TECHNISCHE UNIVERSITÄT WIEN

Operations Research and Control Systems

SWM ORCOS

On the Matthew effect in research careers: Abnormality on the boundary

Gustav Feichtinger, Dieter Grass, Peter Kort, Andrea Seidl

Research Report 2020-01

January 2020

ISSN 2521-313X

Operations Research and Control Systems
Institute of Statistics and Mathematical Methods in Economics
Vienna University of Technology

Research Unit ORCOS
Wiedner Hauptstraße 8 / E105-04
1040 Vienna, Austria
E-mail: orcos@tuwien.ac.at

On the Matthew effect in research careers: Abnormality on the boundary

Gustav Feichtinger[a], Dieter Grass[b], Peter M. Kort[c,d], Andrea Seidl[e]

[a] *Department for Operations Research and Control Systems, Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, Vienna, Austria*
[b] *International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria*
[c] *Tilburg School of Economics and Management, Tilburg University, Tilburg, Netherlands*
[d] *Department of Economics, University of Antwerp, Antwerp, Belgium*
[e] *Department of Business Decisions and Analytics, University of Vienna, Vienna, Austria*

Abstract

The observation that a socioeconomic agent with a high reputation gets a disproportionately higher recognition for the same work than an agent with lower reputation is typical in career development and wealth. This phenomenon, which is known as Matthew effect in the literature, leads to an increasing inequality over time. The present paper employs an optimal control model to study the implications of the Matthew effect on the optimal efforts of a scientist into reputation.

The solution of the model exhibits, for sufficiently low effort costs, a new type of unstable equilibrium at which effort is at its upper bound. This equilibrium, which we denote as Stalling Equilibrium, serves as a threshold level separating success and failure in academia. In addition we show that at the Stalling Equilibrium the solution can be abnormal. We provide a clear economic interpretation for this solution characteristic.

Keywords: optimal control, history-dependent solution, abnormal solution, career planning, Matthew effect

1. Introduction

The quality of his/her published work is certainly an important input to increase a scientist's reputation. However, delivering good work does not guarantee recognition in academia. One of the underlying problems is that reputation is self-enforcing in the sense that contributions of relatively unknown researchers receive disproportionately less attention and recognition than the work of well-established scientists. This self-enforcement effect concerning the reputation for scientific work is known as the *Matthew effect* in the literature. The name of the effect refers to a passage in the Gospel of Matthew, which reads:

"*For unto everyone that hath shall be given, and he shall have abundance; but from him that hath not shall be taken away even that which he hath.*"

This biblical parable has been transferred to scientific production by Merton (25) and has also been applied later to socioeconomic issues; compare Merton (27) and the rich literature cited in this essay.

The Matthew effect is visible, for example, in scientific conferences. The plenary talks are inherently held by prominent speakers, who further increase their visibility by giving these presentations. The Matthew effect can also be present in the event of submitting papers to peer-reviewed journals. In an ideal world the quality of the paper, and not the reputation of the authors, should decide about getting a paper published or not. In reality it is not without reason that some journals enforce a double-blind peer review process to ensure fairness, or check with all coauthors if they contributed to

Email addresses: gustav@eos.tuwien.ac.at (Gustav Feichtinger), dieter.grass@tuwien.ac.at (Dieter Grass), Kort@uvt.nl (Peter M. Kort), andrea.seidl@univie.ac.at (Andrea Seidl)

the paper,[1] see e.g. Wislar et al. (40). A high reputation helps to obtain a position as a full professor, which in turn raises access to resources leading to an increased scientific output. Even just narrowly losing in the job application process could damage a scientific career. In a study analyzing early career academic funding in The Netherlands, Bol et al. (11) find that - despite similar academic background - winners of academic funding accumulate more than twice as much funding as non-winners with near-identical but slightly lower review scores.[2] Azoulay et al. (4) study the impact of becoming a Howard Hughes Medical Institute Investigator on citations and obtain that winning the recognition leads to a citation boost for less known researchers, while the effect for known researchers is small. Merton (25) is obviously the first who translates the biblical parable into the secular accumulation of advantage and disadvantage for scientists. Since then a large number of papers related to the sociology of science has emerged, see for example Merton (26), Dannefer (15), Allison et al. (1), DiPrete and Eirich (17), Mayer (23), Bask and Bask (6, 7).

There are certain factors a scientist cannot influence (at least not immediately), which strengthen the Matthew effect. Such factors include the current position of a scientist (being full professor is more impressive than being assistant or associate professor) and the reputation of the researcher's university. Clauset et al. (13), who studies factors affecting the faculty hiring process, finds that there are reputational benefits associated with working at a prestigious institution. Cole (14) obtains that papers of equal quality, but written by authors of greater scientific reputation, will receive rapid peer recognition, as assessed by the larger number of citations assigned to them. Moreover, he could identify a distinct advantage of scientists located in departments of US - elite universities. Ethnicity, religion, or gender of a scientist can also contribute to the Matthew effect. Women being denied recognition for their work is also known as *Matilda effect* in the literature, see Rossiter (29).

To study the impact of the Matthew effect we use a continuous-time optimal control framework. We consider the scientist's effort as control variable and his/her reputation as state variable, an approach which is analogous to modeling the accumulation of human capital over the life cycle; see Becker (8), Ben-Porath (9), Blinder and Weiss (10). For an application of life cycle models to scientific productivity we refer to McDowell (24), Levin and Stephan (22), and the survey of Stephan (36, sect. 8). A more recent analysis of the development of scientific production over time is provided by Seidl et al. (30) and **?**). Yegorov et al. (41) studies how the market access of a scientist affects his or her success and finds that the outcome can be history-dependent in the sense that it depends on the initial value of the scientist's human capital and the market access.

The present paper also obtains that the development of a research career can be history-dependent, but this time it is due to the Matthew effect and the fact that the scientist's available effort needed to create scientific output is limited. If the initial reputation is low, it is impossible for the scientist to build up a sufficiently high reputation needed to prevail in academia. A scientist can only be successful if the reputation at the beginning of the planning period is already large, caused for instance by his/her well-known institution or if he or she had already built up a substantial reputation before the career actually had started (e.g. during the writing of the Master thesis or during early stages of the PhD studies). We find that in case of low cost of scientific effort there exists a threshold between success and failure, which we will denote as Stalling Equilibrium.[3] At the Stalling Equilibrium the corresponding reputation level has the interesting property that the scientist has to use maximal efforts to stay put. Working less would lead to a loss of reputation from which it is not possible to recover, whereas increasing reputation is not possible as the scientist is already employing maximal effort. Reputation

[1] A form of fraud that has been observed in academia is adding a famous author without his or her agreement or even knowledge to the list of co-authors to improve the chances of being published.

[2] Note, however, that for US data Wang et al. (38) find that near-miss applicants outperform successful candidates in the long-run

[3] The term Stalling Equilibrium originates from aerodynamics and aviation, where it denotes the separation point at which the flow at a foil becomes turbulent and the lift reaches its maximum, see e.g. Stengel (35).

just being slightly above this threshold would guarantee a bright future in academia with less efforts required in the long run due to the Matthew effect.

By the introduction of the concept of the Stalling Equilibrium, we extend the literature on history-dependent solutions, see e.g. Grass et al. (18), Wagener (37), Kiseleva and Wagener (20), Caulkins et al. (12). If, in addition to having low cost of scientific effort, which is a necessary existence condition for the Stalling Equilibrium, also the positive influence of the Matthew effect is relatively large and the scientist is able to put a lot of effort in increasing his or her reputation, the Stalling Equilibrium has the property that the solution is abnormal at this particular point. Abnormality means that the instantaneous contribution to the objective has to be ignored at this particular point to derive the optimal strategy, see e.g. Halkin (19). Abnormal solutions have found little attention in economic problems as they were considered - as the name says - as something abnormal that can only occur in ill-posed models. We find, however, that in a well-posed model, the abnormality results from the upper bound on the control variable, which is the scientific effort in our case. In particular we obtain that in close proximity of the Stalling Equilibrium, the costate variable measuring the marginal value of reputation gets very large whereas it is infinite at the Stalling Equilibrium itself. Intuitively this can be explained as follows. At the Stalling Equilibrium the control effort is at its upper bound, and reputation stays at the same level. If it would have been possible to increase effort further, which is thus not the case, reputation would grow and the decision maker could substantially increase the objective value. Hence, raising its reputation level is not possible for the scientist at the Stalling Equilibrium, but it would have been possible, however, if the state variable has an infinitesimally larger value. A solution with an infinite value of the costate is not an admissible solution by means of the maximum principle. The only way to find a feasible solution is therefore to resort to the abnormal problem.

A history-dependent solution where a Stalling Equilibrium separates solution paths with qualitatively different long-run outcomes, just can occur for relatively low effort costs. To give a complete picture of the solution of the model, we have to study other scenarios as well. To do so, in addition we analytically derive conditions for the admissibility and optimality of a stable interior steady state as well as of the Stalling Equilibrium, and threshold equilibria where effort is not maximal. The latter points to the existence of what is known as a Skiba point in the literature.[4] Furthermore, we discuss the occurring bifurcations, i.e. we determine sets of parameter values at which the optimal solution changes qualitatively.

The paper is structured as follows. In Section 2 the model is presented. Section 3 states the necessary optimality conditions using Pontryagin's maximum principle. Section 4 provides a detailed analysis of the structure of the solutions. The bifurcation analysis presented in Section 5 delivers valuable insights into the qualitative optimal dynamic behavior of the scientist. Section 6 analyzes how certain model assumptions affect the optimal solution. Section 7 concludes.

2. Model description

We consider an optimal control framework with one state variable $K(t)$, denoting the (scientific) reputation of the decision maker, a scientist. Reputation can increase in two ways. First, the scientist can conduct scientific effort to improve his or her reputation. This can be done by doing research, by active participation in scientific events, by networking, by competently contributing to the peer review process, etc. The amount of effort is denoted by the control variable $I(t)$. Second, $K(t)$ is influenced by the so-called Matthew effect. The core idea behind the Matthew effect is that reputation grows on its own when it is high. However, when K is small, this feedback effect is negative.

[4] The point is also sometimes denoted as DNSS or indifference point to acknowledge contributions by Sethi (32, 33), Skiba (34), Dechert and Nishimura (16).

The reputation stock develops according to the usual capital accumulation equation enriched by a term $M(K)$, which we denote Matthew function, and which measures the Matthew effect. The state equation reads

$$\dot{K}(t) = I(t) - \delta K(t) + M(K(t)),$$

where δ denotes the depreciation rate, or rate of forgetfulness, of reputation. It is essential that the Matthew function $M(K)$ is positive for large values of K, but negative for a small reputation. In particular, we define the Matthew function as

$$M(K) := a - \frac{d}{K}. \tag{1}$$

The idea is that recognizable scientists have an advantage in the sense that reputation is self-enforcing when K is large. In such a situation reputation grows by itself, which can even happen without conducting effort. On the other hand, if the reputation is small, it is almost impossible to succeed in the scientific world. Parameter a denotes the self-enforcement rate of reputation and describes how much reputation can grow at most on its own when reputation is large. Parameter d is denoted as discrimination parameter as it measures the extent to which a lack of reputation negatively affects its growth. Both parameters account for factors such as the scientist's position, the reputation of the scientist's university, but also (in some environments) for the scientist's ethnicity or gender. Note that the division of d by K reflects that discrimination is a larger problem for the scientist when his or her reputation is small. Parameters a and d also incorporate individual characteristics of the scientist, as personal appearance can have a substantial impact on whether or not being accepted or finding support within the peer group.

The parameter $I_{\max}$ denotes the capacity of the scientist. It reflects that there exists a maximal number of hours per day that a scientist is able to work:

$$I(t) \leq I_{\max}.$$

Note that this parameter can strongly vary between scientists as it incorporates on the one hand preferences with respect to the work-life balance and on the other hand time restrictions arising from care obligations for children and other dependents.

The scientist's utility increases with his or her reputation. Utility concavely depends on reputation, because when already famous, an additional unit of reputation does not do much. This utility can relate to monetary rewards, as reputation gives the scientist a better bargaining position in negotiations about his or her salary. However, it can also relate to the fact that a higher reputation makes that the scientist is able to publish a larger number of papers, gets invited to conferences, wins awards, gains popularity, etc. A typical concavely increasing function satisfying these characteristics is $g(K) = \ln(K + 1)$.

While certain activities to improve reputation might certainly create satisfaction for the scientific agent on their own (e.g. writing papers, attending conferences), we assume that conducting effort generates (convexly increasing) disutility[5] equal to $c(I) = cI^2$. The parameter c is negatively related to the scientist's talent, i.e., it is less costly to build up reputation for a scientist if he or she is a good researcher.

The objective of the scientist is to maximize the discounted utility stream net of the cost of effort.

[5] In Section 6 we also consider the impact of positive utility of effort.

This implies that the complete model can now be written as

$$V(K_0, I(\cdot), T) := \int_0^T \mathrm{e}^{-rt} \left(\ln(K(t) + 1) - cI(t)^2 \right) \mathrm{d}t, \tag{2a}$$

$$\max_{I(\cdot),T} V(K_0, I(\cdot), T), \tag{2b}$$

$$\text{s.t.} \quad \dot{K}(t) = I(t) - \delta K(t) + a - \frac{d}{K(t)}, \tag{2c}$$

$$0 \leq I(t) \leq I_{\max}, \tag{2d}$$

$$K(T) \geq 0, \tag{2e}$$

$$K(0) = K_0 > 0. \tag{2f}$$

The planning horizon is denoted by T, and is optimally determined, i.e. the scientist has the option to leave academia. For simplicity we assume that a scientist has no benefits from scientific reputation after quitting. We allow for the possibility that the optimal planning horizon is infinite. Despite tertiary education having a positive effect on life expectancy, see Winkler-Dworak (39), no scientist can stay in academia forever. However, it seems reasonable to assume that a scientist does not take leaving because of retirement (or health issues) into account when planning to built up a sustainable career and reputation. The scientist's discount rate is given by r and reflects the scientist's time preference with respect to collecting rewards from reputation net of the cost of effort.

While the next section provides analytical results, we also have to resort to numerical calculations. For this, the Matlab toolbox `OCMat`[6] is used. All parameter values are strictly positive and the base case parameter values are shown in Table 1. Note that it is not possible to generate one set of parameter values that suits all scientists, as the different parameters strongly reflect the individual talent and academic background of a scientist. To generate general insights, we analyze possible outcomes for different scenarios. In particular, the values of cost parameter c, which is related to the scientist's talent, and a, the self-enforcement rate of reputation, are considered as bifurcation parameters.

r	δ	c	d	a	$I_{\max}$
0.03	0.1	$*$	1	$*$	1

Table 1: Base case parameter values.

3. Necessary Optimality Conditions

The appropriate way to deal with a model of this type, namely a deterministic optimal control model in continuous time, is to apply Pontryagin's maximum principle (see, e.g. 18). To do so, we introduce the current-value Hamiltonian, in which we explicitly have to consider the factor λ_0,[7] yielding

$$\mathcal{H}(K, I, \lambda, \lambda_0) = \lambda_0 \left(\ln(K + 1) - cI^2 \right) + \lambda \left(I - \delta K + a - \frac{d}{K} \right),$$

where λ denotes the current-value costate variable. The Lagrangian, with the Lagrangian multiplier ν for the control constraint (2d) can be written as[8]

$$\mathcal{L}(K, I, \lambda, \nu, \lambda_0) = \mathcal{H}(K, I, \lambda, \lambda_0) + \nu(I_{\max} - I).$$

[6]Available at http://orcos.tuwien.ac.at/research/ocmat_software.

[7]An optimal control problem is called *normal* if $\lambda_0 = 1$ and called *abnormal* if $\lambda_0 = 0$. For an in-depth treatment of the topic see Arutyunov (2).

The derivative of the Hamiltonian and Lagrangian with respect to the control I yields

$$\begin{aligned}\partial_I \mathcal{H} &= -\lambda_0 2Ic + \lambda,\\ \partial_I \mathcal{L} &= \partial_I \mathcal{H} - \nu.\end{aligned}$$

From the maximum principle

$$I^*(t) = \underset{0 \le I \le I_{\max}}{\operatorname{argmax}}\ \mathcal{H}(K^*(t), I, \lambda(t), \lambda_0), \tag{3}$$

the following expressions for the control and Lagrangian multipliers can be derived:

$$I^\circ(\lambda, \lambda_0) := \begin{cases} \dfrac{\lambda}{2c\lambda_0} & \text{for } -\lambda_0 2I_{\max}c + \lambda \le 0 \\ I_{\max} & \text{otherwise} \end{cases}, \quad \nu(\lambda, \lambda_0) = \begin{cases} 0 & \text{for } I < I_{\max} \\ -\lambda_0 2I_{\max}c + \lambda & \text{for } I = I_{\max}. \end{cases} \tag{4}$$

The term $I^\circ(\lambda(t), \lambda_0)$ denotes the investments fulfilling the necessary optimality conditions and equals the optimal investments $I^*(t)$ if $\lambda(t)$ is the according costate of the optimal solution. Not surprisingly, if the marginal utility of reputation, denoted by λ, is large, it is optimal to conduct a lot of effort to increase reputation. If it is costly to conduct effort, i.e. c is large, then effort will be low.

The canonical system is given by

$$\dot{K}(t) = I^\circ(\lambda(t), \lambda_0) - \delta K(t) + a - \frac{d}{K(t)}, \tag{5a}$$

$$\dot{\lambda}(t) = r\lambda(t) - \frac{\lambda_0}{K(t)+1} - \lambda(t)\left(\frac{d}{K(t)^2} - \delta\right). \tag{5b}$$

Hence, equilibria of this canonical system satisfy

$$I^\circ(\lambda, \lambda_0) - \delta K + a - \frac{d}{K} = 0, \tag{6a}$$

$$r\lambda - \frac{\lambda_0}{K+1} - \lambda\left(\frac{d}{K^2} - \delta\right) = 0. \tag{6b}$$

An optimal solution $(K^*(\cdot), I^*(\cdot), T^*)$ has to satisfy the (limiting) transversality conditions

$$K^*(T^*) = 0 \quad \text{and} \quad \mathcal{H}(K^*(T^*), I^*(T^*), \lambda(T^*), \lambda_0) = 0, \quad \text{for} \quad T^* < \infty, \tag{7a}$$

or

$$\lim_{t\to\infty} \mathrm{e}^{-rt}\,\lambda(t) = 0, \quad \text{for} \quad T^* = \infty. \tag{7b}$$

To explain expressions (7), we need to distinguish two cases. First, expression (7a) applies when the optimal planning horizon is finite. The academic career stops when reputation equals zero. Note that, when reputation K is declining and approaches zero, the negative feedback effect in the Matthew function $M(K)$ causes K to reach zero at a finite point in time. Furthermore, the Hamiltonian has to be zero at the end of the planning horizon as a positive value would imply that increasing the planning horizon would improve the objective value, and a negative value that shortening the time horizon is beneficial. Second, if the length of the optimal planning period is infinite, we have to consider expression (7b). The discounted marginal utility of reputation has to approach zero over time. A positive/negative discounted costate value, which measures the marginal utility of reputation would imply under- or over-investments into reputation.

[8] We omit taking the control constraint $I \ge 0$ into account, since it never becomes binding.

Solution paths approaching equilibria of the canonical system fulfill the infinite time horizon transversality conditions. We next determine their existence and properties. For notational simplicity we introduce $\underline{I}$, which denotes the minimal admissible effort level such that the roots of the state dynamics (2c) exist for $I \in [\underline{I}, I_{\max}]$:

$$\underline{I} := \begin{cases} 0 & 2\sqrt{d\delta} - a \leq 0 \\ 2\sqrt{d\delta} - a & 0 \leq 2\sqrt{d\delta} - a \leq I_{\max} \\ \infty & \text{otherwise.} \end{cases} \tag{8}$$

Proposition 1 determines when equilibrium condition (6a) can be satisfied for feasible effort levels.

Proposition 1. *Let $I_{\min} := 2\sqrt{d\delta} - a$. Then Eq. (6a) has solutions for $I \in [I_{\min}, \infty)$ and these are related as follows:*

$$K_{1,2}(I) = \frac{I + a \mp \sqrt{(I+a)^2 - 4d\delta}}{2\delta}, \tag{9a}$$

which implies

$$K_1(I_{\min}) = K_2(I_{\min}) = \sqrt{\frac{d}{\delta}}, \quad \frac{\mathrm{d}}{\mathrm{d}I}K_1(I) < 0, \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}I}K_2(I) > 0, \quad I \in [I_{\min}, \infty) \tag{9b}$$

and

$$K_i(I) > 0, \quad i = 1, 2, \quad \lim_{I\to\infty} K_1(I) = 0, \quad \lim_{I\to\infty} K_2(I) = \infty. \tag{9c}$$

For $I = I_{\max}$ we define

$$\tilde{K}_i := K_i(I_{\max}), \quad i = 1, 2. \tag{10}$$

Then, the following cases can be distinguished

$\underline{I} < I_{\max}$**:** *for every $I \in [\underline{I}, I_{\max}]$ there exist two roots of Eq. (9a).*

$\underline{I} > I_{\max}$**:** *no equilibrium exists.*

$\underline{I} = I_{\max}$**:** *the only equilibrium candidate has reputation level $K = \dfrac{I_{\max} + a}{2\delta}$.*

Proof. Simple calculations yield a quadratic equation for the equilibria of (6a) for $0 \leq I \leq I_{\max}$:

$$I - \delta K + a - \frac{d}{K} = 0,$$
$$\delta K^2 - K(I + a) + d = 0,$$

resulting in (9a). The solutions are real valued if the discriminant D fulfills

$$D := (I + a)^2 - 4d\delta \geq 0. \tag{11a}$$

For $I + a \geq 0$ this is equivalent to

$$I \geq 2\sqrt{d\delta} - a =: I_{\min}. \tag{11b}$$

Now it is easy to see that for $I_{\max} > 2\sqrt{d\delta} - a \geq 0$ (9a) has two real valued solutions with

$$I \in [2\sqrt{d\delta} - a, I_{\max}].$$

For $2\sqrt{d\delta} - a < 0$ (9a) has two real valued solutions with $I \in [0, I_{\max}]$. Further details of the proof are shifted to Appendix A. □

Proposition 1 implies that the size of parameters a and d, which govern the Matthew effect, as well as the depreciation rate of reputation have a substantial impact on the possible occurrence of long-run equilibria, and therefore whether a scientist can successfully stay in academia. If the self-enforcement rate of reputation a is large compared to d and δ, it is easy to gain reputation and we get that a steady state can exist with high reputation. However, if the depreciation rate is high (i.e. δ is large) or the academic environment is very discriminating (i.e. d is large), it is harder to maintain a good reputation, and, consequently, no equilibrium exists. This implies that the only possible optimal strategy is to quit academia after some time.[9]

Proposition 1 also implies that the maximum effort a scientist is willing or able to invest into his or her reputation, $I_{\max}$, plays an important role on whether a scientist can stay in academia in the long run. The larger a or the smaller d is, the smaller is the necessary minimum investment I for which a steady state can exist. If this minimum level of investments is larger than the maximal effort of the scientist, no steady state can occur and the scientist has to leave academia sooner or later.

An equilibrium $\tilde{K}_1$ at which the control is at its upper bound and where for lower values of K the state dynamics is negative for $I = I_{\max}$, will be called *Stalling Equilibrium*. The denotation "Stalling" refers to the property that maximum effort has to be applied to stay at least put at this position (see Footnote 2). In the next sections we will see that the Stalling Equilibrium $\tilde{K}_1$ plays an important role in determining the scientist's optimal dynamic behavior.

Proposition 2 shows that the Stalling Equilibrium occurs for a low level of reputation when the scientist has good opportunities to let its reputation grow.

Proposition 2. *Let $(I_{\max}+a)^2-4d\delta > 0$, implying existence of the Stalling Equilibrium (see Eq. (9a)). Then the Stalling Equilibrium $\tilde{K}_1$ decreases in parameters $I_{\max}$ and a and increases in d. Furthermore, $\tilde{K}_1$ is strictly positive.*

Proof. For the proof we calculate the derivatives of $\tilde{K}_1$ given by Eq. (9a) for $a, I_{\max}$ and d. These yield

$$\frac{\partial \tilde{K}_1}{\partial a} = \frac{\partial \tilde{K}_1}{\partial I_{\max}} = \frac{1}{2\delta}\left(1 - \frac{I_{\max}+a}{\sqrt{(I_{\max}+a)^2 - 4d\delta}}\right) < 0,$$

and

$$\frac{\partial \tilde{K}_1}{\partial d} = \frac{1}{\sqrt{(I_{\max}+a)^2 - 4d\delta}} > 0,$$

and hence prove the according statements.

The positivity of $\tilde{K}_1$ follows from expression (9a) and $\delta, d > 0$. □

Later we obtain that the Stalling Equilibrium serves as a threshold, such that for $K > \tilde{K}_1$ it is optimal for the scientist to stay in academia forever. Proposition 2 shows in fact that this is optimal for a larger interval of initial reputation levels, if the scientist's opportunities to increase his or her reputation level are bigger.

4. Optimal Solution

This section presents the optimal solution. We first state under which conditions the necessary conditions are sufficient for optimality.

[9] Note that a diverging solution with $K(t)$ going to infinity can be excluded due to $I \leq I_{\max}$ and $\lim_{K\to\infty} M(K) = a$ yielding $\lim_{K\to\infty} \dot{K} = -\infty$.

Proposition 3. *The optimal control problem* (2) *has an optimal solution. If* $T^* = \infty$ *and* $\lambda_0 = 1$ *there is no other solution over the infinite time horizon.*

For the proof see Appendix B. Note that Proposition 3 does not exclude that a finite and infinite time horizon solution exist.

The existence and uniqueness of the optimal solution for the infinite time horizon problem is theoretically verified by Proposition 3. However, from the state dynamics (2c) it immediately follows that there exists a region in the state space such that even for the maximal control $I_{\max}$ it is not possible to sustain a strictly positive K in the long run. Therefore, the infinite time horizon solution is not always admissible so that the finite time horizon solution needs to be considered as well. In addition, it will turn out that an abnormal solution with $\lambda_0 = 0$ also has to be taken into account.

It follows that the solution can be history dependent. Therefore, the notion of *Skiba point* is crucial, and we proceed with stating the following definition related to the one-state problem (2).

Definition 1. *For every* $K_0 \geq 0$ *we introduce the set of optimal solutions of problem* (2)

$$\mathcal{S}(K_0) := \{(K^*(\cdot), I^*(\cdot), T^*) \colon K^*(0) = K_0 \quad \textit{is an optimal solution of problem } (2)\} \tag{12}$$

and we call an optimal solution $(K^*(\cdot), I^*(\cdot), T^*)$ *with* $T^* < \infty$ *and* $K^*(0) = K_0$ *a* finite time horizon solution at K_0*, otherwise it is called an* infinite time horizon solution at K_0*.*

For an optimal solution $(K^*(\cdot), I^*(\cdot), T^*) \in \mathcal{S}(K_0)$ *with* $T^* = \infty$ *and*

$$\lim_{t\to\infty} K^*(t) = \hat{K} \quad \textit{and} \quad \lim_{t\to\infty} I^*(t) = \hat{I},$$

the point $(\hat{K}, \hat{I})$ *is called the* long run optimal solution (for K_0)*.*

For an optimal solution $(K^*(\cdot), I^*(\cdot), T^*) \in \mathcal{S}(\hat{K})$ *with* $K^*(\cdot) \equiv \hat{K}$ *and* $T^* = \infty$ *we say that the optimal solution* stays at $\hat{K}$ *or we shortly say that* the equilibrium $\hat{K}$ is optimal.

If $|\mathcal{S}(K_0)| = 1$ *the optimal solution is* unique at K_0. *If* $|\mathcal{S}(K_0)| = 2$, K_0 *is called a* Skiba point*and the (two) solutions of* $\mathcal{S}(K_0)$ *are called the* (two) Skiba solutions at K_0*.*

Let for some $\varepsilon > 0$ *and a state value* K_S *the set of optimal solutions satisfy* $|\mathcal{S}(K_0)| = 1$ *for* $|K_S - K_0| < \varepsilon$*, and let for* $|K_S - K_0| < \varepsilon$ *and* $K_0 < K_S$ *the finite time horizon solution be optimal and for* $|K_S - K_0| < \varepsilon$ *and* $K_0 > K_S$ *let the infinite time horizon solution be optimal. Then* K_S *is called* a weak Skiba point*.*

Essentially, three different solution patterns can emerge. The simplest pattern is the controlled movement to zero in finite time for any initial value of $K_0 > 0$. In any other case we encounter history dependence of the solution paths. Then, a point $K_S > 0$ exists that separates the regions of attractions between the finite time solution to zero and the infinite time horizon solution to an equilibrium. In the Skiba case, right at $K(0) = K_0 = K_S$ the scientist is indifferent between the finite-time and the infinite-time solution, which are both optimal. In the weak Skiba case the optimal solution is to stay at K_S forever.[10]

The next proposition presents the different solution patterns. A more detailed overview is provided by Proposition 8 in Appendix D.

Proposition 4. *Depending on the parameter constellation the optimal solution of problem* (2) *can be described by one of the following cases.*

1. *For every* $K_0 \geq 0$ *the finite time horizon solution is optimal, and* $|\mathcal{S}(K_0)| = 1$*.*
2. *For* $K_0 < \tilde{K}_1$ *the finite time horizon solution is optimal. For* $K_0 > \tilde{K}_1$ *the optimal solution converges to some* $\hat{K} > \tilde{K}_1$*, and for* $K_0 = \tilde{K}_1$ *it is optimal to stay at* $\tilde{K}_1$*, and* $|\mathcal{S}(K_0)| = 1$*.*

[10] For a general discussion of the Skiba phenomenon and bifurcations of one state optimal control problems we refer the reader to Grass et al. (18) and Kiseleva and Wagener (20, 21).

3. *For $K_0 < K_S$ with $K_S > \tilde{K}_1$ the finite time horizon solution is optimal. For $K_0 > K_S$ the optimal solution converges to some $\hat{K} > K_S$. The point K_S is a Skiba point and the Skiba solutions are the finite time horizon solution and a solution that converges to some $\hat{K} > \tilde{K}_1$, with $|\mathcal{S}(K_0)| = 1$ for $K_0 \neq K_S$ and $|\mathcal{S}(K_S)| = 2$.*

In the first solution of Proposition 4 it holds that it is not optimal for a scientist to stay in academia forever. The second and the third outcome are state-dependent: if the initial reputation is sufficiently large, the scientist will undertake efforts such that in the long run reputation will be positive and admit the level $\hat{K}$. If the initial reputation is low, the scientist will quit academia in finite time with zero reputation. In the second case, the threshold between success and failure is the steady state $\tilde{K}_1$ where the control is at its upper bound. In the third case, the threshold is a Skiba point K_S, where the scientist is indifferent between a strategy requiring high efforts to success and one with lower overall efforts leading to failure.

The difference between the second and the third case is that in the latter case the scientist actually has the choice between the two options when reputation equals K_S. However, this does not hold in the second case if reputation equals $\tilde{K}_1$. If at $\tilde{K}_1$ the scientist conducts maximal effort, the result is that reputation remains at the level $\tilde{K}_1$. This implies that it is not possible for the scientist to raise his or her reputation level beyond this point, so that growing to the long run steady state level $\hat{K}$ cannot be achieved. We denote the threshold $\tilde{K}_1$ as Stalling Equilibrium.

Let us focus on the second case with the Stalling Equilibrium serving as the threshold. Crucial will be the sign of

$$\tilde{q}_1 := r + \delta - \frac{d}{\tilde{K}_1^2}, \tag{13a}$$

which appears in the denominator for the root of the costate equation (5b) with $\lambda_0 = 1$ at $\tilde{K}_1$ as

$$\tilde{\lambda}_1 := \frac{1}{(\tilde{K}_1 + 1)\tilde{q}_1}. \tag{13b}$$

Corollary 1. *Consider case 2 of Proposition 4.*

1. *For $\tilde{q}_1 < 0$ the costate satisfies*

$$\lim_{K_0 \to \tilde{K}_1, K_0 \neq \tilde{K}_1} \lambda(K_0, 0) = \infty, \tag{14}$$

implying that the optimal solution is abnormal.

2. *For $\tilde{q}_1 > 0$ the costate satisfies*

$$\lim_{K_0 \to \tilde{K}_1, K_0 \neq \tilde{K}_1} \lambda(K_0, 0) = \tilde{\lambda}_1, \tag{15}$$

implying that the optimal solution is normal.

For the proof see a detailed version of Corollary 1 formulated as Corollary 4 in Appendix D.

Figure 1 presents the two cases of Corollary 1. The abnormal solution is depicted in Figures 1a to 1c, whereas the normal solution is shown in Figures 1d to 1f. So, if $\tilde{q}_1$ is negative, the optimal solution is abnormal,[11] meaning that essentially the objective is ignored to determine the optimal strategy. The reason for the optimal solution being abnormal is Figure 1b, where the costate variable λ admits the value of infinity at the Stalling Equilibrium $\tilde{K}_1$. However, the maximum principle requires a finite value for the costate variable. For that reason one has to resort to the abnormal problem with

[11] See Appendix E for a more detailed discussion on abnormality and its occurrence in the present problem.

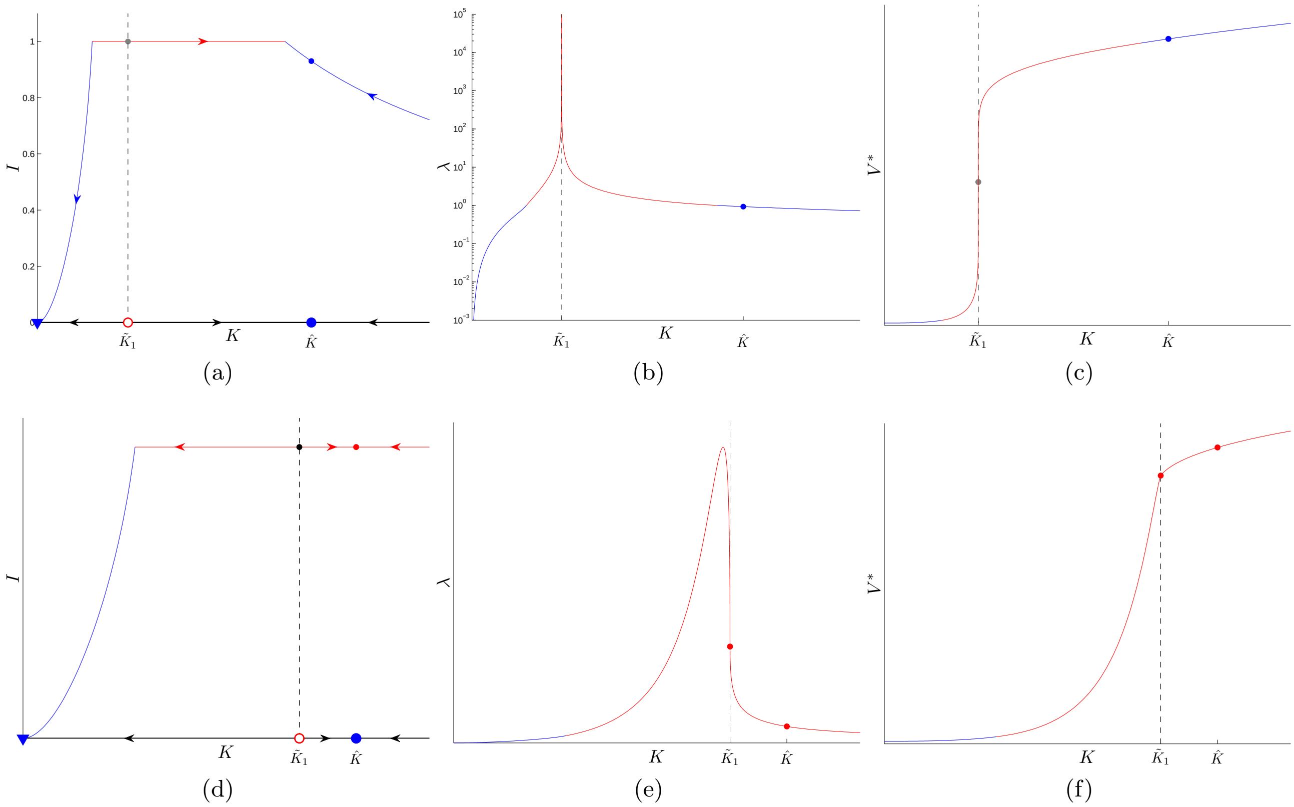


Figure 1: Optimal solutions of the abnormal and the normal case in item 2 as well as the transition between the two cases are plotted in the state-control, state-costate space and showing the objective value. In the state-control space ((a) and (d)) the solutions of the abnormal and the normal case cannot be distinguished. The depiction in the state-costate space reveals the differences of the abnormal case (b), divergence of the costates near $\tilde{K}_1$, and the normal case (e), convergence of the costates near $\tilde{K}_1$. The nearly vertical increase/decrease of the objective value near $\tilde{K}_1$ in the abnormal case (c) yield the heuristic argument for the divergence of the costate, shown in (b). The different colors of the paths (equilibria) refer to the active upper control constraint (red) and the control to be in the interior (blue). The circles refer to the steady states, if the interior is filled it is a saddle, otherwise it is unstable. The triangle refers to the end point of the finite horizon solution.

$\lambda_0 = 0$ to derive the optimal solution. For the economic interpretation of the infinite value of the costate at the Stalling Equilibrium, we first note that the costate variable measures the increase of the objective value due to an infinitesimal increase of the corresponding state variable K. It is important to realize that at the Stalling Equilibrium the scientist cannot increase his or her reputation. In fact, when he or she applies maximal effort, i.e. $I = I_{\max}$, this is just sufficient to keep reputation equal to $\tilde{K}_1$. However, if reputation were infinitesimally larger, reputation would grow when the scientist applies maximal effort, and the optimal saddle point equilibrium $\hat{K}$ can be reached in the long run. This would give a substantial increase of the objective value, which explains why the costate variable diverges at $\tilde{K}_1$.

For comparison the normal case is shown in Figure 1e. It can be seen that at the Stalling Equilibrium the marginal utility of reputation is finite. Although it is not possible to increase reputation any further as the investments are at their maximum level (compare Figure 1d), the additional gains by approaching the higher steady state are not as large as in the abnormal case (compare Figure 1c).

It is important to know for which parameter constellation we can expect the solution to be abnormal. The following proposition states when this happens.

Proposition 5. *Let $(I_{\max} + a)^2 - 4d\delta > 0$ and the equilibrium solution $(K(\cdot), I(\cdot)) \equiv (\tilde{K}, I_{\max})$ at the Stalling Equilibrium be optimal. If*

$$\frac{(I_{\max} + a)^2}{d} > r + 5\delta, \tag{16}$$

the equilibrium solution is abnormal.

For the proof see Appendix C.

Proposition 5 implies that when the scientist is able to assign a lot of effort to increase reputation, and the Matthew effect works out relatively positive on the increase of reputation, the solution is abnormal. This is understandable, because abnormality goes along with infinite value of the costate at the Stalling Equilibrium. And we just stated that the costate being infinite results from the fact that at the Stalling Equilibrium $\tilde{K}_1$ growth towards the saddle point $\hat{K}$ leads to a much higher value of the objective than keeping reputation constant at $\tilde{K}_1$. In the scenario where $I_{\max}$ and a are large, whereas d is low, growth conditions are optimal so that especially then the difference in the objective value will be substantial.

From Proposition 2 we obtain that in the abnormal scenario, i.e. when $I_{\max}$ and a are large while d is low, the threshold $\tilde{K}_1$ will admit a lower value. This makes sense as it will be easier for the scientist to grow towards the optimal long run level $\hat{K}$. This we clearly see in Figure 1 when we compare Figure 1a with Figure 1d.

5. Bifurcation Analysis

This section presents a numerical two-dimensional bifurcation analysis that illustrates the patterns described in the previous section. For the base case parameter values specified in Table 1 the *a*-*c* parameter space is separated into three regions, see Figure 2h. Region I corresponds to Case 2 of Proposition 4, Region II is Case 3, and Region III corresponds to Case 1. So we can conclude that for a Stalling Equilibrium to occur it is needed that the cost of effort is small. This makes sense, because only then it will be optimal to have maximal effort at the threshold level between success and failure.

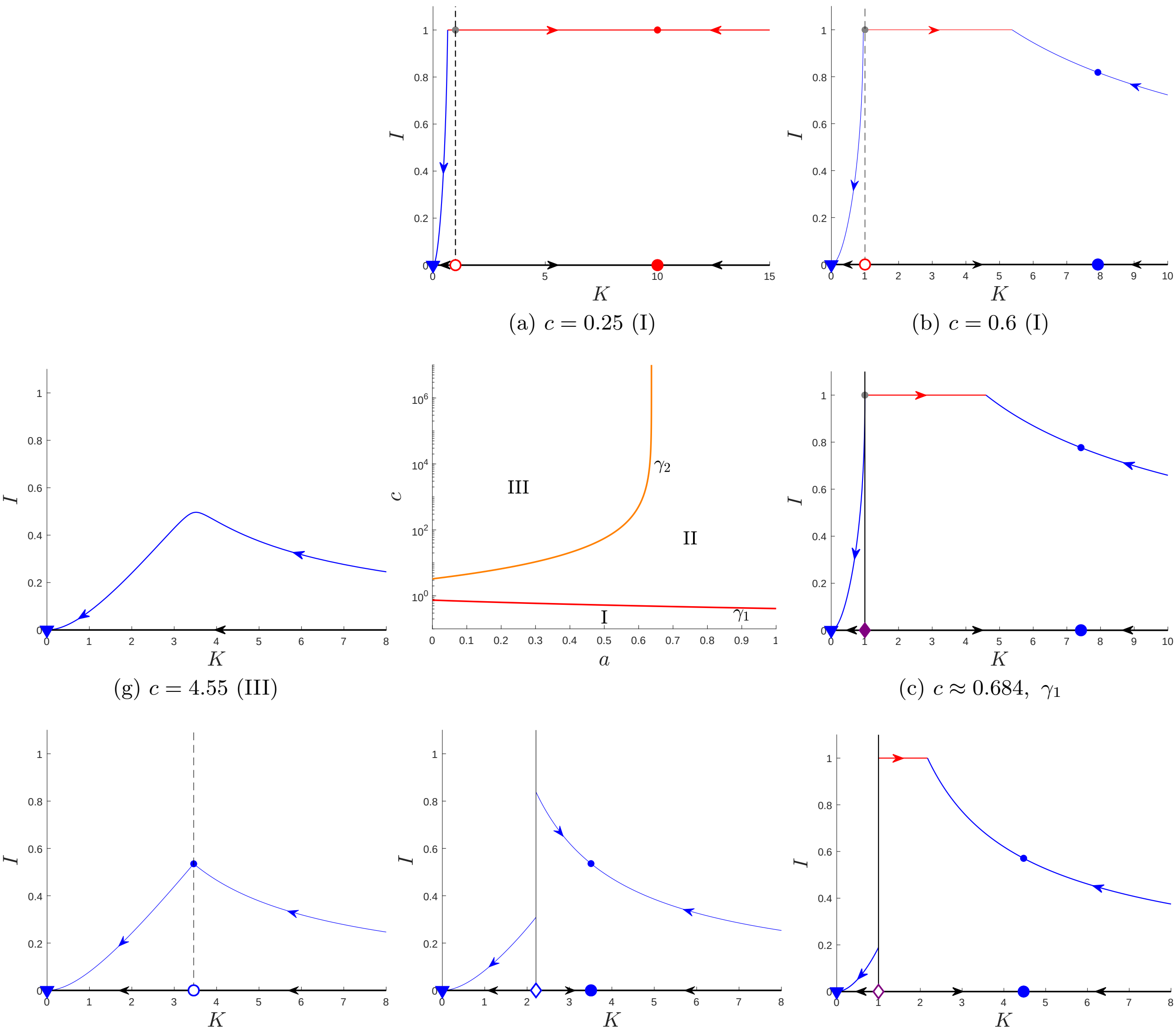
To facilitate the numerical calculations we introduce a small parameter $\tau = 10^{-5}$ in the Matthew-function (1):

$$M(K, \tau) := \frac{aK - d}{K + \tau}. \tag{17}$$

This formulation prohibits a singularity at $K = 0$. The solution $(K(\cdot, \tau), I(\cdot, \tau))$ satisfies

$$\lim_{\tau \to 0} (K(\cdot, \tau), I(\cdot, \tau)) = (K(\cdot, 0), I(\cdot, 0))$$

and the limit is the correct solution $(K(\cdot), I(\cdot))$ for the original Matthew function (1).



(a) $c = 0.25$ (I) (b) $c = 0.6$ (I)

(g) $c = 4.55$ (III) (c) $c \approx 0.684,\ \gamma_1$

(f) $c \approx 4.515,\ \gamma_2$ (e) $c = 4.25$ (II) (d) $c = 2$ (II)

Figure 2: Bifurcation diagram for parameter values a and c is at the center. The examples are calculated for the parameter values taken from Table 1 with $a = 0.1$ and increasing c. The circles refer to the steady states, if the interior is filled it is a saddle, otherwise it is unstable. The triangle refers to the end point of the finite horizon solution, $\diamondsuit$ to a Skiba point.

Region I:. In this region a locally stable and unstable equilibrium exist. The unstable equilibrium is the Stalling Equilibrium $\tilde{K}_1$. Starting with $K(0) < \tilde{K}_1$ the scientist ends his or her career at some optimally chosen finite time. For $K(0) > \tilde{K}_1$, however, the scientist stays in academia and ends up eventually with a reputation level that corresponds with the locally stable equilibrium. For a small cost parameter c this locally stable equilibrium is $\tilde{K}_2$, cf. Figure 2a, where $I = I_{\max}$. For larger values of c this equilibrium $\hat{K}$ lies in the interior of the control region, i.e. $\hat{I} < I_{\max}$, cf. Figure 2b.

In region I the scientist enjoys investing into reputation, because investment costs are low. Therefore, a large domain of initial reputation levels exists for which convergence to a long run equilibrium occurs, at which the scientist spends maximal time on doing activities to maintain a high reputation, e.g. doing research, attending conferences, etc. If, however, the reputation is very low at the beginning of the planning horizon, a scientist cannot gain ground in academia and a scientific career is not viable for the given maximum efforts $I_{\max}$ that the scientist is able to invest into his or her career. It is optimal to put some small efforts into doing research in order to delay leaving academia. At the Stalling Equilibrium, the scientist has to put maximum efforts into staying put at a reputation level that allows staying in academia in the long run. Note, however, that reputation is at a very low level and these efforts are not sufficient to improve it. At the Stalling Equilibrium, the self-enforcement effect of reputation is not strong enough to drive further growth of reputation. Effectively, the maximum amount of effort is required to combat reputation loss due to the negative feedback effect of the Matthew function.

Compared to Figure 2a, the scientist has larger effort costs in Figure 2b. This results in a larger domain in which a scientific career is not viable in the long run. Here the scientist reduces efforts over time and ends his or her scientific career at some point. If reputation is large the scientist does not spend the maximal time possible in order to keep reputation at a very high level. This is because the marginal effect of doing research reduces with K. Also in the long run equilibrium the scientist keeps I below $I_{\max}$.

Bifurcation curve γ_1:. In the transition from Region I to Region II the Stalling Equilibrium becomes a Skiba equilibrium. I.e. for initial values $K(0) = \tilde{K}_1$ the scientist is indifferent between either to stay in the equilibrium with $I_{\max}$ or to choose some $I(0) < I_{\max}$ and end the academic career in finite time, cf. Figure 2c. Both solutions generate an equal value of the objective and hence the indifference property of a Skiba point is satisfied. The threshold property is analogous to the behavior in Region I. For initial reputation levels that exceed $\tilde{K}_1$, it is optimal to approach a locally stable equilibrium with a reputation larger than $\tilde{K}_1$.

Region II:. In this region the Skiba equilibrium from the transition case, where the scientist is indifferent between staying put and quitting academia in the long run, is replaced by a conventional Skiba point and the Stalling Equilibrium vanishes. At the Skiba point the scientist is indifferent between gradually reducing reputation to zero in finite time, or to conduct some extra efforts in order to approach the locally stable equilibrium with a high enough reputation level that guarantees a fruitful stay in academia. The obvious disadvantage of the first strategy is that the scientist has to leave academia at the moment the reputation level reaches zero, and the disadvantage of the second strategy is that costly efforts are required to improve reputation. For lower values of c the scientist either chooses $I_{\max}$ or some $I(0) < I_{\max}$, cf. Figure 2d for the initial investments into reputation. However, if we increase the cost parameter c, it will always hold that $I(0) < I_{\max}$, cf. Figure 2e.

In comparison with Figure 2e, research costs are much smaller in Figure 2d. This results in a smaller domain of initial values of K, where the kind of solution is admitted that a scientific career ends in finite time. Another difference with Figure 2e is that for some (intermediate) values of K the researcher wants to spend maximal time on research to increase his or her reputation.

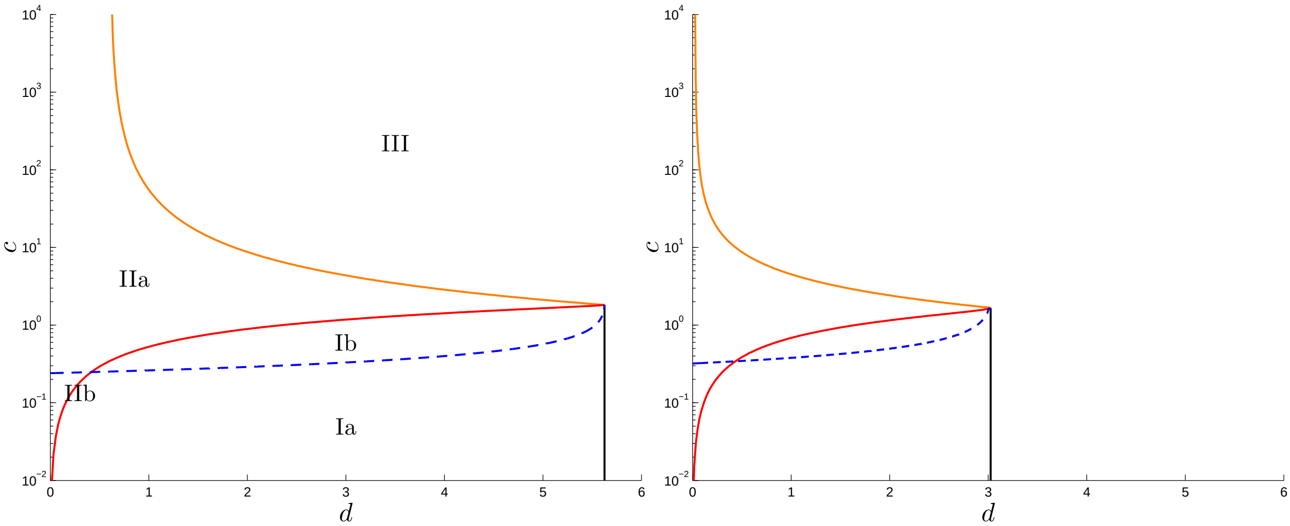
Figure 2e confirms what we have seen before. A large initial stock of reputation results in a lifelong scientific career, and otherwise the research ends in finite time. What is new is that at the Skiba point effort is discontinuous. If the scientist wants to stay active in research forever, at the Skiba point it is necessary to conduct more effort. This is because in this manner the scientist improves the reputation such that after some time reputation begins to reinforce itself due to the Matthew effect. Thus, the reputation improves, the scientist benefits longer from an increase in reputation, making the payoff from research effort higher.

Bifurcation curve γ_2:. In the transition from Region II to Region III the Skiba point coincides with the locally stable equilibrium and becomes a semi-stable equilibrium $\hat{K}$, cf. Figure 2f. Starting with $K(0) < \hat{K}$ the scientist's reputation reaches zero at some optimally chosen finite time and for $K(0) > \hat{K}$ reputation also decreases, but this time it ends at $\hat{K}$.

Region III:. Increasing c above the bifurcation curve γ_2 accomplishes that the semi-stable equilibrium vanishes. Furthermore it holds that Region III only exists for $a \leq 0.637$. Then it is always optimal for the scientist to end his or her career in finite time, cf. Figure 2g. This is not only because c is large, indicating large costs of performing investment into reputation, but also because a is low. The parameter a being low indicates that the self-enforcement effect of reputation is low, which makes it hard and costly to improve reputation. Therefore, it makes sense to conduct low effort and stop the career at some point.

In Figure 2f the cost of doing research is a little smaller than in Figure 2g. In this figure exactly a long run equilibrium exists at the moment effort is at its maximal level. The qualitative implications are enormous, because now the scientist stays active in academia, provided the initial stock of reputation is large enough.

Despite that the scientist will stop his or her career in finite time in Figure 2g, for a large initial reputation level it is still optimal to start increasing research efforts over time. This is because throughout the declining phase K gets smaller, which implies that the marginal efficiency of effort goes up. However, at a final time interval research effort will decrease. This is because the scientist comes closer to the end of his or her career, which reduces the time he or she benefits from reputation investments.



(a) $a = 0.5$

(b) $a = 0.1$

Figure 3: Bifurcation diagram for parameter values d and c. Panel (a) shows the case for $a = 0.5$ and panel (b) that for $a = 0.1$. Three main regions exist. Enclosed by the black and red curve (region I), the Stalling Equilibrium separates solutions going to zero and the high equilibrium. Between the red and orange curve(region II), we find the Skiba region. Outside the black and orange curve (region III) the globally optimal solution is going to zero. The blue dashed line shows the transition from the region where the locally stable equilibrium satisfies $\hat{I} < I_{\max}$ and $\hat{I} = I_{\max}$.

In Figure 3a and Figure 3b it can be seen that also three different regions occur if we consider cost parameter c and discrimination parameter d for the bifurcation analysis. In particular, if investment costs and discrimination are low, whether a scientist is able to succeed in academia is history-dependent, success and failure are separated by a weak Skiba point, i.e. if one starts exactly at this point, one would stay put there. At this steady state the investments are at their maximum level, meaning that here the extent of how much a scientist is able to invest into his or her career at most can be crucial for success. In case of high investment costs, but low discrimination, the outcome is still history-dependent, however, the threshold point is now a conventional Skiba point, where the scientist has the choice between low efforts, which means that he or she will inevitably quit academia after a finite, optimally determined time, or put a lot of effort into the career and succeeds in the long run. A high discrimination rate always means that a scientist is not able to remain in academia, which also holds in case of an intermediate discrimination rate in combination with high investment costs. Comparing Figure 3a and Figure 3b we find that if the self-enforcement rate of reputation is low, discrimination is a larger problem for the success of an academic career than if it is large.

6. Fixed retirement age and positive utility of effort

To analyze how certain model assumptions affect the optimal effort policy of the scientist, we consider how the solution is affected by changes in these assumptions. In our basic model we assume that, besides having a positive effect on reputation, conducting effort only gives disutility. However, the scientist might in fact also like being occupied with writing papers, collaborating with other scientists, networking at conferences, and so on. We therefore consider in Section 6 what happens if we consider the extended cost function $C(I) := cI^2 - eI$, where eI accounts for the utility a scientist obtains from undertaking scientific effort. We find that no matter whether e is small or large, the results are qualitatively the same. For small values of c the optimal solution is history dependent where the Stalling Equilibrium serves as threshold level, for intermediate values of c the optimal solution is still history-dependent but then goes along with the usual Skiba point, and for a large value of c the scientist is bound to fail.

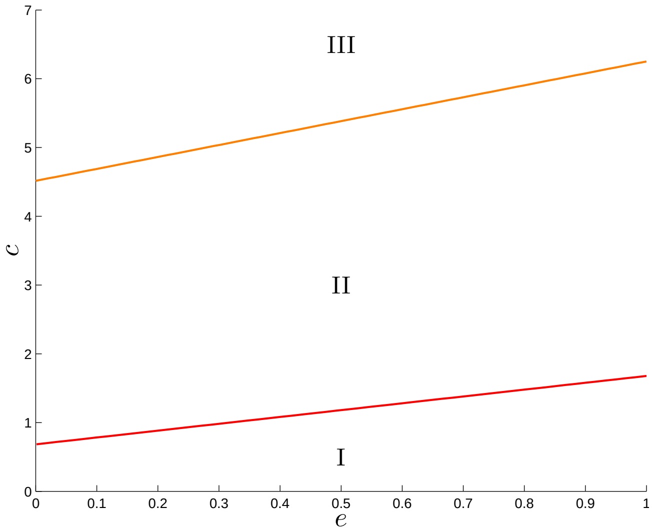


Figure 4: Change of the different regions of solution behavior for changing e from the extended cost function in $C(I) := cI^2 - eI$ from zero to one for the base case parameter values specified in Table 1 and $a = 0.1$.

Figure 5 shows the impact of the assumption that the time horizon is optimally determined and either finite or infinite. Here we consider the optimal strategy of a successful scientist who needs to take into account that at a fixed point in time he or she has to retire. It can be seen that if a fixed time horizon is chosen the optimal strategy is initially the same as in the infinite time horizon case and the researcher has to work hard to improve reputation unless it is already very large in the beginning. Since there is no utility from a high reputation after retiring, a scientist would reduce investment efforts the closer he or she is to the end of the working time. The later is the fixed retirement date, the closer is the optimal strategy to the infinite time horizon solution.

7. Conclusion

The starting point of our analysis is Merton's insight that better-known scientists tend to receive more academic recognition than less-known scientists for similar achievements. The occurrence of this phenomenon is not limited to career planning in academia, but also plays an important role in many other socioeconomic fields like education, health status, income and pension dynamics. Formally, we deal with a feedback mechanism in which the change of a state variable depends on the level of

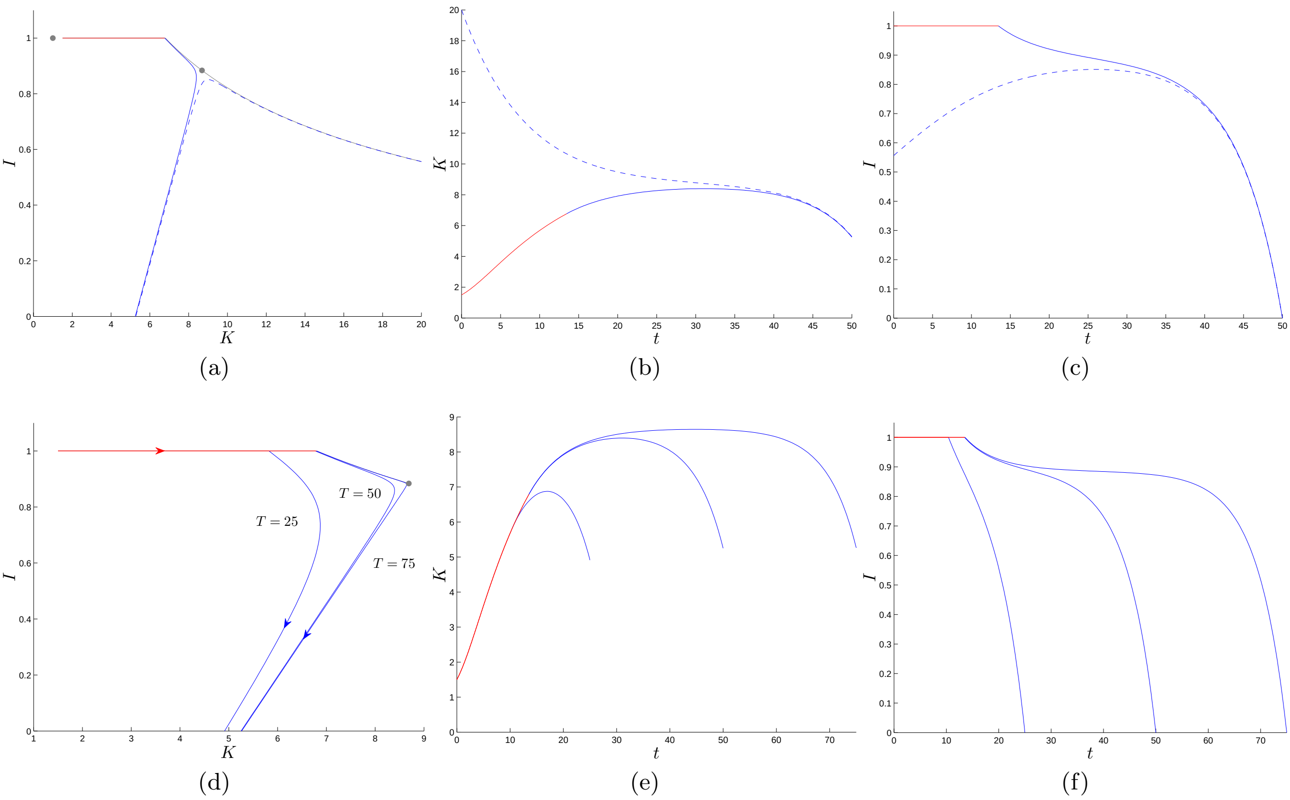


Figure 5: Infinite time horizon solution versus a solution with fixed end-time. In the first row (a)-(c) the solutions for $T = 50$ are depicted starting at $K(0) = 1.5$ and $K(0) = 20$. Panel (a) shows the solutions in the state-control space. The gray elements denote the equilibria and solution paths for the infinite time horizon problem. In panel (b) the state and in panel (c) the control paths for $T = 50$ are plotted. In the second row (d)-(f) solutions for the different end-times $T = 25, 50$ and 75 are depicted starting at $K(0) = 1.5$. Panel (d) shows the solutions in the state-control space and panel (e) the state paths and panel (f) the control paths. In (d) the tunrpike property is particularly well illustrated. For an increasing end-time the solution approaches the equilibrium of the infinite time horizon problem and finally moves away due to the transversality condition.

the state itself. The function describing this influence of reputation is denoted as Matthew function. The feedback effect can both be positive and negative for different regions of the state space. In the case of scientific production, the individual stock of reputation acts as state. For large values of the reputation, the Matthew effect is positive ("once Nobel price winner always Nobel laureate"), while for moderate stocks of reputation, a scientist experiences disadvantages.

We design an optimal control model of the scientist with his or her reputation thus as state variable and the scientist's scientific effort as control variable. We explicitly take into account that the effort level is constrained: there are only a limited number of hours a day that is available for working. Solving the optimal control model essentially leads to two different solutions. In case the scientist is not so talented in the sense that his or her scientific efforts are not too effective and the Matthew feedback effects are not too positive, his or her academic career is doomed to fail. The scientist's reputation decreases over time and after some time the scientific career stops.

For a more gifted scientist the optimal solution is history-dependent. Still for a low level of the initial reputation scientific life stops at some point, but otherwise a fruitful scientific career is awaiting. We identify a threshold level of the reputation that separates these two different solutions. For a moderate level of the cost of scientific effort this threshold is the "usual" Skiba point (Skiba (34)), but if this cost is small we show that the threshold is a steady state where scientific effort is at its maximum. As far as we know this is a new feature and we denote this threshold as Stalling Equilibrium.

Since the Stalling Equilibrium is a steady state with the control at its upper bound, it follows that growth is not possible. So right at the Stalling Equilibrium the scientist has the choice between conducting maximal effort for the rest of his or her academic life and keeping the reputation level fixed at the level corresponding to the Stalling Equilibrium, or undertaking less scientific effort implying that his or her reputation declines and academic life will end at some future point in time. We show that the scientist being indifferent between these options only occurs at a boundary solution. Otherwise, the scientist will always choose for maximal effort and keep the reputation at the same level until the end of the career.

If we consider the reputation level corresponding to the Stalling Equilibrium and we increase it with an infinitesimal amount, it is optimal for the scientist to undertake maximal effort and enter a growth path of reputation. Then the academic career continues indefinitely and reputation level will end up at a saddle point steady state. As we said before, right at the Stalling Equilibrium entering the reputation growth path is not possible so this infinitesimal increase in reputation brings a new possibility: growth. This results in a considerable increase of the objective, which is measured by the costate variable. In some situations, i.e. when the Matthew feedback effect is very positive and the upper bound of the effort level is relatively high, the costate variable gets an infinite value. This is not allowed by the maximum principle, implying that we have to resort to solving the abnormal problem. So this is another new feature of our work: we have a well-posed model in which abnormality results from the upper bound on the control variable.

Up until now abnormal optimal control problems have found little attention in the economic literature. To analyze the occurrence and implications of abnormal solutions in a more general way seems to be a fruitful future research topic.

Acknowledgements

The help of Alexander Bek, Hans-Peter Blossfeld, Valeriya Lykina, Andreas Novak, Fouad El Ouardighi, Vladimir Veliov and Yuri Yegorov is gratefully acknowledged.

References

[1] Allison, P.D., Long, J.S., Krauze, T.K., 1982. Cumulative advantage and inequality in science. American Sociological Review 47, 615–625.

[2] Arutyunov, A.V., 2000. Optimality Conditions: Abnormal and Degenerate Problems. Kluwer Academic Publishers, Dordrecht.

[3] Aseev, S.M., Veliov, V.M., 2015. Maximum principle for infinite-horizon optimal control problems under weak regularity assumptions. Proceedings of the Steklov Institute of Mathematics 291, 22–39.

[4] Azoulay, P., Stuart, T., Wang, Y., 2013. Matthew: Effect or fable? Management Science 60, 92–109.

[5] Barbero-Liñán, M., Muñoz-Lecanda, M.C., 2008. Geometric approach to Pontryagin's maximum principle. Acta Applicandae Mathematicae 108, 429–485.

[6] Bask, M., Bask, M., 2014. Social influence and the Matthew mechanism: The case of an artificial cultural market. Physica A: Statistical Mechanics and its Applications 412, 113–119.

[7] Bask, M., Bask, M., 2015. Cumulative (dis)advantage and the Matthew effect in life-course analysis. PLOS ONE 10, 1–14.

[8] Becker, G.S., 1962. Investment in human capital: A theoretical analysis. Journal of Political Economy 70, 9–49.

[9] Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. Journal of Political Economy 75, 352–365.

[10] Blinder, A.S., Weiss, Y., 1976. Human capital and labor supply: A synthesis. Journal of Political Economy 84, 449–472.

[11] Bol, T., de Vaan, M., van de Rijt, A., 2018. The Matthew effect in science funding. Proceedings of the National Academy of Sciences 115, 4887–4890.

[12] Caulkins, J.P., Feichtinger, G., Grass, D., Hartl, R.F., Kort, P.M., Seidl, A., 2015. Skiba points in free end-time problems. Journal of Economic Dynamics and Control 51, 404 – 419.

[13] Clauset, A., Arbesman, S., Larremore, D.B., 2015. Systematic inequality and hierarchy in faculty hiring networks. Science Advances 1.

[14] Cole, S., 1970. Professional standing and the reception of scientific discoveries. American Journal of Sociology 76, 286–306.

[15] Dannefer, D., 1987. Aging as intracohort differentiation: Accentuation, the Matthew effect, and the life course. Sociological Forum 2, 211–236.

[16] Dechert, W., Nishimura, K., 1983. A complete characterization of optimal growth paths in an aggregated model with a non-concave production function. Journal of Economic Theory 31, 332–354.

[17] DiPrete, T.A., Eirich, G.M., 2006. Cumulative advantage as a mechanism for inequality: A review of theoretical and empirical developments. Annual Review of Sociology 32, 271–297.

[18] Grass, D., Caulkins, J.P., Feichtinger, G., Tragler, G., Behrens, D.A., 2008. Optimal Control of Nonlinear Processes: With Applications in Drugs, Corruption, and Terror. Springer-Verlag, Berlin.

[19] Halkin, H., 1974. Necessary conditions for optimal control problems with infinite horizons. Econometrica 42, 267–272.

[20] Kiseleva, T., Wagener, F.O.O., 2010. Bifurcations of optimal vector fields in the shallow lake system. Journal of Economic Dynamics and Control 34, 825–843.

[21] Kiseleva, T., Wagener, F.O.O., 2015. Bifurcations of optimal vector fields. Mathematics of Operations Research 40, 24–55.

[22] Levin, S.G., Stephan, P.E., 1991. Research productivity over the life cycle: Evidence for academic scientists. The American Economic Review 81, 114–132.

[23] Mayer, K.U., 2004. Whose lives? how history, societies, and institutions define and shape life courses. Research in Human Development 1, 161–187.

[24] McDowell, J.M., 1982. Obsolescence of knowledge and career publication profiles: Some evidence of differences among fields in costs of interrupted careers. The American Economic Review 72, 752–768.

[25] Merton, R.K., 1968. The Matthew effect in science. Science 159, 56–63.

[26] Merton, R.K., 1973. The Sociology of Science: Theoretical and Empirical Investigations. The University of Chicago Press.

[27] Merton, R.K., 1988. The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. Isis 79, 606–623.

[28] Montgomery, R., 1994. Abnormal minimizers. SIAM Journal on Control and Optimization 32, 1605–1620.

[29] Rossiter, M.W., 1993. The Matthew Matilda effect in science. Social Studies of Science 23, 325–341.

[30] Seidl, A., Wrzaczek, S., El Ouardighi, F., Feichtinger, G., 2016. Optimal career strategies and brain drain in academia. Journal of Optimization Theory and Applications 168, 268–295.

[31] Seierstad, A., Sydsæter, K., 1987. Optimal Control Theory with Economic Applications. North-Holland, Amsterdam.

[32] Sethi, S., 1977. Nearest feasible paths in optimal control problems: Theory, examples, and counterexamples. Journal of Optimization Theory and Applications 23, 563–579.

[33] Sethi, S., 1979. Optimal advertising policy with the contagion model. Journal of Optimization Theory and Applications 29, 615–627.

[34] Skiba, A.K., 1978. Optimal growth with a convex-concave production function. Econometrica 46, 527–539.

[35] Stengel, R.F., 2004. Flight Dynamics. Princeton University Press.

[36] Stephan, P.E., 1996. The economics of science. Journal of Economic Literature 34, 1199–1235.

[37] Wagener, F.O.O., 2003. Skiba points and heteroclinic bifurcations, with applications to the shallow lake system. Journal of Economic Dynamics and Control 27, 1533–1561.

[38] Wang, Y., Jones, B.F., Wang, D., 2019. Early-career setback and future career impact. Nature Communications 10.

[39] Winkler-Dworak, M., 2008. The low mortality of a learned society. European Journal of Population / Revue européenne de Démographie 24, 405–424.

[40] Wislar, J.S., Flanagin, A., Fontanarosa, P.B., DeAngelis, C.D., 2011. Honorary and ghost authorship in high impact biomedical journals: a cross sectional survey. BMJ 343.

[41] Yegorov, Y., Wirl, F., Grass, D., Seidl, A., 2016. Dynamic Perspectives on Managerial Decision Making: Essays in Honor of Richard F. Hartl. Springer-Verlag. chapter Economics of Talent: Dynamics and Multiplicity of Equilibria. pp. 37–61.

Appendix A. Proof of Proposition 1

Continuation of the proof for Proposition 1.
For $I_{\min} = 2\sqrt{d\delta} - a$ Eq. (9a) reduces to

$$K_{1,2}(I_{\min}) = \frac{I_{\min} + a}{2\delta}.$$

Next, we consider the derivative $\frac{\partial}{\partial I}K_i(I), i = 1, 2$, which is

$$\begin{aligned}\frac{\partial}{\partial I}K_i(I) &= \frac{\sqrt{(I+a)^2 - 4d\delta} \mp (1+a)}{\delta\sqrt{(I+a)^2 - 4d\delta}} \\ &= \mp\frac{(1+a) \mp \sqrt{(I+a)^2 - 4d\delta}}{\delta\sqrt{(I+a)^2 - 4d\delta}} \\ &= \mp\frac{K_i(I)}{\sqrt{(I+a)^2 - 4d\delta}} \lessgtr 0.\end{aligned}$$

Therefore, the $K_1(I)$ branch is strictly decreasing and the $K_2(I)$ branch is strictly increasing, i.e. for $I_{\min} < I_1, I_2 < I_{\max}$

$$\tilde{K}_1 < K_1(I_1) < K_2(I_2) < \tilde{K}_2.$$

Appendix B. Proof of Proposition 3

For the proof we show that the assumptions stated in Seierstad and Sydsæter (31, Th. 14,15, p. 236–237) are satisfied. The only detail that needs specific attention is the singularity of the state dynamics at the origin.

We start with the existence part. For the existence of an optimal solution we distinguish two cases.

$(I_{\max}+a)^2-4d\delta \geq 0$ *and* $K(0) \geq \tilde{K}_1$. Due to Proposition 1 for every $I \in [0, I_{\max}]$ there exist equilibria $\tilde{K}_{1,2}$ and the state dynamics satisfies

$$-\delta K(t) + a - \frac{d}{\tilde{K}_1} \leq \dot{K}(t) \leq I_{\max} - \delta K(t) + a - \frac{d}{\tilde{K}_2}. \tag{B.1}$$

Thus, for every initial state $K(0) \geq \tilde{K}_1$ and $I(\cdot) \equiv I_{\max}$ the state path $K(\cdot)$ exists for all $t \geq 0$ and stays in the region $\left[\tilde{K}_1, \max\left\{K(0), \tilde{K}_2\right\}\right]$. Using the inequality Eq. (B.1) we find for all admissible pairs $(K(\cdot), I(\cdot))$

$$I(t) - \delta K(t) + a - \frac{d}{K(t)} \leq \Phi(t), \quad \text{with} \quad \int_0^\infty \Phi(t)\,\mathrm{d}t < \infty$$

and for all $K \geq \tilde{K}_1$ and $I \in [0, I_{\max}]$

$$\left|I - \delta K + a - \frac{d}{K}\right| \leq \delta|K| + \beta$$

and

$$\beta := I_{\max} + a - \frac{d}{\tilde{K}_2} > 0.$$

Moreover the set

$$N(K, I, t) := \left\{\left(\mathrm{e}^{-rt}(\ln(K+1) - cI^2) + \gamma, I - \delta K + a - \frac{d}{K}\right) : I \in [0, I_{\max}], \gamma \leq 0, K \geq \tilde{K}_1\right\}$$

is convex. Therefore, under the given assumptions an optimal solution exists over an infinite time horizon.

$(I_{\max}+a)^2-4d\delta<0$ *or* $K(0)<\tilde{K}_1$. Under these assumptions there is no admissible control such that the solution over the infinite time horizon exists. All solutions reach zero in finite time. The critical part is the singularity of the state dynamics Eq. (2c) at $K=0$. W.l.o.g. we assume that $K(0)\le 1$. If $K(0)>1$ we consider the admissible control $I(\cdot)=I_{\max}$ until $K(\tau)=1$ for some τ. Let $s_0>0$ such that for $s\in[0,s_0]$, $K(s)>0$ and $K(s_0)=0$. To handle the singularity at zero we use the following time transformation

$$t=\int_0^s K(l)\,\mathrm{d}l,\quad s\in[0,s_0] \tag{B.2}$$

which is a diffeomorphism from $(0,s_0)$ onto $(0,\int_0^{s_0}K(l)\,\mathrm{d}l)$.

Using time transformation Eq. (B.2) the state dynamics becomes

$$\frac{\mathrm{d}}{\mathrm{d}s}K(s)=I(s)K(s)-\delta K(s)^2+aK(s)-d \tag{B.3a}$$

and the objective reads as

$$\int_0^S \mathrm{e}^{-r\int_0^s K(l)\,\mathrm{d}l}\left(K(s)\ln(K(s)+1)-cK(s)I(s)^2\right)\mathrm{d}s. \tag{B.3b}$$

For $I(\cdot)\equiv I_{\max}$ the ODE (B.3a) has a solution such that for some $S>0$ and $0<K(0)\le 1$ we find $K(S)=0$. Thus, an admissible pair for the finite time problem exists.

Using $K\le 1$ we can estimate Eq. (B.3a)

$$\frac{\mathrm{d}}{\mathrm{d}s}K(s)=I(s)K(s)-\delta K(s)^2+aK(s)-d\le I_{\max}+a-d$$

and hence derive all necessary estimations of the dynamics to show the existence of the finite time horizon solution. Also the convexity of the set $N(K,I,t)$ for the transformed problem is straight forward.

To prove that the infinite time horizon solution satisfies the Arrow sufficiency conditions we note that we previously showed that the state path remains bounded for every admissible control. Let $(K^*(\cdot),I^*(\cdot))$ be an optimal solution and $\lambda_0=1$. Then there exists a costate $\lambda(\cdot)$ satisfying the adjoint equation (5b) and the limiting transversality condition

$$\lim_{t\to\infty}\mathrm{e}^{-rt}\lambda(t)=0.$$

This is an immediate consequence of the linearity of the adjoint equation with respect to λ and the boundedness of the state K. Thus for every admissible pair $(K(\cdot),I(\cdot))$ the following limit exists

$$\lim_{t\to\infty}\mathrm{e}^{-rt}\lambda(t)(K(t)-K^*(t))=0.$$

Since the optimized Hamiltonian is concave in K the Arrow sufficiency conditions are satisfied. The Arrow sufficiency conditions imply the uniqueness of the solutions over an infinite time horizon.

Appendix C. Proof of Proposition 5

We note that $\tilde{q}_1<0$ is a necessary condition for abnormality. Let us introduce

$$\xi:=\frac{4d\delta}{(I_{\max}+a)^2},$$

and rewrite $\tilde{K}_1$ as

$$\tilde{K}_1=\frac{(I_{\max}+a)}{2\delta}(1-\sqrt{1-\xi}).$$

Using Taylors theorem yields

$$\sqrt{1-\xi} = 1 - \frac{\xi}{2\sqrt{1-\bar{\xi}}}, \quad \text{with} \quad 0 < \bar{\xi} < \xi.$$

Since $\sqrt{1-\bar{\xi}}$ is strictly decreasing, it follows that

$$\sqrt{1-\xi} > 1 - \frac{\xi}{2\sqrt{1-\xi}}.$$

Plugging this into the expression for $\tilde{K}_1$ we find

$$\tilde{K}_1 < \frac{(I_{\max}+a)\xi}{4\delta\sqrt{1-\xi}},$$

or equivalently

$$\tilde{K}_1^2 < \frac{(I_{\max}+a)^2\xi^2}{4^2\delta^2(1-\xi)} = \frac{d^2}{(I_{\max}+a)^2 - 4d\delta}.$$

This is equivalent with

$$\frac{d}{\tilde{K}_1^2} > \frac{(I_{\max}+a)^2}{d} - 4\delta.$$

We conclude that if

$$\frac{(I_{\max}+a)^2}{d} - 4\delta > r + \delta \quad \text{or} \quad \frac{(I_{\max}+a)^2}{d} > r + 5\delta,$$

the optimal solution is abnormal.

Appendix D. Results concerning the structure of the optimal solution

In this section we summarize results concerning the structure of the optimal solution.

Lemma 1. *Let $(K^*(\cdot), I^*(\cdot), T^*)$ be an optimal solution of problem (2) with $0 < T^* < \infty$ and $K^*(T^*) = 0$. Then, there exists a σ with $T^* \geq \sigma > 0$ such that*

$$0 < I^*(t) < I_{\max}, \quad t \in (T^* - \sigma, T^*) \quad \textit{and} \quad I^*(T^*) = 0.$$

Proof. Let $\varepsilon > 0$ be small and $K(0) = \varepsilon$. Then $\frac{d}{\varepsilon} \gg 1$ such that the terms a and δK of the state dynamics Eq. (2c) can be neglected. Then we compare solution paths $K_1(\cdot)$ and $K_2(\cdot)$ satisfying

$$\dot{K}_1(t) = I(t) - \frac{d}{K_1(t)}, \quad \text{with} \quad I(t) > 0$$

and

$$\dot{K}_2(t) = -\frac{d}{K_2(t)}, \quad \text{i.e.} \quad I(t) = 0.$$

The latter equation yields the end-time σ, for $K_2(\sigma) = 0$, as

$$\sigma = \frac{\varepsilon^2}{2d}. \tag{D.1a}$$

Moreover we have

$$K_1(t) > K_2(t), \quad 0 \le t < \sigma. \tag{D.1b}$$

Next we compare the difference of objective function Eq. (2a) for $(K_1(\cdot), I(\cdot))$ and $(K_2(\cdot), 0)$ yielding

$$\Delta V := \int_0^\sigma \mathrm{e}^{-rt} \left(K_1(t) - K_2(t) - cI(t)^2\right) \mathrm{d}t$$

Using $K_i(t) = K_i(0) + \int_0^t \dot{K}(s)\,\mathrm{d}s$ yields

$$\Delta V = \int_0^\sigma \mathrm{e}^{-rt} \left(\int_0^t I(s) - d\left(\frac{1}{K_1(s)} - \frac{1}{K_2(s)}\right) \mathrm{d}s \right) \mathrm{d}t - c \int_0^\sigma \mathrm{e}^{-rt}\, I(t)^2\, \mathrm{d}t. \tag{D.1c}$$

Due to Eq. (D.1b) the term

$$-d\left(\frac{1}{K_1(s)} - \frac{1}{K_2(s)}\right) > 0.$$

Using partial integration we find

$$\int_0^\sigma \mathrm{e}^{-rt} \int_0^t I(s)\,\mathrm{d}s\,\mathrm{d}t = \frac{1}{r}\int_0^\sigma \mathrm{e}^{-rt}\, I(t)\,\mathrm{d}t - \frac{1}{r}\mathrm{e}^{-r\sigma} \int_0^\sigma I(t)\,\mathrm{d}t. \tag{D.1d}$$

Now let $I(\cdot) \equiv \bar{I}$. Using Eq. (D.1d) and the last term of Eq. (D.1c) yields

$$\frac{1}{r}\int_0^\sigma \mathrm{e}^{-rt}\, \bar{I}\,\mathrm{d}t - \frac{1}{r}\mathrm{e}^{-r\sigma}\int_0^\sigma \bar{I}\,\mathrm{d}t - c\int_0^\sigma \mathrm{e}^{-rt}\, \bar{I}^2\,\mathrm{d}t = \frac{1}{r^2}\bar{I}\left(1 - \mathrm{e}^{-r\sigma}\right) - \frac{1}{r}\mathrm{e}^{-r\sigma}\,\sigma\bar{I} - \frac{c\bar{I}^2}{r}\left(1 - \mathrm{e}^{-r\sigma}\right).$$

Approximating $\mathrm{e}^{-r\sigma} \approx 1 - r\sigma$

we get

$$= \sigma^2\bar{I} - c\sigma\bar{I}^2 = \sigma\bar{I}(\sigma - c\bar{I}) > 0$$

for $\bar{I} < \dfrac{\sigma}{c}$. Since σ is independent of $I(\cdot)$ (cf. Eq. (D.1a)), this can always be achieved. □

Lemma 1 immediately yields

Corollary 2. *Let* $(K^*(\cdot), I^*(\cdot), T^*)$ *with* $T^* < \infty$ *then the corresponding state and costate variable* $\lambda(\cdot)$ *satisfy*

$$K^*(T^*) = \lambda(T^*) = 0. \tag{D.2}$$

Proof. This immediately follows from the maximizing condition Eq. (4). For $T^* - \sigma < t < T^*$

$$I^*(t) = I^\circ(\lambda(t), \lambda_0) = \frac{\lambda(t)}{2c}.$$

Due to the continuity of the optimal control and costate we find

$$\lim_{t\to T^*} I^*(t) = \lim_{t\to T^*} \frac{\lambda(t)}{2c},$$
$$0 = \frac{\lambda(T^*)}{2c},$$

and hence

$$\lambda(T^*) = 0.$$

□

Next we prove that if we start from the initial conditions $K(0) = \lambda(0) = 0$ and calculate backwards in time we get a unique solution for the canonical system. The according state and control path is the optimal solution at least for $K(t) < \tilde{K}_1$. To determine the optimal solution for every initial state it is important to find the switching point, when the control reaches its maximum $I_{\max}$. Subsequently, we will see that there is a qualitative difference in the structure of the optimal solutions if this switching point is before or after the Stalling Equilibrium $\tilde{K}_1$.

Lemma 2. *The initial value problem* $K(0) = \lambda(0) = 0$ *with* $\lambda_0 = 1$ *for the ODEs*

$$\dot{K}(t) = I^\circ(\lambda(t), \lambda_0) - \delta K(t) + a - \frac{d}{K(t)}, \quad I^\circ(\lambda(t), \lambda_0) < I_{\max}, \tag{D.3a}$$

$$\dot{\lambda}(t) = \lambda(t) q(K(t)) - \frac{1}{K(t)+1}, \tag{D.3b}$$

with

$$q(K(t)) := r + \delta - \frac{d}{K(t)^2},$$

and

$$I^\circ(\lambda, 1) = \frac{\lambda}{2c} \tag{D.3c}$$

has a unique solution $(\bar{K}(t), \bar{\lambda}(t))$ *with* $\bar{K}(t) \geq 0,\ \bar{\lambda}(t) \geq 0$ *and* $\dot{\bar{\lambda}}(t) > 0$ *for all* $t \leq 0$.

Proof. In the neighborhood of $K = 0$ we use the following time transformation

$$t = \int_0^s K(s)\,\mathrm{d}s$$

which transform the ODE Eq. (D.3a) into

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s} K(s) &= \frac{\lambda(s)}{2c} K(s) - \delta K(s)^2 + aK(s) - d \\ \frac{\mathrm{d}}{\mathrm{d}s} \lambda(s) &= \lambda(s) \left((r+\delta)K(s) - \frac{d}{K(s)} \right) - \frac{K(s)}{K(s)+1}. \end{aligned}$$

At the origin these ODEs can be approximated by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}s} K(s) &= aK(s) - d \\ \frac{\mathrm{d}}{\mathrm{d}s} \lambda(s) &= -\frac{d}{K(s)} \lambda(s) \end{aligned}$$

yielding for $s \leq 0$ and $K(0) = \lambda(0) = 0$

$$K(s) = \frac{1}{a}\left(1 - \mathrm{e}^{as}\right) \geq 0, \quad s \leq 0.$$

Hence the adjoint equation yields a singularity of the first kind at $s = 0$ with the solution

$$\lambda(s) = \frac{1}{d}\left(1 - \mathrm{e}^{as}\right) \geq 0, \quad s \leq 0.$$

Thus, for some $\varepsilon > 0$ a unique solution exists for $\|(K(\cdot, \lambda(\cdot))\| < \varepsilon$ and the IVP (D.3). In a region with $K \geq \varepsilon > 0$ the dynamics of Eq. (D.3) is continuously differentiable with respect to K and λ and hence the solution is unique. And the estimation of the ODEs

$$\begin{aligned} \dot{K}(t) &= I^\circ(\lambda(t), \lambda_0) - \delta K(t) + a - \frac{d}{K(t)} \leq \frac{\lambda(t)}{2c} - \delta K(t) \\ \dot{\lambda}(t) &= \lambda(t) q(K(t)) - \frac{1}{K(t)+1} \leq \lambda(t)(r+\delta) \end{aligned}$$

implies the existence of a solution for every $t \leq 0$. □

In the next step we consider the zeros of Eq. (5) in dependence of a (control) value $I \in \mathbb{R}$ with $\lambda_0 = 1$. The properties of these equations yield information about the existence and admissibility of the equilibria for the canonical system Eq. (5). Thus, we analyze the following equations[12]

$$0 = I - \delta K + a - \frac{d}{K}, \tag{D.4a}$$

$$0 = -\frac{1}{K+1} + \lambda \left(r + \delta - \frac{d}{K^2} \right). \tag{D.4b}$$

Remember that $I_{\min}$, defined in Eq. (11b), denotes the minimal value of I such that Eq. (D.4a) has a zero, which is not necessarily positive.

Lemma 3 (Properties of the zeros for Eq. (D.4))**.** *For $I \in [I_{\min}, \infty)$, with*

$$I_{\min} := 2\sqrt{d\delta} - a,$$

the zeros $(K(I), \lambda(I))$ of Eq. (D.4) consist of two branches $i = 1, 2$. The properties of $K_i(I), i = 1, 2$, are stated in Proposition 1. For $\lambda_i(I), i = 1, 2$, the zeros of Eq. (D.4b), we find

$$\lambda_i(I) = \frac{1}{(1 + K_i(I))\, q_i(I)}, \tag{D.5a}$$

with

$$q_i(I) := r + \delta - \frac{d}{K_i(I)^2}.$$

Let

$$I_{\text{crit}} := \sqrt{\frac{d}{r+\delta}}(r + 2\delta) - a, \tag{D.6a}$$

then

$$I_{\text{crit}} > I_{\min}. \tag{D.6b}$$

The two branches $\lambda_i(I), i = 1, 2$, satisfy

$$\lambda_2(I) > 0, \quad I \in [I_{\min}, \infty), \tag{D.7a}$$

$$\lambda_1(I_{\min}) = \lambda_2(I_{\min}) = \frac{\sqrt{d}}{r(\sqrt{\delta} + \sqrt{d})} > 0, \tag{D.7b}$$

$$\lim_{I\to\infty} \lambda_1(I) = 0, \quad \lim_{I\to\infty} \lambda_2(I) = 0, \tag{D.7c}$$

$$\lim_{I\to I_{\text{crit}}^-} \lambda_1(I) = \infty, \quad \lim_{I\to I_{\text{crit}}^+} \lambda_1(I) = -\infty, \tag{D.7d}$$

$$\lim_{I\to I_{\text{crit}}^-} \lambda_2(I) = \lim_{I\to I_{\text{crit}}^+} \lambda_2(I), \tag{D.7e}$$

and the derivatives satisfy

$$\frac{\partial}{\partial I}\lambda_1(I) > 0, \quad I \in [I_{\min}, \infty) \setminus \{I_{\text{crit}}\}, \tag{D.7f}$$

$$\frac{\partial}{\partial I}\lambda_2(I) < 0, \quad I \in [I_{\min}, \infty). \tag{D.7g}$$

[12]In the subsequent we omit the time argument t.

Proof. From Eq. (D.4b) we find the costate value

$$\lambda_i(I) = \frac{1}{(1 + K_i(I))\left(r + \delta - \dfrac{d}{K_i(I)^2}\right)}, \tag{D.8a}$$

or

$$\lambda_i(I) = \frac{1}{(1 + K_i(I))\, q_i(I)}, \tag{D.8b}$$

with

$$q_i(I) := r + \delta - \frac{d}{K_i(I)^2},$$

where $K_i(I)$ is given by Eq. (9a).

Next we derive the value I_{crit}. Since $K_i \geq 0$, the denominator of Eq. (D.8a) becomes zero iff

$$q_i(I) := r + \delta - \frac{d}{K_i(I)^2} = 0 \quad \text{or} \quad K_i(I) = \sqrt{\frac{d}{r+\delta}} =: K_{\text{crit}}. \tag{D.9}$$

The value of the branch $K_2(I)$ can be estimated from above, using $K_2(I) \geq K_2(I_{\min})$

$$K_2(I) = \frac{I + a + \sqrt{(I+a)^2 - 4d\delta}}{2\delta} \geq \sqrt{\frac{d}{\delta}},$$

and hence

$$q_2(I) = r + \delta - \frac{d}{K_2(I)^2} \geq r > 0, \quad \text{for all} \quad I \geq I_{\min},$$

yielding

$$\lambda_2(I) > 0, \quad \text{for all} \quad I \geq I_{\min}.$$

Therefore K_{crit} can only occur on the branch $K_1(I)$. To determine the corresponding control value I_{crit} we follow

$$\begin{aligned} K_{\text{crit}} &= \frac{I + a - \sqrt{(I+a)^2 - 4d\delta}}{2\delta}, \\ (I+a) - 2\delta K_{\text{crit}} &= \sqrt{(I+a)^2 - 4d\delta}. \end{aligned} \tag{D.10a}$$

Using the identity Eq. (D.9), i.e. $d = K_{\text{crit}}^2(r+\delta)$, and squaring Eq. (D.10a) yields

$$\begin{aligned} &-4(I+a)\delta K_{\text{crit}} + 4\delta^2 K_{\text{crit}}^2 = -4K_{\text{crit}}^2(r+\delta)\delta, \\ &I + a = K_{\text{crit}}(r + 2\delta). \end{aligned}$$

Thus, I_{crit} is given by

$$I_{\text{crit}} := K_{\text{crit}}(r + 2\delta) - a = \sqrt{\frac{d}{r+\delta}}(r + 2\delta) - a. \tag{D.10b}$$

Moreover, from the identity $I + a = K_{\text{crit}}(r + 2\delta)$ we find $(I + a) - 2\delta K_{\text{crit}} = K_{\text{crit}} r > 0$ verifying the squaring of Eq. (D.10a).

To show that I_{crit}, given by Eq. (D.10b), satisfies $I_{\text{crit}} > I_{\min}$ we note that

$$(I_{\text{crit}} + a)^2 - 4d\delta = \frac{d}{r+\delta}(r+2\delta)^2 - 4d\delta = \frac{dr^2}{r+\delta} > 0$$

holds.

For $\lambda_2(I)$ the derivative with respect to I exists for all $I \in [I_{\min}, \infty)$, and for $\lambda_1(I)$ the derivative with respect to I exists for all $I \in [I_{\min}, \infty) \setminus \{I_{\text{crit}}\}$. We can derive

$$
\begin{aligned}
\frac{\partial}{\partial I}\lambda_i(I) &= \frac{\partial}{\partial I}\frac{1}{(1+K_i(I))\left(r+\delta-\frac{d}{K_i(I)^2}\right)} \\
&= -\frac{\frac{\partial}{\partial I}K_i(I)}{(1+K_i(I))\left(r+\delta-\frac{d}{K_i(I)^2}\right)}\left(\frac{1}{1+K_i(I)}+\frac{2d}{K_i(I)^3\left(r+\delta-\frac{d}{K_i(I)^2}\right)}\right) \\
&= -\frac{\partial}{\partial I}K_i(I)\frac{K_i(I)^3(r+\delta)+2d+dK_i(I)}{(1+K_i(I))^2K_i(I)^3\left(r+\delta-\frac{d}{K_i(I)^2}\right)^2}.
\end{aligned}
$$

Since the last term

$$
\frac{K_i(I)^3(r+\delta)+2d+dK_i(I)}{(1+K_i(I))^2K_i(I)^3\left(r+\delta-\frac{d}{K_i(I)^2}\right)^2} > 0,
$$

we find

$$
\operatorname{sgn}\left(\frac{\partial}{\partial I}\lambda_i(I)\right) = -\operatorname{sgn}\left(\frac{\partial}{\partial I}K_i(I)\right),
$$

which proves Eqs. (D.7f) and (D.7g). □

From Lemma 3 it follows that the branches $\lambda_i(I), i = 1, 2$, are of the form depicted in Figure D.6a. For a complete characterization of the equilibria of the canonical system Eq. (D.4) we first note that the control values are either $I_{\max}$ or satisfy $2cI = \lambda$. Thus candidates for the equilibria are $(K_i(I_{\max}), \lambda_i(I_{\max})), i = 1, 2$ or where $\lambda_i(I)$ has an intersection point with $2cI$. Moreover, for an equilibrium at $I_{\max}$ the corresponding Lagrange-multiplier has to be positive, i.e. $\lambda_i(I_{\max}) \geq 2cI_{\max}$.

The following proposition states these conditions explicitly.

Lemma 4. *For the equilibria of the canonical system Eq. (D.4) and $\lambda_0 = 1$, we have to distinguish the following cases*

$I_{\text{crit}} \leq 0$: *depending on the cost-parameter c we find (see Figure D.6b)*

- $\lambda_2(I_{\max}) < 2cI_{\max}$: *one interior admissible equilibrium exists.*
- $\lambda_2(I_{\max}) \geq 2cI_{\max}$: *one admissible equilibrium at $I_{\max}$ exists.*

$0 < I_{\text{crit}} \leq I_{\max}$: *depending on the cost-parameter c we find (see Figure D.6c)*

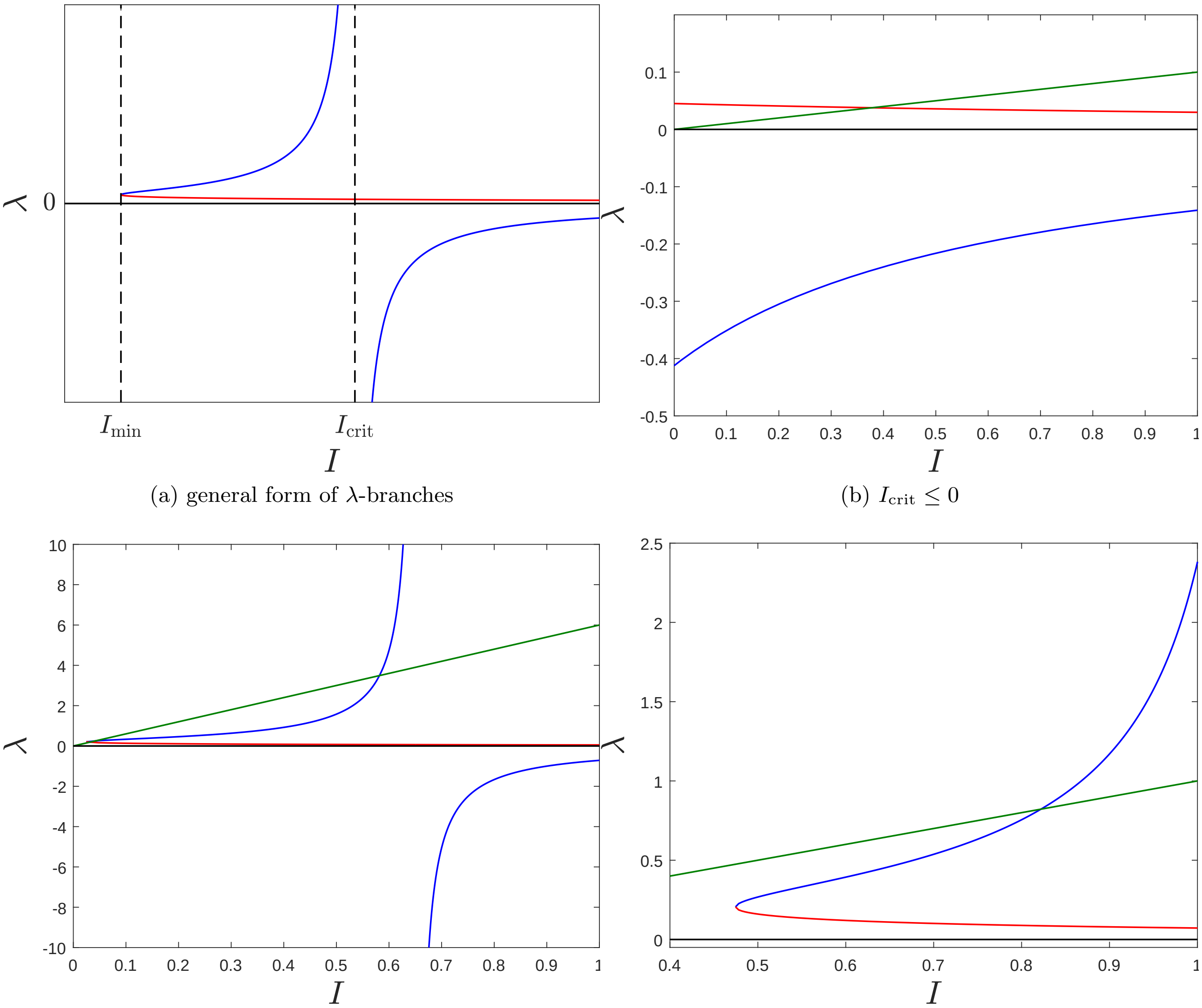
- $\lambda_2(I_{\max}) < 2cI_{\max}$: *one or three interior admissible equilibria exist.*
- $\lambda_2(I_{\max}) \geq 2cI_{\max}$: *one admissible equilibrium at $I_{\max}$ exists.*

$I_{\text{crit}} > I_{\max}$: *depending on the cost-parameter c we find (see Figure D.6d)*

- $\lambda_1(I_{\max}) < 2cI_{\max}$: *none or two interior admissible equilibria exist.*
- $\lambda_1(I_{\max}) = 2cI_{\max}$: *one admissible equilibrium at $I_{\max}$ exists and none or two interior admissible equilibria exist.*
- $\lambda_1(I_{\max}) < 2cI_{\max} < \lambda_2(I_{\max})$: *one or three interior equilibria exist and one admissible equilibrium at $I_{\max}$ exists.*

$\lambda_2(I_{\max}) \le 2cI_{\max}$: *two admissible equilibria at $I_{\max}$ exists.*

The proof is an immediate consequence of the properties of the zero branches $\lambda_i(I)$, $i = 1, 2$, stated in Lemma 3. We omit the tedious details and refer to Figure D.6, where the different cases are depicted.



(a) general form of λ-branches

(b) $I_{\text{crit}} \le 0$

(c) $0 < I_{\text{crit}} < I_{\max}, I_{\min} > 0$

(d) $I_{\text{crit}} \ge I_{\max}, I_{\min} > 0$

Figure D.6: In panel (a) the general form of the costate branches derived in Lemma 3 are depicted. The blue line denotes the first branch $\lambda_1(I)$ and the red curve the second branch $\lambda_2(I)$. The subsequent panels show the three main cases (b) $I_{\text{crit}} \le 0$, (c) $0 < I_{\text{crit}} < I_{\max}$, and (d) $I_{\text{crit}} \ge I_{\max}$. The green line in these figures corresponds to $\lambda = 2cI$ and hence indicates the existence or admissibility of the equilibria of the canonical system, see Lemma 4.

In the following lemma we consider the stability properties of the canonical system at the Stalling Equilibrium $\tilde{K}_1$.

Lemma 5. *Let $I_{\max} \ge 2\sqrt{d\delta} - a$ and $\lambda_0 = 1$, then $(\tilde{K}_1, \tilde{\lambda}_1)$ exists, with the Jacobian $\tilde{J}_1$*

$$\tilde{J}_1 = \begin{pmatrix} \dfrac{d}{\tilde{K}_1^2} - \delta & 0 \\ \tilde{J}_{21} & r + \delta - \dfrac{d}{\tilde{K}_1^2} \end{pmatrix} = \begin{pmatrix} r - \tilde{q}_1 & 0 \\ J_{21} & \tilde{q}_1 \end{pmatrix}, \tag{D.11}$$

and

$$\tilde{J}_{21} := \frac{1}{(\tilde{K}_1+1)^2} + \tilde{\lambda}_1 \frac{2d}{\tilde{K}_1^3} \neq 0,$$
$$\tilde{K}_1 = \frac{I_{\max} + a - \sqrt{(I_{\max}+a)^2 - 4d\delta}}{2\delta},$$
$$\tilde{\lambda}_1 = \frac{1}{(\tilde{K}_1+1)\left(r+\delta-\frac{d}{\tilde{K}_1^2}\right)} = \frac{1}{(\tilde{K}_1+1)\tilde{q}_1},$$
$$\tilde{q}_1 = r + \delta - \frac{d}{\tilde{K}_1^2}.$$

The eigenvalues ξ_i *and eigenvectors* v_i*,* $i = 1,2$ *of* $\tilde{J}_1$ *are*

$$\xi_1 = r - \tilde{q}_1 > 0, \quad v_1 = \begin{pmatrix} -\frac{1}{\tilde{J}_{21}}(2\tilde{q}_1 - r) \\ 1 \end{pmatrix}, \tag{D.12a}$$

$$\xi_2 = \tilde{q}_1 \lesseqqgtr 0, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{D.12b}$$

Proof. Plugging the values for $(\tilde{K}_1, \tilde{\lambda}_1)$ into the linearization of canonical system yields Eq. (D.11). Since the Jacobian $\tilde{J}_1$ is upper triangular the eigenvalues are the diagonal elements, yielding $\tilde{q}_1$ and $r - \tilde{q}_1$.

Next we show that $\tilde{J}_{21} \neq 0$.

$$\begin{aligned}
\tilde{J}_{21} =& \frac{1}{(\tilde{K}_1+1)^2} + \tilde{\lambda}_1 \frac{2d}{\tilde{K}_1^3} \\
=& \frac{1}{(\tilde{K}_1+1)^2} + \frac{1}{(\tilde{K}_1+1)\tilde{q}_1}\frac{2d}{\tilde{K}_1^3} \\
=& \frac{\tilde{q}_1\tilde{K}_1^3 + (\tilde{K}_1+1)2d}{(\tilde{K}_1+1)^2\tilde{q}_1\tilde{K}_1^3} \\
=& \frac{(r+\delta)\tilde{K}_1^3 + d\tilde{K}_1 + 2d}{(\tilde{K}_1+1)^2\tilde{q}_1\tilde{K}_1^3} \neq 0.
\end{aligned}$$

To determine the eigenvector for $\tilde{q}_1$ we solve

$$\begin{pmatrix} r - 2\tilde{q}_1 & 0 \\ \tilde{J}_{21} & 0 \end{pmatrix} \begin{pmatrix} v_1^1 \\ v_1^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

setting $v_1^2 = 1$ we find

$$\begin{aligned}
&(r - 2\tilde{q}_1)v_1^1 = 0, \\
&\tilde{J}_{21}v_1^1 = 0.
\end{aligned}$$

Since $\tilde{J}_{21} \neq 0$, we find

$$v_1^1 = 0.$$

To determine the eigenvector corresponding to $r - \tilde{q}_1$ we solve

$$\begin{pmatrix} 0 & 0 \\ \tilde{J}_{21} & 2\tilde{q}_1 - r \end{pmatrix} \begin{pmatrix} v_2^1 \\ v_2^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Setting $v_2^2 = 1$, we find

$$\tilde{J}_{21} v_2^1 + 2\tilde{q}_1 - r = 0,$$

due to $\tilde{J}_{21} \neq 0$ we hence find

$$v_2^1 = -\frac{1}{\tilde{J}_{21}}(2\tilde{q}_1 - r).$$

□

The following proposition formulates the condition for which the solution starting at the Stalling Equilibrium is optimal.

Proposition 6. *The constant solution $K^* \equiv \tilde{K}_1$ with $T^* = \infty, K^*(\cdot) \equiv \tilde{K}_1$ and $I^*(\cdot) = I_{\max}$ is optimal iff $\tilde{D} := (I_{\max} + a)^2 - 4d\delta \geq 0$ and one of the following cases is satisfied:*

1. *There exists an $\varepsilon > 0$ such that for all K_0 with $0 < \tilde{K}_1 - K_0 < \varepsilon$ and with $(K^*(K_0, \cdot), I^*(K_0, \cdot), T^*(K_0)) \in \mathcal{S}(K_0)$ we have $I^*(K_0, 0) = I_{\max}$.*
2. *The point $\tilde{K}_1$ is a Skiba point and one Skiba solution $(K_S^*(\cdot), I_S^*(\cdot), T_S^*) \in \mathcal{S}(\tilde{K}_1)$ satisfies $T_S^* < \infty$*

$$I_S^*(0) = I_{\max} \quad \text{and} \quad I_S^*(t) < I_{\max}, \; 0 < t \leq T_S^*. \tag{D.13}$$

Proof. We note that for $\tilde{D} \geq 0$ the Stalling Equilibrium $\tilde{K}_1$ exists. In Lemma 2 we proved that the initial value problem of Eq. (D.3) with $K(0) = \lambda(0) = 0$ for $t \leq 0$ has a unique solution. Moreover we showed that the according costate $\lambda(t)$ satisfies $\lambda(t) \geq 0$ and $\dot{\lambda}(t) > 0$ for $t \leq 0$. Therefore a $\sigma < 0$ exists such that one of the following cases applies

1. $K(\sigma) < \tilde{K}_1$ and $\lambda(\sigma) = 2cI_{\max}$.
2. $K(\sigma) = \tilde{K}_1$ and $\lambda(\sigma) \leq 2cI_{\max}$.

In the first case with $\varepsilon := \tilde{K}_1 - K(\sigma)$ we find solutions satisfying item 1.

In the second case for $\lambda(\sigma) < 2cI_{\max}$ the solution staying at $\tilde{K}_1$ cannot be optimal. To prove that we consider the maximized Hamiltonian function with $\lambda_0 = 1$

$$\mathcal{H}^\circ(K, \lambda) := \ln(K+1) - \frac{\lambda^2}{4c} + \lambda\left(\frac{\lambda}{2c} - \delta K + a - \frac{d}{K}\right),$$

yielding the derivatives

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathcal{H}^\circ(K, \lambda) = \frac{\lambda}{2c} - \delta K + a - \frac{d}{K},$$
$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2}\mathcal{H}^\circ(K, \lambda) = \frac{1}{2c} > 0.$$

Thus, the solution of the equation

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathcal{H}^\circ(\tilde{K}_1, \lambda) = 0,$$

yields a strict minimum of the maximized Hamiltonian at $\tilde{K}_1$, with respect to the costate λ. This proves that the minimum is achieved for[13]

$$\lambda = 2cI_{\max}.$$

[13]That the Hamiltonian achieves its minimum at an equilibrium of the canonical system is a general result. A proof can be found in Grass et al. (18).

Hence for every $\lambda \neq 2cI_{\max}$ we find

$$\mathcal{H}^\circ(\tilde{K}_1, \lambda) > \mathcal{H}^\circ(\tilde{K}_1, 2cI_{\max}) = 0.$$

Specifically this holds for $\lambda(\sigma) < 2cI_{\max}$. Since the objective value of a solution that satisfies the necessary optimality conditions is given by

$$\frac{1}{r}\mathcal{H}^\circ(\tilde{K}_1, \lambda(0)),$$

we find that the solution that stays at $\tilde{K}_1$ is not optimal.

For the second case with $\lambda(\sigma) = 2cI_{\max}$ we find that the according finite time solution achieves the same objective value as the solution that stays at $\tilde{K}_1$. These are the only possible solutions. Hence $\tilde{K}_1$ is a Skiba point and the previously described solutions are the according Skiba solutions. □

For normal optimal control problems an equilibrium point usually cannot be a Skiba point as in item 2. The reason that item 2 holds is the abnormality of problem (2) for the infinite time horizon solution.

Corollary 3. *Let the conditions of item 2 of Proposition 6 be satisfied, then the problem for the infinite time horizon solution is abnormal.*

Remark. *Note that as a consequence of Corollary 3 we find that abnormality is not a property of the problem (2) for a specific initial condition $K(0) = K_0$. Rather it is a property of the specifically chosen maximization. In that case maximization with respect to the infinite time or the finite time horizon.*

Proof. Let item 2 of Proposition 6 be satisfied and the problem be normal. Then from the properties of the equilibrium $(\tilde{K}_1, \tilde{\lambda}_1)$ in the canonical system analyzed in Lemma 4 and the admissibility of the equilibrium we find $\tilde{q}_1 > 0$ and hence that $\tilde{\lambda}_1 \geq 2cI_{\max}$. Now let us assume that there exists a finite time solution that is equally optimal. Consequently this has to satisfy $\lambda(0) = 2cI_{\max}$. If $\tilde{\lambda}_1 = 2cI_{\max}$ the choice $\lambda(0) = 2cI_{\max}$ implies $\dot{\lambda}(0) = 0$ and hence the according costate path stays at $\tilde{\lambda}_1$ forever and cannot satisfy $\lambda(t) = 0$ for some $t > 0$. On the other hand if $\tilde{\lambda}_1 > 2cI_{\max}$ repeating the argument of the proof for Proposition 6 we find that the according finite time solution is better than the infinite time horizon solution. The Hamiltonian achieves the minimum at the equilibrium value. This contradicts the assumed optimality of the equilibrium solution. Thus, we find that the problem cannot be normal for the infinite time horizon solution and hence has to be abnormal. □

An immediate consequence of Proposition 6 is that if one of the cases items 1 and 2 applies the optimal solution for $K(0) > \tilde{K}_1$ cannot be a finite time solution. This is formulated in the following proposition

Proposition 7. *Let the conditions of Proposition 6 be satisfied, then the optimal solution for $K(0) = K_0 > \tilde{K}_1$, $(K^*(K_0, \cdot), I^*(K_0, \cdot), T^*(K_0))$ satisfies $T^*(K_0) = \infty$ and*

$$\lim_{t\to\infty} K^*(K_0, t) = \hat{K} \quad \textit{and} \quad \lim_{t\to\infty} I^*(K_0, t) = \hat{I}. \tag{D.14}$$

Proof. Let the conditions of Proposition 6 be satisfied and chose for some (small enough) $\varepsilon > 0$ a K_0 such that $\tilde{K}_1 < K_0 < \tilde{K}_1 + \varepsilon$. Then the according optimal solution $(K^*(K_0, \cdot), I^*(K_0, \cdot), T^*(K_0))$ cannot end at the origin. Let otherwise $(K^*(K_0, \cdot), I^*(K_0, \cdot), T^*(K_0))$ be a finite time horizon solution, then there exists $\sigma > 0$ such that $K^*(K_0, \sigma) = \tilde{K}_1$ and $\dot{K}^*(K_0, t) < 0$ for $0 \leq t \leq T^*(K_0)$. This implies that $I^*(K_0, t) < I_{\max}$ for $0 \leq t \leq \sigma$ contradicting that under the conditions of Proposition 6 the optimal control at $\tilde{K}_1$ has to be at its maximum $I_{\max}$. Therefore for all $K(0) > \tilde{K}_1$ the optimal solution cannot end at the origin.

Next we show that for these initial $K_0 > \tilde{K}_1$ the optimal solutions cannot be finite time solutions. Otherwise with $0 \le T^*(K_0) < \infty$ we have $K^*(K_0, T^*(K_0)) > \tilde{K}_1$. Then consider the solution $\bar{K}(K_0, \cdot)$ corresponding to the control $\bar{I}(K_0, \cdot)$ with

$$\bar{I}(K_0, t) := \begin{cases} I^*(K_0, t), T^*(K_0) & \text{for} \quad 0 \le t \le T^*(K_0) \\ 0 & \text{for} \quad T^*(K_0) < t \le \sigma(K^*(K_0, T^*(K_0)), \end{cases}$$

where $\sigma(K^*(K_0, T^*(K_0)))$ is the time such that $\bar{K}(K_0, T^*(K_0)) = 0$ and $\bar{K}(K_0, t) > 0$ for $t < T^*(K_0)$. In that case

$$V(K_0, \bar{I}(K_0, \cdot), T^*(K_0) + \sigma(K^*(K_0, T^*(K_0))) > V(K_0, I^*(K_0, \cdot), T^*(K_0)),$$

contradicts the optimality of $(K^*(K_0, \cdot), I^*(K_0, \cdot), T^*(K_0))$. Thus, $T^*(K_0)$ has to be infinite. The infinite time horizon problem has a unique optimal solution. A general result for one state infinite time optimal control problems states that the state path is monotonous. Since the optimal state cannot diverge we find Eq. (D.14) satisfied. □

The following proposition provides an exhaustive overview of all possible optimal policies that can occur for problem (2).

Proposition 8. *Let*

$$\tilde{D} := (I_{\max} + a)^2 - 4d\delta,$$

be the discriminant defined in Eq. (11a) at $I_{\max}$, defined in Eq. (13a). Then any optimal solution of $\mathcal{S}(K_0)$ can be described by one of the following cases.

1. $\tilde{D} < 0$: *For every $K_0 \ge 0$ the finite time horizon solution is optimal, and $|\mathcal{S}(K_0)| = 1$.*
2. $\tilde{D} \ge 0$ *and $\tilde{q}_1 < 0$. Therefore, the equilibrium $\tilde{K}_1$ exist. One of the following generic cases applies:*
   - (a) *For $K(0) < \tilde{K}_1$ the finite time horizon solution is optimal. For $K(0) > \tilde{K}_1$ the optimal solution converges to some $\hat{K} > \tilde{K}_1$, and for $K(0) = \tilde{K}_1$ it is optimal to stay at $\tilde{K}_1$, i.e. $|\mathcal{S}(K_0)| = 1$;*
   - (b) *For $K(0) < K_S$ with $K_S > \tilde{K}_1$ the finite time horizon solution is optimal. For $K(0) > K_S$ the optimal solution converges to some $\hat{K} > K_S$. The point K_S is a Skiba point and the Skiba solutions are the finite time horizon solution and a solution that converges to some $\hat{K} > \tilde{K}_1$;*
   - (c) *For every $K_0 \ge 0$ the finite time horizon solution is optimal, and $|\mathcal{S}(K_0)| = 1$.*

   Also, one of the bifurcation case can occur:
   - (a′) *For $K(0) < \tilde{K}_1$ the finite time horizon solution is optimal. For $K(0) > \tilde{K}_1$ the optimal solution converges to some $\hat{K}$. The point $\tilde{K}_1$ is a Skiba point and the Skiba solutions are the finite time horizon solution and the solution that stays at $\tilde{K}_1$.*
   - (b′) *For $K(0) < \hat{K}$ with $\hat{K} > \tilde{K}_1$ the finite time horizon solution is optimal. For $K(0) > \hat{K}$ the optimal solution converges to some $\hat{K}$. For $K(0) = \hat{K}$ the optimal solution stays at $\hat{K}$.*
3. $\tilde{D} \ge 0$ *and $\tilde{q}_1 \ge 0$. Therefore, the equilibrium $\tilde{K}_1$ exists. And one of the following generic cases applies:*
   - (a) *For $K(0) < \tilde{K}_1$ the finite time horizon solution is optimal. For $K(0) > \tilde{K}_1$ the optimal solution converges to some $\hat{K} > \tilde{K}_1$, and for $K(0) = \tilde{K}_1$ it is optimal to stay at $\tilde{K}_1$;*
   - (b) *For every $K_0 \ge 0$ the finite time horizon solution is optimal, and $|\mathcal{S}(K_0)| = 1$.*

   Also, one of the bifurcation case can occur:
   - (a′) *For $K(0) < \tilde{K}_1$ the finite time horizon solution is optimal. For $K(0) > \tilde{K}_1$ the optimal solution converges to $\tilde{K}_1$ and for $K(0) = \tilde{K}_1$ the optimal solution stays at $\tilde{K}_1$.*

We omit the proof since these are the immediate consequences of the results of Proposition 6 and Proposition 7.

The following corollary formulates two structurally different case of item 3a of Proposition 8.

Corollary 4. *Let the optimal solutions of the optimal control problem (2) be of type item 3a of Proposition 8. Then there exists $\varepsilon > 0$ such that for K_0 with $|\tilde{K}_1 - K_0| < \varepsilon$ the following properties are satisfied*

$$I^*(K_0, 0) = I_{\max} \quad \text{for} \quad 0 \leq \tilde{K}_1 - K_0 < \varepsilon, \tag{D.15a}$$

$$I^*(K_0, 0) \begin{cases} = I_{\max} & 0 \leq K_0 - \tilde{K}_1 < \varepsilon, \quad \text{for} \quad \tilde{q}_1 < 0 \quad \text{or} \quad \dfrac{1}{(\tilde{K}_1 + 1)\tilde{q}_1} > 2cI_{\max} \\ < I_{\max} & 0 < K_0 - \tilde{K}_1 < \varepsilon, \quad \text{for} \quad \dfrac{1}{(\tilde{K}_1 + 1)\tilde{q}_1} = 2cI_{\max} \end{cases}, \tag{D.15b}$$

$$T^*(K_0) < \infty \quad \text{for} \quad 0 < \tilde{K}_1 - K_0 < \varepsilon, \tag{D.15c}$$

$$T^*(K_0) = \infty \quad \text{for} \quad 0 \leq K_0 - \tilde{K}_1 < \varepsilon, \tag{D.15d}$$

Let $\lambda(K_0, \cdot)$ be the corresponding costate of the optimal solution for $K(0) = K_0$ and $\tilde{q}_1 = 0$. If $\tilde{q}_1 < 0$, the solutions satisfy

$$\lim_{K_0 \to \tilde{K}_1, K_0 \neq \tilde{K}_1} \lambda(K_0, 0) = \infty, \qquad \lim_{K_0 \to \tilde{K}_1} T^*(K_0) = \infty. \quad \text{(abnormal)} \tag{D.16}$$

If $\tilde{q}_1 > 0$, the solutions satisfy

$$\lim_{K_0 \to \tilde{K}_1, K_0 \neq \tilde{K}_1} \lambda(K_0, 0) = \tilde{\lambda}_1, \qquad \lim_{K_0 \to \tilde{K}_1} T^*(K_0) = \infty. \quad \text{(normal)} \tag{D.17}$$

In both cases the optimal solution for $K_0 = \tilde{K}_1$ is unique and satisfies

$$K^*(\cdot) \equiv \tilde{K}_1, \quad I^*(\cdot) \equiv I_{\max} \quad \text{and} \quad T^* = \infty. \tag{D.18}$$

Therefore, the Stalling Equilibrium $\tilde{K}_1$ is a weak Skiba point.

Proof. In Proposition 6 item 1 we proved that if $\tilde{K}_1$ is an optimal equilibrium and if it is not a Skiba point there exists a $\bar{K} < \tilde{K}_1$ such that for $\bar{K} \leq K_0 < \tilde{K}_1$ the optimal control $I^*(K_0, 0) = I_{\max}$. In Proposition 7 we showed that for $K_0 > \tilde{K}_1$ the optimal solution converges to an equilibrium, i.e. some long run optimal solution.

Let $\lambda(K_0, \cdot)$ be the costate corresponding to the optimal solution in $\mathcal{S}(K_0)$. For K_0 near enough to $\tilde{K}_1$ the optimal solution is unique. We can distinguish different cases, depending on the sign of $\tilde{q}_1$.[14] First we note that for K_0 near enough to $\tilde{K}_1$ the optimal control is positive and hence $\lambda(K_0, 0) > 0$.

- $\tilde{q}_1 < 0$: Due to the adjoint dynamics Eq. (5) we find

$$\dot{\lambda}(K_0, 0) = \lambda(K_0, 0)q(K_0) - \frac{1}{K_0 + 1} < -\frac{1}{\tilde{K}_1 + \varepsilon + 1},$$

  for all $0 < K_0 - \tilde{K}_1 < \varepsilon$, with ε appropriately chosen. This immediately yields the existence of some $\bar{K}$ such that $\lambda(\bar{K}, 0) = 2cI_{\max}$ and hence $I^*(K_0, 0) = I_{\max}$ for $0 < K_0 < \bar{K}$.

- $\tilde{q}_1 > 0$: In that case the equilibrium $(\tilde{K}_1, \tilde{\lambda}_1)$ satisfies $\tilde{\lambda}_1 \geq 2cI_{\max}$. Due to the uniqueness of the optimal solution we find

$$\lim_{K_0 \to \tilde{K}_1} \lambda(K_0, 0) = \tilde{\lambda}_1 \geq 2cI_{\max}.$$

[14] We omit the case $\tilde{q}_1 = 0$ since it needs a specific argumentation and yields no further insight.

If $\tilde{\lambda}_1 = 2cI_{\max}$ we find due to $\dot{\lambda}(K_0, 0) < 0$, that $\lambda(K_0, 0) < 2cI_{\max}$ for $K_0 > \tilde{K}_1$ and hence $I^*(K_0, 0) < I_{\max}$ for $0 < K_0 - \tilde{K}_1 < \varepsilon$. If $\tilde{\lambda}_1 > 2cI_{\max}$ there exists $\bar{K}$ such that $\lambda(\bar{K}, 0) = 2cI_{\max}$ and we can argue as in the case with $\tilde{q}_1 < 0$.

□

Appendix E. Abnormal Case

Halkin (19) gives an example for an infinite time horizon problem, that is abnormal, i.e. $\lambda_0 = 0$. Based on that paper, it seemed that the abnormal case is only of academic interest and showing that $\lambda_0 \neq 0$ is mere a mathematical exercise. In the nineties Montgomery (28) proved the existence of abnormal minimizers in a problem of Riemannian geodesics. Since then the existence and characterization of abnormal solutions is an active topic in optimal control theory and related fields, see e.g. Barbero-Liñán and Muñoz-Lecanda (5).

Anyhow, in economic applications abnormal solutions were of no interest aside from pathological examples, like in Aseev and Veliov (3). Analyzing the example in Halkin (19) in more detail, reveals a characteristic similarity to our problem (2). In Halkin's model, it is optimal to choose zero as the optimal control value, otherwise the state value, which is zero initially, increases and the solution diverges. Importantly, for solutions where the state value is larger than zero, the direction of the drift cannot be changed and it moves uncontrollable to the undesirable state ∞.

In our model the Stalling Equilibrium plays the comparable role of zero in Halkin's model. Only $I_{\max}$ guarantees that the solution stays put, otherwise the direction of the drift cannot be changed and shifts the state into the undesirable position at zero. These analogies provide a plausibility argument on why the occurrence of an abnormal solution is not surprising.

Subsequently, we prove under which conditions the Stalling Equilibrium is admissible for the abnormal problem.

Proposition 9. *Let the problem (2) be abnormal for $K(0) = \tilde{K}_1$ and let $\underline{I} \leq I_{\max}$. Then if $I_{\max} \neq I_{\text{crit}}$ the equilibrium $(\tilde{K}_1, 0)$ with $\tilde{I} = I_{\max}$ is an admissible equilibrium of the canonical system Eq. (5).*

Proof. For the abnormal problem the costate dynamics reduces to

$$\dot{\lambda}(t) = \lambda(t)\left(r - \frac{d}{K(t)^2} + \delta\right), \tag{E.1}$$

and the Lagrangemultiplier ν becomes

$$\nu = \lambda.$$

For $I_{\max} \neq I_{\text{crit}}$ the term

$$r - \frac{d}{\tilde{K}_1^2} + \delta \neq 0,$$

and Eq. (E.1) becomes zero iff $\lambda = 0$. Therefore $(\tilde{K}_1, 0)$ is an equilibrium of the canonical system and the according Lagrange multiplier satisfies $\tilde{\nu} = 0$. □

The following proposition formulates the condition for which the solution starting at the Stalling Equilibrium is abnormal.

Proposition 10. *Let the conditions of Proposition 6 be satisfied. If additionally $\tilde{q}_1 < 0$, then the problem (2) with $K(0) = \tilde{K}_1$ is abnormal. Specifically, item 2 of Proposition 6 can only be satisfied if $\tilde{q}_1 \leq 0$.*

Proof. Due to Proposition 6 the solution, staying at $\tilde{K}_1$ is optimal. Assume that the problem (2) with $K(0) = \tilde{K}_1$ is normal and $\tilde{q}_1 < 0$. For $\tilde{q}_1 < 0$ and $\lambda_0 = 1$ the according costate $\tilde{\lambda}_1 < 0$ yields that the Lagrange multiplier

$$\nu = -\lambda_0 2 I_{\max} c + \tilde{\lambda}_1 < 0$$

and hence the equilibrium $(\tilde{K}_1, \tilde{\lambda}_1)$ of the canonical system is not admissible. Moreover, the properties of the equilibrium analyzed in Lemma 5 show that no other costate value can be chosen, such that the according state costate path is admissible for every $t \geq 0$.

Secondly, assume that item 2 of Proposition 6 holds and that the problem (2) with $K(0) = \tilde{K}_1$ is normal. Let $\tilde{q}_1 > 0$ then the properties of the equilibrium analyzed in Lemma 5 show that $(\tilde{K}_1, \tilde{\lambda}_1)$ is an admissible equilibrium and that the costate dynamics at $\tilde{\lambda}_1$ is zero. Therefore no finite time solution can exist that starts at $\tilde{\lambda}_1$. □