CESifo Working Papers
www.cesifo.org/wp

Optimal Emission-Extraction Policy in a World of Scarcity and Irreversibility

Fabien Prieur
Mabel Tidball
Cees Withagen

CESifo Working Paper No. 3512
Category 9: Industrial Organisation
Original Version: July 2011
This Version: May 2012

CESifo *Working Paper No. 3512*

Optimal Emission-Extraction Policy in a World of Scarcity and Irreversibility

Abstract

This paper extends the classical exhaustible-resource/stock-pollution model with irreversibility of pollution decay, meaning that after reaching some threshold there is no decay of the pollution stock. Within this framework, we answer the question how the potential irreversibility of pollution affects the extraction path. We investigate the conditions under which the economy will optimally adopt a reversible policy, and when it is optimal to enter the irreversible region. In the case of irreversibility it may be optimal to leave a positive amount of resource in the ground forever. As far as the optimal extraction/emission policy is concerned, several types of solutions may arise, including solutions where the economy stays at the threshold for a while. Given that different programs may satisfy the first order conditions for optimality, we further investigate when each of these is optimal. The analysis is illustrated by means of a calibrated example. To sum up, for any pollution level, we can identify a critical resource stock such that there exist multiple optima i.e. a reversible and an irreversible policy that yield exactly the same present value. For any resource stock below this critical value, the optimal policy is reversible whereas with large enough resources, irreversible policies outperform reversible programs. Finally, the comparison between irreversible policies reveals that it is never optimal for the economy to stay at the threshold for a while before entering the irreversible region.

JEL-Code: Q300, Q530, C610.

Keywords: non-renewable resource, irreversible pollution, optimal policy.

Fabien Prieur
LAMETA / University of Montpellier I
& INRA
Montpellier / France
prieur@supagro.inra.fr

Mabel Tidball
LAMETA / University of Montpellier I
& INRA
Montpellier / France
tidball@supagro.inra.fr

Cees Withagen
VU University of Amsterdam
& Tinbergen Institute
Amsterdam / The Netherlands
cwithagen@feweb.vu.nl

The authors like to thank the participants of the Workshop in honour of Gérard Gaudet in Montreal in April 2011, and in particular Jean-Pierre Amigues, Pierre Lasserre and Michel Moreaux for their stimulating remarks.

1 Introduction

In the design of optimal climate policy it should be taken into account that most emissions of CO2 result from burning fossil fuels, which originate from non-renewable resources. It is forcefully argued by e.g., D'Arge and Kogiku (1973) that "the 'pure' mining problem must be coupled with the 'pure' pollution problem". Although not really applicable to the climate change problem they also state that questions like these become relevant: which should we run out first, air to breathe or fossil fuels to pollute the air we breathe? These questions have attracted environmental and resources economists' attention for a long time. An early prototype model can be found in Withagen (1994) where utility is derived from consumption of fossil fuel from a non-renewable resource and where the use of the fossil fuel also contributes to the accumulation of CO2. The accumulated stock of CO2 causes damage, represented by a convex damage function. One of the findings is that the optimal extraction path of the fossil fuel becomes flatter than in the absence of environmental damage. In a similar model Sinclair (1994) and Ulph and Ulph (1994) derive the optimal carbon tax needed to implement the first-best optimum. Still focusing on the exhaustible-resource/stock-pollution model, Tahvonen (1997) fully characterizes the properties of the optimal extraction/emission policy in the case of exponential decay of pollution. He notably shows that, since extraction and pollution necessarily converge toward zero, in the long run, the pollution problem does not have an influence on the total amount of the resource extracted over the planning period. In that sense, the pollution and the resource management problems are independent of each other.

A widely used alternative to capturing environmental damage by a damage function is to impose a ceiling on the total accumulated stock of CO2. Examples of this approach are Chakravorty et al. (2006, 2008). This is usually motivated in the following ways. First one could argue that a ceiling is a political reality. International negotiations are indeed aiming at keeping the temperature rise below 2 degrees ^{0}C and it is widely accepted that in order to accomplish this the CO2 concentration should be no more than 450 ppmv. Second, and related to the first argument, taking a damage function rather than a ceiling, may lead to highly undesirable outcomes because there is substantial uncertainty surrounding the effects of climate change and catastrophes may occur (see e.g., Tsur and Zemel, 2008). With a ceiling, one considers that there is no pollution damage as long as the pollution stock remains below or equal to this critical level whereas the damage occurs, with potentially huge repercussions on the economy, once the ceiling is exceeded. As a consequence, it is never optimal to let pollution accumulate beyond the ceiling in this framework. In our model, we also have an irreversible event related to the crossing of a critical pollution stock. But this event cannot be qualified as a catastrophe from the point of view of the economy since it only affects the natural regeneration of ecosystems.

It is well known that the climate system[1] is extremely complex and that economists' modelling of it is rather rudimentary. The focus of the present contribution is on irreversibility beyond a certain pollution level, but in a different sense. Usually the decay of pollution is modeled as linear, meaning that a constant percentage of the existing stock is diluted per unit of time. This approach has been criticized by many authors including Dasgupta (1982), Fiedler (1992), and Pethig (1993). Representing decay as a constant fraction of the existing stock is far too simplistic and with a large stock of CO2 the absorption capacity of oceans and forests may be reduced considerably. This is what we want to capture by introducing the ceiling or irreversibility threshold. Indeed, experts of the second working group of the IPCC (2007) have identified positive climate feedbacks due to emissions of greenhouse gases (GHG). There is more and more evidence that increasing emission levels and concentrations of GHG disturb the regeneration capacity of natural ecosystems. Oceans, that form the most important carbon sink, display a buffering capacity that begins to get saturated. At the same time, the assimilation capacity of terrestrial ecosystems (lands and forests form the other important carbon sink) will likely peak by mid-century and then decline to become a net source of carbon by the end of the present century. Therefore, the irreversible degradation of the assimilative capacity of Nature does not seem so distant from today.

In economics alternative specifications have been proposed by e.g., Forster (1975). They usually allow for inverted-U shaped decay with the important feature that there exists a critical threshold of pollution above which the assimilation capacity of Nature becomes permanently exhausted, thereby implying an irreversible concentration of pollution. In this way a ceiling is introduced, not on the allowed stock of pollutants but on the stock of pollutants that allows for decay. The decision maker is then faced with the problem whether it is optimal to stay below the ceiling and benefit from decay or going beyond it, because of higher consumption, and then stay in the irreversible region.

Forster (1975) is the very first one who breaks with the assumption of a constant decay rate. He claims that the inverted-U shaped decay may give rise to a multiplicity of solutions but does not deal with the technical issue of non-convexity. This is done by Tahvonen and Withagen (1996) who provide a detailed analysis of the impact of irreversibility of decay on the optimal control of pollution. The inverted-U shaped decay function introduces

[1]We focus on climate change, but one could argue that all ecosystems are complex and fragile. Cropper (1976) has been one of the first contributions in economics to consider the consequences. See Polasky et al. (2011) for a more recent account.

a non-convexity and they show that this may result in the existence of multiple paths satisfying the necessary conditions, starting from the same initial stock values.

The present paper adds to the contribution by Tahvonen and Withagen by introducing the irreversibility of decay in the classical exhaustible-resource/stock-pollution model, without extraction costs and backstop technology. In contrast with Tahvonen and Withagen we take exhaustibility explicitly into account. We also build on Tahvonen (1997) who deals with exhaustibility, but has linear decay. Our approach to modelling irreversibility is to assume that the decay rate is linear for levels of pollution up to some critical level, after which decay is zero and remains zero. This way of modeling decay differs from the quadratic approach adopted by Tahvonen and Withagen (1996). We do so in order to capture the fact that the regenerative capacity may vanish abruptly, and not in a smooth way. Within this framework, the first question is how the potential irreversibility of pollution affects the extraction path. Our aim is also to emphasize the conditions under which the economy will optimally adopt an irreversible or a reversible policy. Our work is also related to a recent contribution by Amigues and Moreaux (2012) who consider a less drastic case, where decay doesn't become zero after a certain threshold, but the decay rate stays positive, but at a lower level. Amigues and Moreaux, however, don't take exhaustibility into account.

Once the situation has turned irreversible, we show that the pollution problem does affect the total amount of resource extracted. In particular, it may be optimal to leave a positive amount of resource in the ground forever. As far the optimal extraction/emission policy is concerned, several types of solutions may arise. We derive a simple condition that guarantees that it is optimal to stay in what is called the reversible region. When this condition does not hold, it is difficult to conclude whether the optimal policy is reversible or irreversible. But we are able to characterize all programs that satisfy the necessary conditions. Moreover we provide some intuition for the optimal choice to be made. We also have a rudimentary non-convexity and it is one of the aims of the present paper to investigate the occurrence of multiplicity in our model. We illustrate our findings using a numerical and calibrated example.

With our numerical example we are able to divide the set of initial conditions, for pollution and the resource stock, into several regions, each associated with one or more optimality candidates. When the initial resource stock is low enough, the upper bound set on the resource stock becoming lower as initial pollution increases, optimal solution is reversible and may feature a period of time staying at the threshold. For larger resource stocks, irreversible and reversible policies co-exist. We illustrate the existence of multiple candidates for optimality. Computation of present values reveals that for any pollution level, the reversible policy yields the highest value for low to medium resource stocks, whereas it is optimal to follow an irreversible policy when the initial endowment of the resource is high enough. In addition, among the irreversible policies, the ones immediately entering the irreversible region always dominate those that stay at the threshold for a period of time. More importantly, for any initial pollution level, one can find the corresponding resource stock such that reversible and irreversible policies yield exactly the same present value. Therefore, we show the existence of multiple optima. This result echoes Tahvonen and Withagen (1996)'s findings, in their model with a quadratic decay and abundant resource. It also raises the question of how a policy maker should decide between these two policy alternatives with so distinct features.

We will also pay attention to the type of decay function employed by Tahvonen and Withagen (1996) in order to see what the effect is of exhaustibility in the context of their model. Since we cannot have sustained extraction at a positive level, it should be clear that, the results differ to a large extent. The crucial difference between the two approaches is that close to the threshold there is large decay in our model, whereas in theirs decay is vanishing. This implies that with quadratic decay it is never optimal to stay at the threshold for a while, neither before returning to the reversible region, nor before entering the irreversible region. However, for low initial fossil fuel stocks the difference are not striking.

The paper is organized as follows. In section 2 we present the formal model. Section 3 characterizes the optimality candidates, reversible as well as irreversible. In section 4 we derive the optimum for a calibrated example and provide some economic intuition about when irreversibility plays a role. Section 5 compares our results with those obtained by Tahvonen and Withagen (1996). Section 6 concludes.

2 The model

We consider a partial equilibrium representation of the global warming problem. Our carbon economy is described by the following set of assumptions. The economy produces one final good. Production of the good is denoted by y. We assume that it is produced with a technology that uses a nonrenewable natural resource of fossil fuels, in such a way that final production equals the rate of extraction. Extraction is costless. Emissions are a one-to-one by-product of production. So they equal y as well. The resource stock x follows the usual law:

$$\dot{x}(t) = -y(t) \text{ with } x(0) = x_0 \text{ given, } x(t) \geq 0 \text{ and } y(t) \geq 0 \quad (1)$$

Hence $x(t) = x_0 - \int_0^t y(u)du$. Let $U(y)$ be the utility derived from consumption of the good. In the same vein as Tahvonen and Withagen (1996) we make:

Assumption 1. *The utility function is such that:* $U(0) = 0$, $U''(y) < 0$, $0 < U'(0) < \infty$ *and there exists* $\bar{y}$ such that $U'(\bar{y}) = 0$.

Remark. The utility function can also be understood as a profit function: $U(y) = py - c(y)$ with p a constant and exogenous price and $c(y)$ a convex production cost function.

Emissions contribute to the accumulation of a pollution stock, z. The initial pollution stock is given: z_0. Pollution accumulation is not innocuous to ecosystems and, in particular, it affects their capacity to regenerate. We assume that pollution turns irreversible if the stock passes a critical threshold $\bar{z}$. We suppose $\bar{z} > z_0$. This irreversibility threshold is known to the policy-maker. We do not consider any uncertainty surrounding $\bar{z}$. To account for irreversibility, the dynamics of the stock of the pollutant are defined piece-wise:

$$\dot{z}(t) = \begin{cases} y(t) - \alpha z(t) & \text{if } z(t) \leq \bar{z} \\ y(t) & \text{else} \end{cases} \tag{2}$$

The natural regeneration or assimilation rate α is constant and positive as long as accumulated emissions are not too high, that is, as long as the stock remains below or is at the irreversibility threshold $\bar{z}$. Once the threshold is surpassed, a new stage is entered where the regeneration capacity is completely and permanently vanished. Thus, pollution becomes irreversible. In section 5 we will consider an alternative specification where decay is not exponential up to the ceiling, but is inverted U-shaped, the case studied by Tahvonen and Withagen (1996), in order to assess the differences between their model and ours. Hereafter, the domain where $z \leq \bar{z}$ is called the reversible region whereas whenever pollution is higher than $\bar{z}$, the economy is said to be in the irreversible region. For the solution it matters whether the reversible region is an open or a closed set. We assume the latter.

Pollution is damaging to the economy. For any level z, pollution damage is denoted by $D(z)$.

Assumption 2. *The damage function is such that:* $D(0) = 0$, $D'(z) > 0$, $D''(z) > 0$ *for all* $z > 0$, $D'(0) = 0$ *and* $\lim_{z \to \infty} D'(z) = \infty$.

The social welfare function reads

$$V = \int_0^\infty e^{-\delta t} \left[U(y(t)) - D(z(t)) \right] dt \tag{3}$$

with δ the discount rate.

3 Optimality candidates

We make a distinction between reversible and irreversible solutions, according to the behavior of the pollution stock. A solution is said to be reversible if the pollution stock remains within the reversible region forever: $z(t) \leq \bar{z}$ for all $t \geq 0$. By contrast, a solution is irreversible if the pollution stock enters the irreversible region in finite time. In this section we provide a first sketch of the characteristics of these solutions by outlining the optimization programs that correspond to the candidates.

A reversible policy solves the problem of maximizing (3) subject to (1), $\dot{z}(t) = y(t) - \alpha z(t)$, and $z(t) \leq \bar{z}$. The latter condition is a pure state constraint. The current value Lagrangian reads $L(y, z, \lambda, \mu) = U(y) - D(z) - \lambda(y - \alpha z) - \mu y + \kappa(\bar{z} - z)$. The co-state λ is to be interpreted as the cost of pollution and μ is the in situ value of the resource stock. The Lagrange multiplier κ is associated with the pure state constraint. The necessary conditions include (omitting the time argument when there is no danger of confusion)

$$\begin{array}{l} U'(y) - \lambda - \mu \leqq 0,\ (U'(y) - \lambda - \mu)y = 0,\ y \geqq 0 \\ \dot{\mu} = \delta\mu \\ \dot{\lambda} = (\delta + \alpha)\lambda - D'(z) - \kappa,\ \kappa \geq 0,\ \kappa(\bar{z} - z) = 0 \\ \lim_{t\to\infty} e^{-\delta t}(\lambda(t)z(t) + \mu(t)x(t)) = 0 \end{array} \tag{4}$$

Irreversible policies solve a two-stage optimal control problem. Suppose that from a given $z_0 < \bar{z}$, it is optimal to reach the threshold in finite time and to go beyond $\bar{z}$ immediately afterwards. Fix, for the time being the date of the transition T and the resource stock at the transition: x_T. Then the optimal path can be decomposed into two parts, each solving an optimization problem. The first problem reads

$$\max \int_0^T e^{-\delta t}(U(y) - D(z))dt$$

subject to

$$\begin{array}{rcl} \dot{z} & = & y - \alpha z, z(0) = z_0, z(T) = \bar{z} \\ \dot{x} & = & -y,\ y(t) \geq 0,\ x(0) = x_0,\ x(T) = x_T \end{array}$$

For this problem we can neglect the pure state constraint and therefore the necessary conditions defined in (4) with $\kappa = 0$ are also valid here (except the transversality condition since this regime is not terminal).

The second problem reads

$$\max \int_T^\infty e^{-\delta t}(U(y) - D(z))dt$$

subject to

$$\begin{aligned} \dot{z} &= y, z(T) = \bar{z} \\ \dot{x} &= -y, \ y(t) \geq 0, \ x(T) = x_T \end{aligned}$$

For the second problem we get the same necessary conditions as in (4), with α put equal to zero. The pure state constraint can be neglected here as well.

The two optimization problems posed above have fixed T and x_T. It will be our task in the sequel to determine the optimal transition time T as well as the optimal resource stock to be left in situ at that instant of time.

There is another optimality candidate. In principle, it may be optimal to stay at the threshold for a non-degenerate period of time before entering the irreversible region. Also this possibility will be investigated in due course.

3.1 Irreversible solutions

In principle four types of irreversible programs exist. One distinction is between staying at the threshold for just one instant of time or for a non-degenerate interval of time. The other distinction is between full or partial exhaustion of the resource. In the present section we characterize each of these options. We show that partial exhaustion can only be optimal if the threshold is passed immediately after reaching it.

Before proceeding, we identify a set of conditions that need to be met for the existence of irreversible optimality candidates. Consider the problem of maximizing social welfare while neglecting the resource constraint as well as the possible irreversibility of decay. This problem has necessary conditions similar to (4) with $\mu = \kappa = 0$. Focus on interior solutions: $y(t) > 0$ for all t.

First, for our problem to be of interest, it should be physically feasible to reach the threshold. A necessary condition for this is that $\bar{y} > \alpha\bar{z}$: If the inequality would not hold then decay at the threshold level is just too high. Put differently, assume that the economy with abundant resource sets extraction to the maximum level for all t: $y(t) = \bar{y}$. Then, the pollution stock monotonically converges toward a level $\bar{\bar{z}} = \bar{y}/\alpha$. If $\bar{\bar{z}} < \bar{z} \Leftrightarrow \bar{y} < \alpha\bar{z}$ then it is clear that the threshold will not be reach by an economy submitted to resource scarcity.

Second, the threshold should be smaller than the steady state in a resource abundant economy that is not threatened by irreversibility. Indeed, it is easily seen that in this setting it is optimal for the economy to converge to (y^*, z^*) defined by

$$U'(\alpha z^*) = \frac{D'(z^*)}{\alpha + \delta}, \ y^* = \alpha z^* \tag{5}$$

The optimal policy consists in appropriately choosing initial emissions so that $(y(0), z_0)$ is located on the stable branch of the saddle point. Hence, when initial pollution is low, the initial extraction rate is high and decreases; pollution monotonically increases until the steady state is reached. For a high z_0, the economy starts with a low extraction rate, that is increasing towards the steady state whereas the pollution stock decreases. Returning to the problem with exhaustibility and irreversibility, we require $\bar{z} < z^*$. The reason is that if $\bar{z} \geq z^*$ and noting that $z_0 < \bar{z}$, the threshold $\bar{z}$ will never be exceeded in an optimum. To see this compare two economies: a resource abundant economy, denoted by hats ($\hat{x}_0 = \infty$) and a resource-poor economy ($x_0 < \infty$). Suppose at some instant of time T we have $\hat{z}(T) = z(T) = \bar{z} > z^*$ and the resource-poor economy enters the irreversible region: $y(T) > \alpha\bar{z}$. The resource rich economy necessarily has $\hat{y}(T) \leq \alpha\bar{z}$ because it monotonically converges to $\hat{z}^*$. Hence, from the combination of necessary optimality conditions, $U'(y(T)) = \lambda(T) + \mu(T) < U'(\hat{y}(T)) = \hat{\lambda}(T)$, so that $\lambda(T) < \hat{\lambda}(T)$. The necessary conditions also yield:

$$\hat{\lambda}(T) = e^{(\alpha+\delta)T} \int_T^\infty e^{-\alpha(\alpha+\delta)s} D'(\hat{z}(s))ds$$

and

$$\begin{aligned}
\lambda(T) &= e^{\delta T} \int_T^\infty e^{-\delta s} D'(z(s))ds \\
&= e^{(\alpha+\delta)T} \int_T^\infty e^{-(\alpha+\delta)s} e^{\alpha(s-T)} D'(z(s))ds \\
&> e^{(\alpha+\delta)T} \int_T^\infty e^{-(\alpha+\delta)s} D'(z(s))ds > \hat{\lambda}(T)
\end{aligned}$$

because $z(t) \geq \hat{z}(t)$ for all $t \geq T$. This is a contradiction. So, $\bar{z} < z^*$, which is equivalent to $U'(\alpha\bar{z}) > D'(\bar{z})/(\alpha + \delta)$, is also necessary for the existence of irreversible solutions.

Third, once the economy finds itself at the threshold, it should at least not be refrained, from a welfare point of view, from passing the threshold. This requires that the marginal utility of the first unit of consumption $U'(0)$ should be larger than total discounted marginal damages: $D'(\bar{z})/\delta$. To see this, focus now on the problem with irreversibility of decay but without resource scarcity. If an economy with abundant resource were to exceed the threshold in finite time, it would achieve in the long run a steady state level of pollution given by $U'(0) = D'(z^\infty)/\delta$. If $U'(0) \leq D'(\bar{z})/\delta$ then we obtain a contradiction because $z^\infty \leq \bar{z}$. So, it must hold that $U'(0) > D'(\bar{z})/\delta$.

With these assumptions in mind, the analysis of irreversible solutions proceeds as follows. These candidates solve an optimization program that can be decomposed into the two sub-problems presented above. In addition, they must satisfy an additional optimality condition related to the instant when pollution exceeds the threshold. Define the value functions, for the given T and x_T

$$\begin{aligned} V(T) &= \int_0^T e^{-\delta t}(U(y) - D(z))dt \\ \hat{V}(T) &= \int_T^\infty e^{-\delta t}(U(\hat{y}) - D(\hat{z}))dt \end{aligned}$$

Here the optimum of problem 2 is denoted by hats. According to Seierstad and Sydsaeter (1987, p. 213) we have for $T > 0$

$$\begin{aligned} \frac{\partial V}{\partial T} &= e^{-\delta T}[U(y(T)) - \lambda(T)(y(T) - \alpha\bar{z}) - \mu(T)y(T) - D(\bar{z})] \\ &= e^{-\delta T}[U(y(T)) - U'(y(T))y(T) + \lambda(T)\alpha\bar{z} - D(\bar{z})] \end{aligned}$$

$$\begin{aligned} -\frac{\partial \hat{V}}{\partial T} &= e^{-\delta T}[U(\hat{y}(T)) - \hat{\lambda}(T)\hat{y}(T) - \hat{\mu}(T)\hat{y}(T) - D(\bar{z})] \\ &= e^{-\delta T}[U(\hat{y}(T)) - U'(\hat{y}(T))\hat{y}(T) - D(\bar{z})] \end{aligned}$$

Hence, if an optimal T exists, the following equation is satisfied:

$$U(y(T)) - U'(y(T))y(T) + \lambda(T)\alpha\bar{z} = U(\hat{y}(T)) - U'(\hat{y}(T))\hat{y}(T) \tag{6}$$

Under assumption 1, the function $U(y) - U'(y)y$ is monotonically increasing. This implies that the entrance in the irreversible region is accompanied by an upward discontinuity in y. The economy compensates for the loss of benefits (from pollution decay) by an increase in consumption.

Next we will consider all possible irreversible solutions in detail, thereby paying attention to the possibility of partial depletion. This cannot occur in Tahvonen's model since the pollution problem has no influence on the extraction policy and full exhaustion occurs in finite time. Within our framework with irreversibility, it may however be possible to leave some resource in the ground forever because of the ever increasing environmental damage. Since the shadow value of the resource stock vanishes due to the transversality condition, necessary conditions for the second problem are given by (4) with $(\mu, \alpha, \kappa) = (0, 0, 0)$.

Along any optimal irreversible path, in the limit for $t \to \infty$, the extraction rate vanishes. This implies that $U'(0) = D'(z(\infty))/\delta$, which uniquely defines $z(\infty)$. Let T be the moment when the irreversible region is entered. Since $z(T) = \bar{z}$, the time path of z is uniquely determined. This implies that the time path of $\hat{y}$, the solution of the second problem, including $\hat{y}(T)$ is also uniquely defined. Expression (6) simplifies to

$$U(y(T)) - U'(y(T))(y(T) - \alpha\bar{z}) = U(\hat{y}(T)) - U'(\hat{y}(T))\hat{y}(T).$$

This yields $y(T)$. For a given z_0 we can uniquely determine the initial extraction rate $y(0)$ that leads the economy to $y(T)$. We can then also find the initial resource stock that makes the proposed path feasible. If the actual initial resource stock is larger than or equal to this critical value, then we can always find a path satisfying the necessary conditions. The higher initial stock will not change social welfare, because the shadow value of the resource stock is zero. Hence, the path satisfying the necessary conditions will not alter. Note that $\hat{y}(T)$ does not depend on α and that it should be larger than $\alpha\bar{z}$ because otherwise an upward jump will not prevail. Therefore, and this is intuitively appealing, we should have α small enough for this case to occur. The general conclusion is that with α large it is optimal to choose the reversible program.

Note that the dynamic system governing irreversible paths without exhaustion is qualitatively the same as the one we would obtain in the pollution problem alone. In particular the isoclines for y and z are identical to those of the problem with an abundant resource and it is thus possible to illustrate the features of this kind of irreversible policy in the (z, y) plane (see the trajectory starting from y_0^i in figure 1).[2]

Next, we address the question whether it could be optimal to stay in $\bar{z}$ for a while before entering the irreversible region and leaving some of the resource unexploited. The answer is negative, as can be seen from condition (6). Suppose that for some interval of

[2]An example of reversible policy is also depicted for illustrative purposes (the one starting from y_0^r). Actually, it should be clear that for policies featuring exhaustion the shadow price is not nil which complicates the dynamics. This notably implies that the $\dot{y} = 0$ locus cannot be represented so simply because its location changes with the evolution of the shadow price. See the next section.

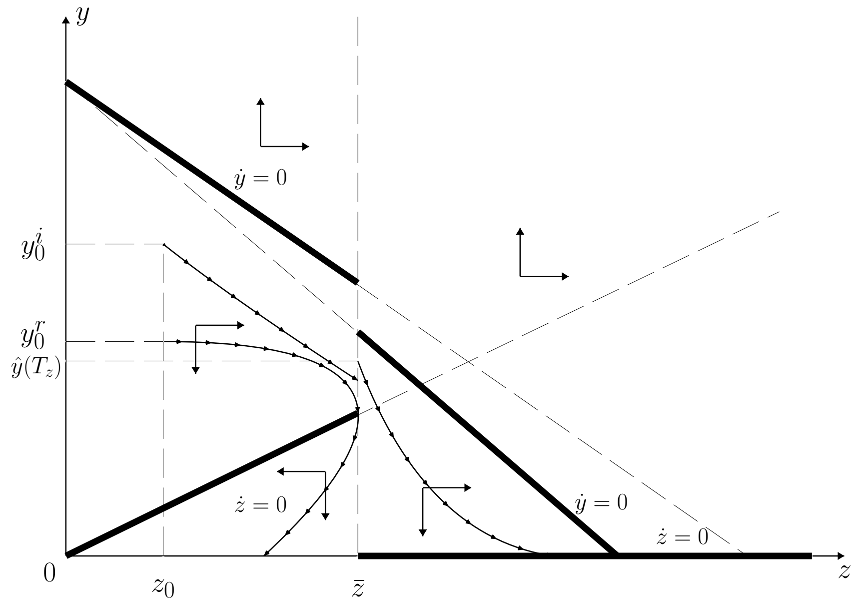

Figure 1: Irreversible path with a discontinuity in extraction.

time $z(t) = \bar{z}$ and $y(t) = \alpha\bar{z}$. Then (6) reduces to

$$U(\alpha\bar{z}) = U(\hat{y}(T)) - U'(\hat{y}(T))\hat{y}(T)$$

With $\hat{y}(T)$ given, this condition does not hold in general. Intuition runs as follows. Take $z_0 = \bar{z}$, thus $T = 0$, and suppose we have $z(t) = \bar{z}$ and $y(t) = \alpha\bar{z}$ until some instant of time T_1 whereas for $t \geq T_1$ we have the optimal path constructed above, leaving some of the resource unexploited. Since the resource is not completely depleted its shadow price equals zero, meaning that adding to the initial stock does not increase social welfare. This implies that the planner is indifferent between the resource stocks at all instants of time between 0 and T_1. Hence, the program followed from time 0 on is welfare equivalent to the program followed from T_1 on. But the welfare values obviously differ. This yields a contradiction. Hence we do not have to worry about programs that do not exhaust the stock and stay in $\bar{z}$ for a while.

The last possibility consists in entering the irreversible region and fully exhausting the resource. In this case, we cannot find simple arguments to rule out a solution featuring a stage at the threshold. So, in general, there are two irreversible candidates leading to exhaustion: one where the system stays in $\bar{z}$ for a while and one where the system passes through $\bar{z}$ at just one instant of time. In addition, one logically expect that irreversible candidates with exhaustion exist for initial resource stocks lower than the ones leading to irreversibility but conservation.

We summarize all the preceding analysis in the following proposition.

Proposition 1 *Suppose $z_0 \leq \bar{z}$. The conditions $\bar{y} > \alpha\bar{z}$, $U'(0) > D'(\bar{z})/\delta$, $U'(\alpha\bar{z}) > D'(\bar{z})/(\alpha+\delta)$ and (6) are necessary for the existence of irreversible solutions. There exist three kinds of irreversible optimality candidates:*

i/ Two irreversible candidates with exhaustion of the resource in finite time: One directly reaching the irreversible region whereas along the other, the economy stays at the threshold for a non-degenerate period of time.

ii/ One irreversible candidate with some amount of resource left in the ground in the long run, immediately crossing the threshold, as soon as it is reached.

In the next section, we examine the features of reversible optimality candidates.

3.2 Reversible solutions

We will show that there are two candidates for optimality with reversibility that is, two different solutions to the necessary optimality conditions (4). One candidate is always in the interior of the reversible region, the other stays at the boundary for a while.

For the case where the constraint $z(t) \leq \bar{z}$ is ignored a priori, it has been shown by Tahvonen (1997) that, given x_0, if z_0 is small enough, $z(t)$ is inverted U-shaped. Otherwise z is monotonically decreasing. In both cases the resource is exhausted in finite time, since it has been assumed that $U'(0) < \infty$. Tahvonen works with fixed x_0 by varying z_0. In the subsequent analysis, it is more convenient to think in terms of fixed z_0 and varying x_0. It turns out that Tahvonen's claim can easily be restated as:

Lemma 1 *If x_0 is small enough, z is monotonicaly decreasing whereas, for high enough x_0, z is inverted U-shaped.*

Proof. See appendix A. ■

Then, we can establish the following:

Proposition 2 *For any $x_0 > 0$ there exists a program satisfying the necessary conditions (4). There exists $\check{x}_0$ such that the unique reversible solution is to have the pollution stock* *increasing initially, hitting the threshold for just an instant of time, and decreasing eventually. For $x_0 < \check{x}_0$ the solution is to stay in the reversible region forever. For $x_0 > \check{x}_0$ the program satisfying the necessary conditions has an interval of time along which the pollution stock is at its threshold level. In all cases the resource stock is depleted eventually.*

Proof. See appendix B. ■

In sum, for any z_0 in the reversible region, there always exists a unique reversible optimality candidate. The nature of the solution mostly depends on the ordering between the initial resource stock and the critical $\check{x}_0$. On the one hand, for all initial resource stocks smaller than $\check{x}_0$ the threshold will never be hit, which means that the economy stays in the interior of the reversible region forever. On the other hand, any initial $x_0 > \check{x}_0$ makes the crossing of the threshold feasible. Hence, when $x_0 > \check{x}_0$, the pure state constraint will be binding for a non degenerate period of time. The resulting solution looks as follows. Initially the rate of extraction is high and pollution approaches the threshold. Then, for a period of time $(z, y) = (\bar{z}, \alpha\bar{z})$ and $\dot{\lambda} = (\delta + \alpha)\lambda - D'(\bar{z}) - \kappa$.[3] There is a final phase where pollution is decreasing.

In all the cases considered here, the resource stock is depleted within finite time, extraction from the stock goes to zero as time goes to infinity and the pollution stock vanishes asymptotically. See the path touching the threshold but remaining in the reversible region in figure 1 for an illustration.

In conclusion, we always have a unique solution to the reversible program. In addition, under the set of conditions summarized in proposition 1, we know that there may also exist irreversible optimality candidates. So, in general, the optimal extraction/emission problem tackled so far can exhibit a multiplicity of optimality candidates with very different features. This raises a series of questions: does our problem always produce multiple solutions? In case of multiplicity, which one of the optimality candidates yields the optimum? Clearly, the answer to the first question is negative. With a very low initial resource stock and a low enough pollution stock, the threshold will never be hit. Therefore, in this case, the optimum necessarily corresponds to the reversible solution. In other circumstances, however, it is not possible to provide a definitive answer. That is why, in the next section, we resort to a numerical example to conduct the optimality analysis.

[3]For the characterization of this solution, we make use of the continuity of z, x, λ and μ in the transition dates. These conditions imply the continuity of y.

4 Optimality: a numerical example

We use a quadratic specification of utility and damage

$$\begin{cases} U(y) = \theta y(2\bar{y} - y), \ \theta > 0 \\ D(z) = \frac{\gamma z^2}{2}, \ \gamma > 0 \end{cases}, \tag{7}$$

and choose the following set of baseline parameters: $\delta = 0.05$, $\alpha = 0.0083$, $\gamma = 0.0022$, $\theta = 26.992$ and $\bar{y} = 12.466$. These values are taken from Karp and Zhang (2012), whose estimates are consistent with the values used by other related papers, dealing with the optimal regulation of GHG, and more specifically CO_2, emissions (see Hoel and Karp, 2002, Liski, 2002, Karp and Zhang, 2005). It is worth mentioning that all these papers ignore resource scarcity and work in discrete time, except Liski (2002) whose numerical analysis relies on an earlier version of Hoel and Karp (2002). One has to be careful when passing from discrete to continuous time models, because for instance, our y corresponds to a rate of emission or extraction, not to a level. Karp and Zhang (2012) have one year periods. So, keeping the same discount rate and decay rate in continuous time makes sense.[4] The damage function depends on the stock of CO_2, so that the issue of continuous versus discrete time is not an issue. We can use the same estimate of γ as Karp and Zhang. Finally, if we want to use their estimates of the parameters of $U(y)$, which in their framework represents the benefit from emission, one must accept some degree of approximation.[5] The parameter $\bar{y}$ corresponds to the maximum emission level. In the literature, it is interpreted as the business-as-usual emission level. Business-as-usual emissions are not constant in Karp and Zhang since they decrease due to investment in abatement capital and also respond to random variables. Here we take $\bar{y}$ equal to Karp and Zhang's estimate of the intercept of the BAU emissions. It allows for a maximum emission level of approximately 12 GtC per year. The initial stock of CO_2 in the atmosphere is $z_0 = 781$ GtC. We assume that the stock of carbon that will likely trigger irreversible changes is situated in the interval $[550, 650]$ ppm. Converting these values into GtC, one obtains an interval for the threshold: $\bar{z} \in [1168, 1380]$. Let us set $\bar{z} = 1200$ GtC. For the moment, the initial stock of resource, also measured in GtC, is left free. A sensitivity analysis with respect to x_0 will be conducted in what follows.

For this set of baseline parameters, we obtain a clear-cut result. Whatever the initial resource endowment and the initial pollution stock (below the threshold), there exists a

[4] See Hoel and Karp, 2002, for the relationship between the discount rate in discrete time and its continuous counterpart, the same kind of relationship can be derived for the decay rate.

[5] For the sake of consistency, we may interpret y as the level of emission per year.

unique reversible optimum. In accordance with our proposition 1, for any initial pollution level, the level x_0 only determines whether the economy will spend some time at the threshold. The frontier depicted in figure 2 is the set of initial stocks $(z_0, \check{x}_0)$ such that starting from a point in this set, there exists a path satisfying the necessary conditions, staying in the reversible region and hitting the threshold exactly once. From any initial condition strictly below this frontier, the unique solution is purely reversible whereas from any x_0 above the frontier, the optimal path is still unique and reversible but stays at the threshold for a non degenerate period of time.

For the parameters values used in our benchmark, the three first necessary conditions for the existence of irreversible solutions (given in proposition 1) are satisfied. However, there is no $T < \infty$ satisfying condition (6). The reason for this is simple. Suppose that the economy is at $\bar{z}$ during a period of time. Then, during this period the extraction level is precisely equal to $\alpha\bar{z} = 9.96$. According to condition (6), for the switch to the irreversible region to be optimal, the loss incurred by the economy, due to the vanishing of pollution decay, must be compensated by a sufficient increase in extraction and hence consumption. But, for the parameter values chosen, the extraction rate is already high (and close to the maximum possible level $\bar{y}$) before entering the irreversible region, which precludes the upward jump. Thus, it is not optimal to cross the threshold.

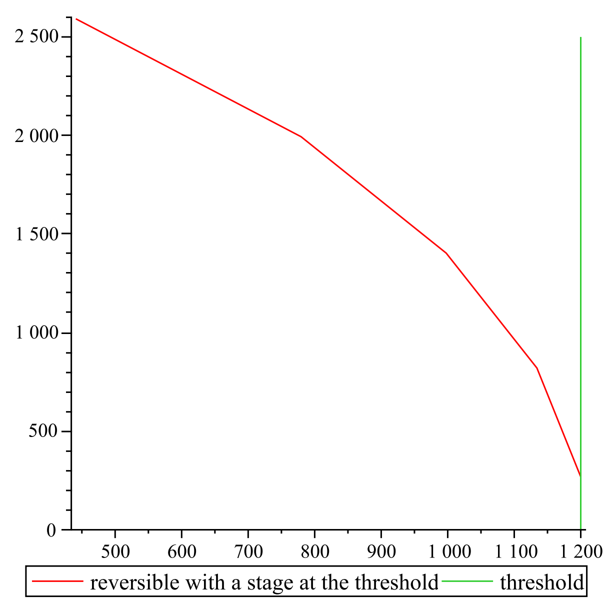

Figure 2: Optimal reversible policies

It goes without saying that the results are sensitive to the assumed maximum emission level. Inspection of IPCC (2010)'s emissions scenarios reveals that projected annual emissions of CO_2 from all sources (energy, industry and land-use change) will achieve levels above our maximum of 12 GtC in most scenarios. This is particularly true for scenarios based on a high-coal and high-oil and gas use). Thus, we find it both reasonable and interesting to compare the results of our benchmark scenario with those in which we allowed for higher business as usual emissions.

For that purpose, let us consider a 30 % higher $\bar{y}$. Then, the results drastically change because of the existence of multiple optimality candidates, some of them featuring irreversible pollution. As an illustration, the following exercise can be conducted. By varying the initial pollution stock and the resource stock we are able to divide the (z_0, x_0) plane into four regions, delimited by three frontiers (see figure 3).

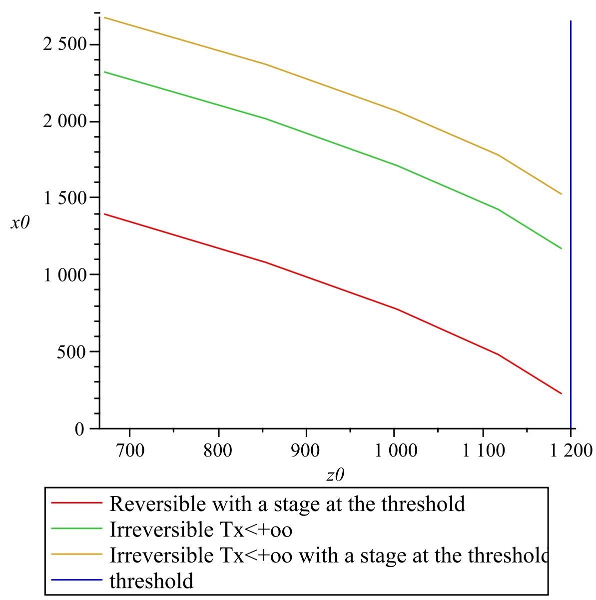

Figure 3: Frontiers

The "Reversible with a stage at the threshold" frontier has the same meaning as in the benchmark. However, we can now identify two other critical loci. Let T_x be the instant when the resource is exhausted. The "irreversible $T_x < \infty$" frontier is the lower bound of existence of irreversible solutions crossing $z = \bar{z}$ with exhaustion of the resource. Originating from any point between the reversible and the "irreversible $T_x < \infty$" frontiers, the optimal solution is still reversible but has a period of time when the system remains in $\bar{z}$ for a while. A third frontier is called "irreversible $T_x < \infty$ with a stage at the threshold".

For any initial point lying between "irreversible $T_x < \infty$" and "irreversible $T_x < \infty$ with a stage at the threshold", there are two optimality candidates: the reversible path staying in $\bar{z}$ for a period of time and an irreversible path that passes through the threshold for just an instant of time, both featuring exhaustion in finite time. Above the latter frontier there exists another candidate along which the economy stays in $\bar{z}$ for a while before entering the irreversible region and exhausting the resource. In other words, from the frontier "irreversible $T_x < \infty$ with a stage at the threshold" onwards, we have three optimality candidates.[6]

So, for any initial pollution stock, the initial resource endowment is crucial to understand both the number and the nature of the optimality candidates. This is a prerequisite for the analysis of the optimal policy. Fixing the stock of pollution at z_0=781 GtC, such an analysis of course relies on the measurement of the stock of available fossil fuels. The least one can say is that the level x_0 varies to a very large extent among studies. For example, Coulomb and Henriet (2010) take a x_0 close to 800 GtC by referring to the estimates of reserves of fossil fuels by the International Energy Agency. By contrast, Dietz and Asheim (2012) use 6000, based on Nordhaus (2008). The differences can be explained by the different definitions of the available stock of fossil fuels they use. Coulomb and Henriet (2010)'s study only considers conventional reserves of fossil fuels, whereas Dietz and Asheim (2012) also include unconventional reserves and (un)conventional resources.[7]

Let us now assess the impact of x_0 on the nature of the optimal policy. It is clear that as long as x_0 is small enough ($\leqq$ 2200 GtC), there exists a unique reversible solution that yields the optimum. Therefore, by focusing only on conventional reserves of fossil fuels, we can argue, as in the benchmark case, that the optimal policy is reversible. Next we have to make the welfare comparison for the cases where there are multiple optimality candidates. For that purpose, we depict, for the initial pollution stock $z_0 = 781$, the present values associated with three optimality candidates, namely the reversible, and the two irreversible with exhaustion differing by the fact that one of them stays in $\bar{z}$ for a while and the other doesn't. Figure 4 illustrates how the values evolve when varying x_0.

First, we can find a critical initial resource stock for which we have multiple optima. The reversible policy staying in $\bar{z}$ for a while and the policy entering directly the irre-

[6]Note that there exists a last frontier above which irreversible candidates, that passes through the threshold for just an instant of time, lead to resource conservation. But, the range of values of x_0 (>19000 GtC) allowing for resource conservation does not seem to be relevant for our analysis. That is why we choose not to draw this frontier.

[7]Kalkuhl and Edenhofer (2011) report the following estimates, summing the stock of oil, coal and gas: conventional reserves: 791 GtC, conventional resources: 11588 GtC, unconventional resources: 1839 GtC.

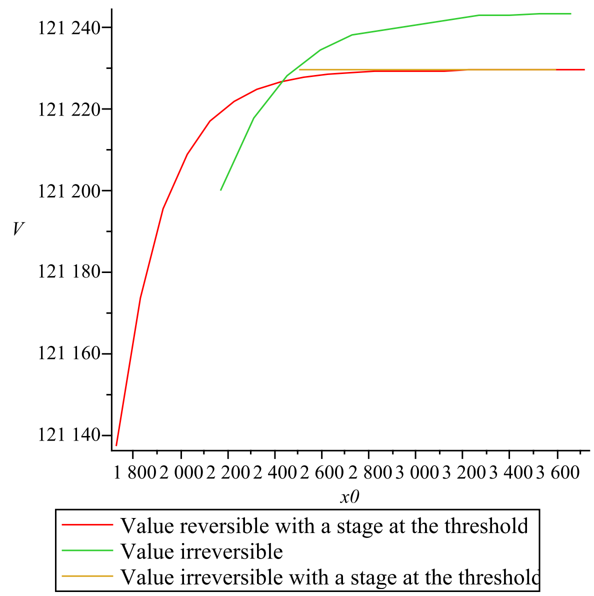

Figure 4: Optimality candidates: comparison of present values

versible region yield exactly the same present value. This feature echoes Tahvonen and Withagen (1996)'s finding in their model with quadratic decay but without exhaustibility. The reversible policy is the optimum for low enough stock of resources. But, once the resource endowment is above a critical level the optimum becomes irreversible. Second, we observe that the irreversible path with the economy not staying in $\bar{z}$ always dominates the other candidate with a stage in $\bar{z}$. So, it seems to be always better to directly enter the irreversible region upon arrival at $\bar{z}$ than to stay at the threshold for a period of time. The reason is the following. Compare these two solutions for the period of time just after the threshold is hit. In the first case, pollution is slightly higher than $\bar{z}$ (due to continuity) but the economy benefits from higher consumption (due to the upward jump) whereas in the second case, pollution is equal to $\bar{z}$ but the extraction rate is much lower and equal to $\alpha\bar{z}$. Therefore, if the former optimality candidate dominates the latter this is mainly because the gain from higher consumption exceeds the loss from higher pollution.

The following conclusion can be drawn from this numerical exercise. If one believes that BAU emissions will remain quite moderate or thinks in terms of conventional reserves of fossil fuels only, then there is a unique reversible optimum. However, when one allows for higher potential emissions and also adds part of the amount of conventional resources to the initial resource stock, the results are substantially modified. One can identify multiple optimality candidates and the optimal policy becomes irreversible for a large enough resource stock. The policy implications of both allowing higher maximum emission levels and higher initial endowments of exhaustible resource are thus far from negligible.

5 Quadratic decay

In the preceding analysis, we have stressed the sensitivity of the results to the maximum emission level and to the initial stock of fossil fuels. In this section, our purpose is twofold. First, we investigate how sensitive the results are to the shape of the decay function. Second, we draw a parallel with Tahvonen and Withagen's conclusions in the quadratic decay case without exhaustible resources.

Tahvonen and Withagen (1996), study a decay function that is zero at zero pollution, increases, reaches a maximum and then declines towards zero, reaching zero at some point $\bar{z}$. This introduces a non-convexity in the problem that leads to the possibility of multiple equilibrium candidates. They do not include exhaustibility. It is the purpose of this section to see what the effect is in their model of being constrained by a nonrenewable resource. It cannot be expected that we reach many general conclusions from a purely theoretical analysis. Some of their results, however, do go through if we include exhaustibility. First, if the discounted marginal damage at $\bar{z}$ is larger than the marginal utility of the first drop of oil, then we will only have reversible solutions. Second, another obvious result is that due to exhaustibility we will never have a positive steady state extraction rate, whereas in Tahvonen and Withagen convergence to a steady state is a possibility. Third, in the case of quadratic decay it is not optimal to stay at the threshold for longer than just an instant of time. This implies that the set of equilibrium candidates is much easier to depict.

To illustrate this we provide a numerical exercise using the same specification and parameters values as in our benchmark scenario. Tahvonen and Withagen approximate the quadratic decay function in a triangular way. Following their approach, we assume $\alpha(z) = \alpha z$ for $z \in [0, \frac{\bar{z}}{2}], = \alpha(\bar{z} - z)$ for $z \in [\frac{\bar{z}}{2}, \bar{z}]$. Note that the initial stock of pollution belongs to the second interval, which conveys the idea that for the current concentration of CO_2, the carbon sinks' ability to store carbon has started to decline. This is consistent with the evidence reported by the IPCC (2007).

The subsets of the (z_0, x_0) plane that delineate the optimality candidates look much simpler that in the previous section. As explained above, the reason is that it is never optimal now to stay at the threshold for longer than an instant of time. Here there is almost no decay close to the threshold, whereas earlier decay is maximal at the threshold. So, now there is no benefit to be found in staying at the threshold.

Generally, there are three regions that can be distinguished, for a given initial pollution stock z_0. For a small initial resource stock it is optimal to stay in the reversible region and to deplete the resource stock. For higher resource stocks it is optimal to enter the irreversible region, with exhaustion in finite time. With still higher resource stock we are back in the their model with only partial exhaustion. See figure 5.

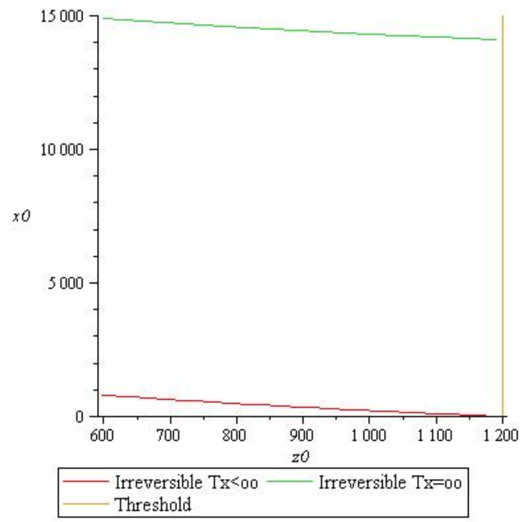

Figure 5: Zones of existence of the different optimality candidates

These results drastically differ from those of our benchmark case with discontinuous decay. In addition, it echoes that of Tahvonen and Withagen. Indeed, they identify a critical $\tilde{z}_0$ which is associated with two optima. This means that for this critical level, there exist both a reversible and an irreversible solution yielding exactly the same present value. Here, we have another state variable, the exhaustible resource, and it appears that for each z_0, there exists a critical $\tilde{x}_0$ such that we observe multiple optima. For example, with our initial value $z_0 = 781$ GtC, this critical value is $\tilde{x}_0 \simeq 514$ GtC and the resulting present value is approximatively equal to $V \simeq 56396$. For any $x_0 < \tilde{x}_0$, the only solution is the reversible one whereas for $x_0 > \tilde{x}_0$, the only solution is irreversible.

The comparison with the discontinuous decay case shows two important differences. First, with a quadratic decay, there exist irreversible solutions whereas in the benchmark scenario of the previous section we only obtain reversible solutions. The reason is that with the discontinuous decay, entering the irreversible region, even though it is feasible, is very costly since it implies a switch from maximum to zero decay that must be compensated by a substantial increase in consumption. Here this is no longer the case because there is almost no decay in the neighborhood of the threshold. In the quadratic decay case, exceeding the threshold is optimal provided that the economy initially owns high enough resources. Second, since there is no opportunity to stay at the threshold for a period of time, we cannot have multiple optimality candidates, except for the critical combination $(z_0, \tilde{x}_0)$. This feature is robust to a change in the maximum emission level $\bar{y}$. In other words, the multiplicity of candidates observed in the preceding section, with a 30 % higher $\bar{y}$, do not hold in this case.

6 Conclusion

This paper has introduced irreversibility of pollution decay in the classical exhaustible-resource/stock-pollution model. Within this framework, we have studied how the potential irreversibility of pollution affects the optimal extraction path. Several results have been obtained. For a small marginal utility of the first unit of the raw material from the non-renewable resource compared to total discounted marginal damage at the threshold level, or for a low business-as-usual scenario or if the threshold is too high compared to the steady state without exhaustibility of the fossil fuel stock, the economy will never enter the irreversible region. For a resource stock not too small it is optimal to stay at the threshold level for a while and then to return to the interior of the reversible region. In any optimum with reversibility, it is optimal to deplete the entire resource stock. With increasing resource stocks it will become optimal to enter the irreversible region, possibly after staying at the threshold for a while. Also in the latter case, all fossil fuel is depleted. These results indicate that staying at the threshold, in order to benefit from the high decay rate, is profitable. For very large fossil fuel stocks it is optimal to cross the threshold without staying there. Partial exhaustion is possibly optimal.

Regimes where the economy stays at the threshold for a period of time make sense. However, there is a danger in following this policy With the introduction of uncertainty on e.g., the threshold level, or the actual business-as-usual emissions, one might easily enter the irreversible region, whereas this is suboptimal, or premature (see Ayong et al., 2011, who introduce uncertainty about the threshold). Hence, the equilibrium during the phase at the threshold is really a knife-edge. There is also evidence that available and existing stock of fossil fuels are not well known. This suggests a first extension of our work, which is motivated by the observation that a lot of uncertainty surrounds both the extent of fossil fuels reserves present in the ground and the concentration of GHG that will initiate irreversible phenomena. Another extension of the present paper naturally comes into mind. The introduction of a backstop technology would be a means to examine how the optimal timing of the backstop adoption and how the optimal combination of technologies are affected by irreversibility.

We have also investigated the introduction of exhaustibility in a model with quadratic decay. It turns out that qualitatively the results are very sensitive. Apart from the obvious fact that no steady state extraction will occur, an important outcome is that with quadratic decay the set of optimality candidates looks much simpler, which also makes it easier to determine the optimum. More importantly, however, a regime where the economy stays at the threshold for a while is not optimal. This strengthens our conclusion that knowledge on actual decay is crucial for determining the best extraction-pollution program.

For future research it would be important to investigate in more detail and for larger parameters sets the differences in welfare arising from different specifications of decay, and the seriousness of making mistakes in estimating the decay schedule.

References

[1] Amigues, J-P. and Moreaux, M. (2012). Potential irreversible catastrophic shifts of the assimilative capacity of the environment. IDEI working paper series 697.

[2] D'Arge, R. and K. Kogiku (1973). Economic growth and the environment, Review of Economic Studies, 40, 61-77.

[3] Asheim, G. and S. Dietz (2012). Climate policy under sustainable discounted utilitarianism, Journal of Environmental Economics and Management doi: 10.1016/j.jeem.2012.01.003.

[4] Ayong Le Kama, A., A. Pommeret and F. Prieur (2011). Optimal emission policy under the risk of irreversible pollution. FEEM working paper 14.2011.

[5] Chakravorty, U. and D.L. Krulce (1994). Heterogeneous demand and order of resource extraction, Econometrica, 62, 6, 1445-1452.

[6] Chakravorty, U., B. Magne and M. Moreaux (2006). A Hotelling modcel with a ceiling on the stock of pollution, Journal of economic Dynamics and Control 30, pp. 2875-2904.

[7] Chakravorty, U., M. Moreaux and M. Tidball (2008). Ordering the extraction of polluting nonrenewable resources, American Economic Review, 98, 3, 1128-1144.

[8] Coulomb, R. and F. Henriet (2010). Carbon price and optimal extraction of a polluting fossil fuel with restricted carbon capture, PSE working papers 2010-11.

[9] Cropper, M. (1976). Regulating activities with catastrophic environmental effects, Journal of Environmental Economics and Management 3, 1-15.

[10] Forster, B. (1975) Optimal pollution control with nonconstant exponential rate of decay. Journal of environmental Economics and Management, 2, 1-6.

[11] Gaudet, G., M. Moreaux and S.W. Salant (2001). Intertemporal depletion of resource sites by spatially distributed users, American Economic Review, 91, 4, 1149-1159.

[12] Hoel, M. (1983). Monopoly resource extractions under the presence of predetermined substitute production, Journal of Economic Theory, 30, 201-212.

[13] Hoel, M. and Karp, L. (2002). Taxes versus quotas for a stock pollutant, Resource and Energy Economics, 24, 367-384.

[14] Hoel, M. and Kverndokk, S. (1996). Depletion of fossil fuels and the impacts of global warming, Resource and Energy Economics, 18, 115-136.

[15] Herfindahl, O.C. (1967). Depletion and economic theory, in M. Gaffney (ed.), Extractive Resources and Taxation, University of Wisconsin Press, Madison, WI.

[16] Intergovernmental Panel on Climate Change, Technical summary:"Climate change 2007: Impact, adaptation and vulnerability. Contribution of working group II to the fourth assessment report of the IPCC (2007)

[17] Intergovernmental Panel on Climate Change: special report emissions scenarios (2010). http://www.ipcc.ch/pdf/special-reports/spm/sres-en.pdf

[18] Kalkuhl, M. and O. Edenhofer (2011). Stocks of carbon in the ground and in the atmosphere. http://www.pik-potsdam.de/ kalkuhl/working-paper/carbon-stocks.pdf

[19] Karp, L. and J. Zhang (2005). Regulation of Stock Externalities with Correlated Abatement Costs, Environmental and Resource Economics, 32, 273-299.

[20] Karp, L. and J. Zhang (2012). Taxes versus quantities for a stock pollutant with endogenous abatement costs and asymmetric information, Economic Theory, 49, 371-409.

[21] Liksi, M. (2002). Taxing average emissions to overcome the shutdown problem, Journal of Public Economics, 85, 363-384.

[22] Nordhaus, W. (2008). A Question of Balance: Economic Models of Climate Change. Yale University Press, New Haven, CT.

[23] Ploeg, F. van der and C. Withagen (2010a). Is there really a Green Paradox?, Research Paper 35, Oxcarre, University of Oxford, Oxford, U.K.

[24] Ploeg, F. van der and C. Withagen (2010b). Economic development, renewables and the optimal carbon tax, Research Paper 55, Oxcarre, University of Oxford, Oxford, U.K.

[25] Plourde, C. (1972). A model of waste accumulationand disposal. Canadian Journal of Economics 5, 119-25

[26] Polasky, S., De Zeeuw, A. and Wagener, F. (2012). Optimal management with potential regime shifts, Journal of Environmental Economics and Management 62, 229-240.

[27] Sinclair, P.J.N. (1994). On the optimum trend of fossil fuel taxation, Oxford Economic Papers, 46, 869-877.

[28] Tahvonen, O. and C. Withagen (1996). Optimality of irreversible pollution accumulation. Journal of Economic Dynamics and Control 20, pp. 1775-1795.

[29] Tahvonen, O. (1997). Fossil fuels, stock externalities, and backstop technology, Canadian Journal of Economics, 30, 4a, 855-874.

[30] Tsur, Y. and A. Zemel (2003). Optimal transition to backstop substitutes for nonrenewable resources, Journal of Economic Dynamics and Control, 27, 551-572.

[31] Tsur, Y. and A. Zemel (2005). Scarcity, growth and R&D, Journal of Environmental Economics and Management, 49, 484-499.

[32] Tsur, Y. and A. Zemel (2008). Regulating Environmental Threats, Environmental and Resource Economics, 39: 297-310.

[33] Tsur, Y. and A. Zemel (2002). The Regulation of Environmental Innovations, Journal of Environmental Economics and Management, 44, 242-260.

[34] Ulph A. and D. Ulph (1994). The optimal time path of a carbon tax, Oxford Economics Papers, 46, 857-868.

[35] Withagen, C. (1994). Pollution and exhaustibility of fossil fuels, Resource and Energy Economics, 16, 235-242.

Appendix

A Proof of lemma 1

From Tahvonen (1997)'s proposition 2, we know that the path of z cannot have U-shaped segments. He also shows that z asymptotically converges to 0. Hence z is either inverted U-shaped or monotonically decreasing.

Suppose $y(0) < \alpha z_0$. Then $U'(y(0)) > U'(\alpha z_0)$ and z is monotonically decreasing. Since

$$\lambda(t) = e^{(\alpha+\delta)t} \int_t^{\infty} e^{-(\alpha+\delta)s} D'(z(s))ds, \tag{8}$$

we then have $\lambda(t) \leq \frac{D'(z_0)}{\alpha+\delta}$ for all $t \geq 0$. Moreover, it has been assumed that

$$\frac{D'(z_0)}{\alpha+\delta} < U'(\alpha z_0). \tag{9}$$

Therefore $\lambda(t) \leq U'(\alpha z_0)$ for all $t \geq 0$. So

$$U'(y(0)) = \lambda(0) + \mu_0 \leq U'(\alpha z_0) + \mu_0. \tag{10}$$

Note that μ_0 depends on x_0. We can make μ_0 arbitrarily small by taking x_0 large enough. This can be seen as follows. Suppose there exists $\varepsilon > 0$ such that $\mu(0; x_0) > \varepsilon$ for all x_0, where we use the convention that for any variable w, $w(t; x_0)$ is the optimal value of w at instant of time t when the initial resource stock is x_0. Then $U'(y(t; x_0)) = \lambda(t; x_0) + \mu(t; x_0) > \varepsilon e^{\delta t}$ for all t and all x_0. But the inequality implies that $\int_0^{\infty} y(t)dt < M$ for some $M < \infty$. This is a contradiction. But then, for x_0 large enough, and therefore μ_0 very small, we cannot have (10). Hence for x_0 large enough, we have $y(0) > \alpha z_0$. Q.E.D.

B Proof of proposition 2

It has been shown in lemma 1 that for large enough x_0 pollution will initially rise. Now we show in a first step that for large enough x_0 pollution will hit the threshold. Recall that (y^*, z^*) is defined as the steady state of the resource abundant economy, not subject to potentially irreversible pollution.

Consider a constant path of extraction: $\tilde{y}(t) = \frac{\alpha}{2}(\bar{z} + \hat{z}^*)$ for all t. With this extraction policy, the threshold will be reached at instant of time t_1 given by

$$\bar{z} = z_0 e^{-\alpha t_1} + \frac{1}{2}(\bar{z} + z^*)(1 - e^{-\alpha t_1})$$

Let us consider the problem with exhaustibility. Suppose that for some x_0 we have $z(t;x_0) \leq \bar{z}$ for all $t \geq 0$. Then

$$\lambda(t;x_0) = e^{(\alpha+\delta)t}\int_t^\infty e^{-(\alpha+\delta)s}D'(z(s);x_0)ds \leq \frac{D'(\bar{z})}{\alpha+\delta}$$

and

$$\begin{aligned} U'(y(t);x_0) &= \lambda(t;x_0) + \mu(t;x_0) < \frac{D'(\bar{z})}{\alpha+\delta} + \mu(0;x_0)e^{\delta t} \\ &< \frac{D'(z^*)}{\alpha+\delta} + \mu(0;x_0)e^{\delta t} = U'(\alpha z^*) + \mu(0;x_0)e^{\delta t} \end{aligned}$$

for all $t \geq 0$. In particular, this inequality is valid at the instant t_1 defined above. Now, choose x_0 such that

$$\mu(0;x_0) = \frac{U'(\alpha\bar{z}) - U'(\alpha z^*)}{e^{\delta t_1}},$$

This equation has a solution, since we can make μ arbitrarily small by a proper choice of x_0 as demonstrated in the previous lemma. Hence, we obtain a contradiction. Then, from a continuity argument, it is clear that there exists a unique initial resource stock $\check{x}_0$ such that the threshold is exactly hit for a single instant of time.

The second step of the proof works while ignoring the pure state constraint $z(t) \leq \bar{z}$. It establishes that for any solution starting from (x_0, z_0), with $x_0 > \check{x}_0$, it holds that $z(t) > \check{z}(t)$ for all t, where $\check{z}$ indicates the program corresponding to $\check{x}_0$. This means that for any solution originating from $x_0 > \check{x}_0$, the pollution stock will be above the threshold for a connected period of time. Hence, adding the pure state constraint $z(t) \leq \bar{z}$, we can argue that a reversible solution with $x_0 > \check{x}_0$ will stay at the threshold for a non-degenerate period of time (during which the constraint $z(t) \leq \bar{z}$ is binding).

To start with, we claim that for any z_0, if $x_0 > \check{x}_0$ then $y_0 > \check{y}_0$ that is, $z(t) > \check{z}(t)$ for an initial period of time.

Assume that $x_0 > \check{x}_0$ and $y_0 < \check{y}_0$. This implies that $z(t) < \check{z}(t)$ for an initial period of time. The resource is more abundant, which means that $\mu_0 < \check{\mu}_0$ and $\mu(t) < \check{\mu}(t)$ for all t. The inequality $y_0 < \check{y}_0$ is equivalent to $\lambda_0 + \mu_0 > \check{\lambda}_0 + \check{\mu}_0$. Hence, $\lambda_0 > \check{\lambda}_0$ and, using (8), there exists an instant of time $t_1 > 0$ such that $z(t) < \check{z}(t)$ for all $t \in [0, t_1)$ and $z(t_1) = \check{z}(t_1)$, with $\dot{z}(t_1) > \dot{\check{z}}(t_1)$, and therefore $y(t_1) > \check{y}(t_1)$. This is turn implies that there exists a date $t_2 \in (0, t_1)$ such that $y(t) < \check{y}(t)$ for $t \in [0, t_2)$, $y(t_2) = \check{y}(t_2)$ and $\dot{y}(t_2) > \dot{\check{y}}(t_2)$. From

$$U''(y)\dot{y} = \delta U'(y) + \alpha\lambda - D'(z), \tag{11}$$

$\dot{y}(t_2) > \dot{\check{y}}(t_2)$ is equivalent to $\alpha(\lambda(t_2) - \check{\lambda}(t_2)) < D'(z(t_2)) - D'(\check{z}(t_2))$. Hence, $\lambda(t_2) < \check{\lambda}(t_2)$. We also have $\mu(t_2) < \check{\mu}(t_2)$ so that $y(t_2) > \check{y}(t_2)$, a contradiction.

So, at least initially $z(t) > \check{z}(t)$. Let us then assume that $x_0 > \check{x}_0$ and $z(t) < \check{z}(t)$ for some interval of time $[t_1, t_2]$, with $0 < t_1 < t_2 \leqq \infty$, with $z(t) > \check{z}(t)$ for all $t \in [0, t_1)$ and $z(t_1) = \check{z}(t_1)$ with $\dot{z}(t_1) < \dot{\check{z}}(t_1)$, so that $y(t_1) < \check{y}(t_1)$. Moreover, $z(t) < \check{z}(t)$ for all $t \in (t_1, t_2)$ and $z(t_2) = \check{z}(t_2)$ with $\dot{z}(t_2) > \dot{\check{z}}(t_2)$ so that $y(t_2) > \check{y}(t_2)$. Finally $z(t) > \check{z}(t)$ for all $t \in (t_2, \infty)$.

First, suppose that $t_2 = \infty$. The t_1 is the unique instant of time where the paths $z(t)$ and $\check{z}(t)$ intersect. We have $z(t) < \check{z}(t)$ for all $t \in (t_1, \infty)$, so that $\lambda(t_1) < \check{\lambda}(t_1)$. But we also have $y(t_1) < \check{y}(t_1)$ and $\mu(t_1) < \check{\mu}(t_1)$. Hence from (4) $\lambda(t_1) > \check{\lambda}(t_1)$, a contradiction.

Next, consider the case $t_2 < \infty$. We have $y(t_1) < \check{y}(t_1)$, $y(t_2) > \check{y}(t_2)$, so that $y(t_3) = \check{y}(t_3)$ and $\dot{y}(t_3) > \dot{\check{y}}(t_3)$ for some $t_1 < t_3 < t_2$. In view of (11) we then have $\lambda(t_3) < \check{\lambda}(t_3)$. Moreover, $\mu(t_3) < \check{\mu}(t_3)$. But this contradicts $\lambda(t_3)+\mu(t_3) = \check{\lambda}(t_3)+\check{\mu}(t_3)$ which is required because $y(t_3) = \check{y}(t_3)$.

In a similar we we can prove that any path originating from (x_0, z_0), with $x_0 < \check{x}_0$ has $z(t) < \check{z}(t))$ for all t. Q.E.D.