Optimal environmental policy for a mine under polluting waste rocks and stock pollution*

Pauli Lappi† and Markku Ollikainen‡

Abstract

This study analyzes socially optimal environmental policy for a mine with a model that takes into account waste or waste rock production and abatement possibilities of the mine. We develop a model, in which the mine produces an externality related to the waste rocks, such as acid mine drainage. We find that the extraction rate tends to be lower in a mine with higher waste rock production, and that the optimal tax on the waste rock production is strictly increasing in time. We extend the model to incorporate an additional externality in the form of a stock pollutant. We analyze the optimal taxes and show that the typical result that the time path of the tax on the stock pollutant is inverted U-shaped may be lost in a mine model with abatement possibility and fixed operation period.

Keywords: Emission taxation; Exhaustible resources; Mining; Optimal control; Stock pollution; Waste rock.

JEL codes: Q30, Q38, Q50, Q58.

1 Introduction

Mining has a large impact on the environment. Clearing land for an open-pit changes dramatically the landscape and this impact is reinforced by piles of waste rocks. Mining is also a source of local and global pollutants. Mining activities are typically regulated:

* **Acknowledgements:** We would like to thank Olli Tahvonen and Anna-Kaisa Kosenius for useful discussions. We thank the referee, the editor, Karen Pittel and other EAERE2016 participants for helpful comments.

† Corresponding author. University of Helsinki, Department of Economics and Management, Latokartanonkaari 5, PO Box 00014; e-mail pauli.lappi@helsinki.fi.

‡ University of Helsinki, Department of Economics and Management.

Opening a mine requires environmental impact assessment and authorities design the environmental permit of mine. When designing regulation authorities search for a balance between economically sound business and environmental and social impacts. Developing and maintaining economically efficient and sustainable mining sector is important for example for the European Union, which has an under-developed mining sector despite of the fact the EU is very dependent on imported raw materials (Tiess, 2010).

Mines are typically regulated in a detailed way, for instance, through environmental permits, which contain technological requirements, limits to various pollutants, rights to use water and requirement to abate waste water, requirements on placing waste rocks and many other issues. Current environmental regulation seldom is based on economic insights or analysis. And as a contrast, economic studies on mines typically do not account for the details of environmental issues related to mines. What is needed, is optimal dynamic environmental policy analysis that internalizes environmental impacts and allows for a simultaneously maximal profits for the mining firms. This study introduces in the analysis some factors characteristic to mining activities, which are not present in the exploitation of some other exhaustible resources. Namely, we extend the previous models by adding polluting waste rocks and their disposal to the analysis.

A large body of literature exists on exhaustible resources, stock externalities and regulation. For example, Ulph and Ulph (1994) study the role of a carbon tax under stock externality and exhaustible polluting resource. In particular, they study the time path of the carbon tax, and find in a special case that the tax is first rising and then declining. Their analysis is generalized by Hoel and Kverndokk (1996) and Tahvonen (1997). Also Farzin (1996) analyzes exhaustible resources, stock externalities and taxation. While in Tahvonen (1997) only nature abates pollution stock, Farzin allows for human abatement but he excludes the natural cleaning process. This literature has then expanded to cover the existence of Green Paradox (Sinn, 2008; van der Ploeg and Withagen, 2012;

van der Meijden, van der Ploeg, and Withagen, 2015), the analysis of backstop technologies (Tahvonen, 1997; Tsur and Zemel, 2003; Michielsen, 2014), the more detailed description of the dynamics of carbon stock in the atmosphere (Farzin and Tahvonen, 1996; Amigues and Moreaux, 2013; Winter, 2014) and the optimality with carbon capture and storage (Moreaux and Withagen, 2015).

What is known on environmental taxation from the above stylized infinite-horizon models is that, under certain simplifying assumptions, the optimal emission tax is the negative of the shadow value of pollution, and that the time path of the tax first increases and then decreases. These results bear important message to mining models, as well. But in contrast to infinite-horizon models, mining is characterized by a finite planning horizon and mine specific features, different to many other exhaustible resources. Indeed, the properties of environmental policy in the mining industry are a largely unstudied issue. Few exceptions can be found, though. Roan and Martin (1996) study mining under command and control policy but they do not study the optimal policy like we will do in this paper. Also, Farzin (1996) and Stollery (1985) have examined environmental policy for a mine. However, Stollery studies flow externality only, when our focus is on stock externalities. Finally, Sullivan and Amacher (2009) studied mine land restoration and the divergence between the costs resulting from the choices of the mining firm compared to the allocation that is socially optimal and the ways to improve the applied bond system.[1]

We examine the optimal policy for a mine when extraction creates polluting waste rocks, producing metals from the extracted ore is a source of emissions and abatement is made by the mining firm and the nature. Production of waste rocks is an essential ingredient of a mining model, because mining operations create many kinds of waste, such as topsoil, waste rocks and tailings.[2] The quantities of waste rocks produced in mines are

[1]White et al. (2012) analyze the optimal mine rehabilitation and Pigovian taxation using a discrete-time model and apply it to Australian context.

[2]Lottermoser (2010) explains, that to reach the ore vein and to obtain eventually the valuable material,

large and their replacement is costly. In Australia the amount of produced waste rock exceeds the amount of mined ore and for example for gold and copper the waste rock to ore ratio is rising over time (Mudd, 2010). Mining industry in Canada produced in 2008 about 473 Mt of waste rock and tailings (Statistics Canada, 2012). In Finland the ratio between waste rocks and ore was about 1.5 in the year 2014.[3] Globally, Lottermoser (2010) approximates that in 2006 the quantity of non-fuel mineral commodities produced was 5 800 Mt and the total production of waste rocks and tailings was 20 000-25 000 Mt.

In this paper we use the term waste rocks to include waste rock and tailings from the mining operations. Proper disposal of waste rocks incurs increased costs, which we include in the model.[4] Often, especially in the case of open-pit mines, the waste rocks are stored in the mine area.[5] The size of the waste rock stock over time depends on the amount of extracted material and the amount of waste rocks disposed. The waste rock stock is a source of a stock pollution, which is caused by acid mine drainage. Acid mine drainage (AMD) is one of the most serious environmental problems related to mining (Dold, 2014). In short, AMD means the flow of potentially toxic or harmful substances that are separated from wastes by (often) acidic waters. These acidic waters themselves are a product of oxidation of sulfide minerals, such as pyrite. (Lottermoser, 2010; Nordstrom, 2011; Dold, 2014; Lèbre, Corder, and Golev, 2017)

We analyze also another externality in addition to AMD: Valuable metals are produced from extracted material using production process, that releases harmful materials from

say metal, the topsoil and waste rocks must be first removed. Then the ore is mined and processed for example in a mill from which additional mine wastes called tailings are produced. Afterwards further processing is still needed. See Lottermoser (2010) for additional information.

[3]More precisely, the mines in Finland produced 29 544 560 tons of ore (plus valuable stones) and 45 299 631 tons of waste rocks in 2014 (Tukes, 2014).

[4]In this paper disposing the waste rocks is always costly, although we acknowledge that some types of waste rocks can include valuable side products for the mine.

[5]In underground mining the waste rocks are often used to fill old mining tunnels, but presumably not without cost.

the extracted material.[6] These emissions can be abated both by the mining firm and by natural processes, as in practice both abatement processes exist simultaneously, but their interplay has not been examined simultaneously in mining context this far.[7] In our model pollution can be for example discharges to the local water body.

We pose two sets of research questions. First, what are the effects on the extraction rate of adding polluting waste rocks and their disposal to a mining model? What is the socially optimal way for to handle the waste rocks? What kind of time path does optimal tax on waste rock production have? These questions are answered in a model with waste rock production, in which the regulator maximizes the total discounted welfare. This model allows us to examine the evolution of waste rock disposal and waste rock stock and the effect of modeling waste rocks on the extraction rate. The model is then extended to cover an additional externality. Given that the mine pollutes, the firm may abate and nature abates in all circumstances, how should the regulator design the optimal dynamic emission tax to control pollution? This extension to the model outlines the optimal design of environmental policies under two different externalities calling for two instruments and facilitates the analysis of the first-best choice of extraction, waste rock disposal and abatement, and simultaneously, we produce understanding of the optimal dynamic emission tax. In this extension we also examine the extraction paths of two mines with different waste rock production and emission coefficients, when the regulator is unable to apply optimal taxes, but is forced to use time invariant taxes.

To anticipate the main results, we show in the basic model that while the rate of extraction decreases on the interval where it is positive, the rate of waste rock disposal always increases. The waste rock stock is strictly increasing or has a single peak value, so

[6] Hence the paper is related also to the literature with multiple pollutants, which includes Moslener and Requate (2007) and Ambec and Coria (2013) among others.

[7] An exception is carbon capture and storage, which is essentially abatement (Moreaux and Withagen, 2015). Compared to our model, they have and infinite time horizon model. In addition to this, our model has two externalities and is related specifically to mining.

that the optimal stock is either increased throughout the planning interval or is increased to its highest value and then decreased. This implies that it is not economically optimal to dispose of waste rocks immediately after they have been produced. An intuition for this is that it is optimal to move a part of the waste rock disposal cost to the more distant future due to discounting of the costs. The cost of replacing waste rocks and eventual damages from the left-over waste rock stock impacts also the extraction rate. The extraction rate is either the same or lower in a mine with high waste rock production rate than in an identical mine with a lower waste rock production rate. In the extension of the model, we analyze the optimal taxes. We show that the optimal tax on waste rock production is strictly increasing and outline three possibilities for the dynamically optimal emission tax. First, the emission tax may be qualitatively similar to the ones found in the previous literature, that is, it may have an inverted U-shaped time path. Second, the tax may be always strictly increasing. Third, the tax may increase at the beginning and at the end, but be decreasing during the middle part of the operation period. However, without further structure on the model we are unable to say, which of these time paths is actually optimal. Regarding the extraction paths under time invariant taxes on waste rock production and emissions, we show that a mine with a lower marginal tax burden is exhausted before a mine with higher marginal tax burden. This implies, that the mine with a lower marginal tax burden begins its extraction path at a higher level.

We continue by outlining the model set-up and assumptions regarding the technology and preferences. The regulator's problem with the waste rocks is studied next. We then add the second pollution stock in the model and characterize the optimal taxes and offer intuitive explanations for the results. The final section concludes our analysis.

2 Model

The planning interval is $[0, T]$. We denote by x_0 the amount of extractable material in the mine at time zero, and we let $q(t)$ be the extraction rate. We assume that the ore body is homogenous. The amount of material left in the mine at time t is defined by equation

$$X(t) = x_0 - \int_0^t q(\tau)\ \mathrm{d}\tau, \tag{1}$$

so that $\dot{X} = -q$.[8] The extraction cost is captured by function C, with $C'(q) > 0$ and $C''(q) > 0$. Let $S(t)$ be the size of the waste rock stock at time t, and let $z(t)$ be the waste rock disposal rate. Suppose $S(0) = 0$ and that a unit of extraction produces α units of waste rocks. Then

$$S(t) = \int_0^t \alpha q(\tau) - z(\tau)\ \mathrm{d}\tau, \tag{2}$$

and hence $\dot{S} = \alpha q - z$. We let the cost function of waste rock disposal be C^s, with value $C^s(z)$, and assume that $C^{s\prime}(z) > 0$ and $C^{s\prime\prime}(z) > 0$ for $z > 0$. Section 3 analyzes the optimal policy for a mine that produces waste rocks, which cause pollution. The related damage function $\hat{D}$, with value $\hat{D}(S(T))$ satisfies properties $\hat{D}' > 0$ and $\hat{D}'' > 0$ for $S(T) > 0$. Hence the stock causes damages only after the mine has been shut down. This assumption is admittedly a simplification, but it is based on natural sciences: Dold (2014) argues that the chemical processes leading to AMD formation begin after the production stage of the mine given that the mine is properly managed. In particular, the active tailings imboundment should not cause AMD, if the tailings dam is built properly and if the tailings imboundment is kept water saturated.[9] This is the main reason for assuming that the waste rock stock causes damages after the production stage.[10]

[8] We often drop the time argument from the notation.

[9] According to Dold (2014, page 635), the AMD formation takes several years even after the mine shut-down.

[10] If we would assume that acid mine drainage causes damages during the production stage, then many of the paper's results would change. For example, the exponential development of the shadow value of the

The optimal emission taxation of the stock pollutant is analyzed in Section 4. The size of the pollution stock is $N(t)$ and it represents for example the level of turbidity of the local water bodies or the toxic discharges from gold or copper mines such as arsenic. We denote by parameter δ the rate at which the pollution stock decreases by natural processes. The rate of change in the pollution stock is positively related to the extraction rate and negatively to the size of the pollution stock. If the firm also abates at rate $a(t)$, the time derivative of the pollution stock is given by

$$\dot{N}(t) = \beta q(t) - a(t) - \delta N(t). \tag{3}$$

Abatement is costly, so we let the abatement cost function be A, with value $A(a)$. We assume that $A'(a) > 0$ and $A''(a) > 0$ for $a > 0$. Equation (3) is similar to the one in Farzin (1996) except that we allow natural processes to clean the environment. The pollution stock causes damages. We assume that they are captured by a function D, with $D'(N) > 0$ and $D''(N) > 0$ for $N > 0$. We do not allow the damages depend on any flow externality, such as noise and odorous emissions, although these may be important local impacts of mining.

Literature contains different assumptions about the price of the extracted good. For example Farzin (1992) assumes, that the price is some (regular enough) function of time. Others, like Caputo (1990), give less structure to price by assuming that the price is constant throughout the planning interval.[11] Our modeling choice is to use a constant price. In addition, we make the following assumption:

Assumption 1.

waste rock stock plays a crucial role in the proof of propositions 4, 5, 12 and 13. If the waste rock stock causes damages during the production stage, the shadow value would not develop exponentially, which might imply that these propositions would not hold in this case. Similarly, if the waste rock production coefficient α depends on the remaining stock and $\alpha'(X) < 0$ (exploitation implies higher amount of waste rocks per extracted unit), then the results would change since the shadow value of the resource stock would not develop exponentially.

[11] Also Roan and Martin (1996) make this assumption although they are implicit about it.

(*i*) $p - C'(0) > 0$;

(*ii*) ${C^s}'(0) = 0$ and $A'(0) = 0$;

(*iii*) $D'(0) = 0$ and $\hat{D}'(0) = 0$.

Assumptions $p - C'(0) > 0$, ${C^s}'(0) = 0$ and $A'(0) = 0$ ensure that the rate of extraction is positive at some point in time, and assumption is $D'(0) = 0$ is used to show among other things that it is not optimal to keep the environment in the pristine state, which we define as the state where pollution stock is zero. Assumptions in (*iii*) mean that the first unit of pollution causes approximately zero damages.

3 Regulation under polluting waste rock production

This section analyzes the optimal extraction and waste rock disposal choice of the regulator when the pollution stock is absent. We attach a scrap value to the terminal waste rock stock, which describes the damages from the left-over waste rocks after the mine has been shut down. We assume following Caputo and Wilen (1995) that the terminal time T is fixed.[12] The regulator aims to find the extraction and waste rock disposal rates that maximize the difference between the total discounted value of the instantaneous profit and the scrap value subject to constraints, that is, the regulator's problem is to

$$\max_{\{q(t),z(t)\}} \int_0^T e^{-\rho t}\left[pq(t) - C(q(t)) - C^s(z(t))\right]\,\mathrm{d}t - \int_T^\infty e^{-\rho t}\hat{D}(S(T))\,\mathrm{d}t \tag{4}$$

$$\text{s.t}\quad \dot{X}(t) = -q(t), \quad X(0) = x_0, \quad X(T) \geq 0, \tag{5}$$

$$\dot{S}(t) = \alpha q(t) - z(t), \quad S(0) = 0, \quad S(T) \geq 0, \tag{6}$$

$$q(t) \geq 0, \quad z(t) \geq 0. \tag{7}$$

[12]Note that the fixed horizon does not exclude the physical or economical exhaustion of the mine. The choice of fixed horizon simplifies the analysis. If T is a choice variable, then the problem may not have a solution, since increasing T postpones the damages from AMD. This postponement is the benefit of waiting, and it may not have a corresponding cost unless there are fixed costs in production.

We simplify the notation by defining

$$G(S(T)) := \int_T^\infty e^{-\rho t} \hat{D}(S(T)) \, dt. \tag{8}$$

Note that $G' > 0$ for $S(T) > 0$ and $G'(0) = 0$. This model is a modification of the "conventional mining model", in which the waste rocks, the waste rock disposal and the related costs are absent. Our goal is to analyze the consequences of adding the waste rocks to the model by comparing the extraction rates under different waste rock production coefficients α. We also analyze the time paths of the waste rock stock and of the tax on the waste rock production rate $\alpha q - z$.

The current value Hamiltonian related to this problem is

$$H(q, z, X, S, \lambda, \mu) = pq - C(q) - C^s(z) - \lambda q + \mu(\alpha q - z), \tag{9}$$

where λ is the shadow value of resource on the ground and μ is the shadow value of waste rock stock. The Hamiltonian measures the total net profit at a given instant of time. Applying the Maximum Principle gives us conditions that are necessary for optimality:[13]

$$H_q = p - C'(q) - \lambda + \mu\alpha \leq 0, \quad q \geq 0, \quad qH_q = 0, \tag{10}$$

$$H_z = -C^{s\prime}(z) - \mu \leq 0, \quad z \geq 0, \quad zH_z = 0, \tag{11}$$

$$\dot{X} = -q, \tag{12}$$

$$\dot{S} = \alpha q - z, \tag{13}$$

$$\dot{\lambda} = \rho\lambda, \tag{14}$$

$$\dot{\mu} = \rho\mu, \tag{15}$$

$$\lambda(T) \geq 0, \; X(T) \geq 0, \; \lambda(T)X(T) = 0, \tag{16}$$

$$\mu(T) + G'(S(T)) \geq 0, \; S(T) \geq 0, \; [\mu(T) + G'(S(T))]S(T) = 0. \tag{17}$$

[13]Theorem 5 in Chapter 3 of Seierstad and Sydsæter (1987).

We obtain from (14) and (15) that $\lambda(t) = \lambda(0)e^{\rho t}$ and $\mu(t) = \mu(0)e^{\rho t}$. The value of λ at time t is positive or zero.[14] Variable μ is the shadow value of the disposable waste rocks that have been extracted: It is approximately the change in the regulator's optimal value function from a unit (jump) increase in the waste rock stock. We will show, that it is strictly negative. This implies that, at optimum, when the rate of waste rock disposal is positive, the marginal cost of waste rock disposal equals the marginal benefit of having one unit of waste rock less on the stock.

Proposition 1 presents some properties of the optimal controls in our mining model.

Proposition 1.

(i) *The rate of extraction and the rate of rock removal are continuous in time.*

(ii) *The rate of extraction is positive on some interval* $[0, t_1]$ *with* $t_1 \leq T$ *and is strictly decreasing on the interval* $[0, t_1]$.

(iii) *The waste rock disposal rate is positive and strictly increasing over the whole planning interval, and the shadow value of the waste rock stock is strictly negative.*

(iv) *Whenever the rate of extraction is positive, the instantaneous marginal profit of extraction increases at the rate of interest.*

Proof. See Appendix A.1. □

The key results of Proposition 1 are that the rate of extraction declines in time (when extraction is positive) and that the rate of waste rock disposal increases in time. That the waste rock disposal always increases by (iii), is intuitive. The regulator needs to restore some of the waste rocks by the end of the mining. Because the disposal is costly, and the future costs are discounted, it is optimal to start the disposal at a low disposal rate and

[14] By $0 \leq \lambda(T) = \lambda(0)e^{\rho T}$, $\lambda(0) \geq 0$, so that $\lambda(t) \geq 0$.

to increase the rate until the end. When the controls are positive, we obtain in the case of differentiable controls that

$$\dot{q} = \frac{-\rho\lambda + \rho\alpha\mu}{C''(q)} < 0, \tag{18}$$

$$\dot{z} = -\frac{\rho\mu}{C^{s\prime\prime}(z)} > 0, \tag{19}$$

confirming the properties (ii) and (iii) in Proposition 1.

These results can be applied for the analysis of the optimal tax on the waste rock production. Since the waste rock disposal rate is always strictly positive, it readily follows from (11) that the optimal tax, defined as $-\mu$, is strictly increasing in time. We formulate this result as a proposition.

Proposition 2. *The time path of the optimal tax on the waste rock production is strictly increasing.*

Hence, to give the private mining firm correct incentives to dispose waste rocks, the tax on waste rock production $\alpha q - z$ should be strictly increasing.

We next analyze the behavior of state variables and, in particular, provide a set of sufficient conditions, which guarantee that the time path of the waste rock stock is either strictly increasing or has an inverted U-shape. These sufficient conditions are already shown to hold in this problem (see property (ii) of Proposition 1).

Proposition 3. *Let* $\bar{t} \in (0, T]$ *and suppose that* $q > 0$ *and* $\dot{q} < 0$ *on some interval* $[0, \bar{t})$*, and* $q = 0$ *on* $[\bar{t}, T]$*. Then* $S(T) > 0$ *and the time path of the waste rock stock has a unique maximum value.*

Proof. See Appendix A.2. □

The unique maximum value is either at $t = T$ or on the interval $(0, T)$. Hence the time path is either strictly increasing or has an inverted U-shape. When the waste rock stock has a unique maximum on the interval $(0, T)$, it is optimal to increase first the waste rock stock to some maximal size and then start disposing the waste rocks at a higher rate than they are produced by mining operations. Disposing later leads simply to lower costs due to discounting, thus, disposing all waste rocks immediately when they are produced is clearly suboptimal. Intuitively speaking, whether the time path has a unique maximum or is strictly increasing, seems to depend on the steepness of the damage function.

We next characterize the extraction rates under different waste rock production rates α. In particular, we compare the extraction rates of two mines which are identical except that one has a high rate α^H and the other has a low rate α^L. As this suggests, we set $\alpha^H > \alpha^L$.[15] With a higher waste rock production rate, extracting a unit of ore entails a larger waste rock amount and a higher cost of waste rock disposal. Since extraction is therefore more costly for every unit, one may expect that the extraction rates differ between high and low coefficient mines. This need not to be the case, as Proposition 4 suggests. It is assumed for this result, that both mines are exhausted during the mines' operation period.[16]

Proposition 4. *Consider two mines, one with higher rate of waste rock production than the other, and assume that the mines are otherwise identical. Suppose that both mines are exhausted. Then they are exhausted at the same instant of time, and the rate of extraction is the same for both mines.*

Proof. See Appendix A.3. □

Proposition 4 shows that not only are the extraction rates the same at the exhaustion

[15] We also use superscript H for extraction rates and shadow values of the mine with a high coefficient value and superscript L in case of the mine with a low coefficient value.

[16] It is a plausible possibility that both mines are exhausted. Tilton and Guzmán (2016) argue on page 142, that a mine often continues to operate until the ore body is exhausted due to high capital cost of mining. However, because these costs are not included in the model and because the terminal time is fixed, we study all the cases including 'both mines are exhausted', 'only one of the mines is exhausted' and 'neither of the mines are exhausted'.

time, but they are also the same over the whole planning interval. The case of Proposition 4 can be interpreted as a mine with a large operation period (or scarce resource). Here, the regulator has sufficient time to move the additional cost of waste rock disposal to the distant future without affecting the rate of extraction.

The next proposition modifies this result by relaxing the assumption regarding exhaustion of both of the mines and includes two cases. In the first case the mine with a low waste rock production rate is exhausted and the mine with a high rate is not, and in the second case neither of the mines is exhausted. In these cases extraction rate is lower with a high waste rock production rate than with a low one.

Proposition 5. *The rate of extraction is always lower in a mine with a high waste rock production rate than with a low production rate, if either*

(i) *the mine with a low waste rock production rate is exhausted and the one with a high rate is not, or*

(ii) *neither of the mines are exhausted.*

Proof. See Appendix A.4. □

This result agrees with the intuition that a mine with higher costs has a lower extraction rate compared to a low cost mine. Consider finally the case where only the mine with a high waste rock production rate is exhausted. We demonstrate that this case is impossible.

Proposition 6. *It is impossible that the mine with a high waste rock production rate is exhausted but the mine with a low rate is not.*

Proof. See Appendix A.5. □

This is an expected result. Exhaustion of the resource implies that it has been profitable to extract it. If the extraction is profitable with a high waste rock production rate, it surely is also profitable for a identical mine with a low production rate. Hence assuming in this case that only the high production rate mine is exhausted violates this basic intuition.

To conclude, with the exception of the case where both mines are exhausted, higher waste rock production rate implies lower extraction rate. This is because the cost of extracting one unit of material becomes larger since some of the extracted material is costly to dispose of. Note that if T is a choice variable, then at the optimal T the Hamiltonian equals the partial derivative of the damages from AMD with respect to T. The waste rock disposal rate is still strictly positive throughout the planning interval, and it is difficult to rule out the possibility that the mine is exhausted before T. Hence, also when T is a choice variable, the comparisons of high and low waste rock coefficient mines must consider the different exhaustion orderings as in the case with fixed T in propositions 4, 5 and 6.

4 Regulation under polluting waste rocks and stock pollution

The model is now extended to cover also the regulation of a stock pollutant. The regulator accounts for the negative stock externality created by the accumulative emissions. Therefore, the regulator's problem is to find the rates of extraction, waste rock disposal and abatement that maximize the total discounted social welfare consisting of the sum of gross profits, mine costs and damages from the pollution stock. The price of the extracted good is constant.[17] The regulator's problem is to

$$\max_{\{q(t),z(t),a(t)\}} \int_0^T e^{-\rho t}\left[pq(t) - C(q(t)) - C^s(z(t)) - A(a(t)) - D(N(t))\right] \,\mathrm{d}t$$

$$- G(S(T)) - \int_T^\infty e^{-\rho t} D(N(t)) \,\mathrm{d}t \tag{20}$$

$$\text{s.t} \quad \dot{X}(t) = -q(t), \quad X(0) = x_0, \quad X(T) \geq 0, \tag{21}$$

$$\dot{S}(t) = \alpha q(t) - z(t), \quad S(0) = 0, \quad S(T) \geq 0, \tag{22}$$

$$\dot{N}(t) = \beta q(t) - a(t) - \delta N(t), \quad N(0) = 0, \quad N(T) \geq 0, \tag{23}$$

$$q(t) \geq 0, \quad z(t) \geq 0, \quad a(t) \geq 0. \tag{24}$$

Now the regulator's problem contains an additional state variable, which is the pollution stock N. This stock grows with extraction and decreases through abatement and natural processes. After the terminal time T only natural abatement processes are active. Hence, after T the size of the pollution stock is given by equation $N(t) = N(T)e^{-\delta(t-T)}$. We simplify the notation by introducing function R defined as

$$R(N(T)) := \int_T^\infty e^{-\rho t} D\Big(N(T)e^{-\delta(t-T)}\Big) \,\mathrm{d}t. \tag{25}$$

Note that $R' > 0$ for $N(T) > 0$ and $R'(0) = 0$. The current value Hamiltonian related to regulator's problem is

$$H(q, z, a, X, S, N, \lambda, \mu, \gamma) = pq - C(q) - C^s(z) - A(a) - D(N) - \lambda q +$$

$$\mu(\alpha q - z) + \gamma(\beta q - a - \delta N), \tag{26}$$

[17] This modeling choice is different than for example in Chakravorty and Krulce (1994), Farzin (1996) and Chakravorty, Krulce, and Roumasset (2005), who use an inverse demand function. We are modeling the optimal policy of a single mine, which takes the price as given and therefore so does the regulator. This price can be interpreted as the world market price.

where γ is the shadow value of pollution stock. Applying the Maximum Principle gives

$$H_q = p - C'(q) - \lambda + \mu\alpha + \gamma\beta \leq 0, \quad q \geq 0, \quad qH_q = 0, \tag{27}$$

$$H_z = -C^{s\prime}(z) - \mu \leq 0, \quad z \geq 0, \quad zH_z = 0, \tag{28}$$

$$H_a = -A'(a) - \gamma \leq 0, \quad a \geq 0, \quad aH_a = 0, \tag{29}$$

$$\dot{X} = -q, \tag{30}$$

$$\dot{S} = \alpha q - z, \tag{31}$$

$$\dot{N} = \beta q - a - \delta N, \tag{32}$$

$$\dot{\lambda} = \rho\lambda, \tag{33}$$

$$\dot{\mu} = \rho\mu, \tag{34}$$

$$\dot{\gamma} = D'(N) + (\delta + \rho)\gamma, \tag{35}$$

$$\lambda(T) \geq 0,\ X(T) \geq 0,\ \lambda(T)X(T) = 0, \tag{36}$$

$$\mu(T) + G'(S(T)) \geq 0,\ S(T) \geq 0,\ [\mu(T) + G'(S(T))]S(T) = 0, \tag{37}$$

$$\gamma(T) + R'(N(T)) \geq 0,\ N(T) \geq 0,\ [\gamma(T) + R'(N(T))]N(T) = 0. \tag{38}$$

Equations (29), (32), (35) and (38) as well as the new shadow value in Equation (27) convey the role of pollution and abatement in the optimal solution. The shadow value of the pollution stock is approximately the change in the regulator's optimal value function from a unit (jump) increase in the pollution stock. Whenever $a(t) > 0$, this shadow value is strictly negative, because then $\gamma = -A'(a) < 0$ by Equation (29). Therefore, as the pollution stock grows by one unit, the optimal value decreases approximately by the amount $-\gamma$. Since equation $A'(a) = -\gamma$ holds for positive abatement, abatement is made such that along the optimal path the marginal abatement cost equals the marginal benefit of abatement. Function μ has the same interpretation as before.

The results characterizing the regulator's choice are the counterpart of Proposition 1.

Proposition 7.

(i) *The extraction rate, the waste rock disposal rate and the abatement rate are continuous in time.*

(ii) *The extraction rate is positive on some interval.*

(iii) *The waste rock disposal rate is positive and strictly increasing over the whole planning interval, and the shadow value of waste rock stock is strictly negative.*

(iv) *If* $a(t) > 0$ *on any time interval, the time path of abatement can be in there decreasing, increasing or both:*

  (a) *if the shadow value of pollution stock is increasing, the time path of abatement is decreasing;*

  (b) *if the shadow value of pollution stock is decreasing, the time path of abatement is increasing.*

Proof. See Appendix A.6. □

Note first that unlike in Proposition 1, we cannot argue for the case that the rate of extraction is strictly decreasing. When the optimal controls are differentiable,

$$\dot{q} = \frac{-\rho\lambda + \rho\alpha\mu + \beta\dot{\gamma}}{C''(q)}, \tag{39}$$

$$\dot{z} = -\frac{\rho\mu}{C^{s\prime\prime}(z)} > 0. \tag{40}$$

The interpretation of the result that $\dot{z} > 0$ is the same as before. The relationship between abatement and the shadow value of the pollution stock is interesting: The time path for abatement can be increasing, decreasing or both. Differentiating $H_a = 0$ with respect to time, we obtain

$$-A''(a)\dot{a} - \dot{\gamma} = 0. \tag{41}$$

Using this equation and (35), we obtain

$$\dot{a} = -\frac{D'(N) + (\delta + \rho)\gamma}{A''(a)}. \tag{42}$$

Equation (42) suggests that abatement is increasing, provided the instantaneous marginal damage from an extra unit of pollution is smaller than the discounted future total damage (multiplied by $\delta + \rho$), that is when $D'(N) < -(\delta + \rho)\gamma$. Condition (41) also allows for an interpretation of a case when the mining firm is taxed for pollution. Define $-\gamma$ as the tax on the flow of net emissions. Then the abatement is increasing exactly when the tax is increasing. From the viewpoint of (42), abatement is increasing, when the tax multiplied by $\delta + \rho$ is greater than the marginal damage.[18]

The following short-run comparative statics of the model are derived using conditions (27)-(29). Here we are interested on how a change either in the waste rock production coefficient α or in the emission coefficient β affects the optimal choice of the control variables in the short-run. It is straightforward to derive that

$$\frac{\partial q}{\partial \alpha} = -\frac{\mu}{-C''(q)} < 0, \quad \frac{\partial q}{\partial \beta} = -\frac{\gamma}{-C''(q)} < 0, \tag{43}$$

$$\frac{\partial z}{\partial \alpha} = \frac{\partial z}{\partial \beta} = \frac{\partial a}{\partial \alpha} = \frac{\partial a}{\partial \beta} = 0. \tag{44}$$

Hence, only the short-run extraction rate is affected by a change either in the waste rock production coefficient or in the emission coefficient. An increase in either of these parameters increases the cost of extraction and hence decreases the short-run extraction rate.

We now turn to the waste rock stock and its taxation. Part (*iii*) of Proposition 7 and condition (28) imply that the time path of the optimal tax on waste rock production is strictly increasing.

[18] Also, whenever $\dot{\gamma} < 0$, the induced decrease in the optimal value of the program from a unit increase in the pollution stock is increasing. Intuition suggests and Part (*iv*) of Proposition 7 confirms, that in this case it is optimal to increase abatement.

Proposition 8. *The time path of the optimal tax on the waste rock production is strictly increasing.*

However, in the current model the extraction rate is not guaranteed to be decreasing, and therefore the result in Proposition 3 regarding the time path of the waste rock stock may not hold. Of course, the result continues to hold in this model also, if it is assumed that extraction is decreasing.

4.1 Analysis of the optimal emission tax

We next turn to study the time path of the shadow value of the pollution stock and hence the emission tax. We divide the analysis of the model into two cases by analyzing first the case in which $N(T) = 0$ is optimal and then the case in which $N(T) > 0$ is optimal. We begin with a lemma concerning the exclusion of special cases related to the time paths of the pollution stock and the shadow value.

Lemma 1.

(*i*) *The pollution stock is not zero on the whole planning interval.*

(*ii*) *The shadow value γ cannot be a non-zero constant on any interval.*

Proof. See Appendix A.7. □

That the pollution stock is not zero on the planning interval is reasonable since by assumption $D'(0) = 0$ the first units of pollution in the stock cause approximately zero damages. The second part of the lemma tells us that the shadow value cannot have non-zero “flat parts” in its time path.

Case $N(T) = 0$ is optimal. To analyze the possible time paths for the emission tax, we first define $\hat{t} \leq T$ as the time instant, whose successors have zero pollution stock but (possibly) close predecessors positive stock, that is, $\hat{t}$ satisfies conditions

$$\begin{cases} N(t) > 0, & \text{for all } \ t \in [t_0, \hat{t}), \\ N(t) = 0, & \text{for all } \ t \in [\hat{t}, T], \end{cases} \tag{45}$$

for some $t_0 > 0$. This is well-defined, since $N(t) \equiv 0$ is impossible as shown in Lemma 1. We need the following lemma for the proof of Proposition 10 below.

Lemma 2.

(i) *There exists a time instant* $t^+ < \hat{t}$ *such that* $a(t) > 0$ *for every* $t \in (t^+, \hat{t})$.

(ii) $\gamma(t) \leq 0$ *for all* $t \geq \hat{t}$.

Proof. See Appendix A.8. □

Part (i) says that the abatement rate is positive just before the time instant at which the pollution stock obtains a zero value. This is reasonable, since the natural decrease of the stock is exponential and is therefore unable by itself to bring the stock to zero. We continue the analysis on the role of $\hat{t}$ for the time path of the pollution stock by considering two cases: one with $\hat{t} < T$ and the other with $\hat{t} = T$. We begin with the case $\hat{t} < T$. The following lemma shows that when the pollution stock becomes zero, also extraction, abatement and the shadow value of the pollution stock become zero.

Lemma 3. *Suppose that* $\hat{t} < T$. *Then*

(i) $\beta q(t) = a(t) = 0$ *for all* $t \geq \hat{t}$,

(ii) $\gamma(t) = 0$ *for all* $t \geq \hat{t}$.

Proof. See Appendix A.9. □

In particular, the mine has been exhausted by time instant $\hat{t}$ in the economic sense, since extraction is zero afterwards. Since $\gamma(\hat{t}) = 0$, we obtain that

$$\gamma(t) = \begin{cases} -e^{(\delta+\rho)t} \int_t^{\hat{t}} D'(N(\tau))e^{-(\delta+\rho)\tau} \, \mathrm{d}\tau, & \text{for } \ t \in [0, \hat{t}), \\ 0, & \text{for } \ t \in [\hat{t}, T]. \end{cases} \tag{46}$$

So $\gamma(t) < 0$ for all $t \in [0, \hat{t})$. Now we can state the main result for the particular case $\hat{t} < T$.

Proposition 9. *Suppose that $N(T) = 0$ and $\hat{t} < T$. Then the shadow value of the pollution stock, γ, has a U-shaped time path.*

Proof. See Appendix A.10. □

For $\hat{t} < T$, the optimal tax on the rate of pollution flow from the mine can be given as

$$-\gamma(t) = \begin{cases} e^{(\delta+\rho)t} \int_t^{\hat{t}} D'(N(\tau))e^{-(\delta+\rho)\tau} \, \mathrm{d}\tau, & \text{for } \ t \in [0, \hat{t}), \\ 0, & \text{for } \ t \in [\hat{t}, T], \end{cases} \tag{47}$$

where $N(t)$ is the first-best pollution stock at time t. The optimal tax has an inverted U-shape, which is the result obtained for example by Tahvonen (1997) in an infinite-horizon problem without abatement. Adding abatement and waste rock disposal, or tinkering with the time horizon do not change his result at least when $\hat{t} < T$.

We now characterize the shadow value of pollution stock, when $\hat{t} = T$, so that the pollution stock is exhausted at the end of the mining operations, not before. The second main result of this section is the following:

Proposition 10. *Suppose that $N(T) = 0$ and $\hat{t} = T$. Then $a(T) = 0$ and the shadow value of the pollution stock, γ, has a U-shaped time path.*

Proof. See Appendix A.11. □

Note that the shadow value of the pollution stock is given by the equation

$$\gamma(t) = -e^{(\delta+\rho)t} \int_t^T D'(N(\tau)) e^{-(\delta+\rho)\tau} \, \mathrm{d}\tau \quad \text{for all} \quad t \in [0, T]. \tag{48}$$

We have shown that when $N(T) = 0$ is optimal, the optimal emission tax has an inverted U-shaped time path. Now we turn to the other possibility of the optimal value for the terminal pollution stock.

Case $N(T) > 0$ is optimal. The analysis and results of this case mirror the ones obtained above. Namely, we exclude strictly increasing time path for γ as well as the time path with an inverted U-shaped section.

Proposition 11. *Suppose $N(T) > 0$. The shadow value of the pollution stock, γ, cannot have a strictly increasing time path, or a time path, where the first derivative is zero more than twice.*

Proof. See Appendix A.12. □

This result rules out many forms of the time path of the shadow value of pollution stock. Proposition 11 leaves two possibilities in addition to U-shaped time path for the shadow value. In the first one, the shadow value is strictly decreasing (although it is possible $\dot{\gamma} = 0$ for a single time instant), and in the second one exactly two instants of time exist, where $\dot{\gamma} = 0$. Moreover, in the second possibility the first instant of time, where $\dot{\gamma} = 0$, must give a local minimum for γ, and the second instant of time must give a local maximum for γ.[19]

This completes the analysis of optimal emission tax. We have shown that there are three possibilities for the time path of the optimal tax, $-\gamma$. In the first one the time path of the tax increases at the beginning and then decreases until the end of operations.

[19] Note that the points, in which $\dot{\gamma} = 0$, cannot be inflection points of γ. If they were, we would obtain a contradiction as in the proof of Proposition 9, because $\dot{\gamma} < 0$ between these two points.

In the second one the time path is always increasing, and in the third one it has a local maximum followed by a local minimum. We provide the following interpretation for the different taxes.

When the operation period is relatively short, it is plausible that the emissions tax is always increasing as is also abatement. Equation (35) reveals, that the augmented interest on the value of discounted total future damages is greater than the instantaneous marginal damage, that is

$$(\delta + \rho) \int_t^T D'(N(\tau)) e^{-(\delta+\rho)(\tau - t)} \, \mathrm{d}\tau > D'(N(t)). \tag{49}$$

The pollution stock never reaches such a high value to change the direction of this inequality, and it is always the future damages that dominate instantaneous damages. This tax path, as well as the others, involves a build-up phase for the pollution stock. As it is always optimal to increase abatement, it is also optimal to move the abatement costs to the future due to discounting.

When the operation period is longer, the emissions tax may increase at the beginning and at the end, but is decreasing during the middle part. Again, abatement follows this pattern. Now, the augmented interest on the value of discounted total future damages is smaller than the instantaneous marginal damage during the middle part, that is, it holds during the middle part that

$$(\delta + \rho) \int_t^T D'(N(\tau)) e^{-(\delta+\rho)(\tau - t)} \, \mathrm{d}\tau < D'(N(t)). \tag{50}$$

Why do tax and abatement decrease? During some initial phase, the pollution stock has reached such a high value, in which the marginal damages are large relative to future damages in the sense revealed by (50). A high N means high natural decrease of the pollution stock. Since the natural process cleans the stock at rate δN, it is beneficial for the regulator to exploit this by decreasing abatement and thus moving abatement costs to the end of planning period. As time approaches T, the pollution stock becomes smaller and smaller, as well as the rate of abatement by natural processes. The inequality in (50) changes direction, and abatement starts to increase in order to push the pollution stock to a low enough value, which continues to cause damages after the mine has been shut down. Although abatement costs are increased by this, the increase is offset by the benefits of prior decreased abatement.

When time path of the tax is an inverted U-shape, the future damages are initially larger than marginal damages as emphasized by (49), so that abatement increases. The operation period is divided into two parts. In the first one, inequality $-(\delta+\rho)\gamma > D'(N)$, and in the second one inequality $-(\delta + \rho)\gamma < D'(N)$ hold. Two aspects are relevant for this time path. It is again beneficial to move abatement costs to the more distant future during some initial time interval. Also, the regulator can exploit the high natural abatement rate before the terminal time to decrease abatement. Abatement can decrease until the end, because it is enough to clean the pollution stock, which has been dispersed for a longer period of time.

4.2 A comparison of extraction rates of different mines under time invariant taxes

We compared in Section 3 two mines with different waste rock coefficients and showed in Proposition 4 that if both the high coefficient mine and the low coefficient mines are exhausted, then extraction rates are the same. Doing the similar comparisons as in Section 3 for mines with different emission coefficients β and waste rock production coefficients α in the model of Section 4 is analytically challenging, since the proofs of many propositions in Section 3 are based on the exponential development of the costate variables.[20] When the emission stock $N(t)$ is added to the model, costate variable γ may

[20]See for example the proof of Proposition 4.

develop non-exponentially or even non-monotonically, which implies that the methods of proof in Section 3 become infeasible.

Therefore we perform a similar comparison as in Proposition 4 for the case where the regulator chooses time invariant (i.e. constant with respect to time) taxes. There are two reasons for doing this. First, the general comparison with first-best taxes is analytically challenging for the reasons mentioned above. Second, it may be that choosing two time variant taxes is infeasible in practice perhaps due to complexity or due to some political reasons. Hence we assume that the regulator attempts to correct the externalities by setting time invariant taxes. This assumption is similar to the one made in van der Ploeg and Withagen (2012).[21]

Let v_S and v_N be the taxes for the net waste rock production $\alpha q - z$ and for the net emissions $\beta q - a$, respectively. For the comparison we consider two mines, which have waste rock production coefficients α^H and α^L, and emission coefficients β^H and β^L. Define

$$K := v_S(\alpha^H - \alpha^L) + v_N(\beta^H - \beta^L). \tag{51}$$

Since $v_S\alpha^H + v_N\beta^H$ is the additional tax burden per extraction unit for mine H, K is the difference between the marginal tax burdens of the mining firms. A strictly positive K means that the marginal tax burden of mine H is larger than the burden of mine L. Denote the exhaustion times of the mines with t^H and t^L. In the next proposition we show that constant taxes imply that a mine with a lower marginal tax burden is exhausted before a mine with higher marginal tax burden.

Proposition 12. *Suppose that the regulator applies time invariant taxes and consider two mines with different waste rock production and emission coefficients, and assume that the mines are otherwise identical. Suppose that both mines are exhausted. Then*

(*i*) $K > 0$ *and* $t^H < T$, *imply* $t^L < t^H$,

[21] See their Section 3.2.

(*ii*) $K = 0$ *implies* $t^L = t^H$.

Proof. See Appendix A.13. □

When the regulator applies time invariant taxes and $t^H < T$, the mining firm with a lower marginal tax burden exhausts the resource stock faster than the mine with a higher marginal tax burden.

It is intuitively clear that since both of the mines are exhausted, the mine with a lower tax burden has an extraction profile that begins at a higher level than the profile of the mine with a higher tax burden. In addition, it seems the extraction profiles cross exactly once. Note that Proposition 12 is silent about the possibility that $t^H = T$ when $K > 0$. In this case it of course holds that $t^L \leq T$. Whether $t^L < T$ or $t^L = T$ holds, it must be that the extraction paths of the mines cross once and that the initial extraction level is higher in mine L than in mine H. This is the contents of the last proposition, which says that the extraction path q^L crosses path q^H from the above when $K > 0$.

Proposition 13. *Suppose that the regulator applies time invariant taxes and consider two mines with different waste rock production and emission coefficients, and assume that the mines are otherwise identical. Suppose that both mines are exhausted. Let* $K > 0$. *Then* $q^L(t) > q^H(t)$ *on* $[0, \hat{t})$ *and* $q^L(t) < q^H(t)$ *on* $[\hat{t}, T)$ *for some* $\hat{t} \in (0, T)$.

Proof. See Appendix A.14. □

5 Conclusions and discussion on further research

We examined the optimal environmental policy for a mine when some of the special features of mining are taken into account. In our model, extraction creates waste rocks causing pollution such as acid mine drainage, producing metals from the extracted ore is a source of emissions and abatement is made by the mine owner and the nature. Production of waste rocks alone as an essential ingredient of mining operations represents a change towards more realistic modeling of mines.

Our results show that the higher waste rock production rate tends to decrease the rate of extraction compared to an identical mine with a lower waste rock production rate. However, it can also be that the extraction rates are the same. Regarding taxation, we found that the optimal tax on the waste rock production to correct the externality caused by them is strictly increasing in time. Hence, we find that to correct the externality caused for example by acid mine drainage calls for an increasing tax. In addition, if the mine produces another pollutant and the damages depend on its stock, there are three possibilities for the optimal emission tax. The first one is that the emission tax is qualitatively similar to the ones found in the previous literature such as Ulph and Ulph (1994) and Tahvonen (1997). The second possibility is that the emission tax is always strictly increasing, and the third is that the tax increases at the beginning and at the end, but is decreasing during the middle part of the operation period.[22]

Our work represents a step towards a more concrete and policy relevant modeling of mines to provide more economic insights, for instance, in the determination of environmental permit for mines. The current model includes one stylized fact of mining operations, namely the waste rocks that pollute and their costly disposal. Incorporation of other stylized facts, such as capacity constraints on extraction and cut-off grade selection, constitute a possible extension.[23] Although there is an implicit upper bound on extraction in the current model (defined by price and marginal cost), a more detailed modeling of the capacity choice and the relationship between it and the environmental policy is needed.[24]

[22]Our model has not analyzed the actual implementation of the two taxes, which is an important practical aspect.

[23]There already exists literature on capacity choice in mining and capacity constraints (for example Campbell 1980; Lozada 1993; Cairns 2001; Holland 2003) and on the analysis of cut-off grades in a heterogeneous ore body (for example Krautkraemer 1988; Shinkuma 2000; Cairns and Shinkuma 2003).

[24]Stollery (1985) studies this in a set-up with flow pollution.

Furthermore, since for some metals the waste rock to ore ratio is increasing over time, a reasonable modification of the model is to let the waste rock production depend on the amount of extractable material left in the mine.

The monitoring and enforcement of a mine is also an important issue, which has not received much attention in the literature. This is surprising given that noncompliance with the environmental policy happens in the industry from time to time. Therefore, a sensible research question is that given a chosen regulative instrument, how should the monitoring and enforcement of the mine be designed? Also the current model does not analyze mine rehabilitation. Monitoring and enforcement are also related to rehabilitation of the mine area, since after the mining operations have ceased, monitoring is often still needed.[25]

Another relevant extension would be to analyze uncertain damages due to acid mine drainage. Furthermore, waste dams can be subject to catastrophic accidents, which can cause massive damages. One option to analyze the relation of optimal taxation with environmental accidents in mining would be let the terminal time of the planning interval be a random variable as for example in Boukas, Haurie, and Michel (1990), Clarke and Reed (1994) and Schumacher (2011). This variable would be date of the accident, and the probability of the accident would depend on the waste rock stock.

In our current model the price of the resource is given without emphasis on market structure.[26] We have basically only asked how should a single mining firm be taxed to correct one or two externalities. To obtain better picture for policy advices regarding environmental policy, a more detailed model should contain multiple resource producers and the price formation should be endogenous.

[25] Rehabilitation of mine sites and bond design has been analyzed in Sullivan and Amacher (2009) and White et al. (2012).

[26] Market structure and exhaustible resources have received much attention, see for example Benchekroun, Halsema, and Withagen (2009).

A Appendix

A.1 Proof of Proposition 1

(i) Since the control set $U := \{(q,z) \in \mathbb{R}^2 \mid q \geq 0, z \geq 0\}$ is convex and H is strictly concave in (q,z), the control functions are continuous functions.

(ii) First we show that $q(t) > 0$ at some t. Assume not, that is $q(t) \equiv 0$. Then $X(T) = x_0 > 0$, and therefore by (16) $\lambda(T) = 0$. This implies $\lambda(t) \equiv 0$. Equation (13) together with boundary condition $S(T) \geq 0$, implies $z(t) \equiv 0$. Then by assumption $C^{s\prime}(0) = 0$ (recall Assumption 1) and by condition (11), $\mu(t) \geq 0$ for every t. This yields a contradiction since (10), that is, condition

$$p - C'(0) + \mu\alpha \leq 0, \tag{A.1}$$

and assumption $p - C'(0) > 0$, imply $\mu(t) < 0$. Therefore $q(t) > 0$ at some t. Since $q(t)$ is continuous, it must be that $q(t) > 0$ at some interval $[t_0, t_1]$.

Now we show that the rate of extraction is strictly decreasing, when the rate of extraction is positive. We first show that $\mu(0) < 0$. Note that $z > 0$ somewhere: Suppose not, that is, suppose $z \equiv 0$. Then $S(T) > 0$ and $G'(S(T)) > 0$. By (11) and assumption $C^{s\prime}(0) = 0$, $\mu \geq 0$, which implies by (15) that $\mu(0) \geq 0$. Hence $\mu(T) \geq 0$. Because $\mu(T) \geq 0$ and $G'(S(T)) > 0$, also $\mu(T) + G'(S(T)) > 0$, which implies by condition (17) that $S(T) = 0$. This contradicts $S(T) > 0$. Therefore $z > 0$ somewhere. Because $C^{s\prime}(z) = -\mu(0)e^{\rho t}$ for some t, $\mu(0) < 0$. To show that the rate of extraction is strictly decreasing let $t_0^+, t_1^+ \in [t_0, t_1]$ with $t_0^+ < t_1^+$. We obtain, after using (10) twice and subtracting, that on the interval where $q(t) > 0$,

$$p - C'(q(t_0^+)) - p + C'(q(t_1^+)) = \tag{A.2}$$

$$\lambda(0)[e^{\rho t_0^+} - e^{\rho t_1^+}] - \mu(0)[e^{\rho t_0^+} - e^{\rho t_1^+}] < 0. \tag{A.3}$$

This implies $q(t_1^+) < q(t_0^+)$ by assumption $C''(q) > 0$.

Continuity of $q(t)$ implies that we can choose $t_0 = 0$, because otherwise $q(t)$ would be discontinuous or increasing somewhere.

(iii) Recall that $z(t) > 0$ at some interval $[t_2, t_3]$. At this interval, control $z(t)$ is strictly increasing. This is true because on this interval, $C^{s\prime}(z) = -\mu(0)e^{\rho t}$, which clearly implies that $z(t)$ must increase: Let $t_2^+, t_3^+ \in [t_2, t_3]$ with $t_2^+ < t_3^+$. Using (11) twice and subtracting, we get

$$C^{s\prime}(z(t_3^+)) - C^{s\prime}(z(t_2^+)) = \mu(0)[e^{\rho t_2^+} - e^{\rho t_3^+}] > 0. \tag{A.4}$$

This implies $z(t_3^+) > z(t_2^+)$ by assumption $C^{s\prime\prime}(q) > 0$. Note in addition that because $\mu(0) < 0$, $\mu(t) < 0$ for all t. Because $z(t)$ is strictly increasing and continuous, there can be only one interval, where $z(t) > 0$. In addition, since $\mu(0) < 0$, $z(0) > 0$: Suppose not. By condition (11)

$$-C^{s\prime}(0) - \mu(0) \leq 0. \tag{A.5}$$

Then by assumption $C^{s\prime}(0) = 0$, $\mu(0) \geq 0$, which contradicts $\mu(0) < 0$.

(iv) When $q(t) > 0$,

$$p - C'(q) = (\lambda(0) - \mu(0)\alpha)e^{\rho t}. \tag{A.6}$$

Because $\lambda(0) > 0$ and $\mu(0) < 0$, the instantaneous marginal profit rises at the rate of interest.

A.2 Proof of Proposition 3

Note first that $\mu(T) < 0$ and (17) imply $G'(S(T)) > 0$. This implies $S(T) > 0$. For the second part of the result, we show first that $\alpha q(0) \leq z(0)$ is generally impossible at optimum. Suppose that $\alpha q(0) \leq z(0)$ holds. Then $\dot{S}(0) \leq 0$, and because $\dot{q} < 0$ for every $t \in [0, \bar{t})$ and $\dot{z} > 0$ for every $t \in [0, T]$, $\dot{S}(t) < 0$ for all $t \in (0, T]$. Then $S(T) > 0$ cannot be satisfied. Therefore, inequality $\beta q(0) > z(0)$ holds implying that $\dot{S}(0) > 0$. Note that the time path of S cannot have a minimum: Suppose it has a minimum. At such a point extraction is positive and $\ddot{S} \geq 0$, which is equivalent with $\alpha\dot{q} \geq \dot{z}$ by Equation (13). But since $\dot{z} > 0$, we obtain $\dot{q} > 0$, which contradicts Part (ii) Proposition 1.

Hence the time path is either strictly increasing or has a maximum value on $(0, T)$. If the path is strictly increasing, the maximum value is at $t = T$. We show next, that if the path has a maximum value on $(0, T)$, the maximum value is unique. To this end, suppose that the stock has a maximum value at $t_0 < T$, which is denoted by $S(t_0)$. The stock cannot have another local maximum: Suppose it does. Two possibilities exists in this case: either there exists t_1 with a local maximum satisfying $S(t) < \max\{S(t_0), S(t_1)\}$ for all $t \in (t_0, t_1)$, or there exists a local maximum t_2 satisfying $S(t) = S(t_0)$ for all $t \in (t_0, t_2)$. The first case is ruled out, because it implies that S has a local minimum on (t_0, t_1), which contradicts the above reasoning. To rule out the second possibility, we note that $\dot{S}(t_0) = 0$ and $\dot{S}(t) > 0$ for every $t \in (0, t_0)$. Equation $\dot{S}(t_0) = 0$ is equivalent to $\beta q(t_0) = z(t_0)$. Since $\dot{q} \leq 0$ and $\dot{z} > 0$, it must hold that after t_0 we have $\dot{S}(t) < 0$. This contradicts $S(t) = S(t_0)$ for all $t \in (t_0, t_2)$.

A.3 Proof of Proposition 4

Let t^L be the time instant when the mine is exhausted with the waste rock production rate α^L and let t^H be the respective time instant with coefficient α^H. We show first that $t^L > t^H$ is impossible, so assume that $t^L > t^H$ holds. In both cases $\dot{q} < 0$ whenever $q > 0$. Because of this and because extraction is zero after exhaustion, q^H hits zero before t^L. Since the mine is exhausted in both cases,

$$\int_0^{t^H} q^H(\tau)\ \mathrm{d}\tau = x_0 = \int_0^{t^L} q^L(\tau)\ \mathrm{d}\tau. \tag{A.7}$$

It follows from these, that there exists a time instant $\hat{t} < t^H$ for which

$$q^L(\hat{t}) = q^H(\hat{t}) > 0. \tag{A.8}$$

From this and from the necessary conditions of the models' problems, we obtain the following equations

$$p - C'(q^H(\hat{t})) = \lambda^H(\hat{t}) - \alpha^H \mu^H(\hat{t}), \tag{A.9}$$

$$p - C'(q^L(\hat{t})) = \lambda^L(\hat{t}) - \alpha^L \mu^L(\hat{t}). \tag{A.10}$$

From (A.8)-(A.10) we obtain that

$$\lambda^L(\hat{t}) - \alpha^L \mu^L(\hat{t}) = \lambda^H(\hat{t}) - \alpha^H \mu^H(\hat{t}). \tag{A.11}$$

Furthermore, by the necessary conditions,

$$\lambda^L(t) = \lambda^L(0)e^{\rho t}, \quad \lambda^H(t) = \lambda^H(0)e^{\rho t}, \quad \mu^H(t) = \mu^H(0)e^{\rho t} \quad \text{and} \quad \mu^L(t) = \mu^L(0)e^{\rho t}. \tag{A.12}$$

Define now function f with $f(t) := \lambda^L(t) - \alpha^L \mu^L(t) - \lambda^H(t) + \alpha^H \mu^H(t)$, and note that $f(t) = f(0)e^{\rho t}$ and $f(\hat{t}) = f(0)e^{\rho \hat{t}}$. Since by (A.11) $f(\hat{t}) = 0$, we obtain that $f(0) = 0$. Hence $f(t) \equiv 0$. Since (A.9) and (A.10) hold also for other time instants with positive extraction, we obtain equation

$$p - C'(q^L(t)) - p + C'(q^H(t)) = f(t) = 0. \tag{A.13}$$

Hence $q^L(t) = q^H(t)$ for all t with positive q^L and q^H. But then

$$\int_0^{t^L} q^L(\tau)\ \mathrm{d}\tau = \int_0^{t^H} q^L(\tau)\ \mathrm{d}\tau + \int_{t^H}^{t^L} q^L(\tau)\ \mathrm{d}\tau \tag{A.14}$$

$$= \int_0^{t^H} q^H(\tau)\ \mathrm{d}\tau + \int_{t^H}^{t^L} q^L(\tau)\ \mathrm{d}\tau \tag{A.15}$$

$$= x_0 + \int_{t^H}^{t^L} q^L(\tau)\ \mathrm{d}\tau > x_0, \tag{A.16}$$

which contradicts (A.7).

By the same principles and after minor changes, $t^L < t^H$ is also impossible. Hence $t^L = t^H$. Similar arguments as used above show that the extraction rates are the same.

A.4 Proof of Proposition 5

(i) Since the mine with coefficient α^H is not exhausted, $\lambda^H(T) = 0$, and therefore $\lambda^H(t) \equiv 0$. The rate of extraction with coefficient α^H, q^H, cannot be anywhere greater than the rate of extraction with coefficient α^L, q^L: Suppose otherwise, that is, suppose that $q^H(t) > q^L(t)$ for some $t \in [0, T]$. Then there exists a time instant $\hat{t} < T$ for which

$$q^L(\hat{t}) = q^H(\hat{t}), \tag{A.17}$$

and we obtain using similar arguments as in the proof of Proposition 4 that $\lambda^L(t) - \alpha^L \mu^L(t) = -\alpha^H \mu^H(t)$ for every t with positive extraction levels. This means that $q^H(t) = q^L(t)$ (when extraction is positive), which yields a contradiction with $q^H(t) > q^L(t)$.

Since the mine with α^L is exhausted and the mine with α^H is not, the extraction rates cannot be equal at any time instant, because similar arguments as above show that the extraction rates are equal for all time instants in such a case.

(ii) We will show that $q^H < q^L$ for all $t \in [0, T]$ by contradiction. To this end, we assume $q^H \geq q^L$ for some $t \in [0, T]$. Note then that $q^H \geq q^L$ for all $t \in [0, T]$: Otherwise (that is, if $q^H < q^L$ for some t) there would exist a time instant $\hat{t}$ such that $q^H = q^L$, which implies using similar arguments as in the proof of Proposition 4 that $q^H \equiv q^L$, which contradicts inequality $q^H < q^L$ for some t.

Because neither of the mines are exhausted, $\lambda^H = \lambda^L \equiv 0$. Since $q^H \geq q^L$, $p - C'(q^H) \leq p - C'(q^L)$. Combining this with (10) gives us $-\alpha^H \mu^H \leq -\alpha^L \mu^L$ when extraction is positive. Then

$$-\mu^H \leq -\frac{\alpha^L}{\alpha^H} \mu^L < -\mu^L, \tag{A.18}$$

since $\alpha^H > \alpha^L$. This implies together with equation $C^{s\prime}(z) = -\mu$ from (11), that $z^L > z^H$ for all $t \in [0, T]$. This and $q^H \geq q^L$ for all $t \in [0, T]$ imply by (13) that $\dot{S}^H > \dot{S}^L$ for all $t \in [0, T]$. Combining this with $S^H(0) = S^L(0) = 0$ gives

$$S^H(T) > S^L(T). \tag{A.19}$$

Hence $G'(S^H(T)) > G'(S^L(T))$. We obtain using these and (17) that

$$-\mu^H(T) = G'(S^H(T)) > G'(S^L(T)) = -\mu^L(T), \tag{A.20}$$

but $-\mu^H(T) > -\mu^L(T)$ contradicts (A.18). Hence $q^H < q^L$ for all $t \in [0, T]$.

A.5 Proof of Proposition 6

Since the mine with a high waste rock production rate is exhausted and the one with a low rate is not, we have

$$x_0 = \int_0^{t^H} q^H(\tau)\ \mathrm{d}\tau > \int_0^T q^L(\tau)\ \mathrm{d}\tau. \tag{A.21}$$

This clearly implies that there exists a time instant at which $q^H > q^L$. Then $q^H > q^L$ for all $t \in [0, T]$ (otherwise we would again obtain a contradiction). Note that $\lambda^L \equiv 0$. When $q^L > 0$, condition (10) gives us after subtraction that

$$p - C'(q^H) - p + C'(q^L) = \lambda^H - \alpha^H\mu^H + \alpha^L\mu^L. \tag{A.22}$$

Since the left-side of this is strictly negative, we obtain

$$\lambda^H - \alpha^H\mu^H + \alpha^L\mu^L < 0. \tag{A.23}$$

This implies

$$-\mu^H < -\frac{\lambda^H}{\alpha^H} - \frac{\alpha^L}{\alpha^H}\mu^L < -\frac{\alpha^L}{\alpha^H}\mu^L < -\mu^L. \tag{A.24}$$

Because $q^H > q^L$ for all $t \in [0, T]$ and $-\mu^H < -\mu^L$ for all $t \in [0, T]$, we obtain a contradiction using similar lines of reasoning as in the proof of Part (*ii*) of Proposition 5. Hence it is impossible that a mine with a high waste rock production rate is exhausted, but the mine with a lower rate is not.

A.6 Proof of Proposition 7

(i) Similar to the proof of Part (i) of Proposition 1.

(ii) Assume that q is always zero. With similar reasoning as in the proof of Part (ii) of Proposition 1, it follows from this that $\mu(t) \geq 0$ for all t. (Note: Because of $\gamma(t)$, we cannot argue directly for a contradiction using condition $H_q \leq 0$ in (27).) Since $q(t) \equiv 0$, (32) becomes

$$\dot{N} = -a - \delta N, \tag{A.25}$$

whose solution using the given initial condition is

$$N(t) = \int_0^t -a(\tau)e^{-\delta(t-\tau)} \, \mathrm{d}\tau. \tag{A.26}$$

Then, because $a(t) \geq 0$,

$$N(T) = \int_0^T -a(\tau)e^{-\delta(T-\tau)} \, \mathrm{d}\tau \leq 0, \tag{A.27}$$

implying that $a(t) \equiv 0$. Condition (29) and assumption $A'(0) = 0$, imply $-\gamma \leq 0$ or equivalently $\gamma \geq 0$. Condition $H_q \leq 0$ becomes now

$$p - C'(0) + \mu\alpha + \gamma\beta \leq 0, \tag{A.28}$$

implying that $\mu < 0$. This contradicts $\mu \geq 0$. Therefore there exists a time instant where $q(t) > 0$. Again, by continuity of q, there is an interval, where $q(t) > 0$.

(iii) Similar to the proof of Part (iii) of Proposition 1.

(iv) Let $t_1, t_2 \in I \subset [0, T]$ with $t_1 < t_2$, where I is an interval with $a(t) > 0$. Using (29) twice, we obtain equation

$$A'(a(t_2)) - A'(a(t_1)) = \gamma(t_1) - \gamma(t_2). \tag{A.29}$$

The conclusion follows from this, since $A'' > 0$.

A.7 Proof of Lemma 1

(i) Suppose on the contrary to the claim that $N(t) = 0$ on $[0, T]$. Recall that by Proposition 7 there exists an interval $I \in [0, T]$, where $q > 0$. Then on that interval $\beta q = a > 0$ (by (32)) and $\dot{\gamma} = (\delta + \rho)\gamma$ (by (35) and by $D'(0) = 0$). By (29), $A'(a) = -\gamma > 0$ on I, implying that $\dot{\gamma} < 0$. Then a grows on I by Proposition 7. But by (27), $p - C'(q) = \lambda - \mu\alpha - \gamma\beta$ on I, which implies that q is decreasing on I. This contradicts equation $\beta q = a$ for all $t \in I$.

(ii) Suppose on the contrary to the claim that there exists an interval $I \subset [0, T]$ such that $\dot{\gamma} = 0$ and $\gamma \neq 0$ for every $t \in I$. Then it holds on the interval I that

$$\gamma = -\frac{D'(N)}{\delta + \rho}, \tag{A.30}$$

$$\dot{a} = 0, \tag{A.31}$$

$$\dot{q} = \frac{-\rho\lambda + \rho\alpha\mu}{C''(q)} < 0. \tag{A.32}$$

Since $\dot{\gamma} = 0$ and (A.30) holds, $\dot{N} = 0$. This implies that $\beta q - a = \delta N$ for every $t \in I$. Differentiating this with respect to time yields $\beta\dot{q} - \dot{a} = 0$. This is equivalent with $\beta\dot{q} = 0$, since (A.31) holds. But this contradicts (A.32).

A.8 Proof of Lemma 2

(i) Suppose on the contrary to the claim that such a time instant does not exists, that is, suppose that $a(t) = 0$ for every $t \in (t^0, \hat{t})$, where $|\hat{t} - t^0| < \epsilon$ for any small $\epsilon > 0$. By the definition of $\hat{t}$, $N(t) > 0$ on the interval $(t^0, \hat{t})$ for sufficiently small ϵ. Let $t^1 \in (t^0, \hat{t})$. Then we obtain that

$$N(\hat{t}) = \int_0^{\hat{t}} q(\tau)e^{-\delta(\hat{t}-\tau)}\,\mathrm{d}\tau - \int_0^{\hat{t}} a(\tau)e^{-\delta(\hat{t}-\tau)}\,\mathrm{d}\tau \tag{A.33}$$

$$= \int_0^{\hat{t}} q(\tau)e^{-\delta(\hat{t}-\tau)}\,\mathrm{d}\tau - \int_0^{t^1} a(\tau)e^{-\delta(t^1-\tau)}\,\mathrm{d}\tau, \tag{A.34}$$

by the above assumption about abatement. Because $q(t) \geq 0$, we obtain that $N(\hat{t}) \geq N(t^1)$. However, the definition of $\hat{t}$ implies that $N(t^1) > N(\hat{t})$. These yield a contradiction $N(\hat{t}) > N(\hat{t})$.

(*ii*) Since $\gamma(t) < 0$ for every $t \in (t^+, \hat{t})$, $\gamma(\hat{t}) \leq 0$ by the continuity of the costate. The solution to (35) is

$$\gamma(t) = \int_0^t D'(N(\tau))e^{-(\delta+\rho)(\tau-t)} \, \mathrm{d}\tau + e^{(\delta+\rho)t}\gamma(0). \tag{A.35}$$

This can be used to rewrite $\gamma(\hat{t}) \leq 0$ as

$$\begin{aligned} \gamma(\hat{t}) &= \int_0^{\hat{t}} D'(N(\tau))e^{-(\delta+\rho)(\tau-\hat{t})} \, \mathrm{d}\tau + e^{(\delta+\rho)\hat{t}}\gamma(0) && \text{(A.36)} \\ &= e^{(\delta+\rho)\hat{t}}\left[\int_0^{\hat{t}} D'(N(\tau))e^{-(\delta+\rho)\tau} \, \mathrm{d}\tau + \gamma(0)\right] && \text{(A.37)} \\ &\leq 0. && \text{(A.38)} \end{aligned}$$

Let $t \geq \hat{t}$. Then

$$\begin{aligned} \gamma(t) &= \int_0^t D'(N(\tau))e^{-(\delta+\rho)(\tau-t)} \, \mathrm{d}\tau + e^{(\delta+\rho)t}\gamma(0) && \text{(A.39)} \\ &= e^{(\delta+\rho)t}\left[\int_0^t D'(N(\tau))e^{-(\delta+\rho)\tau} \, \mathrm{d}\tau + \gamma(0)\right] && \text{(A.40)} \\ &= e^{(\delta+\rho)t}\left[\int_0^{\hat{t}} D'(N(\tau))e^{-(\delta+\rho)\tau} \, \mathrm{d}\tau + \gamma(0)\right] && \text{(A.41)} \\ &\leq 0, && \text{(A.42)} \end{aligned}$$

where the last equality follows from assumption $D'(0) = 0$, since after $\hat{t}$ we know that $N(t) = 0$ by definition, and the inequality follows from (A.37)-(A.38).

A.9 Proof of Lemma 3

(*i*) Since $N(t) = 0$ for all $t \in [\hat{t}, T]$, it must be that $\beta q(t) = a(t)$ there. Assume contrary to the claim that there exists $I \subset [\hat{t}, T]$ such that $a(t) > 0$ on I. Then $\gamma(t) < 0$ on I and therefore $\dot{a} > 0$ on I, because $\dot{\gamma} = (\delta + \rho)\gamma < 0$ on I. But by $p - C'(q) = \lambda - \mu\alpha - \gamma\beta$, $\dot{q} < 0$ on I. This contradicts condition

$$\beta q(t) = a(t) \quad \text{for all} \quad t \in [\hat{t}, T]. \tag{A.43}$$

(*ii*) Let $t \geq \hat{t}$. Part (*ii*) of Lemma 2 says that $\gamma(t) \leq 0$. Part (*i*) of the current Lemma says that $a(t) = 0$. Condition (29) and assumption $A'(0) = 0$ give then that $-\gamma(t) \leq 0$. These imply $\gamma(t) = 0$ for all $t \in [\hat{t}, T]$.

A.10 Proof of Proposition 9

The proof has similarities with the proof of Proposition 1 in Tahvonen (1997).

Note first that $\dot{\gamma}(t) > 0$ for all $t \in [0, \hat{t})$ is not true, since $\dot{\gamma}(0) = D'(N(0)) + (\delta + \rho)\gamma(0) = (\delta + \rho)\gamma(0) < 0$.

Recall that $\gamma(T) = 0$ and suppose that the time path has an inverted U-shape somewhere. There must exist a closed interval $I = [t_1, t_2]$, where $\dot{\gamma} < 0$ on the interior of I and $\dot{\gamma}(t_1) = \dot{\gamma}(t_2) = 0$. This implies that $\dot{q} < 0$ and $\dot{a} \geq 0$ on I, so that

$$\beta\dot{q} - \dot{a} < 0 \quad \text{for all} \quad t \in I. \tag{A.44}$$

Note that $\ddot{\gamma}(t_1) \leq 0$ and $\ddot{\gamma}(t_2) \geq 0$. By Equation (35), $\ddot{\gamma}(t) = D''(N(t))\dot{N}(t)$ at $t = t_1$ and at $t = t_2$, implying that

$$\dot{N}(t_1) \leq 0 \quad \text{and} \quad \dot{N}(t_2) \geq 0. \tag{A.45}$$

The derivative $\ddot{\gamma} = D''(N(t))\dot{N}(t) + (\delta + \rho)\dot{\gamma}$ is positive somewhere on the interval $[t_1, t_2]$, implying that $\dot{N}(t) > 0$ there, since $\dot{\gamma} < 0$ on (t_1, t_2). Then there exists $t_3 \in [t_1, t_2)$ such that $\dot{N}(t_3) = 0$. But then

$$\ddot{N}(t_3) = \beta\dot{q}(t_3) - \dot{a}(t_3) \geq 0, \tag{A.46}$$

which is a contradiction with (A.44). Hence the time path is U-shaped.

A.11 Proof of Proposition 10

We argue that $a(T) = 0$. Assume on the contrary that $a(T) > 0$. Then $\gamma(T) < 0$. However, condition (38) together with $N(T) = 0$ and $R'(0) = 0$ imply $\gamma(T) \geq 0$, which contradicts $\gamma(T) < 0$. Hence $a(T) = 0$. Since $\hat{t} = T$, $N(t) > 0$ for all $t \in (T - \epsilon, T)$ for some (small) $\epsilon > 0$. Note that

$$\dot{N}(T) = \beta q(T) - a(T) - \delta N(T) = \beta q(T) - a(T), \tag{A.47}$$

which is non-positive at optimum. Since $\beta q(T) \leq a(T) = 0$, we have that $q(T) = 0$. By Lemma 2, we have $a(t) > 0$ for all $t \in (T - \epsilon, T)$. For these instants of time, $\gamma(t) < 0$, implying that $\gamma(T) \leq 0$. This and $\gamma(T) \geq 0$ obtained using condition (38) as above, imply that $\gamma(T) = 0$. Thus, when $\hat{t} = T$ and $N(T) = 0$, the shadow value of the pollution stock, γ, has a U-shaped time path by the same principles as in the proof of Proposition 9.

A.12 Proof of Proposition 11

Note that in this case condition (38) implies $-\gamma(T) = R'(N(T))$. Since the right-side is strictly positive, we obtain $\gamma(T) < 0$. This implies together with

$$\gamma(T) = \int_0^T D'(N(\tau))e^{-(\delta+\rho)(\tau-T)}\, \mathrm{d}\tau + e^{(\delta+\rho)T}\gamma(0), \tag{A.48}$$

that also $\gamma(0) < 0$. By (35) and assumption $D'(0) = 0$,

$$\dot{\gamma}(0) = D'(N(0)) + (\delta + \rho)\gamma(0) = (\delta + \rho)\gamma(0) < 0. \tag{A.49}$$

Note that $\dot{\gamma}(0) < 0$ excludes the possibility that $\dot{\gamma}(t) > 0$ for all $t \in [0, T]$. This leaves four possibilities:

1. $\dot{\gamma}(t) < 0$ for all $t \in [0, T]$;

2. $\dot{\gamma}(t) < 0$ for all $t \in [0, T]$ except at an isolated time instant at which $\dot{\gamma}(t) = 0$;

3. The time path is U-shaped;

4. The time path is first decreasing, then increasing and lastly decreasing again.

There are no other possibilities: Suppose on the contrary that the time path has three points with $\dot{\gamma}(t) = 0$. Then the path has an inverted U-shape somewhere, and the same arguments as in the proof of Proposition 9 produce a contradiction.

A.13 Proof of Proposition 12

(i) Assume on the contrary to the claim that $t^L \geq t^H$. Given the time invariant taxes and strictly positive extraction rates, the necessary conditions of the mining firms' problems include the following conditions:

$$p - C'(q^H(t)) = \lambda^H(t) + \alpha^H v_S + \beta^H v_N, \tag{A.50}$$

$$p - C'(q^L(t)) = \lambda^L(t) + \alpha^L v_S + \beta^L v_N. \tag{A.51}$$

Since both mines are exhausted and $t^L \geq t^H$, the time paths of the mines' extraction rates must cross at least once.

Consider first the possibility that $t^L > t^H$. Since there exists a smallest time instant $\hat{t}$ such that $q^H(\hat{t}) = q^L(\hat{t})$, equations (51), (A.50) and (A.51) imply that

$$\lambda^H(\hat{t}) - \lambda^L(\hat{t}) = -K < 0. \tag{A.52}$$

Note that the function $t \mapsto \lambda^H(t) - \lambda^L(t)$ is then an exponentially decreasing function, obtains strictly negative values, and

$$\lambda^H(t) - \lambda^L(t) \in (-K, 0) \quad \text{for all} \quad t \in [0, \hat{t}). \tag{A.53}$$

This and equations (A.50) and (A.51) imply

$$-C'(q^H(t)) + C'(q^L(t)) = \lambda^H(t) - \lambda^L(t) + K > 0 \quad \text{for all} \quad t \in [0, \hat{t}). \tag{A.54}$$

Hence

$$q^L(t) > q^H(t) \quad \text{for all} \quad t \in [0, \hat{t}). \tag{A.55}$$

If $\hat{t}$ is unique, then $q^L(t) > q^H(t)$ for all t except at $t = \hat{t}$. This clearly implies that the total extraction for mine L is strictly greater than x_0, which is impossible. Suppose that $\hat{t}$ is not unique. Then Equation (A.52) holds for at least two time instants, which contradicts the fact that the function $\lambda^H(t) - \lambda^L(t)$ decreases exponentially.

Consider the possibility that $t^L = t^H$. Note that continuity implies that equations (A.50) and (A.51) hold also at t^L. Because $t^L = t^H$ and $t^H < T$, there are two possibilities: either the extraction paths cross at some instant of time smaller than t^L, or the paths don't cross. If they cross, then Equation (A.52) holds for two time instants, which again contradicts the fact that the function $\lambda^H(t) - \lambda^L(t)$ decreases exponentially. If they don't cross, then one of the mines is not exhausted, which is a contradiction.

(ii) Assume on the contrary to the claim that $t^L > t^H$ or $t^L < t^H$. First, let $t^L > t^H$. Following the reasoning of Part *(i)*, Equation (A.52) implies that $\lambda^H(\hat{t}) = \lambda^L(\hat{t})$. Hence $q^L(t) = q^H(t)$ for all t. But $q^L(t)$ is strictly positive on $[t^H, t^L)$. These imply that the total extraction under profile $q^L(t)$ is strictly greater than x_0, which is a contradiction. The proof for the other case, $t^L < t^H$, is similar.

A.14 Proof of Proposition 13

The proof is divided into two cases, one with $t^H < T$ and the other with $t^H = T$. Consider the first possibility. The proof of Proposition 12 shows that $t^L < t^H$ and that there exists $\hat{t}$ such that $q^L(\hat{t}) = q^H(\hat{t})$. There is only one such time instant, since otherwise $\lambda^H(t) - \lambda^L(t)$ would not decrease exponentially. Since $t^L < t^H$, $q^L(t) < q^H(t)$ for all $t \in (\hat{t}, t^H]$. This implies that $q^L(t) > q^H(t)$ for all $t \in [0, \hat{t})$.

Consider then the second possibility. If $t^H = T$, then either $t^L < T$ or $t^L = T$. If $t^L < T$, the extraction profiles of the mines must cross, since otherwise mine L is not exhausted. This can occur only once, since otherwise the function $\lambda^H(t) - \lambda^L(t)$ does not evolve exponentially. Then it must be that the extraction path q^L crosses path q^H from the above. Next it is shown that if $t^L = T$, then $q^H(T) > q^L(T)$. To see this, suppose $q^H(T) \leq q^L(T)$, and note that q^H cannot always equal q^L, since $K > 0$. Hence the paths are not the same, but they nevertheless must cross. As in the proof of Proposition 12, there exists $\hat{t}$ such that $q^L(t) > q^H(t)$ for all $t \in [0, \hat{t})$. But since $q^H(T) \leq q^L(T)$, the paths cross twice, which again means that the function $\lambda^H(t) - \lambda^L(t)$ does not decrease exponentially. Hence $q^H(T) > q^L(T)$, and the extraction path q^L crosses path q^H from the above.

References

Ambec, Stefan and Jessica Coria. 2013. “Prices vs Quantities with Multiple Pollutants.” *Journal of Environmental Economics and Management* 66:123–140.

Amigues, Jean-Pierre and Michael Moreaux. 2013. “The Atmospheric Carbon Resilience Problem: A Theoretical Analysis.” *Resource and Energy Economics* 35:618–636.

Benchekroun, Hassan, Alex Halsema, and Cees Withagen. 2009. “On Nonrenewable Resource Oligopolies: The Asymmetric Case.” *Journal of Economic Dynamics & Control* 33:1867–1879.

Boukas, El-Kebir, Alain Haurie, and Phillippe Michel. 1990. “An Optimal Control Problem with a Random Stopping Time.” *Journal of Optimization Theory and Applications* 64:471–480.

Cairns, Robert D. 2001. “Capacity Choice and the Theory of the Mine.” *Environmental and Resource Economics* 18:129–148.

Cairns, Robert D and Takayoshi Shinkuma. 2003. “The Choice of the Cutoff Grade in Mining.” *Resources Policy* 29:75–81.

Campbell, Harry F. 1980. “The Effect of Capital Intensity on the Optimal Rate of Extraction of a Mineral Deposit.” *Canadian Journal of Economics* 13:349–356.

Caputo, Michael R. 1990. “A Qualitative Characterization of the Competitive Nonrenewable Resource Extractive Firm.” *Journal of Environmental Economics and Management* 18:206–226.

Caputo, Michael R and James E Wilen. 1995. “Optimal Cleanup of Hazardous Wastes.” *International Economic Review* 36:217–243.

Chakravorty, Ujjayant and Darrell L Krulce. 1994. "Heterogenuous Demand and Order of Resource Extraction." *Econometrica* 62:1445–1452.

Chakravorty, Ujjayant, Darrell L Krulce, and James Roumasset. 2005. "Specialization and non-renewable Resources: Ricardo Meets Ricardo." *Journal of Economic Dynamics & Control* 29:1517–1545.

Clarke, Harry R and William J Reed. 1994. "Consumption/pollution Tradeoffs in an Environment Vulnerable to Pollution-related Catastrophic Collapse." *Journal of Economic Dynamics & Control* 18:991–1010.

Dold, Bernhard. 2014. "Evolution of Acid Mine Drainage Formation in Sulphidic Mine Tailings." *Minerals* 4:621–641.

Farzin, Y H. 1992. "The Time Path of Scarcity Rent in the Theory of Exhaustible Resources." *The Economic Journal* 102:813–830.

———. 1996. "Optimal Pricing of Environmental and Natural Resource Use with Stock Externalities." *Journal of Public Economics* 62:31–57.

Farzin, Y H and Olli I Tahvonen. 1996. "Global Carbon Cycle and the Optimal Time Path of a Carbon Tax." *Oxford Economic Papers* 48:515–536.

Hoel, Michael and Snorre Kverndokk. 1996. "Depletion of Fossil Fuels and the Impacts of Global Warming." *Resource and Energy Economics* 18:115–136.

Holland, Stephen P. 2003. "Extraction Capacity and the Optimal Order of Extraction." *Journal of Environmental Economics and Management* 45:569–588.

Krautkraemer, Jeffrey A. 1988. "The Cut-Off Grade and the Theory of Extraction." *Canadian Journal of Economics* 21:146–160.

Lèbre, Éléonore, Glen D Corder, and Artem Golev. 2017. “Sustainable Practices in the Management of Mining Waste: A Focus on the Mineral Resource.” *Minerals Engineering* 107:34–42.

Lottermoser, Bernd G. 2010. *Mine Wastes: Characterization, Treatment and Environmental Impacts*. Springer, 3. ed.

Lozada, Gabriel A. 1993. “The Conservationist’s Dilemma.” *International Economic Review* 34:647–662.

Michielsen, Thomas O. 2014. “Brown Backstops Versus the Green Paradox.” *Journal of Environmental Economics and Management* 68:87–110.

Moreaux, Michel and Cees Withagen. 2015. “Optimal Abatement of Carbon Emission Flows.” *Journal of Environmental Economics and Management* 74:55–70.

Moslener, Ulf and Till Requate. 2007. “Optimal Abatement in Dynamic Multi-pollutant Problems when Pollutants can be Complements or Substitutes.” *Journal of Economic Dynamics & Control* 31:2293–2316.

Mudd, Gavin M. 2010. “The Environmental Sustainability of Mining in Australia: Key Mega-trends and Looming Constraints.” *Resources Policy* 35:98–115.

Nordstrom, Darrell K. 2011. “Mine Waters: Acidic to Circumneutral.” *Elements* 7:393–398.

Roan, Philip F and Wade E Martin. 1996. “Optimal Production and Reclamation at a Mine Site with an Ecosystem Constraint.” *Journal of Environmental Economics and Management* 30:186–198.

Schumacher, Ingmar. 2011. “When Should We Stop Extracting Non-renewable Resources?” *Macroeconomic Dynamics* 15:495–512.

Seierstad, Atle and Knut Sydsæter. 1987. *Optimal Control Theory with Economic Applications.* North-Holland, 2. ed.

Shinkuma, Takayoshi. 2000. “A generalization of the Cairns–Krautkraemer Model and the Optimality of the Mining Rule.” *Resource and Energy Economics* 22:147–160.

Sinn, Hans-Werner. 2008. “Public Policies Against Global Warming: A Supply Side Approach.” *International Tax and Public Finance* 15:360–394.

Statistics Canada. 2012. “Human Activity and the Environment: Waste Management in Canada.” `http://www.statcan.gc.ca/pub/16-201-x/16-201-x2012000-eng.pdf`. [Online; accessed 5-April-2017].

Stollery, Kenneth R. 1985. “Environmental Controls in Extractive Industries.” *Land Economics* 61:136–144.

Sullivan, Jay and Gregory S Amacher. 2009. “The Social Costs of Mineland Restoration.” *Land Economics* 85:712–726.

Tahvonen, Olli I. 1997. “Fossil Fuels, Stock Externalities, and Backstop Technology.” *Canadian Journal of Economics* 30:855–874.

Tiess, Günter. 2010. “Minerals Policy in Europe: Some Recent Developments.” *Resources Policy* 35:190–198.

Tilton, John E and Juan Ignacio Guzmán. 2016. *Mineral Economics and Policy.* RFF Press, 1. ed.

Tsur, Yacov and Amos Zemel. 2003. “Optimal Transition to Backstop Substitutes for Nonrenewable Resources.” *Journal of Economic Dynamics & Control* 27:551–572.

Tukes. 2014. "Tilastotietoja vuoriteollisuudesta 2014." http://www.tukes.fi/Tiedostot/kaivokset/tilastot/Tilastotiedot_vuoriteollisuus_2014 [Statistics from Finnish Safety and Chemicals Agency TUKES. In Finnish; Online; accessed 5-April-2017].

Ulph, Alistair and David Ulph. 1994. "The Optimal Time Path of a Carbon Tax." *Oxford Economic Papers* 46:857–868.

van der Meijden, Gerard, Frederick van der Ploeg, and Cees Withagen. 2015. "International Capital Markets, Oil Producers and the Green Paradox." *European Economic Review* 76:275–297.

van der Ploeg, Frederick and Cees Withagen. 2012. "Is There Really a Green Paradox?" *Journal of Environmental Economics and Management* 64:342–363.

White, Ben, Graeme J Doole, David J Pannell, and Veronique Florec. 2012. "Optimal Environmental Policy Design for Mine Rehabilitation and Pollution with a Risk of Non-compliance Owing to Firm Insolvency." *The Australian Journal of Agricultural and Resource Economics* 56:280–301.

Winter, Ralph A. 2014. "Innovation and the Dynamics of Global Warming." *Journal of Environmental Economics and Management* 68:124–140.