OPTIMAL BEHAVIOR UNDER POLLUTION IRREVERSIBILITY RISK AND DISTANCE TO THE IRREVERSIBILITY THRESHOLDS: A GLOBAL APPROACH

R.Boucekkine, W.Ruan and B.Zou

LIDAM Discussion Paper IRES
2024 / 01

UCLouvain

LIDAM
IRES | Louvain Institute of Data Analysis and Modeling in economics and statistics

Optimal behavior under pollution irreversibility risk and distance to the irreversibility thresholds: A global approach

R.Boucekkine *　　W.Ruan†　　B.Zou‡

January 30, 2024

Abstract

We study optimal behavior under irreversible pollution risk. Irreversibility comes from the decay rate of pollution sharply dropping (possibly to zero) above a threshold pollution level. In addition, the economy can instantaneously move from a reversible to an irreversible pollution mode, following a Poisson process, the irreversible mode being an absorbing state. The resulting non-convex optimal pollution control is therefore piecewise deterministic. First, we are able to characterize analytically and globally the optimal emission policy using dynamic programming. Second, we prove that for any value of the Poisson probability, the optimal emission policy leads to more pollution with the irreversibility risk than without in a neighborhood of the pollution irreversibility threshold. Third, we find that this local result does not necessarily hold if actual pollution is far enough from the irreversibility threshold. Our results enhance the importance of the avoidability of the latter threshold in the optimal economic behavior under the irreversibility risk.

Keywords: Irreversible pollution, uncertainty, piecewise deterministic, optimal behavior under risk, avoidability of the irreversible regime

JEL classification: Q52, C61, D81

*Centre for Unframed Thinking, Rennes School of Business (France), and Corresponding Member, IRES-UCLouvain (Belgium). E-mail: raouf.boucekkine@rennes-sb.com

†Purdue University Northwest, USA. E-mail: Wruan@pnw.edu

‡Corresponding author. DEM, University of Luxembourg. E-mail: benteng.zou@uni.lu

1 Introduction

Climate change has been recently proclamated as *the ultimate challenge for economists*, by Nordhaus (2019) in his Nobel lecture, and indeed it's actually the ultimate challenge for the whole community of scientists given the depth and multidimensionality of the problem. Among the inherent topics tackled, irreversible environmental change is one of hottest nowadays as the ecological crisis originating in global warming is getting increasingly acute. An already abundant scientific literature points at unequivocal signals of irreversible climate change. An example of such a literature is in the growing evidence on the weakening of the AMOC (Atlantic Meridian Overturning Circulation), and more specifically on the Gulf stream. See Bonnet et al (2021) on the former and Dong et al (2019) on the latter for example. The main idea behind irreversibility is that continuous accumulation of pollutants (GHG in the case of global warming) can at a certain point in time reach a threshold level such that beyond this level, key regulating and vital mechanisms become permanently partially or totally defective. In the case of climate change, this translates into an upward destabilizing temperature path.

The economic literature has been concerned by this type of irreversible pollution for a while. Many highly interesting papers have been published along this research line, from the 90s essentially. A significant part of these papers is actually connected to the literature of tipping points in dynamic systems as in the so-called the shallow lake problem (see Maler (2000) or Wagener and de Zeeuw (2021) for a more recent example): the emergence of tipping points in the lake ecosystem dynamics follows small variations in phosphorus loads, ultimately leading to significant losses in ecosystem services.[1] A key question turns out to (legitimately) be to which extent irreversible environmental changes can be avoided. In a seminal paper, Tahvonen and Withagen (1996) have demonstrated in a quite general convex-concave optimal pollution control problem that avoiding the threat of irreversible pollution thresholds is not granted. In their model, irreversibility occurs if pollution reaches a critical level above which the decay rate of pollution drops permanently to zero. Tahvonen and Withagen (TW hereafter) prove that even a benevolent central planner cannot always avoid crossing the critical pollution level, which formulated into a game-theoretic frame means that

[1]The irreversible pollution literature also significantly intersects with the broader stream on regime shifts, see Boucekkine et al. (2013) for an application including regime shifts of the irreversible pollution type.

even full cooperation among players cannot always prevent this unpleasant outcome (see the dynamic game extension of TW in Boucekkine et al, 2023). Moreover, even though optimal maintenance/abatement are introduced for pollution control, optimal paths with irreversible pollution paths may still emerge under some non-extreme parametric conditions (Prieur, 2009).

Another aspect appears to be potentially important for the ultimate impact of irreversible pollution: uncertainty. Indeed, uncertainty surrounding this event can have multiple sources, the most obvious being that the inherent irreversibility threshold levels are not known with certainty. This is crystal clear if one explores, for example, the hottest science literature on the AMOC expected collapse: for Ditlevsen and Ditlevsen (2023), this may occur between 2025 and 2095. Of course, uncertainty is not only on the level of the irreversibility thresholds (resulting in the poor accuracy of the dating of the regime shift) but it also bears on the extent of damages caused when the irreversibility thresholds are reached. The literature is quite diverse in this respect, we summarize here below very briefly some of the essential avenues taken.

There is indeed a strong research line which associates irreversibility levels crossing with catastrophic damage, therefore entailing a catastrophic risk component. Clarke and Reed (1994) and Tsur and Zemel (1998) are two salient representatives of this research. In the former, the value function associated with the underlying optimal control problem is zero at the threshold irreversibility level while it is minus infinity in the latter. This goes with the idea that as the irreversibility levels are reached, this will come with catastrophic events (such as extreme climate consequences). It's unclear whether these scenarios are the most relevant. Alternatively one could assume that reaching the irreversible regime will produce a permanent sharp drop in the value function, not necessarily to minus infinity (see for example, Le Kama et al., 2014, within a different theoretical framework that however keeps some analogies with TW).[2] This is also somehow consistent with the view implicit in the related hard science publications mentioned above that catastrophic damage will not necessarily occur immediately after crossing the irreversibility thresholds but may well materialize according to a more progressive process.

[2]Many other insightful stochastic models with irreversibility or related features (such as tipping points or regime shifts problems) can be found in the economic literature. See in particular Bretschger and Vinagradova (2019), Diekert (2017) or Ren and Polasky (2014).

Putting technicalities apart, the essential issue tackled in this literature is whether the irreversibility risk with the associated sharp drop in environmental quality, be it catastrophic or not, will induce more conservative or more aggressive behavior, for example in terms of pollutants emissions. Intuitively, one would think that subject to the irreversibility risk, the economic agents would prefer to behave in a more conservative way and pollute less. Indeed, a key point is whether the irreversibility risk depends or not on pollution. In the catastrophic risk model studied by Clarke and Reed (1994), where the catastrophic event nullified the value function, pollution and consumption are bigger at the steady state equilibrium in presence of the irreversibility risk than without if the latter does not depend of the pollution stock. On the contrary, if the risk is *strongly enough* increasing with the pollution stock, optimal behavior is more conservative and one gets lower pollution and consumption at the steady state in the presence of risk. Of course, mere dependence of the risk on pollution is not enough to generate the conservative behavior under the irreversibility risk. As excellently explained in Clarke and Reed, "*...The rationale for these results hinges on the distinction between 'avoidable' and 'unavoidable' risk*". (page 1009). The same rationale is at work in many of the related papers, including Tsur and Zemel (1998), van der Ploeg (2014) or Le Kama et al. (2014) cited above. [3]

A remarkable feature of this literature stream is that the sensitivity analysis with respect to the presence/absence of the irreversibility risk is performed at the respective steady state equilibria. In the case of Clarke and Reed (1994) or Tsur and Zemel (1998), tractability issues can be invoked to justify such an exclusive long-term focus. However, a probably equally important question is how contemporaneous or short-term optimal behavior is shaped by the irreversibility risk. Said differently, **it might be equally relevant to explore optimal behaviour at any given level of pollution in the presence or absence of the latter risk even if (and especially if) the *event does not occur*** to use the terminology of the related literature (i.e. even if the bad shock does not occur). In this paper, we propose a global analytical approach to the problem which does allow to respond the latter questions. To this end, we consider a piecewise deterministic extension of the TW model: the model has two modes, a reversible vs an irreversible

[3] One can find a similar discussion in the context of optimal management of renewable resources under the risk of regime shift in Ren and Polasky (2014). See Diekert (2017) for a game-theoretic extension in the same context.

pollution mode, and the probability to move from the former to the latter is a Poisson process, the irreversible mode being an absorbing state. The objective function is standard, increasing in consumption/production (and ultimately, in emissions) and decreasing in the stock of pollution. More importantly, the law of motion of pollution is of the hard irreversibility type as in TW, the pollution decay rate drops to, possibly, zero above a certain pollution level, featuring a non-convex problem, which makes it nontrivial to solve (see again the original paper of Tahvonen and Withagen, 1996).

To deal properly with this non-convexity, we apply a dynamic programming approach as in the deterministic problem studied in Boucekkine et al. (2023). We partly rely on Dockner et al. (2000) to deal with the stochastic extension examined here. Specifying a linear-quadratic objective function additionally permits the derivation of a comprehensive analytical characterization of the optimal policy paths (i.e, that's the paths representing the optimal emission policy as a function of the state variable, the stock of pollution). This in turn allows to tackle the main research question and deeply explore the contemporaneous or short-term optimal behavior in terms of the irreversibility risk. In particular, we prove a couple of new results. One is local: for any value of the Poisson transition probability, optimal emission policy is more aggressive with the irreversibility risk than without in a neighborhood of the pollution irreversibility threshold. This is not so counterintuitive if one keeps in mind the rationale outlined by Clarke and Reed (1994) in their analysis: the 'avoidability' of the irreversible regime argument. Thanks to our global solution, we are also able to conduct the same exercise at any level of pollution and at any value of the Poisson probability. As we will see, the local result showcased above does no longer hold in several parametric cases when the pollution level is far enough from the irreversibility threshold. Incidentally and in contrast to previous work along this line of research, we are also able to study how the optimal policy changes for different level of risks, corresponding to different values of the Poisson probability. This significantly enriches the analysis.

The rest of the paper is organized as following. Section 2 describes the piecewise-deterministic version of the TW model we solve. Section 3 highlights an important preliminary results on the impact of uncertainty on optimal emissions in the neighborhood of the irreversibility pollution threshold. Section 4 provides with the global solution to our optimal control problem and establishes a general characterization of irreversibility thresholds reachability conditions for given risk (as captured by the Poisson arrival rate).

Section 5 uses the global solution of Section 4 to study "globally" the impact of uncertainty on optimal emission, in particular when actual pollution is far from the irreversibility thresholds. Section 6 concludes.

2 The model

Following Tahvonen and Withagen (1996), we investigate a situation where the decision maker faces irreversible pollution accumulation. For simplicity, the pollution emission, $y(t)$, is used to measure the output level. The objective of the decision maker is to maximize social welfare:

$$\max_{y(\cdot)} W = \int_0^{+\infty} (U(y) - D(z))e^{-rt}dt, \tag{1}$$

where r is time preference, $z(t)$ is accumulated pollution, $U(y)$ is the utility from enjoying final output generated with pollution $y(t)$, and $D(z)$ is the damage function due to the aggregate pollution stock z. For analytical tractability, we take linear-quadratic functional forms :

$$U(y) = ay - \frac{y^2}{2}, \qquad D(z) = \frac{cz^2}{2}. \tag{2}$$

Pollution stock $z(t)$ may decay at rate $\delta(z)$. However, the decay rate drops irreversibly to zero as $z(t)$ reaches a threshold value $\bar{z}$. After the drop, no decay is possible. In other words, the pollution accumulation is given by the following:

$$\dot{z} = \begin{cases} y - \delta(z) & \text{if } z < \bar{z}, \\ y & \text{if } z \geq \bar{z}, \end{cases}, \quad z(0) = z_0 \text{ given.} \tag{3}$$

In addition, while $z(t)$ has not reached $\bar{z}$ the decay can still drop to zero due an exogenous shock such as a major ecological accident or a climate singularity bringing the economy instantaneously to the irreversible pollution regime. Quite naturally, we model this occurrence as a piecewise deterministic process (see Davis, 1984, or Dockner et al., 2000): there are two modes, with and without pollution decay, denoted by $m = 1$, the reversible regime, and 0, the irreversible regime, respectively. The jump from mode 1 to 0 occurs at the constant rate

$$\lambda = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\{m(t + \Delta t) = 0 | m(t) = 1\}.$$

In other words, the probability of the mode change during the interval $(t, t+\Delta t]$, given that the mode at t is 1, is proportional to Δt, that is, the arrival of the irreversible regime follows Poisson process with intensity parameter $\lambda \geq 0$. Obviously, when $\lambda = 0$, no regime change happens as long as $z(t) < \bar{z}$.

Before handling the optimization part of our piecewise-deterministic problem, it's worth comparing with the stochastic specifications adopted in the related literature. The closest works to ours are Clarke and Reed (1994), Tsur and Zemel (1998), and to a lower extent, Le Kama et al (2014), which mainly builds on Tsur and Zemel for their stochastic specifications. In Clarke and Reed, the hazard rate, corresponding to the Poisson probability in our setting, depends on the contemporaneous level of pollution while it's taken constant in our fully dynamic approach for analytical tractability. Another common characteristic with the latter work is the fact that the irreversible pollution mode is absorbing, no more jumps are supposed to happen after this mode is reached. Of course, as outlined in Tsur and Zemel (1998) and as also transparent in the environmental literature of tipping points (described in the Introduction), this need not be the case. Of course, one can be interested in other empirical and theoretical contexts requiring further adjustments. For example, hazard rates may also depend on the history of pollution accumulation in additon to its contemporaneous trend (see also Tsur and Zemel, 1998).

Here, we basically study a problem similar to Clarke and Reed's with the simplifying assumption of a constant hazard rate in order to derive the full dynamic implications of the model analytically. This is enough to make analytically our point on the crucial importance of the avoidability of the irreversibile regime (as captured by the distance of actual pollution to the irreversibility threshold) in the shape of optimal economic behavior under risk. This does not mean that the dependence of the hazard rates upon pollution is not a relevant feature of the irreversibility problem,[4] it simply renders analytical study outside the steady states roughly unfeasible. Indeed, as outlined in the Introduction, we solve for the global dynamics, which will allow us to highlight new relevant aspects of the problem. As often, dynamic programming is the natural candidate to conduct a global analysis, especially

[4]In a study of optimal carbon tax framework while considering the probability of irreversible regime switches, van der Ploeg (2014) provides different scenarios under different hazard rates, which includes the precautions behavior before the tipping point, the so-called 'raising the stakes effect.

when the problem is non-convex as ours. We use this method here taking advantage of the piecewise-deterministic nature of the underlying process.

Indeed, along with the piecewise-deterministic stochastics of our model, the planner's optimal control problem can be decomposed into (connected) two sub-problems, or using a more proper terminology, into Periods I and II, corresponding to modes (or regimes) 1 and 0, respectively. Precisely, the per-Period state dynamics are given by:

Period I: $t < T$, where T is the time of mode switching (either by spontaneous jump or as $z(t)$ reaches $\bar{z}$). During this period the state is governed by

$$\dot{z} = y - \delta(z) \qquad \text{for } 0 < t < T; \qquad z(0) = z_0, \tag{4}$$

Period II: $t \geq T$. During this period the state is governed by

$$\dot{z} = y \qquad \text{for } t > T; \;\; z(T) = z\left(T^-\right). \tag{5}$$

We now get to the optimization part (dynamic programming).

Optimal control of the piecewise deterministic process The optimal control of a piecewise deterministic process is far from a new topic in optimization theory as outlined above. Here, we mostly rely on the dynamic programming approach developed in Dockner et al (2000). Let $V_m(z)$ denote the value function in mode m for $m = 0, 1$.

<u>In mode 0</u> there is no possibility of mode change. The optimal control problem is determined by the system dynamic equation (5) and utility and damage functions $U(y)$ and $D(z)$. The usual dynamic programming method leads to the HJB equation

$$rV_0(z) = \max_{y \geq 0} \left\{ U(y) - D(z) + V_0'(z)\, y \right\} = \max_{y \geq 0} \left\{ U(y) + yV_0'(z) \right\} - D(z).$$

Since mode 0 is an absorbing state, the HJB above is the standard deterministic one which only depends on the value function $V_0(z)$. This is not the case in mode 1.

Indeed <u>in mode 1</u>, there is a probability that the mode changes at a given Poisson probability λ. Accordingly, the corresponding HJB equation should be written as (see Theorem 8.1 in Dockner et al. 2000)

$$rV_1(z) = \max_{y \geq 0} \{U(y) - D(z) + V_1'(z)[y - \delta(z)] + \lambda [V_0(z) - V_1(z)]\}.$$

A new term emerges compared to mode 0: the change in value induced by a possible Poisson jump to mode 1 enters the HJB as $\lambda[(V_0(z) - V_1(z))]$, which is nonzero as long as the jump risk is nonzero ($\lambda > 0$). Notice that this mode HJB now includes the two value functions $V_m(z)$, $m = 0, 1$, which features the intertemporal (or inter-period) nature of the problem. The HJB above can be rewritten into a more practical form:

$$(r + \lambda) V_1(z) = \max_{y \geq 0} \{U(y) + yV_1'(z)\} - \delta(z) V_1'(z) - D(z) + \lambda V_0(z).$$

Obviously, $r + \lambda$ plays the role of an effective discount rate.

HJB equations in final form We now use the precise functional specifications given in the beginning of this Section, to write down the HJB equations into the final forms we will handle in our analytical part. Using (2), we find

$$\begin{aligned} \max_{y \geq 0} \{U(y) + yV_0'(z)\} &= \max_{y \geq 0} \left\{ay - \frac{y^2}{2} + yV_m'(z)\right\} \\ &= \begin{cases} (a + V_m'(z))/2 & \text{if } a + V_m'(z) > 0, \\ 0 & \text{if } a + V_m'(z) \leq 0, \end{cases} \end{aligned}$$

for $m = 0, 1$. Therefore, the HJB equations take the form

$$2rV_0(z) = \begin{cases} (a + V_0'(z))^2 - cz^2 & \text{if } a + V_0'(z) \geq 0, \\ -cz^2 & \text{if } a + V_0'(z) < 0, \end{cases} \qquad \text{for } z > 0, \tag{6}$$

and

$$2(r + \lambda) V_1(z) = \begin{cases} \begin{array}{c} (a + V_1'(z))^2 - 2\delta(z) V_1'(z) \\ -cz^2 + 2\lambda V_0(z) \end{array} & \text{if } a + V_1'(z) \geq 0, \\ -2\delta(z) V_1'(z) - cz^2 + 2\lambda V_0(z) & \text{if } a + V_1'(z) < 0, \end{cases} \tag{7}$$

for $0 < z < \bar{z}$. Note that for $z \geq \bar{z}$ there is no Mode 1. Hence, $V_1(z) = V_0(z)$ for $z \geq \bar{z}$. In particular,

$$V_1(\bar{z}) = V_0(\bar{z}). \tag{8}$$

Note that in any mode the emission rate y_m^* that maximizes $U(y) + yV_m'(z)$ is

$$y_m^*(z) = \max\{a + V_m'(z), 0\} \qquad \text{for } m = 0, 1. \tag{9}$$

The corresponding net pollution emission rates ("pollution rates" thereafter) are

$$f_0(z) = y_0^*(z), \qquad f_1(z) = y_1^*(z) - \delta(z). \tag{10}$$

3 Optimal behavior under irreversibility risk: A local result

In this section, we establish one of two main results of this paper: we will show that for any level of uncertainty, that's for any value of λ, optimal (polluting) behaviour is less conservative under the irreversibility risk ($\lambda > 0$) than without ($\lambda = 0$) for an actual pollution level close enough to the threshold level, $\bar{z}$. More concretely, we examine how the irreversibility risk affects $f_1(z)$, the pollution rate in Period I compared to $f_1^d(z)$ which denotes the optimal pollution rate in Mode 1 with $\lambda = 0$. The result is first obtained for $z = \bar{z}$) (Proposition 2), then a more general (local) result is generated around $\bar{z}$ (Proposition 3).

To establish these properties, a careful analysis of the HJB equations and a few intermediate results are needed. We develop some of the essential steps in this Section for transparency, the heaviest mathematical developments are reported in the Appendix. We first derive the value function and the pollution rate in Mode 0 which is needed in this section.

3.1 Value function and emission rate in Mode 0

We first find the value function, $V_0(z)$, using (6).

Note that the right-hand side of (6) is piecewise quadratic in z and V_0'. We seek the solution V_0 to be piecewise quadratic. By (6),

$$V_0(z) = -\frac{cz^2}{2r} \qquad \text{if } a + V_0'(z) < 0.$$

Since $V_0'(z) = -cz/r$, it follows that $a + V_0'(z) < 0$ if and only if

$$z > \frac{ar}{c} = \bar{z}_0.$$

It can be seen from (5) and (9) that $\bar{z}_0$ is the steady state in mode 0. This steady state $\bar{z}_0$ will play a role in our results here below, especially in the analysis of global and local behavior in Section 4. Notice it has a simple and economically meaningful structure: it's proportional to productivity (parameter a) and inversely proportional to the pollution cost (parameter c), the proportionality factor being the discount rate, r: the larger this factor, the bigger the impact of productivity and pollution cost on $\bar{z}_0$.

For $z \leq \bar{z}_0$ we assume

$$V_0(z) = \frac{A_0}{2} z^2 + B_0 z + C_0.$$

Substituting the right-hand side into (6) and comparing coefficients, we find

$$rA_0 = A_0^2 - c, \quad rB_0 = A_0 (B_0 + a), \quad 2rC_0 = (B_0 + a)^2. \tag{11}$$

The quadratic equation for A_0 has two roots, one negative and the other positive. We use the negative one since V_0 is decreasing and concave. Thus

$$A_0 = \frac{r - \sqrt{r^2 + 4c}}{2} \equiv h_0 \tag{12}$$

and consequently,

$$B_0 = \frac{h_0 a}{r - h_0}, \quad C_0 = \frac{(B_0 + a)^2}{2r}. \tag{13}$$

As a result,

$$a + V_0'(z) = h_0 z + \frac{h_0 a}{r - h_0} + a = h_0 z + \frac{ar}{r - h_0}.$$

Using (12) we find

$$\frac{ar}{h_0 (h_0 - r)} = \frac{ar}{c} \equiv \bar{z}_0. \tag{14}$$

Hence,

$$V_0'(z) + a = h_0 \left(z - \frac{ar}{c} \right) = h_0 (z - \bar{z}_0).$$

Since $h_0 < 0$, the above quantity is positive if and only if $z < \bar{z}_0$. Hence

$$V_0(z) = \begin{cases} \frac{1}{2r} \left[h_0^2 (z - \bar{z}_0)^2 - cz^2 \right] & \text{for } z \leq \bar{z}_0, \\ -\frac{c}{2r} z^2 & \text{for } z > \bar{z}_0. \end{cases} \tag{15}$$

By differentiation,

$$V_0' = \begin{cases} \frac{1}{r}\left[h_0^2\left(z-\bar{z}_0\right)-cz\right] & \text{for } z \leq \bar{z}_0, \\ -\frac{c}{r}z & \text{for } z > \bar{z}_0. \end{cases}$$

By the definitions of h_0 and $\bar{z}_0$ in (12) and (14), respectively, one can derive

$$\frac{1}{r}\left[h_0^2 - c\right] = h_0, \qquad \frac{c}{r}\bar{z}_0 = a.$$

Hence,

$$V_0'(z) = \frac{1}{r}\left[h_0^2\left(z-\bar{z}_0\right) - c\left(z-\bar{z}_0\right)\right] - \frac{c}{r}\bar{z}_0 = h_0\left(z-\bar{z}_0\right) - a$$

for $z < \bar{z}_0$. It follows that

$$f_0(z) = \max\left\{a + V_0'(z), 0\right\} = \begin{cases} h_0\left(z-\bar{z}_0\right) & \text{if } z \leq \bar{z}_0, \\ 0 & \text{if } z > \bar{z}_0 \end{cases} \tag{16}$$

The computations above allow to obtain the following result, which is instrumental in the derivation of the main properties of this Section.

Proposition 1 *The pollution rate in Mode 0 never exceeds a. i.e., $0 \leq f_0(z) < a$ for all $z \geq 0$.*

In short, production (and consumption) are strictly bounded by the productivity level of the economy, a. This might seem as an automatic implication of the linear-quadratic utility function, it's not. It depends primarily on the structure of our problem. Since $h_0 < 0$, by (12)

$$0 \leq f_0(z) \leq -h_0\bar{z}_0.$$

By the definitions of h_0 and $\bar{z}_0$ in (12) and (14),

$$-h_0\bar{z}_0 = \frac{\sqrt{r^2+4c}-r}{2}\frac{ar}{c} = \frac{2ra}{\sqrt{r^2+4c}+r} < a.$$

We next derive the pollution rates f_1 and f_1^d in Mode 1, which is definitely much more complicated.

3.2 Pollution rates in Mode 1

We first derive the rates at the threshold level, $\bar{z}$. There are two cases, either $\bar{z} < \bar{z}_0$ or $\bar{z} \geq \bar{z}_0$. In the first case, recall that the transition condition at $\bar{z}$ is the continuity of the value functions given by (8). Using (7) and (8) we find,

$$\begin{aligned} 2rV_0(\bar{z}) + c\bar{z}^2 &= y_1^*(\bar{z})^2 - 2\delta(\bar{z})\left[y_1^*(\bar{z}) - a\right] \\ &= \left[y_1^*(\bar{z}) - \delta(\bar{z})\right]^2 - \left[\delta(\bar{z}) - a\right]^2 + a^2 \\ &= f_1(\bar{z})^2 - \left[\delta(\bar{z}) - a\right]^2 + a^2. \end{aligned} \tag{17}$$

Thus

$$f_1(\bar{z}) = \sqrt{2rV_0(\bar{z}) + c\bar{z}^2 + \left[\delta(\bar{z}) - a\right]^2 - a^2}.$$

We show that

$$2rV_0(\bar{z}) + c\bar{z}^2 = f_0(\bar{z})^2. \tag{18}$$

By (16) and (6), $f_0(\bar{z}) > 0$ and

$$f_0(\bar{z})^2 = \left[V_0'(\bar{z}) + a\right]^2 = 2rV_0(\bar{z}) + c\bar{z}^2.$$

thus, (18) is true. This proves that

$$f_1(\bar{z}) = \sqrt{f_0(\bar{z})^2 + \left[\delta(\bar{z}) - a\right]^2 - a^2}. \tag{19}$$

This equation holds for any $\lambda \geq 0$.

In the latter case where $\bar{z} \geq \bar{z}_0$, $f_0(\bar{z}) = 0$. By (7) and (8),

$$\left(a + V_1'(\bar{z})\right)^2 - 2\delta(\bar{z})\, V_1'(\bar{z}) = 0$$

if $a + V_1'(\bar{z}) > 0$. The above quadratic equation has the positive solution

$$a + V_1'(\bar{z}) = \delta(\bar{z}) + \sqrt{\left[\delta(\bar{z}) - a\right]^2 - a^2}.$$

Hence,

$$f_1(\bar{z}) = \sqrt{\left[\delta(\bar{z}) - a\right]^2 - a^2}$$

if $a + V_1'(\bar{z}) > 0$. If $a + V_1'(\bar{z}) \leq 0$, by (10), $f_1(\bar{z}) = -\delta(\bar{z})$.

We first prove this preliminary important property.

Proposition 2 *Suppose*

$$\bar{z} < \bar{z}_0, \qquad \delta(\bar{z}) < a - \sqrt{a^2 - f_0(\bar{z})^2}. \tag{20}$$

Then $f_1(\bar{z})$ *is given by (19) and satisfies*

$$0 < f_1(\bar{z}) = f_1^d(\bar{z}) \leq f_0(\bar{z}) - \delta(\bar{z}) \qquad \text{for any } \lambda > 0. \tag{21}$$

Furthermore, the inequality is strict if and only if $\delta(\bar{z}) > 0$.

The proof requires some tedious algebra, it's given in the Appendix. At the irreversibility threshold $\bar{z}$, the jump has already occurred and therefore: $f_1(\bar{z}) = f_1^d(\bar{z})$. More interestingly, in case the decay drop at $\bar{z}$ is maximal, net emissions are necessarily discontinuous at the threshold, with an upward jump to the latter emission after the switch to the irreversible regime. In general whatever the drop in the decay rate at $\bar{z}$, the emissions are non-decreasing at this threshold which is a reasonable outcome. Notice that the case covered by Proposition 2 stipulates that $\bar{z} < \bar{z}_0$ which holds true for productivity large enough and/or for a small enough pollution cost by definition of $\bar{z}_0$. In the next section, we shall see that the condition $\bar{z} < \bar{z}_0$ is indeed key for reachability of the irreversibility threshold for given risk, λ.

We next show one of the main results of this paper, the impact of the irreversibility risk on the polluting behavior in the neighborhood of the threshold $\bar{z}$.

Proposition 3 *Suppose (20) holds. Suppose also that*

$$\delta'(\bar{z}) f_0(\bar{z}) \left[a^2 - f_0(\bar{z})^2\right] + 4a^2 \delta(\bar{z})(r - h_0) \leq 0. \tag{22}$$

Then for any $\lambda > 0$ *there is an* $\varepsilon > 0$ *such that* $f_1(z) \geq f_1^d(z)$ *if* $\bar{z} - \varepsilon < z < \bar{z}$.

Corollary 1 *Proposition 3 is true if* $\delta(\bar{z}) = 0$.

The proof is long and tricky, it's given in the Appendix.

Let us elaborate more on the Corollary case which is the reference case. If $\delta(\bar{z}) = 0$, then Condition (22) becomes

$$\delta'(\bar{z}) f_0(\bar{z}) \left[a^2 - f_0(\bar{z})^2\right] \leq 0.$$

This inequality holds since by Proposition 1, $a > f_0(\bar{z})$, and given that $\delta(z) \geq 0$ and $\delta(\bar{z}) = 0$, we necessarily have $\delta'(\bar{z}) \leq 0$. The main conclusion to draw from Proposition 3 is therefore that under mild conditions on the decay rate and **whatever the level of the irreversibility risk as captured by** λ, the optimal polluting behavior is more aggressive under the irreversibility risk than without as soon as the actual pollution stock gets close enough to the threshold $\bar{z}$. The rationale behind has already been pointed out by Clarke and Reed (1994) in its generic form: avoidability. As the economy becomes close enough to the irreversibility threshold, so that the irreversibility regime sounds as unavoidable, the optimal decisions taken by the central planner, taking into account the pollution benefit/cost tradeoff and the law of motion of pollution, will result clearly less pro-environmental. The last section of this paper will show that this is indeed a local property which need not hold far from the irreversibility threshold.

4 Optimal behavior and reachability of the threshold for given risk: Global results

In this Section, we provide with a comprehensive analysis of the solutions to the HJB equation (7) to characterize optimal behavior more globally. This will allow us to explore analytically the reachability of the irreversibility threshold for any level of risk (or value of λ). The latter is fully addressed in Proposition 4 below. The proof of the proposition together with the characterization of the optimal decisions into the form of policy functions (which is itself required for the proof of Proposition 4) is quite long, it's given in the Appendix. The next section will use the policy function analysis in the proof of Proposition 4 to illustrate numerically the variety of optimal behavior under risk depending notably on the actual value of pollution (not necessarily close to $\bar{z}$).

Let's concentrate here on the reachability of the threshold for any level of uncertainty $\lambda \geq 0$ in the absence of any instantaneous Poisson jump from mode 1 to mode 0. We define two types of reachability as follows.

Definition 1 *The threshold $\bar{z}$ is called globally reachable if any pollution stock $z(t)$ with the initial value $z(0) < \bar{z}$ grows across $\bar{z}$ in finite time. The threshold $\bar{z}$ is called locally reachable if there is a stable steady state in mode 1, denoted $\bar{z}_1$, and a critical value, $\bar{z}_1^*$ that satisfy $0 \leq \bar{z}_1 < \bar{z}_1^* < \bar{z}$, such that* *the pollution stock* $z(t)$ *converge to* $\bar{z}_1$ *if* $z(0) < \bar{z}_1^*$ *and* $z(t)$ *grows across* $\bar{z}$ *in finite time if* $z(0) > \bar{z}_1^*$.

The next proposition gives a necessary and sufficient condition for the threshold $\bar{z}$ to be reachable, and shows that it is globally reachable if it is sufficiently small.

To prove this result, we need to specify more precisely the decay function, we impose the following Assumption

(A) $\delta(z)$ is a linear function

$$\delta(z) = \alpha - \beta z.$$

and $\delta(\bar{z}) = 0$.

Clearly $\bar{z} = \alpha/\beta$ under this condition.

Proposition 4 *Suppose (A) holds. Then* $\bar{z}$ *is (globally or locally) reachable if and only if* $\bar{z} < \bar{z}_0$. *Furthermore, there is* $\hat{z} \in (0, \bar{z}_0]$ *such that* $\bar{z}$ *is globally reachable if and only if* $\bar{z} < \hat{z}$.

The proof is in the Appendix. From the proof one can see that for $z < \bar{z}$ and in a neighborhood of $\bar{z}$, the decision maker has two options, one is to produce more, consume more, and as a result, pollute more. The other is the opposite. In the case where $\bar{z}$ is globally reachable, the first option always yields a higher total benefit. On the other hand, if $\bar{z}$ is locally reachable, then there is a pollution level $\bar{z}_1^* < \bar{z}$, such that if $z < \bar{z}_1^*$ the decision maker would be better off restricting production to lower the pollution level, but if $z > \bar{z}_1^*$ the decision maker would rather consume more and thus pollute more. As we shall see in our numerical exercises below, the pollution rate, f_1, may be low enough or even negative for $z < \bar{z}_1^*$ leading to asymptotic convergence to the steady state value in mode 1 (provided it's lower than $\bar{z}_1^*$).

Remark 1 *Of course, the threshold value* $\bar{z}_1^*$ *is endogenously determined, and therefore it may depend on all the parameters of the optimal control problem. Unfortunately, we cannot obtain this threshold in closed form (see the proof of Proposition 4 in the Appendix, Case 2-b). In the next section we show via numerical exercises that it does depend on the pollution cost parameter,* c*, on the irreversibility threshold,* $\bar{z}$*, and on the level of uncertainty,* λ*. In particular, we show that* $\bar{z}_1^*$ *need not be monotonic in* λ *for given* c*. See Example 2 in Section 5.*

Referring to Definition 1, the presence of $\hat{z}$ ensures that under mode 1, either a stable steady state, $\overline{z}_1$, does not exist, or the initial pollution level surpasses $\overline{z}_1$, indicating $z(0) > \overline{z}_1^*$. Consequently, pollution accumulation invariably increases until it exceeds the threshold, $\bar{z}$. Put differently, in a more robust ecological system where the threshold level is high, reaching the irreversible regime can be prevented if the reversible stable steady state in mode 1 has not been already significantly exceeded. However, in a fragile ecological system characterized by a low irreversible threshold or when the pollution level is already near the threshold, irrespective of uncertainties, transitioning into the irreversible regime becomes inevitable. This final statement further elucidates the local behavior of the pollution rate in Proposition 3 - the inevitability of crossing the threshold. In such a scenario, the irreversible stable steady state $\bar{z}_0$ acts as an attractor.

We can now examine the impact of uncertainty on optimal emission rates for all $z \in (0, \bar{z}]$, complementing our local analysis (around the threshold $\bar{z}$) in Section 3. This will be done in the next Section.

5 Irreversibility, uncertainty level and the optimal emission rate

In the following in order to use the analytical characterization of the optimal solutions obtained and used in the proof of Proposition 4 (see the Appendix), we keep on relying on Assumption (A). We compare the optimal emission rates for different levels of pollution unit costs (parameter c), irreversibility thresholds ($\bar{z}$) and uncertainty (parameter λ), keeping the following parameters fixed:

$$r = 0.2, \qquad \beta = 0.1, \qquad a = 18.$$

These parameter values are used in the numerical example of Tahvonen and Withagen (1996). In the first example, we use $c = 0.002$ and $\bar{z}$ taking values of 100, 180, and 300 respectively. In the second example we use $c = 0.02$ and $\bar{z} = 100$, 120, and 140 respectively. Also, in each example, we compare the optimal pollution rates $f_1(z)$ with zero, small and large values of λ. Recall that function $f_1(z)$ are optimal feedback functions in the sense of dynamic programming. Specifically, we use $\lambda = 0, 0.1$ and 1.0, respectively in each numerical illustration. Also recall that $\lambda = 0$ corresponds to the absence of the irreversibility risk: with the notations of Section 3, we have in such a case, $f_1(.) \equiv f_1^d(.)$.

Example 1 *-small damage with* $c = 0.002$. *In this case* $\bar{z}_0 = 1800$. *For* $\bar{z} = 100$, *the threshold* $\bar{z}$ *is globally reachable for all three values of* λ. *For* $\bar{z} = 180$, *the threshold is globally reachable for* $\lambda = 0$ *and it is locally reachable for* $\lambda = 0.1$ *and* 1.0, *with* $\bar{z}_1^* = 4.8$ *and* 8.3, *respectively. In the latter case, for any initial pollution level below* $\bar{z}_1^*$, $z(t)$ *decreases to 0 as* $t \to \infty$. *The graphs are shown in Fig. 1.*

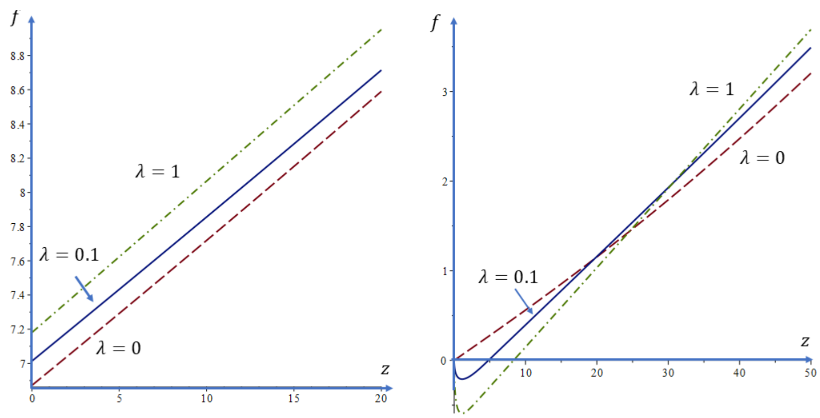

Figure 1: Small damage with $c = 0.002$ and $\bar{z} = 100$ (left) and $\bar{z} = 180$ (right).

For $\bar{z} = 300$, *the threshold is locally reachable for the three values of* λ, *with* $\bar{z}_1^* = 150$ *for* $\lambda = 0$, $\bar{z}_1^* = 145.6$ *for* $\lambda = 0.1$ *and* $\bar{z}_1^* = 142.4$ *for* $\lambda = 1.0$. *In all cases, for any initial pollution below* $\bar{z}_1^*$, $z(t) \to 0$ *as* $t \to \infty$. *The graphs are shown in Fig. 2.*

In Fig. 1 and Fig. 2, one first important conclusion is that the value of the irreversibility threshold is crucial in the ranking of the optimal feedback functions for the three different Poisson rate values. In Fig. 1, when $\bar{z} = 100$, the optimal emission rate is the highest at the largest Poisson rate, the lowest corresponding to the case of zero irreversibility risk, for any value of the pollution stock z. This seems to extend the local result established in Proposition 3 globally. However, when $\bar{z} = 180$, one can clearly see in the

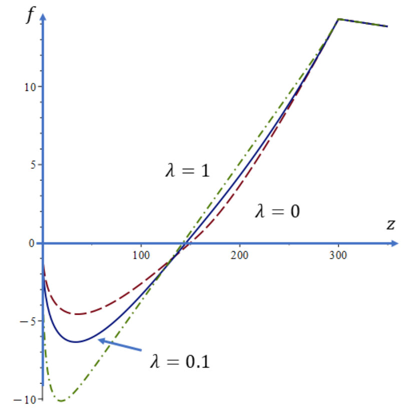

Figure 2: Small damage with $c = 0.002$ and $\bar{z} = 300$.

graphic at the right side of Fig. 1 that while the latter picture holds when z is large enough (consistently with Proposition 3), it's no longer the case when z is low enough: in this z-values interval (say z below the threshold value $z_1^* = 8.3$ corresponding to $\lambda = 1$), the optimal emission rate is the highest at the lowest Poisson rate (that's the zero risk case), the lowest corresponding to the highest irreversibility risk. Of course, this ranking is reversed as the pollution level increases enough in accordance with Proposition 3. The very same outcome arises from Fig. 2 where the irreversibility threshold is even higher ($\bar{z} = 300$), and also in the exercises conducted in Example 2 with a much higher pollution unit cost (See Fig. 3 and 4).

Our numerical exercises deliver more findings. One interesting outcome is that for given pollution unit cost, and for large enough irreversibility threshold values, $\bar{z}$, the latter may only be locally reachable for certain Poisson rate values: for $\lambda = 0.1, 1$ in Fig. 1, and for all λ values in Fig. 2 when the irreversibility threshold is very large. Notice that in these cases the endogenous threshold z_1^* is increasing in λ for $\bar{z} = 180$ but it's decreasing in λ when $\bar{z} = 300$ (of course at the discrete λ-values considered). It's neither monotonic as we will see in the Example 2 below and associated Figures 3 and 4, when the unit pollution cost is much higher.

Another remarkable property we can see in Fig. 1 (right one), and Fig. 2 is the possible non-monotonicity of the policy functions, $f_1(.)$. This is also apparent in the case of large pollution cost (Fig. 3 and 4). One frequent case of non-monotonicity is when reachability is only local but this is not a sufficient condition for non-monotonicity (see Example 2 below, Fig. 4 for $\lambda = 1$). Indeed, when reachability is local and the actual pollution stock is below the endogenous threshold, $\bar{z}_1^*$, the optimal net emission rates may even be negative when the actual pollution stock is low enough, leading the pollution stock to converge to zero asymptotically, corresponding to the steady state in mode 1, that's $\bar{z}_1 = 0$ (Example 1, Figure 2). Figure 3 documents similar cases when the pollution cost is much larger: in these cases, when pollution is below $\bar{z}_1^*$, the optimal net emission rate may decrease sharply initially for certain values of the Poisson rate, without being negative, then converging to a strictly positive steady state in mode 1, $\bar{z}_1 > 0$.

Example 2 *-large damage with $c = 0.02$. In this case $\bar{z}_0 = 180$. For $\bar{z} = 100$, the threshold is globally reachable for all three values of λ. For $\bar{z} = 120$, it is globally reachable with $\lambda = 1.0$, but is locally reachable with $\lambda = 0$ and 0.1. In the latter cases, $\bar{z}_1^* = 113.5$ for $\lambda = 0$ and $\bar{z}_1^* = 108.6$ for $\lambda = 0.1$. In addition, for any initial pollution below $\bar{z}_1^*$ the pollution stock converges to the steady state $\bar{z}_1 = 60$ for $\lambda = 0$ and $\bar{z}_1 = 98$ for $\lambda = 0.1$. The pollution rates are shown in Fig. 3. For $\bar{z} = 140$, the threshold is locally reachable for*

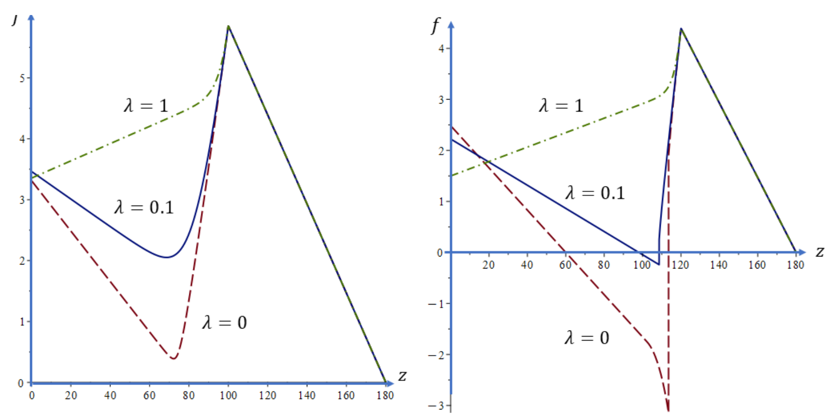

Figure 3: Large damage with $c = 0.02$ and $\bar{z} = 100$ (left) and $\bar{z} = 120$ (right).

all three values of λ, with $\bar{z}_1^* = 138.1$ for $\lambda = 0$, $\bar{z}_1^* = 138.7$ for $\lambda = 0.1$ and $\bar{z}_1^* = 25.2$ for $\lambda = 1.0$. For an initial value of z below $\bar{z}_1^*$, $z(t)$ converges to $\bar{z}_1 = 40$ for $\lambda = 0$, $\bar{z}_1 = 43.4$ for $\lambda = 0.1$, and $\bar{z}_1 = 0$ for $\lambda = 1.0$. The graphs are shown in Fig. 4.

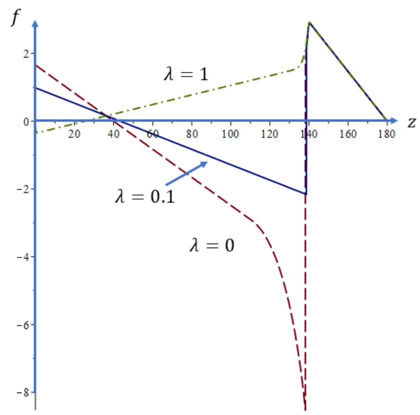

Figure 4: Large damage with $c = 0.02$ and $\bar{z} = 140$.

6 Concluding remarks

In this paper, we have enriched the seminal irreversible pollution model *à la* Tahvonen and Withagen (1996) with an exogenous source of switching to irreversibility following the simplest process, a Poisson process with constant arrival probability. With a linear-quadratic objective function, and while keeping a general non-convex specification of pollution decay, we have been able to extract new results, which are far from inconsistent with the generic "avoidability" argument put forward by Clarke and Reed (1994): more aggressive economic behavior (in terms of emissions) under the irreversibility risk when actual pollution is close enough to the irreversibility threshold counterpart whatever the level of the risk (as measured by the Poisson arrival rate), and possible reversals when actual pollution is far enough from this threshold. These reversals depends also on the other parameters of the model, for example the unit pollution cost and, more crucially, on the value of the irreversibility threshold iself.

All the results are generated analytically with the exceptions of the numerical illustrations of the last section (which do derive directly from closed-form solutions). While the Poisson arrival rate does not depend on pollution, the analytical case we have constructed permits to highlight complementary highly relevant aspects which cannot be obtained from a steady state-based approach. Also the possibility to study in depth the implications of sudden changes in the value of the Poisson arrival rate does somehow mimick the situations where, for many obvious reasons, the irreversibility risk may rise or decrease sharply. Of course, this is not equivalent to endogenizing the risk but it helps estimating analytically the consequences of exogenously moving risk.

References

[1] Bonnet, R., Swingedouw, D., Gastineau, G. et al (2021). Increased risk of near term global warming due to a recent AMOC weakening. *Nature Communications*, 12, 6108 (2021). https://doi.org/10.1038/s41467-021-26370-0

[2] Boucekkine, R., W. Ruan and B. Zou (2023). The irreversible pollution game. *Journal of Environmental Economics and Management*, 120, Article 102841.

[3] Boucekkine, R., A. Pommeret and F. Prieur (2013). Technological vs. ecological switch and the Environmental Kuznets Curve. *American Journal of Agricultural Economics*, 95, 252-260.

[4] Bretschger, L., and Vinogradova (2019). Best policy response to environmental shocks: Applying a stochastic framework. *Journal of Environmental Economics and Management*, 97, 23-41.

[5] Clarke, H., and W. Reed (1994). Consumption/pollution tradeoofs in an environemental vulnerable to pollution-related catastrophic collpase. *Journal of Economic Dynamics and Control*, 18, 991-1010.

[6] Davis, M. H. A. (1984). Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society (B)*, 46, 353-388.

[7] Diekert, F. (2017). Threatening Thresholds? The effect of disastrous regime shifts on the non-cooperative use of environmental goods and service. *Journal of Public Economics*, 147, 30-49.

[8] Ditlevsen, P., and S. Ditlevsen (2023). Warning of a forthcoming collapse of the Atlantic meridional overturning circulation. *Nature Communications*, 14, 4254. https://doi.org/10.1038/s41467-023-39810-w

[9] Dockner, E. J., S. Jϕrgensen, N. V. Long and G. Sorger (2000). *Differential Games in Economics and Management Science*, Cambridge, Cambridge University Press.

[10] Dong, S., M. Baringer, and G. Goni (2019). Slow Down of the Gulf Stream during 1993-2016. *Scientific Reports*, 9, 6672. https://doi.org/10.1038/s41598-019-42820-8

[11] Le Kama A., A. Pommeret and F. Prieur (2014). Optimal emission policy under the risk of irreversible pollution. *Journal of Public Economic Theory*, 16(6), 959-980.

[12] Maler, K.G. (2000). Development, ecological resources and their management: A study of complex dynamic systems. *European Economic Review*, 44(4-6), 645-665.

[13] Nordhaus, W. (2019). Climate change: The ultimate challenge for economics. *American Economic Review*, 109(6), 1991-2014.

[14] Prieur, F. (2009). The environmental Kuznets curve in a world of irreversibility. *Economic Theory*, 40, 57-90.

[15] Ren, B., and S. Polasky (2014). The optimal management of renewable resources under the risk of potential regime shift. *Journal of Economic Dynamics and Control*, 40, 195-212.

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

[17] Tsur, Y., and A. Zemel (1998). Pollution control in an uncertain environment. *Journal of Economic Dynamics and Control*, 22, 967-975.

[18] van der Ploeg, F. (2014). Abrupt positive feedback and the social cost of carbon. *European Economic Review*, 67, 28-41.

[19] Wagener, F., and A. de Zeeuw (2021). Stable partial cooperation in managing systems with tipping points. *Journal of Environmental Economics and Management*, 109, Article 102499.

Appendix

Proof of Proposition 2

Since the right-hand side of (19) is independent of λ, it follows that $f_1(\bar{z}) = f_1^d(\bar{z})$ for any $\lambda \geq 0$. Either $\delta(\bar{z}) = 0$ or $\delta(\bar{z}) > 0$. In the first case, since $f_0(\bar{z}) > 0$ because $\bar{z} < \bar{z}_0$, it follows that $f_1(\bar{z}) = f_0(\bar{z}) > 0$. In the second case, the second relation in (20) implies that

$$f_0(\bar{z})^2 > a^2 - [\delta(\bar{z}) - a]^2 .$$

Hence, by (19) we again have $f_1(\bar{z}) > 0$ for any $\lambda \geq 0$.

It remains to show $f_1^d(\bar{z}) \leq f_0(\bar{z}) - \delta(\bar{z})$. We define g_1 and g_2 as functions of δ by

$$g_1(\delta) = \sqrt{f_0(\bar{z})^2 + (\delta - a)^2 - a^2}, \qquad g_2(\delta) = f_0(\bar{z}) - \delta$$

Then

$$g_1(0) = g_2(0) = f_0(\bar{z}) .$$

By differentiation,

$$g_1'(\delta) = \frac{\delta - a}{\sqrt{f_0(\bar{z})^2 + (\delta - a)^2 - a^2}} < 0, \qquad g_2'(\delta) = -1.$$

Since by Proposition 1, $f_0(z) < a$, it follows that $g_1'(\delta) < -1$. Therefore $g_1(\delta) < g_2(\delta)$, which implies

$$f_1^d(\bar{z}) = g_1(\delta(\bar{z})) \leq g_2(\delta(\bar{z})) = f_0(\bar{z}) - \delta(\bar{z}) .$$

Furthermore, the strict inequality holds if and only if $\delta(\bar{z}) = 0$.

This completes the proof.

Proof of Proposition 3

Define functions

$$\phi_0(z) = V_0'(z) + a, \qquad \phi_1(z) = V_1'(z) + a - \delta(z) . \tag{23}$$

It can be seen that $\phi_m(z) = f_m(z)$ if $V_m'(z) + a \geq 0$. Also, since $V_0'(z) + a \geq 0$ if and only if $z \leq \bar{z}_0$, it follows that

$$\phi_0(z) = f_0(z) \qquad \text{for } z \leq \bar{z}_0.$$

Differentiate the both sides of (7) with respect to z. We obtain

$$\phi_1(z)\,\phi_1'(z) = (r+\lambda)\,\phi_1(z) + (r+\delta'(z))\,(\delta(z)-a) + cz + \lambda\,[\delta(z) - f_0(z)]$$

if $\phi_1(z) > -\delta(z)$ and

$$\begin{aligned} -\delta(z)\,\phi_1'(z) &= (r+\lambda)\,\phi_1(z) + (r+\delta'(z))\,(\delta(z)-a) + cz + \lambda\,[\delta(z) - f_0(z)] \\ &\quad +\delta'(z)\,(\phi_1(z)+\delta(z)) \end{aligned}$$

if $\phi_1(z) \le -\delta(z)$. The two cases can be combined into

$$\begin{aligned} \max\{\phi_1(z), -\delta(z)\}\,\phi_1'(z) &= (r+\lambda)\,\phi_1(z) + (r+\delta'(z))\,(\delta(z)-a) + cz \\ &\quad +\lambda\,[\delta(z) - f_0(z)] + \delta'(z)\min\{\phi_1(z)+\delta(z), 0\}. \end{aligned} \tag{24}$$

By (7) and (8),

$$\phi_1(\bar{z})^2 = f_0(\bar{z})^2 + [\delta(\bar{z}) - a]^2 - a^2. \tag{25}$$

By (20) the right-hand side is positive. Therefore equation (24) is non-singular. Using the existence and uniqueness theorem for differential equation, we see that $\phi_1(z)$ exists and is positive for $z < \bar{z}$ and is near $\bar{z}$. Hence, $\phi_1(z) = f_1(z)$ for such z.

To examine the values of $f_1(z)$ we use expansions

$$\begin{aligned} \delta(z) &= \delta_0 + \delta_1(z-\bar{z}) + \frac{\delta_2}{2}(z-\bar{z})^2 + o\left((z-\bar{z})^2\right), \\ f_1(z) &= A_2 + C_2(z-\bar{z}) + \frac{D_2}{2}(z-\bar{z})^2 + o\left((z-\bar{z})^2\right) \end{aligned}$$

where

$$\delta_0 = \delta(\bar{z}), \qquad A_2 = f_1(\bar{z}) = \sqrt{f_0(\bar{z})^2 + (\delta_0 - a)^2 - a^2} > 0.$$

Substituting the right-hand side of the first equation into (24), we obtain

$$\begin{aligned} & [A_2 + C_2(z-\bar{z})]\,[C_2 + D_2(z-\bar{z})] \\ = \; & (r+\lambda)\,[A_2 + C_2(z-\bar{z})] + (r+\delta_1+\delta_2(z-\bar{z}))\,(\delta_0 - a + \delta_1(z-\bar{z})) \\ & + c(z-\bar{z}) + c\bar{z} + \lambda\,(\delta_0 + \delta_1(z-\bar{z}) - h_0(z-\bar{z}) - h_0(\bar{z}-\bar{z}_0)) + o(|z-\bar{z}|) \end{aligned}$$

Comparing the constant terms on the two sides, we find

$$A_2 C_2 = rA_2 + (\delta_0 - a)(r+\delta_1) + c\bar{z} + \lambda\,(A_2 + \delta_0 - h_0(\bar{z}-\bar{z}_0)) \tag{26}$$

for all $\lambda \geq 0$. Since $A_2 > 0$, it follows that $C_2 > C_2^d$ for $\lambda > 0$ if and only if

$$A_2 + \delta_0 - h_0 (\bar{z} - \bar{z}_0) > 0 \tag{27}$$

We show that (27) cannot hold. Suppose it does. Note that

$$A_2 = f_1^d (\bar{z}) = f_1^d (\bar{z}),$$

and by (16),

$$h_0 (\bar{z} - \bar{z}_0) = f_0 (\bar{z}).$$

Hence, (27) holds if $f_1^d (\bar{z}) + \delta_0 > f_0 (\bar{z})$. This contradicts Proposition 2. As a result, $C_2 \leq C_2^d$.

In the case where the strict inequality, $C_2 < C_2^d$ holds, we have $f_1 (z) > f_1^d (z)$ for $z < \bar{z}$ and is near $\bar{z}$. If $C_2 = C_2^d$, we compare the first order terms in the expansion, which leads to

$$C_2^2 + A_2 D_2 = (r + \lambda) C_2 + (r + \delta_1) \delta_1 + (\delta_0 - a) \delta_2 + c + \lambda (\delta_1 - h_0).$$

It follows that

$$D_2 = \frac{1}{A_2} \left[-C_2^2 + (r + \lambda) C_2 + (r + \delta_1) \delta_1 + (\delta_0 - a) \delta_2 + c + \lambda (\delta_1 - h_0) \right].$$

Setting $\lambda = 0$, we obtain

$$D_2^d = \frac{1}{A_2} \left[-C_2^2 + r C_2 + (r + \delta_1) \delta_1 + (\delta_0 - a) \delta_2 + c \right].$$

Hence,

$$D_2 = D_2^d + \frac{\lambda}{A_2} \left[C_2^d + \delta_1 - h_0 \right]. \tag{28}$$

We show

$$C_2^d + \delta_1 - h_0 \geq 0. \tag{29}$$

By (26),

$$C_2^d = r + \frac{1}{A_2} \left[(\delta_0 - a) (r + \delta_1) + c\bar{z} \right].$$

Hence,

$$\left[C_2^d + \delta_1 - h_0 \right] A_2 = (r + \delta_1 - h_0) A_2 + (\delta_0 - a) (r + \delta_1) + c\bar{z}.$$

It suffices to show that the right-hand side of the above equation is positive. Using $ra = c\bar{z}_0$ we can write the right-hand side as

$$-\left(a-\delta_0-A_2\right)\delta_1+\left(r-h_0\right)A_2-c\left(\bar{z}_0-\bar{z}\right)+r\delta_0. \tag{30}$$

Observe that by (19),

$$A_2=f_1\left(\bar{z}\right)=\sqrt{f_0\left(\bar{z}\right)^2+\left(\delta_0-a\right)^2-a^2}$$

and by Proposition 1 and (20), $a>\delta_0$ and

$$A_2<a-\delta_0.$$

Since $\delta_1\leq 0$, it follows that the first term in (30) is nonnegative. We next use (12) to derive

$$\left(r-h_0\right)h_0=\left(r-\frac{r-\sqrt{r^2+4c}}{2}\right)\frac{r-\sqrt{r^2+4c}}{2}=-c.$$

Hence,

$$\begin{aligned}\left(r-h_0\right)A_2-c\left(\bar{z}_0-\bar{z}\right) &= \left(r-h_0\right)\left[A_2+h_0\left(\bar{z}_0-\bar{z}\right)\right]\\ &= \left(r-h_0\right)\left[A_2-f_0\left(\bar{z}\right)\right].\end{aligned}$$

$$\left[C_2^d+\delta_1-h_0\right]A_2=-\left[a-\delta_0-A_2\right]\delta_1+\left(r-h_0\right)\left[A_2-f_0\left(\bar{z}\right)\right]+r\delta_0.$$

Since $\delta_1\leq 0$ and $A_2\leq a-\delta_0\leq a$, it follows that

$$-\left[a-\delta_0-A_2\right]\delta_1=\frac{-\delta_1\left[a^2-f_0\left(\bar{z}\right)^2\right]}{a-\delta_0+A_2}\geq\frac{-\delta_1}{2a}\left[a^2-f_0\left(\bar{z}\right)^2\right].$$

Also, since $\delta_0\geq 0$ and $A_2\geq 0$, by (25),

$$A_2-y_0^*\left(\bar{z}\right)=\frac{\delta_0\left[\delta_0-2a\right]}{f_0\left(\bar{z}\right)+A_2}\geq\frac{-2a\delta_0}{f_0\left(\bar{z}\right)}.$$

Therefore, by (22),

$$\begin{aligned}\left[C_2^d+\delta_1-h_0\right]A_2 &\geq \frac{-\delta_1}{2a}\left[a^2-f_0\left(\bar{z}\right)^2\right]+\delta_0\left[\frac{-2a\left(r-h_0\right)}{f_0\left(\bar{z}\right)}+r\right]\\ &\geq \frac{-\delta_1}{2a}\left[a^2-f_0\left(\bar{z}\right)^2\right]+\frac{-2\delta_0 a\left(r-h_0\right)}{f_0\left(\bar{z}\right)}\geq 0.\end{aligned}$$

This completes the proof.

Proof of Proposition 4

We first prove that the threshold, $\bar{z}$, is reachable if and only if $\bar{z} < \bar{z}_0$.

Suppose $\bar{z} \geq \bar{z}_0$. Then, by (8) and (15),

$$V_1(\bar{z}) = V_0(\bar{z}) = -\frac{c\bar{z}^2}{2r}.$$

Hence, by (7) and Assumption (A),

$$f_1(\bar{z}) = a + V_1'(\bar{z}) = 0. \tag{31}$$

Differentiate the both sides of (7) with respect to z, we obtain

$$f_1(z) f_1'(z) = (r+\lambda) f_1(z) + (r + \delta'(z))(\delta(z) - a) + cz + \lambda[\delta(z) - V_0'(z) - a].$$

This equation is equivalent to the dynamical system

$$\begin{aligned} \dot{x} &= (r+\lambda) x + (r + \delta'(z) + \lambda)\delta(z) - a(r + \delta'(z)) + cz - \lambda[V_0'(z) + a], \\ \dot{z} &= x, \end{aligned} \tag{32}$$

where $x = f_1(z)$ and the differentiation is with respect to t. If $\bar{z}$ is reached in finite time, by (31), the trajectory passes through the point $(0, \bar{z})$ at certain time $\bar{t}$. Hence, $(0, \bar{z})$ cannot be an equilibrium of the dynamical system and so, $\dot{x}(\bar{t}) < 0$. From the first equation in (32), we find

$$-a(r + \delta'(\bar{z})) + c\bar{z} - \lambda[V_0'(\bar{z}) + a] < 0.$$

However, since $\bar{z} \geq \bar{z}_0$, it follows that $V_0'(\bar{z}) + a \leq 0$. Also, since $\delta(z) \geq \delta(\bar{z}) = 0$ for $z < \bar{z}$, it follows that $\delta'(\bar{z}) \leq 0$. The above inequality implies

$$-ar + c\bar{z} < a\delta'(\bar{z}) + \lambda[V_0'(\bar{z}) + a] \leq 0.$$

Using $ar = c\bar{z}_0$, we obtain $\bar{z} < \bar{z}_0$. This contradicts the assumption of $\bar{z} \geq \bar{z}_0$. Hence, $\bar{z}$ cannot be reached in finite time.

Suppose $\bar{z} < \bar{z}_0$. By (16), $f_0(z) > 0$. As shown in the proof of Proposition 3, $\phi_1(z)$ defined by (23) exists at least locally for z near $\bar{z}$. Furthermore, $\phi_1(\bar{z}) = f_0(\bar{z}) > 0$. Hence, $\phi_1(z) > 0$ for z near $\bar{z}$. With $\phi_1(z)$ solved, one finds the value function $V_1(z)$ by (7), which leads to

$$\begin{aligned} V_1(z) = \tfrac{1}{2(r+\lambda)} \{ & [\max\{\phi_1(z), -\delta(z)\} - \delta(z)](\phi_1(z) + \delta(z)) \\ & + 2a\delta(z) - cz^2 + 2\lambda V_0(z)\}. \end{aligned} \tag{33}$$

Therefore, the optimal strategy leads to z reaching $\bar{z}$ in finite time.

We now prove the second part of the proposition. That is, there is $\hat{z} \in (0, \bar{z}_0]$ such that $\bar{z}$ is globally reachable if and only if $\bar{z} < \hat{z}$. For this purpose, we derive the solution of the optimal control problem.

We first notice that any solution ϕ_1 of (24)-(25) defines a value function V_1 by (33). In general, at any pollution level, z, the decision maker has two choices, either to produce more and pollute more, or produce less and pollute less. We use $\tilde{f}_1$ to denote the solution ϕ_1 of (24)-(25) which is positive near $\bar{z}$, and use $\hat{f}_1$ to denote the solution which is negative near $\bar{z}$. Substituting $\tilde{f}_1(z)$ and $\hat{f}_1(z)$ for $\phi_1(z)$ in (33), we obtain respective functions $\hat{V}_1(z)$ and $\tilde{V}_1(z)$. Clearly the decision maker chooses the strategy based on the larger solution. Hence

$$V_1(z) = \max\left\{\tilde{V}_1(z), \hat{V}_1(z)\right\} \qquad \text{for } z \in (0, \bar{z}). \tag{34}$$

At points where only $\tilde{f}_1$ or $\hat{f}_1$ exists, there is no ambiguity in the definition of V_1. If at a point z^* the decision maker switches from one strategy to the other, it is necessary that $\tilde{V}_1(z^*) = \hat{V}_1(z^*)$ holds. By (33), $\tilde{f}_1(z^*)$ and $\hat{f}_1(z^*)$ are related by

$$\hat{f}_1(z^*) = \begin{cases} -\tilde{f}_1(z^*) & \text{if } \tilde{f}_1(z^*) < \delta(z^*), \\ -\tilde{f}_1(z^*)^2 / [2\delta(z^*)] - \delta(z^*)/2 & \text{if } \tilde{f}_1(z^*) \geq \delta(z^*). \end{cases} \tag{35}$$

To construct a solution to (34), we first construct solutions $\tilde{f}_1$ and $\hat{f}_1$ of Eq. (24) on the interval $(0, \bar{z})$, then compare the corresponding functions $\tilde{V}_1$ and $\hat{V}_1$. This is achieved by converting Eq. (24) into dynamical systems.

For $\tilde{f}_1$, since it is positive in a neighborhood of $\bar{z}$, Eq. (24) is equivalent to the dynamical system

$$\begin{aligned} \dot{x} &= (r+\lambda)x + (r-\beta+\lambda)\delta(z) + (\beta-r)a + cz - \lambda f_0(z), \\ \dot{z} &= x, \end{aligned} \tag{36}$$

where $x(t) = \tilde{f}_1(z(t))$. Note that the right-hand sides of the equations in (36) is linear in x and z, and can be written as

$$\dot{x} = (r+\lambda)x + B_1 z + C_1$$

where

$$\begin{aligned} B_1 &= \beta(\beta - r - \lambda) + c - \lambda h_0, \\ C_1 &= (r-\beta+\lambda)\alpha + (\beta-r)a + \lambda h_0 \bar{z}_0. \end{aligned} \tag{37}$$

There is an equilibrium at $(\bar{z}_1, 0)$ where

$$\bar{z}_1 = -C_1/B_1. \tag{38}$$

The Jacobian matrix takes the form

$$J = \begin{pmatrix} r+\lambda & B_1 \\ 1 & 0 \end{pmatrix}.$$

The eigenvalues are

$$\begin{aligned} h_1 &= \tfrac{1}{2}\left[r+\lambda-\sqrt{(r+\lambda)^2+4B_1}\right], \\ h_2 &= \tfrac{1}{2}\left[r+\lambda+\sqrt{(r+\lambda)^2+4B_1}\right] \end{aligned} \tag{39}$$

Depending on whether $B_1 > 0$ or $B_1 < 0$, the equilibrium is a saddle point or a repeller.

As for $\hat{f}_1$, we first define the zx-plane regions

$$R_1 = \{(z,x) : x > -\delta(z)\}, \qquad R_2 = \{(z,x) : x \leq \delta(z)\}.$$

Eq. (24) is equivalent to the dynamical system (36) in R_1 and

$$\begin{aligned} \dot{x} &= (r+\lambda-\beta)\,x + B_1 z + C_1 - \beta\delta(z), \\ \dot{z} &= -\delta(z) \end{aligned} \tag{40}$$

in R_2, where $x(t) = \hat{f}_1(z(t))$. Observe that by the assumption $\delta(\bar{z}) = 0$ and relation (35), there is no solution $\hat{f}_1$ that satisfies $\hat{V}_1(\bar{z}) = V_0(\bar{z})$. Indeed, by Proposition 2,

$$\tilde{f}_1(\bar{z}) = y_1^*(\bar{z}) > 0 = \delta(\bar{z}).$$

Hence, by (35),

$$\hat{f}_1(\bar{z}) = -\left[\tilde{f}_1(\bar{z})^2 + \delta(\bar{z})^2\right] / (2\delta(\bar{z}))$$

which does not exist. We also notice from (35) that $\hat{f}_1(\bar{z}) \to -\infty$ as $z \to \bar{z}$. Furthermore, we observe that if $\hat{f}_1(0)$ exists and $\hat{f}_1(0) < 0$, then $z(t) = 0$ for all $t > 0$. This is possible only if $\hat{f}_1(0) = 0$. Hence the solution curve $\left(z, \hat{f}_1(z)\right)$ starts from the point $(0,0)$. If $\hat{f}_1(0)$ exists and $\hat{f}_1(0) > 0$, then

$z(t)$ is increasing as long as $\hat{f}_1(z) > 0$. Hence, $\hat{f}_1(z)$ can only vanish at the steady state $\bar{z}_1$. This can only happen if $(\bar{z}_1, 0)$ is a saddle point and $\left(z, \hat{f}_1(z)\right)$ is on its stable manifold.

There are two cases, $B_1 > 0$ and $B_1 \leq 0$. In each case there are three subcases: (a) $\bar{z}_1 \leq 0$, (b) $0 < \bar{z}_1 \leq \bar{z}$ and (c) $\bar{z}_1 > \bar{z}$. In each subcase we show the existence of a positive $\hat{z}$ such that $\bar{z}$ is globally reachable if and only if $\bar{z} < \hat{z}$.

Case 1. $B_1 > 0$**.**

In this case $(0, \bar{z}_1)$ is a saddle point. In addition, $\langle h_i, 1\rangle$ is an eigenvector of the Jacobian J corresponding to the eigenvalue h_i. Hence, the stable manifold at the equilibrium has the negative slope and the unstable one has the positive slope.

Case 1-a. $\bar{z}_1 \leq 0$**.** Since $(\bar{z}_1, 0)$ is a saddle point, there is a trajectory passing through $(0,0)$. Then, $\hat{f}_1(z)$ is defined by the part of this trajectory below the z-axis. The part of this trajectory above the z-axis intersects the vertical line $z = \bar{z}$. Let $\bar{f}_1 > 0$ be the x-coordinate of this intersection. Then, the trajectory starting at the point $(\bar{z}, f_0(\bar{z}))$ lies above the z-axis for all $z \in [0, \bar{z}]$ if

$$f_0(\bar{z}) = h_0(\bar{z} - \bar{z}_0) \geq \bar{f}_1, \tag{41}$$

and if the above inequality does not hold, the trajectory starting at $(\bar{z}, f_0(\bar{z}))$ intersects the z-axis within the interval $(0, \bar{z})$. In the latter case, we let $\bar{z}_1^*$ be a solution to the equation (35). Note that at the zero of $\tilde{f}$ the right-hand side of the above equation vanishes while the left-hand side is negative. We show that for $z < \bar{z}$ and is sufficiently close to $\bar{z}$, the right-hand side approaches $-\infty$ faster than the left-hand side. This would imply the existence of solution (35). Note that $\tilde{f}_1(\bar{z}) = f_0(\bar{z}) > 0$ and

$$\delta(z) = \beta(\bar{z} - z).$$

Hence

$$\tilde{f}_1(z)^2 / [2\delta(z)] + \delta(z)/2 = O\left(|\bar{z} - z|^{-1}\right). \tag{42}$$

To estimate $\hat{f}_1$ as $z \to \bar{z}^-$, we solve (24) which which takes the form

$$\begin{aligned} -\delta(z)\hat{f}_1'(z) &= (r+\lambda)\hat{f}_1(z) + (r-\beta)(\delta(z) - a) + cz \\ &\quad +\lambda[\delta(z) - f_0(z)] - \beta\left[\hat{f}_1(z) + \delta(z)\right]. \end{aligned}$$

This is a linear equation. Multiplying the integrating factor

$$\mu(z) = (\bar{z} - z)^{1-(r+\lambda)/\beta}$$

to the both sides of the equation and integrate from $\bar{z}$ to z, we obtain

$$\hat{f}_1(z)(\bar{z} - z)^{1-\frac{r+\lambda}{\beta}} = \int_{\bar{z}}^{z} (\bar{z} - s)^{1-\frac{r+\lambda}{\beta}} [2\beta - r - \lambda - (\beta - r) a - cs + \lambda f_0(s)] \, ds. \tag{43}$$

If $2\beta > r + \lambda$, then the right-hand side approaches zero as $z \to \bar{z}$. Thus, by l'Hôpital's rule,

$$\begin{aligned} \lim_{z \to \bar{z}^-} \frac{\hat{f}_1(z)}{\bar{z} - z} &= \lim_{z \to \bar{z}^-} \frac{\int_{\bar{z}}^{z} (\bar{z} - s)^{1-\frac{r+\lambda}{\beta}} [2\beta - r - \lambda - (\beta - r) a - cs + \lambda f_0(s)] \, ds}{(\bar{z} - z)^{2-(r+\lambda)/\beta}} \\ &= -\frac{2\beta - r - \lambda - (\beta - r) a - c\bar{z} + \lambda f_0(\bar{z})}{2 - (r + \lambda)/\beta}. \end{aligned}$$

This implies that

$$\hat{f}_1(z) = O(|z - \bar{z}|).$$

If $2\beta < r + \lambda$, the right-hand side of (43) diverges. We have

$$\begin{aligned} \lim_{z \to \bar{z}} \hat{f}_1(z)(\bar{z} - z) &= \lim_{z \to \bar{z}^-} \frac{\int_{\bar{z}}^{z} (\bar{z} - s)^{1-\frac{r+\lambda}{\beta}} [2\beta - r - \lambda - (\beta - r) a - cs + \lambda f_0(s)] \, ds}{(\bar{z} - z)^{-(r+\lambda)/\beta}} \\ &= \lim_{z \to \bar{z}^-} (\bar{z} - z)^2 [-2\beta + r + \lambda + (\beta - r) a + cz - \lambda f_0(z)] = 0. \end{aligned}$$

Hence,

$$\hat{f}_1(z) = o\left(|\bar{z} - z|^{-1}\right).$$

In any case, in view of (42),

$$\hat{f}_1(z) > -\frac{\tilde{f}_1(z)^2}{2\delta(z)} - \frac{\delta(z)}{2}$$

if $z < \bar{z}$ and is sufficiently close to $\bar{z}$. Therefore, there is at least one solution $\bar{z}_1^*$. Let

$$\tilde{V}_1(z) = \frac{1}{2(r + \lambda)} \left\{ \tilde{f}_1(z)^2 - \delta(z)^2 + 2a\delta(z) - cz^2 + 2\lambda V_0(z) \right\} \qquad \text{for } 0 \le z \le \bar{z} \tag{44}$$

and

$$\hat{V}_1(z) = \frac{1}{2(r+\lambda)} \cdot \begin{cases} \hat{f}_1(z)^2 - \delta(z)^2 + 2a\delta(z) - cz^2 + 2\lambda V_0(z) & \text{if } \hat{f}_1(z) > -\delta(z), \\ 2\delta(z)\left[a - \hat{f}_1(z) - \delta(z)\right] - cz^2 + 2\lambda V_0(z) & \text{if } \hat{f}_1(z) \leq -\delta(z). \end{cases} \tag{45}$$

Then there is at least one point at which $\tilde{V}_1(z) = \hat{V}_1(z)$. Define $V_1(z)$ by (34), we obtain a continuous value function. (See Fig. 5.) As a conclusion,

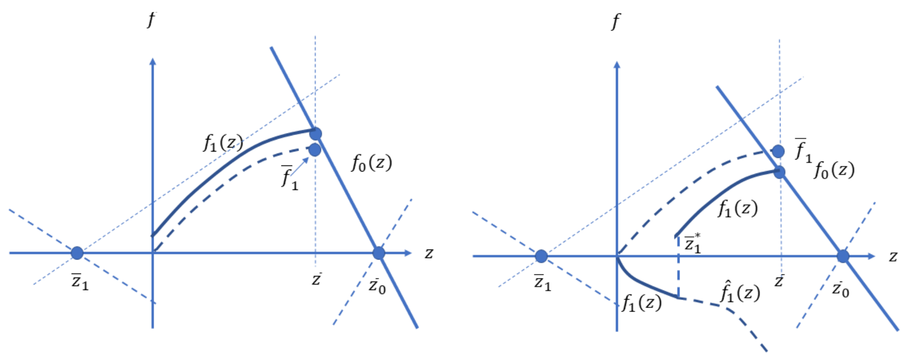

Figure 5: Case 1-a with $f_0(\bar{z}) \geq \bar{f}_1$ (left) and $f_0(\bar{z}) < \bar{f}_1$ (right).

$\bar{z}$ is globally reachable if and only if (41) holds. Thus,

$$\hat{z} = \bar{z}_0 + \frac{\bar{f}_1}{h_0} \tag{46}$$

in this case.

Case 1-b. $0 < \bar{z}_1 \leq \bar{z}$**.** In this case the trajectory that passes through $(0,0)$ stays in the region where $z \leq 0$. Hence, $\hat{f}_1(0) > 0$. Therefore, $\left(z, \hat{f}_1(z)\right)$ is on the stable manifold of the equilibrium $(\bar{z}_1, 0)$. This implies that

$$\hat{f}_1(z) = h_1(z - \bar{z}_1) \qquad \text{if } h_1(z - \bar{z}_1) > -\delta(z).$$

For z that satisfies

$$h_1(z - \bar{z}_1) \leq -\delta(z),$$

we solve (24) for $z > z'$ with the initial condition

$$\hat{f}_1(z') = -\delta(z'),$$

where

$$z' = \bar{z}_1 - \frac{\delta(z')}{h_1}.$$

Let

$$\bar{f}_1 = h_2 (\bar{z} - \bar{z}_1).$$

Then the trajectory starting at the point $(\bar{z}, f_0(\bar{z}))$ lies above the z-axis if (41) holds, and it intersects the z-axis on the interval $(0, \bar{z})$ if the reversed inequality holds. In the latter case the intersection point is greater than $\bar{z}_1$. (See Fig. 6.)

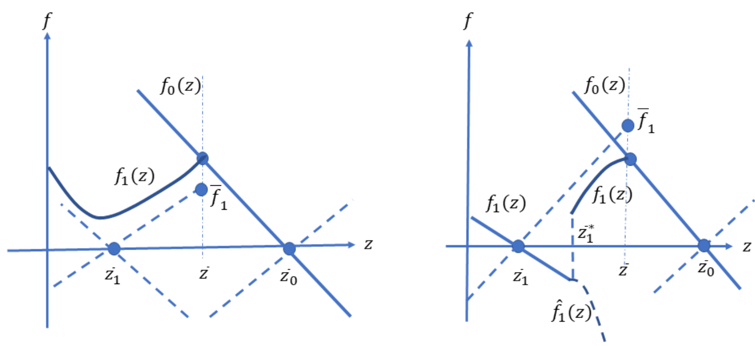

Figure 6: Case 1-b with $f_0(\bar{z}) \geq \bar{f}_1$ (left) and $f_0(\bar{z}) < \bar{f}_1$ (right).

From the above discussion we see that $\bar{z}$ is globally reachable if $\bar{z} < \hat{z}$ where $\hat{z}$ is given by (46).

Case 1-c. $\bar{z}_1 > \bar{z}$**.** In this case any trajectory starting at $(0, 0)$ does not enter the region $z > 0$, and any trajectory below the z-axis for $z \in (0, \bar{z})$ does not intersect the z-axis on this interval. Thus $\hat{f}_1(z)$ is undefined for the entire interval $[0, \bar{z}]$. For any positive value of $f_0(\bar{z})$, $\tilde{f}_1(z)$ is defined for all $z \in [0, \bar{z}]$. Thus only $\tilde{V}_1(z)$ exists. (See Fig. 7.)

Since any $\bar{z} < \bar{z}_0$ is globally reachable, it suffices to choose $\hat{z} = \bar{z}_0$.

Case 2. $B_1 \leq 0$**.**

In this case $(0, \bar{z}_1)$ is a repeller and both unstable manifolds have position slopes in the fz-plane. As $z \to \bar{z}_1$, points $\left(z, \tilde{f}_1(z)\right) \to (\bar{z}_1, 0)$ along the unstable manifold $x = Y_1(z - \bar{z}_1)$. There are three subcases.

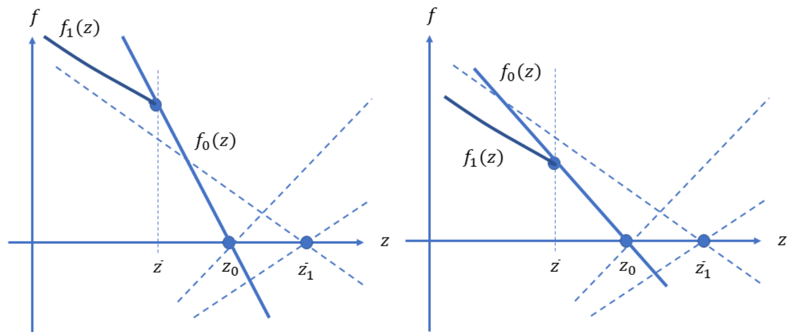

Figure 7: Case 1-c with $f_0(\bar{z}) \geq h_1(\bar{z} - \bar{z}_1)$ (left) and $f_0(\bar{z}) < h_1(\bar{z} - \bar{z}_1)$ (right).

Case 2-a. $\bar{z}_1 \leq 0$**.** In this case the trajectory that passes through $(0, 0)$ does not enter the region below the z-axis for $z > 0$. Hence, $\hat{f}_1(z)$ is not defined for any $z \in (0, \bar{z})$. On the other hand, $\tilde{f}_1(z)$ is defined for all $z \in [0, \bar{z}]$ with any value of $f_0(\bar{z})$. (See Fig. 8.)

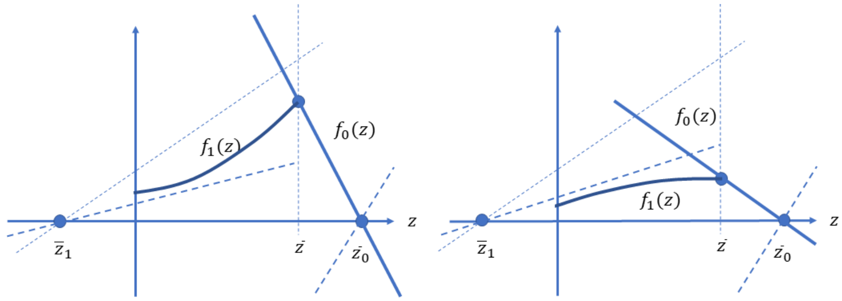

Figure 8: Case 2-a with $f_0(\bar{z}) \geq h_1(\bar{z} - \bar{z}_1)$ (left) and $f_0(\bar{z}) < h_1(\bar{z} - \bar{z}_1)$ (right).

In this case we again see that $\hat{z} = \bar{z}_0$ since $\bar{z}$ is globally reachable for any $\bar{z} < \bar{z}_0$.

Case 2-b. $0 < \bar{z}_1 \leq \bar{z}$**.** In this case the trajectory passing through $(0, 0)$ enter into the region $z > 0$ both above and below the z-axis. The one below the z-axis joins the equilibrium $(0, \bar{z}_1)$. Hence, $\hat{f}_1(z) < 0$ for $0 < z < \bar{z}_1$. The one above the z-axis intersects the vertical line $z = \bar{z}$ at a point, denoted by $\bar{f}_1$. In the case where $f_0(\bar{z}) \geq \bar{f}_1$, $\tilde{f}_1(z)$ is defined and positive for all $z \in [0, \bar{z}]$. See the left graph in Fig. 9). On the other hand, if

$$h_2(\bar{z} - \bar{z}_1) < f_0(\bar{z}) < \bar{f}_1,$$

$\tilde{f}_1(z') = 0$ at some $0 < z' < \bar{z}_1$. It is clear that

$$\tilde{f}_1(\bar{z}_1) > 0 = \hat{f}_1(\bar{z}_1), \quad \tilde{f}_1(z') = 0 > \hat{f}_1(z').$$

By (44)-(45), the first chained inequality of the above implies that

$$\tilde{V}_1(\bar{z}_1) = \frac{1}{2(r+\lambda)} \left\{ -\delta(\bar{z})^2 + 2a\delta(\bar{z}) - c\bar{z}^2 + 2\lambda V_0(\bar{z}) \right\} < \hat{V}_1(\bar{z}_1),$$

and the second implies that

$$\begin{aligned} \tilde{V}_1(z') &= \frac{1}{2(r+\lambda)} \left\{ \tilde{f}_1(z') - \delta(z')^2 + 2a\delta(z') - c(z')^2 + 2\lambda V_0(z') \right\} \\ &> \frac{1}{2(r+\lambda)} \left\{ 2\delta(z')[a - \delta(z')] - c(z')^2 + 2\lambda V_0(z') \right\} = \hat{V}_1(z'). \end{aligned}$$

Hence, there is a point $\bar{z}_1^*$ such that $\hat{V}_1(\bar{z}_1^*) = \tilde{V}_1(\bar{z}_1^*)$. Define $V_1(z)$ by (34) and

$$f_1(z) = \begin{cases} \hat{f}_1(z) & \text{if } z < \bar{z}_1^*, \\ \tilde{f}_1(z) & \text{if } z > \bar{z}_1^*. \end{cases} \tag{47}$$

We obtain a continuous value function. See the right graph in Fig. 9.

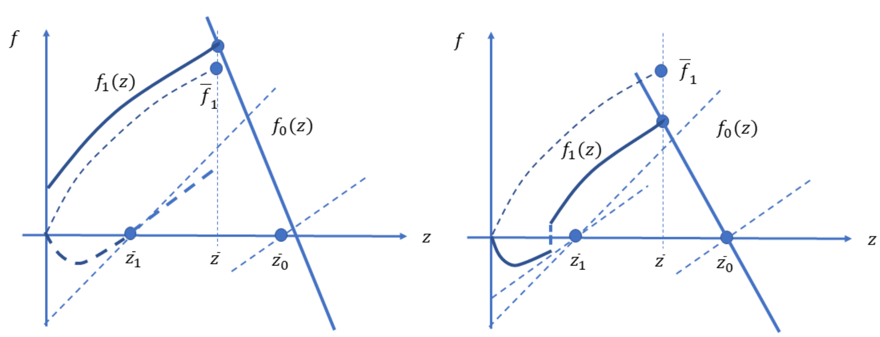

Figure 9: Case 2-b with $f_0(\bar{z}) \geq \bar{f}_1$ (left) and $h_2(\bar{z} - \bar{z}_1) \leq f_0(\bar{z}) < \bar{f}_1$ (right).

In the case where

$$f_0(\bar{z}) \leq h_2(\bar{z} - \bar{z}_1),$$

The trajectory that passes through $(\bar{z}, f_0(\bar{z}))$ approaches $(\bar{z}_1, 0)$ along the unstable manifold $(z, h_1(z - \bar{z}))$. Thus, $\tilde{V}_1(z)$ is defined for $z > \bar{z}_1$ and $\hat{V}_1(z)$ is defined for $z < \bar{z}_1$. We define

$$V_1(z) = \begin{cases} \hat{V}_1(z) & \text{for } z < \bar{z}_1, \\ \tilde{V}_1(z) & \text{for } z > \bar{z}_1, \end{cases}$$

and define $f_1(z)$ similarly. The graph of $f_1(z)$ is shown in Fig. 10.

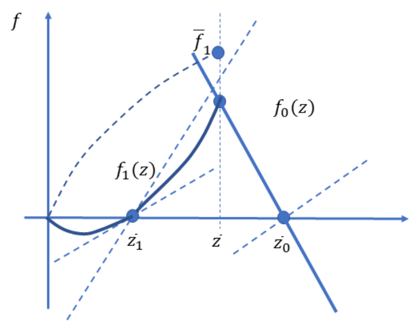

Figure 10: Case 2-b with $f_0(\bar{z}) < h_2(\bar{z} - \bar{z}_1)$.

As can be seen, $\bar{z}$ is globally reachable if and only if $\bar{z} < \hat{z}$ with $\hat{z}$ defined by (46).

Case 2-c. $\bar{z}_1 > \bar{z}$**.** Similar to Case 2-b, $\tilde{f}(z) \geq 0$ for all $z \in (0, \bar{z})$ if $f_0(\bar{z}) \geq \bar{f}_1$ and $\tilde{f}(z) = 0$ on the interval $(0, \bar{z})$. In the former case, $V_1(z) = \tilde{V}_1(z)$ on $[0, \bar{z}]$. In the latter case, there is a point $\bar{z}_1^*$ such that $\hat{V}_1(\bar{z}_1^*) = \tilde{V}_1(\bar{z}_1^*)$. We define $V_1(z)$ and $f_1(z)$ by (34) and (47), respectively. See Fig. 11.

It is clear that $\bar{z}$ is globally reachable if and only if $\bar{z} < \hat{z}$ with $\hat{z}$ given by (46).

This completes the proof.

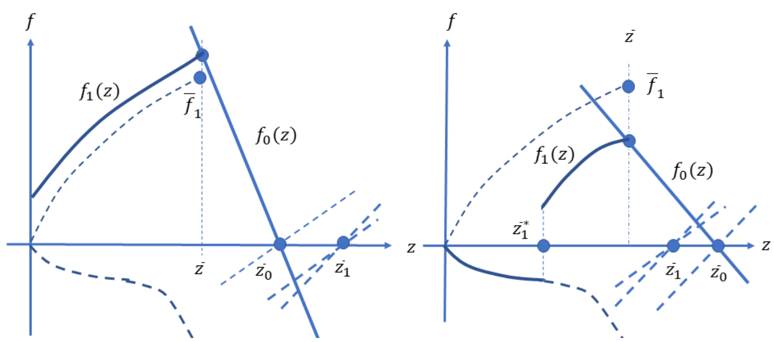

Figure 11: Case 2-c with $f_0(\bar{z}) \geq \bar{f}_1$ (left) and $f_0(\bar{z}) < \bar{f}_1$ (right).

INSTITUT DE RECHERCHE
ÉCONOMIQUES ET SOCIALES

Place Montesquieu 3
1348 Louvain-la-Neuve

ISSN 1379-244X D/2024/3082/01

UCLouvain

IRES | LIDAM Louvain Institute of Data Analysis and Modeling in economics and statistics