text{if } S \le \mathcal{S}_2 \\ R_2\,S - C_2 + \alpha_2\,\eta_2^{\beta}\,S^{\beta}\,\mathcal{M}_1 & \text{if } \mathcal{S}_2 \le S \le \mathcal{S}_1/\eta_2 \\ (R_2 + \alpha_2\,\eta_2\,R_1)\,S - (C_2 + \alpha_2\,C_1) & \text{if } S \ge \mathcal{S}_1/\eta_2. \end{cases} \tag{26}$$

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 $\mathcal{S}_1 \le \eta_2\,\mathcal{S}_2$. 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

$$\mathcal{F}_2(S) = \begin{cases} \mathcal{M}_2\,S^{\beta} & \text{if } S \le \mathcal{S}_2 \\ (R_2 + \alpha_2\,\eta_2\,R_1)\,S - (C_2 + \alpha_2\,C_1) & \text{if } S \ge \mathcal{S}_2. \end{cases} \tag{27}$$

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

As we can see from this derivation of $\mathcal{F}_2(S)$, the solution depends on the relationship between $\mathbb{C}_1$ and $\eta_2\,\mathbb{C}_2$. Interestingly, for the case $\mathbb{C}_1 \le \eta_2\,\mathbb{C}_2$ the value of $\mathcal{F}_2(S)$ in (27) is analogous to the value of $\mathcal{F}_1(S)$ in (25). Indeed, in this case we can combine the two blocks into a single one so that the solution in (27) 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 $\mathcal{S}_1 \le \eta_2\,\mathcal{S}_2$ is equivalent to $S_1^{\mathrm{L}} \le \exp((r-\rho)\,T_2)\,S_2^{\mathrm{L}} = \mathbb{E}^{\mathbb{Q}}[S_{T_2}|S_0 = S_2^{\mathrm{L}}]$. 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 $\tilde{\jmath}$-block project with characteristics $\{(\widetilde{C}_k, \widetilde{R}_k, \widetilde{k}, \widetilde{\eta}_k),\ k = 1,\dots,\tilde{\jmath}\}$ and $\tilde{\jmath} \le j$. The purpose of this algorithm is to aggregate blocks using the same criteria discussed above. The resulting sequence $\{(\widetilde{C}_k, \widetilde{R}_k, \widetilde{k}, \widetilde{\eta}_k),\ k = 1,\dots,\tilde{\jmath}\}$ satisfies some properties that greatly simplify the computation of $\mathcal{F}_j(S)$. (See section 2.5 for notation)

Proposition 2 *Consider a single-sector mining project with j blocks 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 $\tilde{\jmath} = j$.*

Step 1: *Compute the auxiliary variables*

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

Step 2: *Find $\tilde{k} = \min\{2 \le k \le \tilde{\jmath} \ :\ \widetilde{\mathbb{C}}_{k-1} < \widetilde{\eta}_k\,\widetilde{\mathbb{C}}_k\}$. 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: $\tilde{\jmath} = \tilde{\jmath} - \xi$ and*

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

(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

$$\widetilde{S}_k = \left(\frac{\beta}{\beta-1}\right)\widetilde{\mathbb{C}}_k \qquad \text{and} \qquad \widetilde{M}_k = \widetilde{\alpha}_k\,\widetilde{\eta}_k^{\beta}\,\widetilde{M}_{k-1} + \left(\frac{\widetilde{C}_k}{\beta-1}\right)(\widetilde{S}_k)^{-\beta}, \quad k \le \tilde{\jmath}, \tag{28}$$

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

$$\mathcal{F}_j(S) = \left(\widetilde{\theta}^{\times}\,\widetilde{R}\right)^{+}_{h,\tilde{\jmath}} S - \left(\widetilde{\alpha}^{\times}\,\widetilde{C}\right)^{+}_{h,\tilde{\jmath}} + \widetilde{M}_h\,\widetilde{\alpha}^{\times}_{h,\tilde{\jmath}}\left(\widetilde{\eta}^{\times}_{h,\tilde{\jmath}}\right)^{\beta} S^{\beta}, \tag{29}$$

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

PROOF: See the appendix at the end. □

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

The discount factor is $r = 0.12$ and the convenience yield is $\rho = 0.06$. Let us specialize the result in Proposition 2 to the case of the lower bound $F^{\mathsf{L}}(S)$. To do this, we set $\alpha_k = e^{-r\,T_k}$ and $\eta_k = e^{(r-\rho)\,T_k}$, $k = 1, \ldots, 6$.

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 $\widetilde{T}_k$.

AFTER ITERATION 1

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

As we can see, the algorithm finishes after four iteration 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, 3, 4, 5 and 6. From equation (28), we derive the threshold price for block 6 in the original block configuration which is

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

The interpretation of this price is as follows: As soon as the spot price goes above $\widetilde{S}_6$ 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 block is of little 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 $\widetilde{S}_{\tilde{j}}$ using equation (28).
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 $\widetilde{S}_{\tilde{j}} = \mathcal{S}_j$, 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. □

As we can see, the previous policy uses the artificial configuration of blocks proposed by proposition 2 only to compute the threshold price that determines when the last 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 $S_t \geq \frac{33.79\,\beta}{(\beta-1)}$.

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 its implicitly built. The most common design rule is to select the initial extraction front in the sector 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 j 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 that has $\eta_k = 1$. 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.

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 3 (left panel) plots the value function $F_j(S)$ (numerically computed), the upper bound $F_j^U(S)$ and the lower bound $F_j^L(S)$ as a function of S using the data in Example 1. We note that both bounds

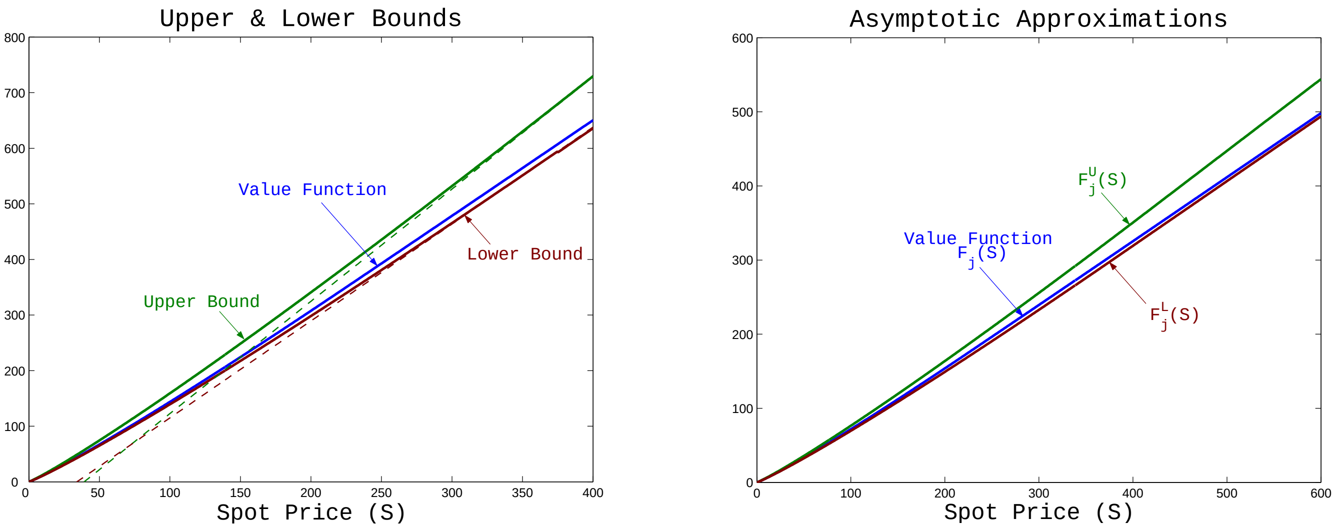


Figure 3: Left Panel: Value function ($F_j(S)$), upper bound ($F_j^U(S)$) and lower bound ($F_j^L(S)$) as a function of the spot price (S) using the data in Example 1. The dashed lines correspond to linear asymptotes for the upper and lower bound approximations. Right Panel: Asymptotic approximations of the value function based on equations (31) and (32).

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 $F_j^U(S) - F_j(S)$ 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 3 (left panel) represent these linear asymptotes. Based on the results in proposition 2 we have the following corollary (see §2.5 and Proposition 2 for notation).

Corollary 1 *Consider a mining project with j blocks and let $\mathcal{F}_j(S)$ be the approximation in (24) for some pair (α_k, η_k), $k = 1, \ldots, j$. Let $(\widetilde{R}_k, \widetilde{C}_k, \widetilde{S}_k, \widetilde{M}_k, \widetilde{\alpha}_k, \widetilde{\eta}_k, \widetilde{\theta}_k)$, $k = 1, \ldots, \tilde{\jmath}$ 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,*

$$\mathcal{F}_j(S) = \big(\widetilde{\theta}^{\times}\,\widetilde{R}\big)^+_{\tilde{\jmath}}\, S - \big(\widetilde{\alpha}^{\times}\,\widetilde{C}\big)^+_{\tilde{\jmath}}, \qquad \text{for all } S \geq \widetilde{S}_1/\widetilde{\eta}^{\times}_{\tilde{\jmath}}. \tag{30}$$

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

$$F_j^U(S) = R_j^+\, S - C_j^+, \qquad \text{for all } S \geq \widetilde{S}_1.$$

Similarly, if we set $\alpha_k = e^{-r\,T_k}$ and $\eta_k = e^{(r-\rho)\,T_k}$ we recover the lower bound $F_j^L(S)$ and equation (30) reduces to

$$F_j^L(S) = \mathcal{R}_j\, S - \mathcal{C}_j, \qquad \text{for all } S \geq \widetilde{S}_1\, e^{-(r-\rho)\,T^+_{1,j}}.$$

Proof: The proof of this corollary follows directly from proposition 2 and it is omitted. □

It is interesting to note that $F_j^{\mathrm{L}}(S)$ 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 $F_j^{\mathrm{L}}(S)$ depicted in Figure 3 (left panel).

Corollary 1 suggests a simple approximation for $F_j(S)$ based on these linear asymptotes. Recall from condition (16) that $F_j(S)$ satisfies

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

with $F_0(S) = 0$. As we mentioned before, the difficult part of solving this recursion is computing the expectation $\mathbb{E}^{\mathbb{Q}}[F_{j-1}(S_{T_j}) \big| S_0 = S]$. Since 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 1.

Using the upper bound asymptote $F_{j-1}^{\mathrm{U}}(S) = R_j^+\, S - C_j^+$ and the martingale property (3) we get that $\mathbb{E}^{\mathbb{Q}}[R_j^+\, S - C_j^+ \big| S_0 = S] = R_j^+\, e^{(r-\rho)\,T_j}\, S - C_j^+$. It follows that we can approximate $F_j(S)$ by

$$\widehat{F}_j^{\mathrm{U}}(S) \triangleq \begin{cases} \widehat{M}_j^{\mathrm{U}}\, S^{\beta_1} & \text{if } \ S \le \widehat{S}_j^{\mathrm{U}} \\ \left(R_j + e^{-\rho\,T_j}\, R_{j-1}^+\right) S - \left(C_j + e^{-r\,T_j}\, C_{j-1}^+\right) & \text{if } \ S \ge \widehat{S}_j^{\mathrm{U}}. \end{cases} \tag{31}$$

(We will use a hat '∧' to denote quantities that are derived using the asymptotic approximation.) Using the value matching and smooth pasting conditions we can approximate the values of S_j^* and M_j by

$$\widehat{S}_j^{\mathrm{U}} = \frac{\beta\,\left(C_j + e^{-r\,T_j}\, C_{j-1}^+\right)}{(\beta-1)\,\left(R_j + e^{-\rho\,T_j}\, R_{j-1}^+\right)} \quad \text{and} \quad \widehat{M}_j^{\mathrm{U}} = \left(\frac{C_j + e^{-r\,T_j}\, C_{j-1}^+}{\beta - 1}\right) \left(\widehat{S}_j^{\mathrm{U}}\right)^{-\beta}.$$

Using exactly the same steps, we can get an alternative approximation for $F_j(S)$ based on the lower bound asymptote in Corollary 1.

$$\widehat{F}_j^{\mathrm{L}}(S) \triangleq \begin{cases} \widehat{M}_j^{\mathrm{L}}\, S^{\beta} & \text{if } S \le \widehat{S}_j^{\mathrm{L}} \\ \mathcal{R}_j\, S - \mathcal{C}_j & \text{if } S \ge \widehat{S}_j^{\mathrm{L}}. \end{cases} \tag{32}$$

and

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

Figure 3 (right panel) plots the values of $\hat{F}_j^{\mathrm{U}}(S)$ and $\hat{F}_j^{\mathrm{L}}(S)$ as well as the value function $F_j(S)$ (numerically computed). As we can see, $\widehat{F}_j^{\mathrm{L}}(S)$ performs quite well over the entire range of prices. This in part due to the fact that by construction $\widehat{F}_j^{\mathrm{L}}(S)$ has exactly the same linear behavior than $F_j(S)$ and $F_j^{\mathrm{L}}(S)$ as S goes to infinity.

In an effort to support the conclusions that we have drawn from Figure 3, we conclude this section comparing numerically the performance of the upper bound $F_j^{\mathrm{U}}(S)$, the lower bound $F_j^{\mathrm{L}}(S)$ and the asymptotic approximations $\widehat{F}_j^{\mathrm{U}}(S)$ and $\widehat{F}_j^{\mathrm{L}}(S)$. 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

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

We choose the interval of spot prices $[S_{\min}, S_{\max}]$ large enough so that it includes the entire range of historical spot prices of copper (see Figure 2). In particular, we chose $S_{\min} = 0.6$ [\$/lb] and $S_{\max} = 6$ [\$/lb].

Table 1 presents the value of $\mathcal{P}(\mathcal{F}_j)$ for the four approximations using the data of Example 1. In the far leftmost table 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 far rightmost table we vary the extraction capacity K.

In all three cases, we can see that F_j^{L} and $\widehat{F}_j^{\mathrm{L}}$ have a significantly better performance than F_j^{U} and $\widehat{F}_j^{\mathrm{U}}$. This is consistent with our previous discussion based on Figure 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: *Performance measure ($\mathcal{P}$) for the approximations F^U, F^L, $\widehat{F}^U$ and $\widehat{F}^L$ 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 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 $F_j(S,K)$, $F_j^U(S,K)$ and $F_j^L(S,K)$ be the value function and upper and lower bounds, respectively, for sector 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,

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

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: $\lim_{K\to\infty} F_j^{\mathrm{U}}(S,K) = \lim_{K\to\infty} F_j^{\mathrm{L}}(S,K)$. Finally, since $F_j(S,K)$ is bounded above and below by $F_j^{\mathrm{U}}(S,K)$ and $F_j^{\mathrm{L}}(S,K)$, respectively, the result follows. □

We conclude this section, we 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 $\mathcal{R}_j^{\pi^i} S - \mathcal{C}_j^{\pi^i}$, $i = 1, 2$. 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* $\widehat{F}_i^L(S) = \mathcal{R}_j^{\pi^i} S - \mathcal{C}_j^{\pi^i}$ *be the (lower bound) asymptotic approximation for the value function if sequence* π^i *is used,* $i = 1, 2$. *Then* $\widehat{F}_1^L(S) \geq \widehat{F}_2^L(S)$ *for all* $S \geq 0$ *if and only if the following two conditions are satisfied:*

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

PROOF: See the appendix at the end. □

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 these bounds 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.

With a slight abuse in notation, we will denote by $F_j(S,K)$ 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. In this case, however, K can increase over time so $F_j(S,K)$ is not necessarily equal to the value function of the previous section.

Recall that we have assumed that there is an upper bound $\bar{K}$ on the maximum level of capacity. Hence, we find convenient to define $\bar{F}_j(S) = F_j(S,\bar{K})$. We use a similar convention to denote other quantities that depend on K such as $\bar{W}_j(S) = W_j(S,\bar{K})$, $\bar{R}_j = R_j(\bar{K})$, $\bar{C}_j = C_j(\bar{K})$ and $\bar{T}_j = T_j(\bar{K})$, all defined in equation (12).

Similar to equation (13), the dynamic programming recursion in this case takes the form

$$F_j(S,K) = \sup_{\tau>0,\ K\leq\tilde{K}\leq\bar{K}} \mathbb{E}^{\mathbb{Q}}\left[e^{-r\,\tau} W_j(S_\tau,\tilde{K}) + e^{-r(\tau+T_j(\tilde{K}))} F_{j-1}(S_{\tau+T_j(\tilde{K})},\tilde{K}) - e^{-r\,\tau}\gamma\,(\tilde{K}-K) \,\Big|\, S_0 = S\right], \tag{33}$$

where $F_0(S,K) = 0$ and γ is the is the marginal cost rate of expanding capacity.

We can attempt a solution to (33) in two steps. First, we determine the optimal capacity expansion for a given pair (S, K) and then we determine the optimal extraction time. For a given S and K let us define the auxiliary function

$$G_j(S,K) \triangleq \sup_{K\leq\tilde{K}\leq\bar{K}} \mathbb{E}^{\mathbb{Q}}\left[W_j(S,\tilde{K}) + e^{-r\,T_j(\tilde{K})}\,F_{j-1}(S_{T_j(\tilde{K})},\tilde{K}) - \gamma\,(\tilde{K}-K)\,\middle|\, S_0 = S\right] \tag{34}$$

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. It follows that once $G_j(S,K)$ is determined, the unrestricted value function satisfies

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

Using a similar line of arguments to the one used to derive equation (16) from equation (13), we can show that there exist two functions $M_j(K)$ and $S_j^*(K)$ such that

$$F_j(S,K) = \begin{cases} M_j(K)\,S^\beta & \text{if } S \leq S_j^*(K) \\ G_j(S,K) & \text{if } S \geq S_j^*(K). \end{cases} \tag{35}$$

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

$$\beta\, G(S_j^*(K),K) = G'(S_j^*(K),K)\,S_j^*(K) \qquad \text{and} \qquad M^j(K) = G(S_j^*(K),K)\,\big(S_j^*(K)\big)^{-\beta}.$$

Hence, most of the difficulty of computing the value function $F_j(S,K)$ in (35) boils down to determining the auxiliary function $G_j(S,K)$. Equation (35) also suggests that we only need to compute the value of $G_j(S,K)$ for S sufficiently large, that is, greater than the threshold $S_j^*(K)$. As in the previous section, we will use an asymptotic approximation to estimate $G_j(S,K)$ 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,*

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

where

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

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 (35), we get the following estimates of $S_j^*(K)$ and $M_j(K)$.

$$\begin{aligned} S_j^*(K) &\approx \left(\frac{\beta}{\beta-1}\right)\left(\frac{\bar{\mathcal{C}}_j + \gamma\,(\bar{K}-K)}{\bar{\mathcal{R}}_j}\right) \qquad \text{and,} \\ M_j(K) &\approx \left(\frac{\bar{\mathcal{R}}_j}{\beta}\right)^{\beta}\left(\frac{\beta-1}{\bar{\mathcal{C}}_j + \gamma\,(\bar{K}-K)}\right)^{\beta-1}. \end{aligned} \tag{36}$$

According to these solutions, the threshold price $S_j^*(K)$ is a linear and deceasing 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 characterization of an optimal strategy in this case where capacity expansion are possible, we need to identify a rule that specifies $K_j^*(S, K)$, that is, how the decision maker should expand capacity over time. The optimization in equation (34) 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

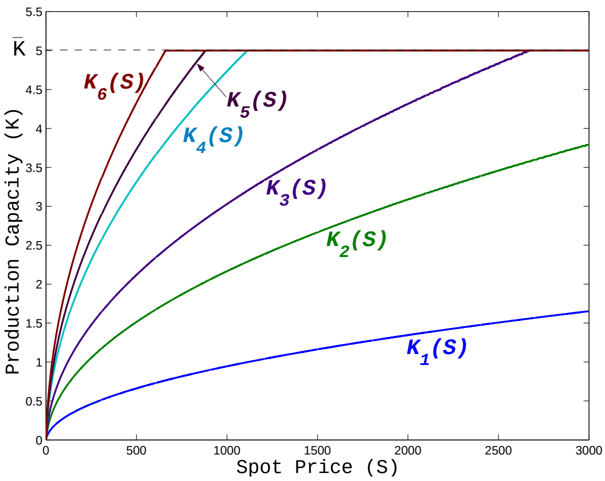
$$\begin{aligned} K_j^*(S,K) &\approx \operatorname*{argmax}_{K \le \tilde{K} \le \bar{K}} \Big\{ \mathcal{R}(\tilde{K})\, S - \mathcal{C}(\tilde{K}) - \gamma\, (\tilde{K} - K) \Big\} \\ &= \max\{K \;;\; \mathcal{K}_j(S)\}, \end{aligned} \tag{37}$$

where

$$\mathcal{K}_j(S) \triangleq \operatorname*{argmax}_{0 \le \tilde{K} \le \bar{K}} \Big\{ \mathcal{R}(\tilde{K})\, S - \mathcal{C}(\tilde{K}) - \gamma\, \tilde{K} \Big\}. \tag{38}$$

The function $\mathcal{K}_j(S)$ represents the optimal capacity expansion if there is no installed capacity and the spot price is S. This is an increasing function of S and we denote by $\bar{S}_j$ the lowest price at which its maximum $\bar{K}$ is achieved. That is, we define $\bar{S}_j \triangleq \inf\{S \ge 0 \;:\; \mathcal{K}_j(S) = \bar{K}\}$. From the decision maker's point of view, this price $\bar{S}_j$ is the threshold above which it is always optimal to expand capacity to its maximum possible level $\bar{K}$ independently of the original level of capacity.

The left panel in Figure 4 plots the values of $\mathcal{K}_j(S)$ for $j = 1, \ldots, 6$ 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 $\bar{S}_6 \le \bar{S}_5 \le \cdots \le \bar{S}_1$.



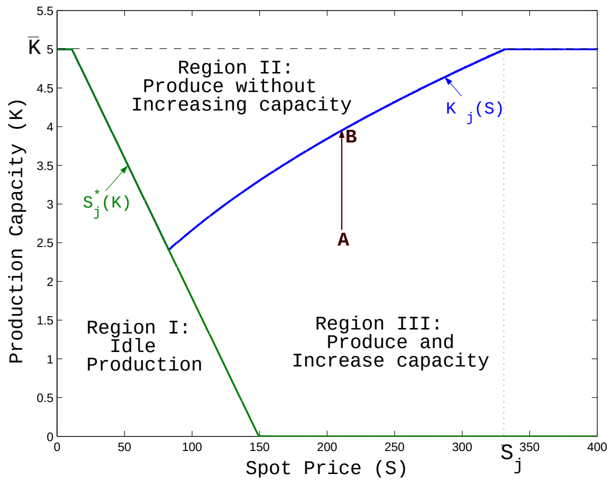


Figure 4: LEFT PANEL: Capacity expansion function $\mathcal{K}_j(S)$ for $j = 1, \ldots 6$ using the data in Example 1, $\gamma = 10$ and a maximum capacity $\bar{K} = 5$. RIGHT PANEL: Optimal production and capacity expansion decisions based on the switching curves $S_j^*(K)$ and $\mathcal{K}_j(S)$ using the data of Example 1, $\gamma = 5$ and $\bar{K} = 5$.

Based on the thresholds $S_j^*(K)$ and $\mathcal{K}_j(S)$ we can divide the state space $\mathcal{S} \triangleq \{(S, K) \;:\; 0 \le K \le \bar{K}$ and $S \ge 0\}$ in three subregions depicted in the right panel in Figure 4. In Region I $\triangleq \{(S, K) \in \mathcal{S} \;:\; 0 \le S \le S_j^*(K)\}$ 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 $\triangleq \{(S, K) \in \mathcal{S} \;:\; S_j^*(K) \le S$ and $\mathcal{K}_j(S) \le K \le \bar{K}\}$ the spot price and capacity are both high and production should start but no capacity expansion is required. Finally, in Region III $\triangleq \{(S,K) \in \mathcal{S} \ : \ S_j^*(K) \le S$ 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 point A in Figure 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 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.

5 Application to a Real Instance

In this section, we use the methodology proposed in the previous sections to estimate the economic value of a mining project at *El Teniente*.

El Teniente, located in the central region of Chile, at 2,500 meters over the sea level, is nowadays the largest underground copper mine in the world. With a processing capacity of almost 47.5 [million tons/year], it produces annually more than 400,000 metric tons of refined copper.

Opened in 1904, this mine have more than 10 active sectors and it is in a continuous expansion with a number of new sectors beginning operations in the following years. One of these sectors, called *El Diablo*, has proven to be rather challenging in term of sequencing design. Because of its unusual spatial distribution, several extraction sequences have been proposed, each one requiring a different economic evaluation. Since *El Diablo* is scheduled to start production in just a few more years, mining operators are pressured to evaluate a large number of potential extraction sequences in a limited amount of time.

In what follows, we show how we can use the methodology proposed in this paper to tackle the sequencing problem for *El Diablo*. Based on the original extraction sequence, we divide the almost 230 million tons of mineral in this sector into ten blocks. Figure 5 shows schematically the spatial distribution of these blocks. Table 2 (left panel) summarizes mineral content, grade and extraction time for the ten blocks in *El Diablo*. 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 for *El Diablo*.

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 these section, we use a simplified method to estimates these extraction costs. If we let $\pi = (\pi_1, \pi_2, \dots, \pi_{10})$ be any of the six sequences that we consider, then the marginal extraction cost (in US$ per lb) for the j^{th} block in this sequence π is

$$A_j^\pi = 4.857 + 0.0162 \cdot d_{\pi_1 \pi_j}$$

where d_{ij} denotes the distance between blocks i and j (see Table 3). In other words, we are modeling the marginal extraction cost of a block as an affine function of the distance from the block to the initial extraction front. The intercept 4.857 and slope 0.0162 were estimated using current production costs at *El Teniente*.

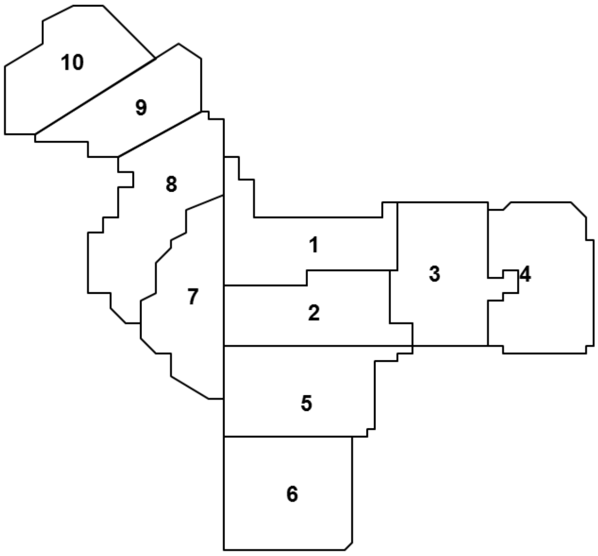


Figure 5: Block spatial distribution at *El Diablo*.

Block	$Q_j[Ton]$	$L_j[\%]$	$T_j[years]$
1	21415510	0.827	2.93
2	21268610	0.915	2.91
3	29526438	0.823	4.04
4	28351480	0.881	3.88
5	24854221	0.845	3.40
6	23931346	0.848	3.28
7	21476937	0.768	2.94
8	26110339	0.727	3.58
9	14913691	0.694	2.04
10	13126378	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 2: LEFT PANEL: Mineral content Q_j, copper grade L_j and extraction time T_j for the ten blocks in *El Diablo*. RIGHT PANEL: Six feasible extraction sequences.

Finally, we considered a fixed production capacity of 7.3 [million tons/year] for this sector. Additional parameters are: a recovery factor of 85%, $r = 12\%$, $\rho = 6\%$, $\sigma = 0.5$.

For each of the six sequences we compute the value function (numerically) and the asymptotic approximation based on the lower bound in equation (32). Table 4 summarizes the results. We conclude that the best extraction sequence (as measured 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 farmost right column in Table 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 1. 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 the six sequences.

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 3: Distance (in meters) among the blocks at *El Diablo*.

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 4: Value function for the six extraction sequences in *El Diablo*. The F columns represent the numerically computed value function and the $\widehat{F}^{\mathrm{L}}$ columns represent the asymptotic approximation using the lower bound in equation (32).

We conclude this section discussing how to use the results of section 4 to estimate an optimal capacity expansion policy for *El Diablo*. Supposing that initially there is no installed capacity we can use equations (36) and (37) 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 (35) and the expansion rule in equation (37). Table 5 summarizes the results.

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]. Since sequence N1 is the one considered in the original design of *El Diablo*, it seems that management at *El Teniente* has misvalued 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$/lb]. This is a rather small value compared to the current spot price which is above 300 [cUS$/lb].

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 5: Optimal capacity K^* and expected value F (numerically computed) for the six sequences of extraction at *El Diablo* as a function of the price S in [cUS$/lb]. The value of K^* is computed using the asymptotic approximation $\widehat{F}^{\text{L}}$ and equation (37).

6 Conclusions and Future Research

In this paper, we have developed a tractable continuous-time model of an underground mining operation and have proposed a methodology to compute near-optimal production and capacity expansion strategies.

On the modeling side, we have represented the mining project as a finite collection of basic cubication units or blocks. These blocks can be arbitrarily defined and differ in terms of ore content, mineralogical composition and extraction costs. In this setting, an optimal production strategy defines the sequence in which blocks should be extracted as well as the timing and rate of extraction. Our discrete block representation of the mine is consistent with current practices 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 practices in the mining industry.

We use a two-step approach to solve the production problem. First, in section 3, we fixed the sequence and production capacity and solved for the optimal timing of extraction contingent upon the evolution of the spot price. In Proposition 1, we derived general properties of an optimal solution and showed that the value of the project is asymptotically equal to an affine function of the price. Unfortunately, for moderate values of the price we were not able to get a simple characterization of the solution. For this reason, we derived in §3.1 and §3.2 upper and lower bounds for the value function, respectively, and used them to propose two simple extraction policies. In addition, we used these bounds in §3.4 to derive a pair of asymptotic approximations. Out of these two approximations, the one derived from the lower bound in equation (32), $\widehat{F}^{\text{L}}(S)$, turned out to be asymptotically equal to the true value function. Moreover, a set of numerical computations reported in Table 1 show that $\widehat{F}^{\text{L}}(S)$ performs extremely well for a wide range of prices and other parameters with an average error of 2%. We concluded section 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 4, we undertook the second step of our solution approach. Here, we showed how to extend the models of the previous section to identify optimal capacity expansion strategies. Our discussion was 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 strategy 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 4). In region I, the spot price is relatively small and the optimal strategy is to idle production. In region II both the price and the capacity are large and so production is running with no increase in capacity. Finally, in region III the spot price is high but capacity is relatively small. The strategy in this case is to increase capacity to a level that depends on the current spot price and then produce. Our analysis provides a set of simple equations that characterize the threshold functions that separate these three regions. As a general rule, we noted that as the capacity of a project increases the option to idling 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 paper, in section 5, with an application of our methodology to a real instance of the problem at *El Teniente*. The example is based on a specific project, called *El Diablo*, that is scheduled to start production in few years. Management at *El Teniente* has proposed a preliminary extraction sequence (N1 in Table 2). Our analysis showed that this sequence is optimal if production capacity is fixed at its nominal value of 7.3 (million tons/year). However, if we allow the capacity to be optimally chosen then it turns out that sequence N5 maximizes the economic value of *El Diablo*.

There are a number of possible extensions to our model. First of all, an important component of an optimal production strategy is the sequence in which blocks are extracted. In this paper, we did not handle explicitly the question of how to dynamically choose this sequence. Instead, we took a scenario-based approach and assumed 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 noted in section 2.4, the problem of dynamically adjusting the sequence of extraction has a combinatorial structure which makes it extremely hard to solve. We think that extending our methodology to explicitly include this dynamic sequencing problem is a challenging research problem that is important from both a 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) 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 the case for a company like CODELCO that produces 18% of the world's copper production. In this situation, we should expect some correlation between output and spot (and futures) price trajectories (see Pincheira (1999)). This type of *large investor* effect has received some attention in the mathematical finance literature (*e.g.* Cuoco and Cvitanić 1998, Frey 1998) but it seems to have been overlooked in the operational context of real options.

Acknowledgment This research has been in part supported by FONDEF grant #D03I1064 and CODELCO.

Appendix

1. Proof of Proposition 1

Let us first prove the monotonicity and convexity of $F_j(S)$ using induction over j. For $j = 1$ these two properties follow directly from equation (17). So let us assume that $F_{j-1}(S)$ is increasing and convex in S for some $j \geq 2$ and let us prove that $F_j(S)$ is also increasing and convex.

Equation (16) together with the fact that $W_j(S) = R_j\, S - C_j$ and $S_t = S_0 \exp\Big((r - \rho - \frac{1}{2}\sigma^2)\, t + \sigma\, \tilde{B}_t\Big)$ (this follows after integrating (4)) imply that $F_j(S)$ satisfies

$$F_j(S) = \begin{cases} M_j\, S^{\beta} & \text{if } S \leq S_j^* \\ R_j\, S - C_j + e^{-r\,T_j}\, \mathbb{E}^{\mathbb{Q}}[F_{j-1}(S \exp\Big((r - \rho - \frac{1}{2}\sigma^2)\, T_j + \sigma\, \tilde{B}_{T_j}\Big))] & \text{if } S \geq S_j^*. \end{cases}$$

The constants M_j and S_j^* are chosen so that $F_j(S)$ is continuous and differentiable at S_j^*. Since $\beta \geq 1$ and $F_{j-1}(S)$ is increasing and convex it follows that $F_j(S)$ is also increasing and convex.

We now prove the linear asymptotic behavior of $F_j(S)$. In order to do so, 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

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

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

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

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 $\{S_k^* \,:\, 1 \leq k \leq j\}$ to control production.

It is not hard to see that $\mathcal{S}_t \geq S_t$ pathwise, hence it follows that $F_j(S) \leq \mathcal{F}_j(S)$ 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, $\mathcal{S}_{T_{k,j}^+} \geq S_k^*$ (a.s.) for all $1 \leq k \leq j$. Therefore, we have that

$$\begin{aligned} F_j(S) \leq \mathcal{F}_j(S) \;\; &= \;\; \sum_{k=1}^{j} e^{-r\,T_{k,j}^+}\, \big(R_k\, \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{\mathcal{T}_{k,j}} | S_0 = S] - C_k\big) \\ &= \;\; \mathcal{R}_j\, S - \mathcal{C}_j + \sum_{k=1}^{j} e^{-r\,T_{k,j}^+}\, R_k\, \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T_{k,j}^+} - S_{T_{k,j}^+} | S_0 = S]. \end{aligned} \quad \text{(A2)}$$

Combining (A1) and (A2), we get that

$$0 \leq F_j(S) - \big(\mathcal{R}_j\, S - \mathcal{C}_j\big) \leq \sum_{k=1}^{j} e^{-r\,T_{k,j}^+}\, R_k\, \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T_{k,j}^+} - S_{T_{k,j}^+} | S_0 = S].$$

To complete the proof, we need to show that the term on the right is bounded above the expression in right-hand side in (19) for S sufficiently large. In order to see this, we first note that by the definition of $\mathcal{S}_t$ we have

$$\mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T^+_{k,j}} - S_{T^+_{k,j}} | S_0 = S] = \sum_{n=k}^{j} \mathbb{E}^{\mathbb{Q}} \left[(S^*_k - \mathcal{S}_{T^+_{k,j}-})^+ | S_0 = S \right] \leq \sum_{n=k}^{j} S^*_k \, \mathbb{Q}(\mathcal{S}_{T^+_{k,j}-} \leq S^*_k | S_0 = S)$$
$$\leq \sum_{n=k}^{j} S^*_k \, \mathbb{Q}(S_{T^+_{k,j}} \leq S^*_k | S_0 = S),$$

where the last inequality uses the (a.s.) facts that $\mathcal{S}_t \geq S_t$ 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 (4)). So, it is not hard to show that

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

where $\tilde{B}_t$ is a standard $\mathbb{Q}$-Brownian motion. Hence, for $S \geq S^m_j \exp\left(- \min_k \left\{ (r - \rho - \frac{\sigma^2}{2}) \, T^+_{k,j} \right\}\right)$, we can bound the tail probability by (*e.g.*, Asmussen 2003, Theorem XIII-2.1)

$$\mathbb{Q} \left(\tilde{B}_{T^+_{k,j}} \geq \frac{1}{\sigma} \left[\ln \left(\frac{S}{S^*_k} \right) + (r - \rho - \frac{\sigma^2}{2}) \, T^+_{k,j} \right] \right) \leq \exp \left(-\frac{1}{\sigma^2} \left[\ln \left(\frac{S}{S^*_k} \right) + (r - \rho - \frac{\sigma^2}{2}) \, T^+_{k,j} \right]^2 \right)$$
$$\leq \exp \left(-\frac{1}{\sigma^2} \min_{1 \leq k \leq j} \left\{ \left[\ln \left(\frac{S}{S^m_j} \right) + (r - \rho - \frac{\sigma^2}{2}) \, T^+_{k,j} \right]^2 \right\} \right).$$

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

$$\sum_{k=1}^{j} e^{-r \, T^+_{k,j}} \, R_k \, \mathbb{E}^{\mathbb{Q}}[\mathcal{S}_{T^+_{k,j}} - S_{T^+_{k,j}} | S_0 = S] \leq S^m_j \, R^m_j \, \frac{j \, (j+1)}{2} \exp \left(-\frac{1}{\sigma^2} \min_{1 \leq k \leq j} \left\{ \left[\ln \left(\frac{S}{S^m_j} \right) + (r - \rho - \frac{\sigma^2}{2}) \, T^+_{k,j} \right]^2 \right\} \right),$$

which competes the proof. □

2. Proof of Propositions 2

The following intermediate result will be used at some point during the proof.

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

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

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

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

Proof Lemma 1: The proof follows using backward induction over $k = j-1, j-2, \ldots$ 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 (28) and (29). 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 '$\sim$' 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

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

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 over 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 (28) and the approximation $\mathcal{F}_j(S)$ satisfies equation (29), that is,

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

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

- For $j = 1$ the result follows directly from equation (25).
- 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 (28) 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 (28) 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 $\max\left\{0 \leq k \leq j-1 \,|\, \mathcal{S}_k \geq \eta^\times_{k,j-1} \, (\eta_j \, \mathcal{S}_j)\right\}$ or equivalently $\max\left\{0 \leq k \leq j-1 \,|\, \mathcal{S}_k \geq \eta^\times_{k,j} \, \mathcal{S}_j\right\}$. This is $\hat{k}$ by definition and we get that

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

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

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

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

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

The inequality contradicts (A6) and we conclude that $\mathcal{S}_{j-1} \geq \eta_j \, \mathcal{S}_j$ 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

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

proving equation (28) 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 (29), 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 $(R_k, C_k, \alpha_k, \eta_k)$, $k = 1, \dots, j$, the values of $\mathcal{F}_j(S)$ and $\mathcal{S}_j$ are given by equations (28) and (29). 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 over the number of blocks, j.

- For $j = 1$ the result follows directly from equation (25).
- 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 what we need to know about the characteristics of blocks $\{j-1, j_2, \dots, 1\}$ to compute $\mathcal{F}_j(S)$ is contained in $\mathcal{F}_{j-1}(S)$. Hence, we can assume without lost of generality that the sequence of blocks $\{j-1, j_2, \dots, 1\}$ does satisfy condition (A3), that is,

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

If this condition is also satisfied for block j then the entire sequence satisfied condition (25) 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 remaining of this proof, we will apply the algorithm in Proposition 2 to a sequence of block 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 (28) and (29).

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 pooled together blocks $\tilde{h}+1, \tilde{h}+2, \dots, j$ 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

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

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

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

Therefore, after the first iteration the algorithm will stop. Using this modified sequence, equation (28) 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 (24).

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

Recall that the value of $\mathcal{F}_{j-1}\big(S\big)$ is known by the induction hypothesis and it is given by equation (29). This induction hypothesis, and in particular condition (A8) imply that

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

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 implies and equation (29) imply that

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

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

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

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

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

Hence, $\bar{h} = \tilde{h}$, which proves that the value of $\mathcal{S}_j$ derived in the proposition 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 $\mathcal{F}_(S)$ are exactly those reported in equations (28) and (29). □

3. Proof of Proposition 4

The proof follows directly from equation (32). Indeed, suppose first that $F^{\mathrm{L}}_1(S) \ge F^{\mathrm{L}}_2(S)$ 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 (32) implies that $M^{\pi_1}_j \ge M^{\pi_2}_j$. This is equivalent to

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

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

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

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

Then, equation (32) immediately implies that $F^{\mathrm{L}}_1(S) \ge F^{\mathrm{L}}_2(S)$ for S sufficiently small and sufficiently large. Finally, the inequality extends to all $S \ge 0$ by the convexity of $F^{\mathrm{L}}_1(S)$ and $F^{\mathrm{L}}_2(S)$. □

References

Alford, C., M. Brazil, D. Lee. 2005. Optimisation on underground mining. Working Paper, Forthcoming on the Handbook on Operations Research in Natural Resources to be published by Kluwer.

Asmussen, S. 2003. *Applied Probabaility and Queues*. Springer-Verlag, New York.

Baker, T.E., L.S. Ladson. 1985. Successive linear programming at exxon. *Management Science* **31** 264–274.

Bertsekas, D. 1995. *Dynamic Programming and Optimal Control (volume I).*. Athena Scientific, Massachusetts.

Black, F., M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* **81** 637–659.

Brennan, M., E. Schwartz. 1985. Evaluating natural resources investments. *Journal of Business* **58** 135–157.

Brennan, M., L. Trigeorgis. 2000. *Optimal Extraction of Nonrenewable Resources When Costs Cumulate*. Oxford University Press, NY.

Caldentey, R., S. Mondschein. 2003. Policy model for pollution control in the copper industry, including a model for the sulfuric acid market. *Operations Research* **51**(1) 1–16.

Casassus, J., P. Collin-Dufresne. 2005. Stochastic convenience yield implied from commodity futures and interest rates. *The Journal of Finance* **LX**(5) 2283–2331.

Cochilco. 1976. *The Copper Manual*. Chilean Copper Commission and Chilean National Copper Corporation, Santiago, Chile.

Cuoco, D., J. Cvitanić. 1998. Optimal consumption choices for a "large" investor. *Journal of Economic Dynamics and Control* **22** 401–436.

Dixit, A., R. Pindyck. 1994. *Investment Under Uncertainties*. Princeton University Press, NJ.

Dyer, J.S., R. Lund, J.B. Larsen, R.P. Leone. 1990. A decision support system for prioritizing oil and gas exploration activities. *Operations Research* **38**(3) 386–396.

Engel, E., R. Valdés. 1990. Prediciendo el precio del cobre: Más allá del camino aleatorio? Working Paper (In Spanish).

Epstein, R., R. Morales, J. Seron, A. Weintraub. 1999. Or models in the chilean forest industry. *Interfaces* **43**(1) 64–79.

Frey, R. 1998. Perfect option hedging for a large trader. *Finance and Stochastics* **2** 115–148.

Gersovitz, M., C.H. Paxson. 1990. The economics of africa and the prices of their exports. Tech. Rep. 68, Princeton University, Princeton, NJ.

Gibson, R., E. Schwartz. 1990. Stochastic convenience yield and the pricing of oil contingent clams. *The Journal of Finance* **45** 959–976.

Hull, C. 1993. *Options, Futures and Other Derivative Securities.*. Prentice-Hall, International, New Jersey.

Lane, K. 1988. *The Economic Definition of Ore: Cut-Off Grades in Theory and Practice*. Mining Journal Books Limited.

Lumley, R., M. Zervos. 2001. A model for investing in the natural resource industry with switching costs. *Mathematics of Operations Research* **26** 637–653.

McDonald, R., D.R. Siegel. 1985. Investment and the valuation of firms when there is an option to shut down. *International Economic Review* **26** 331–349.

Merton, R.C. 1973. The theory of rational option pricing. *Bell Journal of Economics and Management Science* **4** 141–183.

Mondschein, S., A. Schilkrut. 1997. Optimal investment policies for pollution control in the copper industry. *Interfaces* **27**(6) 69–87.

Myers, S.C. 1987. Finance theory and financial strategy. *Midland Corporate Finance Journal* **5**(1) 6–13.

Øksendal, B. 2003. *Stochastic Differential Equations*. Springer-Verlag NewYork Inc., New York.

Pincheira, M. 1999. The price-stock relationship in the copper market, a surprising approach. *Anales del Seminario Copper 99* **1** 135–145.

Pindyck, R. 1978. The optimal exploitation and production of nonrenewable resources. *Journal of Political Economy* 841–862.

Pindyck, R. 1981. The optimal production of an exhaustible resource when price is exogenous and stochastic. *Scandinavian Journal of Economics* 277–288.

Samis, M., D. Laughton, R. Polulin. 2001. Valuing a multi-zone mine as a real asset portfolio a modern asset pricing (real options) approach. *Corporate Finance Abstracts: Valuation, Capital Budgeting and Investment Policy Working Paper Series* **4-29**.

Samuelson, P.A. 2001. Rational theory of warrant pricing. *Industrial Management Review* **6**.

Santibañez, P. 2000. Metodología para la evaluación y optimización de planes mineros de largo plazo. Trabajo de Memoria de Título, Departamento de Ingeniería Industrial, Universidad de Chile (In Spanish).

Schwartz, E. 1997. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance* **52** 923–973.

Schwartz, E., J. Cassasus, G. Cortazar. 2001. Optimal exploration investments under price and geological-technical uncertainty: A real options model. *R & D Management* **31** 181–189.

Schwartz, E., J. Smith. 2000. Short term variations and long-term dynamics in commodity prices. *Management Science* **46** 893–911.

Shreve, S.E. 2004. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Verlag, New York.

Smith, J.E., K.F. McCardle. 1999. Optionsin the real world: Lessons laerned in evaluating oil and gas investments. *Management Science* **47** 1–15.

Trigeoris, L. 1996. *Real Options: managerial Flexibility and Strategy in Resource Allocation*. MIT Press, Cambridge, MA.