Poverty traps and disaster insurance in a Bi-level decision framework

Raimund Kovacevic, Willi Semmler

Research Report 2020-02

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

POVERTY TRAPS AND DISASTER INSURANCE IN A BI-LEVEL DECISION FRAMEWORK

RAIMUND M. KOVACEVIC AND WILLI SEMMLER

ABSTRACT. In this paper we study mechanisms of poverty traps that can occur after large disaster shocks. Our starting point is a stylized deterministic dynamic model with locally increasing returns to scale possibly generating multiple equilibria paths wit finite upper equilibrium. The deterministic dynamics is then overlayed by random dynamics where catastrophic events happen at random points of time. The number of events follows a Poisson process, whereas the proportional capital losses (given a catastrophic event) are beta distributed. In a setup with fixed insurance premium per unit of insured capital, a fraction of the capital might be insured, and this fraction may change after each event. We seek for an optimal strategy with respect to the insured fraction. Falling below the instable equilibrium of the deterministic dynamics introduces the possibility of ending up in a poverty trap after the disaster shocks. We show that the trapping probability (over an infinite time horizon) is equal to one when the stable upper equilibrium of the deterministic dynamics is finite. This is true regardless of the chosen amount of insured capital. Optimization then is done with the expected discounted capital after the next catastrophic event as the objective. Our model may also be useful to assess risk premia and creditworthiness of borrowers when a sequence of shocks at uncertain times and of uncertain size occurs.
JEL classification: C 61, C 63, L 10, L 11 and L 13

1. INTRODUCTION

There is a long tradition of economic literature where it is argued that economic agents can be driven into long lasting poverty traps - or even into ruin - as a result of large negative shocks or disaster events. This often involves random catastrophic losses, leading into absorbing states from which an escape is not possible. Large negative shocks can result from large income or wealth contractions, such as economic and financial melt-downs, from disasters such as earth quakes and climate related disasters. In each case a significant percentage of GDP, public

Raimund Kovacevic at Vienna University of Vienna.

Willi Semmler at New School for Social Research, Bielefeld University, and IIASA.

We want to thank Stefan Mittnik, Elena Rovenskaya and Georg Pflug for valuable discussions. Willi Semmler would also like to thank IMF researchers for stimulating discussions while he was a visiting scholar at the IMF.

and private capital, essential infrastructure, as well as regional damages and life losses are occurring.

As to the economic strand of literature that has studied the likelihood of countries and regions to fall into poverty traps recently, the new growth theory[1] has redirected our attention to important mechanisms that can generate multiple equilibria, among them an attractor which has been called a poverty trap. Those could arise from externalities and increasing returns to scale, constraints in the financial and credit markets, as well as from population movements after disaster shocks that, in the long run, can give rise to diverse per capita incomes across regions and countries. Such mechanisms may be able to explain the forces that bring about a twin-peak distribution of per capita income in the long run, namely the convergence of the size distribution to countries and regions of small per capita income and countries with large per capita income, predicting some twin peak distribution.[2]

Financial studies have focused on large negative shocks that can result from large income or wealth shocks and contractions, with effects on credit markets, income flow, consumption and investment. In particular financial melt-downs and the destruction of capital as well as jumps in risk premia after rare large economic and financial crises are investigated in great detail. For example, Rietz (1988) studies rare market crashes and their effect on equity risk premia. Barro (2006) uses as disaster measure the decline of GDP growth, while Barro and Ursua (2008) and Gabaix (2011) investigate the decline of consumption spending due to large financial and economic disasters.

Climate change and weather extremes studies explore disaster effects from large scale floods, storms, landslides, heat waves and droughts, and forest fires. This literature also stresses nonlinearitites and tipping points, leading the a phase shift, and long period lock-ins. This work goes back to extreme event studies initiated by Gumbel (1938, 1958). Recent important contributions are the one by Burke et al. (2010) and

[1]One approach of the new growth theory views persistent economic growth arising from learning by doing, externalities in investment and increasing returns to scale. This idea had been formalized by Arrow (1962) and rediscovered by Romer (1990), who argues that externalities – arising from learning by doing and knowledge spillover – positively affect the productivity of labor and thus the aggregate level of income of an economy. Lucas (1988), whose model goes back to Uzawa (1965), stresses education and the creation of human capital, Romer (1990) focus on the creation of new technological knowledge as important sources of economic growth. And others emphasize productive public capital and investment in public infrastructure. For a more extensive survey, see Greiner, Semmler and Gong (2005).

[2]An early theoretical study of this problem can be found in Skiba (1978). Further theoretical modeling is in Azariadis and Drazen (1990) Aziaridis (2001) and Aziaridis and Stachurski (2004). For recent empirical studies see, for example, Quah (1996) and Kremer, Onatski and Stock (2001). A more recent study on this issue is Semmler and Ofori (2007) where also empirical evidence is provided.

Burke et al (2011), Hochrainer (2009, 2014), Independent Evaluation of the ADB (2015), Yumashev et al (2019), and Mittnik et al. (2019). Much research work published in the IPCC assessment reports since 1988 have elaborated on those issues.

A further strand of literature is the insurance work on this topic. Here the question is pursued whether and for which type of shocks, insurance against large random shocks can aid to reduce the risk of large capital and income losses leading to dynamics to fall below the poverty trap. An overview of those models is given in Paulsen (2008) and in Kovacevic and Pflug (2011). In the latter work a critical level of capital is introduced that can dampen the losses. Above the critical level, which depends on the fraction covered by insurance, the expansion of capital is feasible after a disaster shock. On the other hand, close to a critical point – at some cliff – insurance might be too expensive, because it disturbs the deterministic growth.

Our paper is related to the above literature. Our deterministic dynamics also has three equilibria: the outer two are stable while the middle one is unstable. The deterministic dynamics is driven by optimizing behavior, similar as in dynamic growth models, as in Semmler and Ofori (2007). As many recent growth models do, we start with a capital accumulation model with a mechanism that gives rise to multiple equilibria. It represents a basic model of the dynamic decision problem of countries where the capital stock is the state variable and investment is the decision variable. We explore mechanisms that may lead to thresholds and the separation of domain of attractions, predicting a twin-peak distribution of per capita income in the long run. We show that only countries that have passed certain thresholds may enjoy a rise of per capita income.[3]

The deterministic dynamics is overlayed by random dynamics, modeling the waiting times between catastrophic events. In this setup the possibility of insurance is analyzed. While Kovacevic and Pflug (2011) consider only the possibility of a fixed retention rate (respectively a fixed proportion of insured capital) we allow a change of the retention rate after each catastrophic event. The search for an optimal process of retention rates gives rise to a bi-level decision problem. In our context of finite upper equilibria of the deterministic dynamics it turns out that it is not meaningful to concentrate on the trapping probability of falling below the cliff (which was the approach in Kovacevic and Pflug (2011), where the upper equilibrium was infinite). It turns out that this probability is always one - that is in the long run ruin happens for

[3]The working of the above mechanisms are then empirically explored by applying a kernel estimator and Markov transition matrices to an empirical data set of per capita income across countries, see Semmler and Ofori (2007).

sure. Therefore we aim at maximizing the expected discounted capital after the next jump and develop a numerical algorithm in order to analyze the optimal retention rate depending on the starting capital.

The remainder of the paper is organized as follows. Section 2 presents a deterministic model with stable outer equilibria and an unstable middle equilibrium. Moreover, the economic mechanisms that make such thresholds plausible are discussed. Section 3 introduces the stochastic process of catastrophic events, stylized by a Poisson process for the number of events and a beta distribution for the proportion of destroyed capital. Moreover insurance is introduced in this part. In section 4 we analyze the long term trapping probability. Section 5 describes the numerical procedure used to calculate optimal decisions in the expected capital framework. Moreover, we present numerical results and implications for a stylized example. Section 6 concludes the paper. In the appendix we sketch the numerical solution procedure.

2. The deterministic dynamic model

The basic economic mechanisms to explain poverty traps frequently refer to technological traps. The idea of a technological trap is based on the work by Rosenstein-Rodan (1943, 1961), Singer (1949), Nurske (1953) and others. The starting-point is a modified production function that has both increasing and decreasing returns to scale. Increasing returns can only be realized if a country is capable to build up a capital stock that is above a certain threshold. If this threshold is passed, and sufficient externalities are generated, the production function exhibits increasing returns. Countries converge to a higher steady state as compared to countries that have fallen short of the threshold. With reference to the technological trap the so called "Big Push Theory" proceeds from the idea that industrial countries had in their past a massive capital inflow and therefore can converge to a steady state with a high income level. In contrast less developed countries have a shortage of such massive capital inflow and accordingly stagnate at a low income level.

A related explanation is given by Myrdal (1957) who points out that a tendency towards automatic stabilization in social systems does not exist and that any process which causes an increase or decrease of interdependent economic factors including income, demand, investment and production will lead to a circular interdependence. Thus this would lead to a cumulative dynamic development that strengthens the effects of up - or downward movement. On this ground poor countries are in a *vicious circle*, becoming poorer. This is in contrast to rich countries who will profit by a positive feedback effect, the so-called "Backwash Effects" arising from capital movement and migration to get richer.[4]

[4]Scitovsky's work in the 1950s is another example predicting poverty traps, thresholds and take-offs, see Scitovsky (1954).

2.1. **The Deterministic Model.** As previously mentioned the idea of externalities and increasing returns to scale has been extensively employed in growth theory recently. It is shown that a variety of positive externalities arising from scale economies, learning by using, increasing returns to information and skills are set in motion if a country enjoys, for example, by historical accident, a "big push" and take-off advantages.

Our proposed variant of a model of dynamic investment decisions of countries builds on locally increasing returns to scale arising from externalities. Locally increasing returns due to positive externalities may be approximated by a convex-concave production function as proposed by Skiba (1978) to illustrate those effects.

To present this idea of a convex-concave production function resulting from externalities and locally increasing returns to scale we use a model similar to Azariadis and Drazen (1990).[5] With capital stock denoted by k, we can write a production function such as

$$
\begin{aligned}
y(k(t)) &= ak(t)^{\alpha_k(t)} \\
\alpha_k(t) &= \begin{cases} \overline{\alpha}_k & \text{if } k(t) > \overline{k}(t) \\ \underline{\alpha}_k & \text{otherwise} \end{cases}
\end{aligned}
\tag{2.1}
$$

where the coefficients $\alpha_k(t)$, vary with the underlying state (k) and the quantity $\overline{k}(t)$ is the threshold for the capital.

We consider an optimal control problem

$$V(k) = \underset{\{u(t)\}}{Max} \int_0^T e^{-\rho t}((y(t)u(t))^{(1-\sigma)}/(1-\sigma))dt \tag{2.2}$$

$$\dot{k}(t) = y(t)(1-u(t)) - \delta k(t), \quad k(0) = k, \tag{2.3}$$

which describes the economy based on the optimal allocation of income between consumption and investment. The control is the consumed fraction of income $u(t)$.

Equation (2.2) represents the related value function and equation (2.3) the evolution of capital stock, whereby the first term $y(t)(1 - u(t))$ is gross investment and the second term $\delta k(t)$ is the depreciation of capital which will be augmented by the loss of capital due to the insurance premium. Finally, σ is the parameter of risk aversion of the economic agents which is in the literature assumed to be between 0.5

[5] See furthermore Durlauf and Quah (1999), Azariadis (2001) and Azariadis and Stachurski (2004).

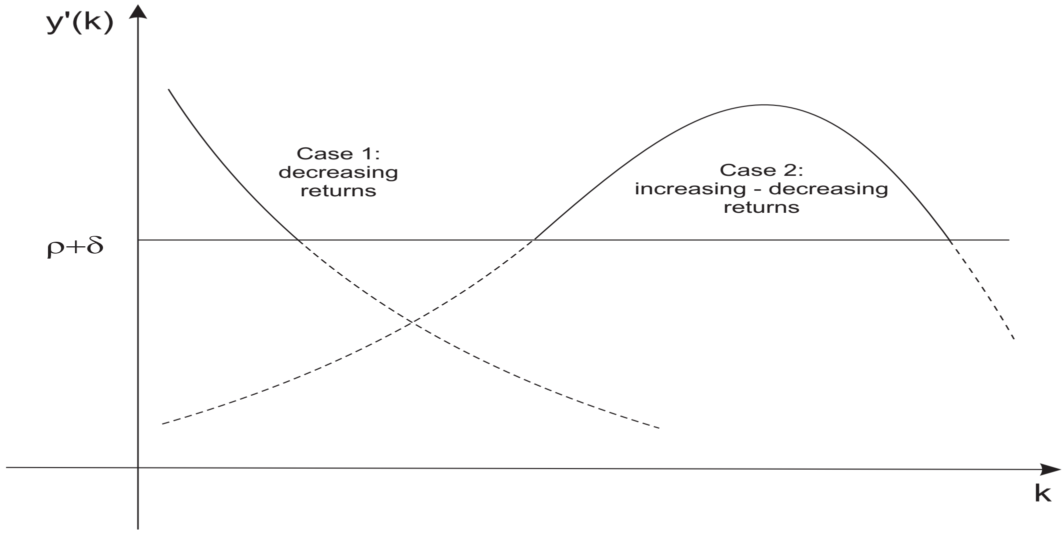

FIGURE 2.1. Increasing and decreasing returns

and 4. (We want to assume the risk aversion on the low side, so the dynamics are not much impacted by this parameter itself.[6]

One can show, using Dechert and Nishimura (1983) that if $\alpha_k < 1$ in equation (2.1) holds forever, the marginal product of capital, $y'(k)$ would approach the line given by the discount rate ρ plus capital depreciation, δ, if depreciation is allowed, from above or below, see case (1) in Figure 2.1.

On the other hand, presuming that the parameter α_k is state dependent and approximating the convex-concave production function by a smooth function one obtains the case 2 in Figure 2.1.

For locally increasing returns to scale, case 2, the return on capital $y'(k)$ will first approach $\rho+\delta$ from below, then move above this line and eventually decrease again. In the first case, the return on capital below $\rho + \delta$, because of externalities, too small a capital stock will generate a too low return for the economy so that the capital stock will shrink.

Thus the case 2 has three equilibria, one unstable equilibrium where the horizontal axis $\rho + \delta$ intersects with the case 2 curve, and the other two equilibria are somewhere above and below the $\rho + \delta$ line.

As Figure 2.1 demonstrates, increasing returns can be assumed to hold, as Greiner et al. (2005, ch. 3) show, only up to a certain level of the capital stock. A region of a concave production function may be dominant thereafter where $y'(k)$ might start falling again.

[6]Population growth could also be included, requiring a slight modification of the model.

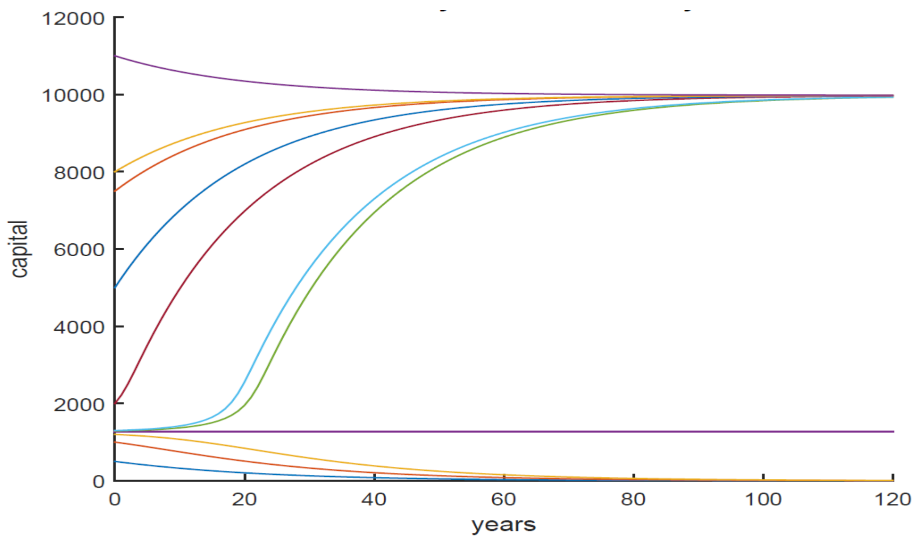

FIGURE 2.2. Deterministic Dynamics for several starting points

If we compute the investment strategy for a model variant with a convex-concave production function as suggested above, the convex-concave production function is for our numerical purpose specified as a logistic function of k

$$y(k) = \frac{a_0 \; exp(a_1 k)}{exp(a_1 k) + a_2} - \frac{a_0}{1 + a_2} \tag{2.4}$$

This convex-concave production function specifies the production function $y(k)$ in equation (2.1). We refer to the model (2.2)-(2.4) as the "deterministic dynamics $DP(\delta)$" in order to emphasize the dependence on the depreciation rate, which later will be augmented by the insurance premium per unit of capital.

For the numerical computation of the solution of model (2.1)-(2.4) we employed the NMPC procedure of Gruene et al. (2015), which is sketched in the appendix of this paper.

In a stylized numerical example, we use parameter values $a_0 = 2500$, $a_1 = 0.0034$, $a_2 = 500$. for the production function and set $\delta = 0.05$ and $\sigma = 0.5$. Figure 2.2 shows that there are indeed three equilibria, the middle one is unstable and the lower and upper equilibria are stable ones.

3. Stochastic Shocks and Insurance

We assume now that the capital normally evolves according to the dynamics described in the previous section, but is reduced at random points in time by shocks of random size. Such catastrophic events may imply substantial damages to the economy, but can be considered as

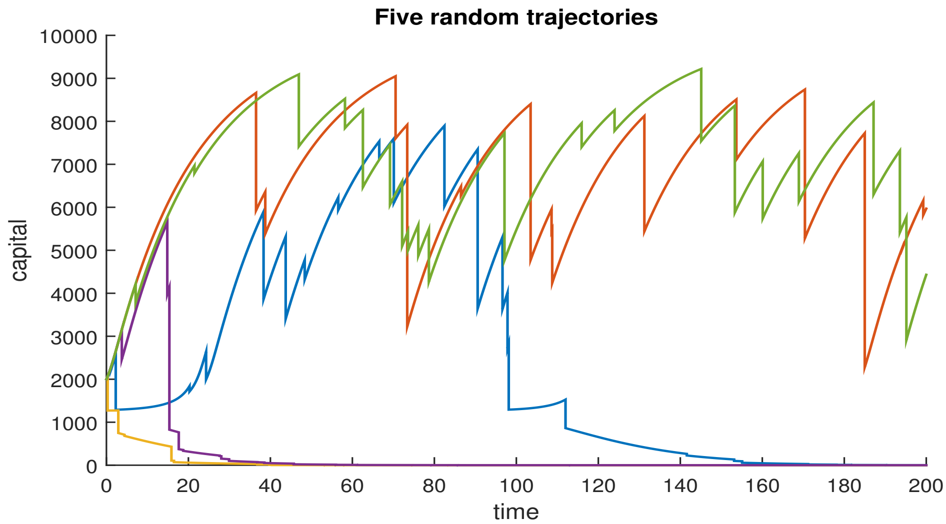

FIGURE 3.1. Stochastic dynamics without insurance ($\eta = 1$)

rare events. Because of this rareness, we assume that all economic decisions by economic agents are made without accounting for the possibility of catastrophic events by the economic decision makers.

Figure 3 provides an example of such dynamics, based on the numerical setup. A detailed description will be given below. Yet, we can already observe that even shocks far away from the poverty trap (the middle unstable equilibrium) can lead to dynamics ending up in the poverty trap.

Next a second layer of decisions is added at this point: Governments e.g. might be well aware of the possibility of catastrophic losses. One way to deal with such events is the introduction of a mandatory insurance scheme. Hereby the economic agents would pay insurance premium (maybe as an additional tax) and receive insurance benefits in form of a reduction of the capital loss, in case of a catastrophic event. We use the terminus "insurance" here, although the "insurance premium" might be also implemented as some kind of earmarked tax and the "insurance benefits" might be just payments of the government in case of catastrophic events.

The insurance premium per unit of capital depends on the retention rate η, i.e. the proportion of capital that is not insured[7] and is added to the deterministic model in form of additional capital depreciation. This slows down capital growth and shifts the equilibrium points of the deterministic dynamics, see Figure 3.2. On the other hand, some capital is recovered after each event. In view of this tradeoff we ask the

[7]If the capital is not insured at all, the retention rate equals one, if the whole capital is insured, the retention rate equals zero.

question, how much insurance - or which retention rate - is “optimal” depending on the start capital.

Dependence of η on capital size is critical, because insensitive introduction of insurance premium ($\eta > 0$ fixed) has the unwanted effect that small start capital just above the unstable middle equilibrium (Skiba point) without insurance would be below the Skiba point after introduction of insurance. This leads to a deterministically shrinking capital exposed to random shocks, hence accelerated extinction. This effect was observed for simpler dynamics in Kovacevic and Pflug (2011).

The overall decision problem has the form of a bi-level optimization problem. The upper level decision (insured fraction of capital) is taken by the government, whereas the lower level decisions (the solution of the control problem) are taken by the economic agents in view of the prescribed amount of premium payments. While bi-level problems are generally difficult to solve, in the present context the task is facilitated by the fact that the control problem has unique solutions.

In the present paper we take the size of the insurance premium (per insured unit of capital) as dependent on the loss distribution but exogenously given for the agents. Questions of exact financing of the insurance system and of adequate premium amounts are left to future research. We here concentrate on proportional insurance, which is the simplest insurance scheme.

3.1. **Catastrophic Events, Insurance and the Modified Deterministic Dynamics.** The catastrophic events happen at random points in time T_i . We assume that the related waiting times $\tau_i = T_i - T_{i-1}$ are i.i.d. according to an Exponential distribution with (constant) parameter $\lambda > 0$. This means that the probability density of each τ_i is $g_\tau(t) = \lambda e^{-\lambda t}$ and the expected waiting time is given by $\frac{1}{\lambda}$. Therefore, the number of events up to time t follows a homogeneous Poisson process.

In the stochastic model, we denote capital by $K(\cdot)$. When a catastrophic event i happens, the instantaneous capital $K(T_i^-)$ before the jump is reduced by a random fraction Z_i such that the Z_i are i.i.d. distributed according to a cumulative distribution function $G_Z(\cdot)$ and probability density $g_Z(\cdot)$, hence

$$K(T_i) = (1 - Z_i)K(T_i^-), \tag{3.1}$$

where $K(T_i^-)$ denotes the capital immediately before the event.

If $G_Z(\cdot)$ is differentiable, we denote the related probability density by $g_Z(\cdot)$. The support of the loss distribution is the interval $[0, 1]$. In addition it is assumed that the Z_i are jointly independent of the waiting times τ_i.

With insurance, (3.1) is adjusted in the following way

$$K(T_i) = (1 - \eta Z_i)K(T_i^-), \tag{3.2}$$

where the retention rate $0 \leq \eta \leq 1$ is the proportion of damage beared by the insured entity.

Using the expectation premium calculation principle (see e.g. Mikosch (2009)), the insurance premium per capital unit, $c(\eta)$, then can be expressed as:

$$c(\eta) = \lambda(1+\gamma)(1-\eta)\mathbb{E}[Z_1]. \tag{3.3}$$

Here $\gamma > 0$ is some risk loading parameter. Recall that the cession rate $1-\eta$ is the proportion of damage beared by the insurer.

The distribution parameter λ, and the expectation $\mathbb{E}[Z_1]$ are assumed to be known throughout the paper. In real world applications they have to be estimated from data. The risk adjustment parameter γ is also given exogenously in the present paper. Basically it is has to be chosen e.g. by an insurance company taking into account the riskiness of the loss distribution. Finally the retention rate η has to be decided by the government.

We assume that a new value of η can be chosen after each catastrophic event, dependent on the remaining capital $K(T_i)$.This justifies the notation $\eta_i = \eta(K(T_i))$.

After each jump, the capital starts with a value of $K(T_i)$. We assume $T_0 = 0$ and so

$$K(0) = k \tag{3.4}$$

is the start value of the whole process. Between two events $i-1$ and i, capital $k(t)$ develops according to the solution of the deterministic control problem (2.2)-(2.3), but the original depreciation rate δ is replaced by $\delta + c(\eta)$ - which we denote by $DP(\delta + c(\eta))$. In consequence we may write

$$K(t) = F(t - T_{i-1}, K(T_{i-1}); \eta) \text{ if } T_{i-1} < t < T_i, \tag{3.5}$$

where F denotes the optimal state $k(\cdot)$ for the deterministic problem $DP(\delta + c(\eta))$ with $k(0) = K(T_{i-1})$ and call the resulting control problem the modified deterministic problem $DP(\delta + c(\eta))$. In order to shorten notation, we may write $F(t - T_i, X_{T_i})$, when the dependence on η is not in the focus.

Altogether, $K(t)$ is a piecewise deterministic Markov process. The stochastic evolution of capital is described by the distribution assumptions on waiting times and proportional jump size, the premium principle (3.3) together with the resulting modification of the deterministic dynamics (2.2)-(2.3) and the defining equations(3.4), (3.2) and (3.5).

In the modified problem all the equilibria depend on the retention rate. With $\widetilde{k}^\eta$ we denote the unstable equilibrium of the dynamics as described above and $\hat{k}^\eta$ denotes the stable upper stationary equilibria of the modified deterministic system. It turns out that the upper equilibrium increases with the retention rate, whereas the Skiba point decreases with increasing retention rate, see Figure (3.2).

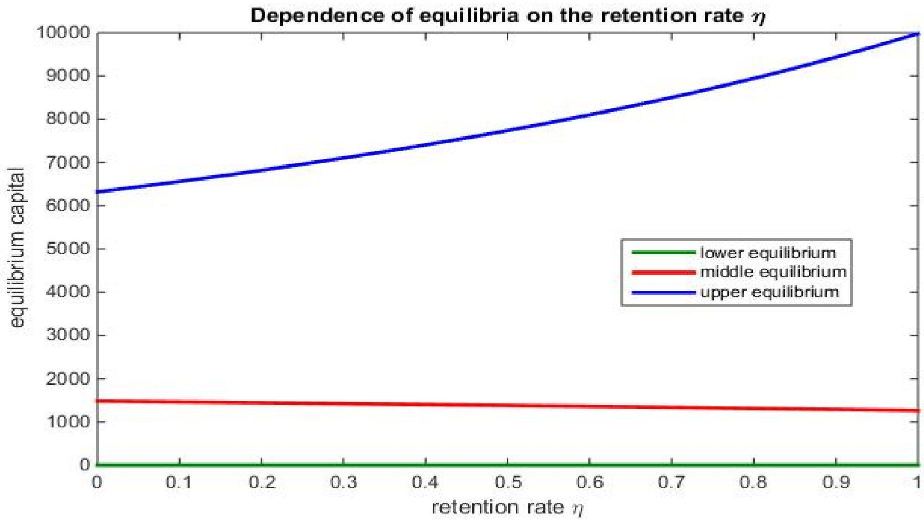

FIGURE 3.2. Dependence of equilibrium values on the retention rate η

The lowest Skiba point $\tilde{k} = \tilde{k}^1$ is the boundary between amounts of capital for which growth is possible with appropriately chosen η, and the region where capital shrinks for any η without a chance for recovery. We also will consider the largest upper equilibrium $\hat{k} = \hat{k}^1$. In principle the upper equilibrium of some dynamics might not exist, i.e. $\hat{k}^\eta = +\infty$. This was e.g. the case for the dynamics used in Kovacevic and Pflug (2011). However the deterministic problem $DP(\delta + c(\eta))$ definitely leads to finite upper equilibria.

For the deterministic part of our stylized example we set the parameter of the exponential distribution (waiting times) to $\lambda = 0.1$. The parameters of the Beta distribution (proportional loss given a catastrophic event) are $\alpha = 1.92$ and $\beta = 3.39$. Moreover we use $\gamma = 0.05$ for the risk adjustment parameter. These values were already used for producing Figure 3.1.

3.2. **The Remaining Capital and Its Transition Distribution.** In order to simplify notation we introduce the random variables $V_i = 1 - \eta Z_i$ (which are also i.i.d.), taking values in $[0, 1]$. They have cumulative distribution function

$$H(v;\eta) = 1 - G_Z\left(\frac{1-v}{\eta}\right)$$

and probability density

$$h(v;\eta) = \frac{1}{\eta} g_Z\left(\frac{1-v}{\eta}\right).$$

V_i models the remaining fraction of capital after event i occurred and the insurer already repaid the insured sum. Equation (3.1) can be rewritten as

$$K(T_i) = V_i K(T_i^-).$$

In order to analyze the remaining capital after the jumps, we may consider the discrete time process

$$K_i = K(T_i),$$

by sampling immediately after the occurrence times of catastrophic events. This is also a Markov process and given the above specification it is possible to characterize the related transition density $p(k_1, k_0; \eta)$, i.e. the conditional probability density for reaching capital level k_1 after the next catastrophic event, when the process starts with capital level k_0 after the last event. If a remaining capital $K_i = k_0$ is observed, it is possible to neglect all previous observations when calculating the density of capital $K_{i+1} = k_1$ (after the next jump). Because η_i is reconsidered after each jump, it will never be chosen such that the new Skiba point would be below the start capital K_i, because this would lead to a decreasing deterministic dynamics. Therefore it suffices to consider the case $k_0 > \widetilde{k}^\eta$.

When $V_i = v$ is given, then the capital after the next loss fulfills

$$K_{i+1} = k_1 = vF(\tau, k_0; \eta),$$

which means that the waiting time τ until the next event can be calculated as a function of k_0, k_1, v by

$$\tau = F^{-1}\left(\frac{k_1}{v}, k_0\right),$$

where $F^{-1}(\cdot, k_0; \eta)$ is the inverse function of $F(\cdot, y_0\eta)$ with respect to the first argument. Inversion is possible because $F(\cdot, y_0)$ is strictly increasing (recall the assumption $k_0 > \widetilde{k}^\eta$). Note also that we have $F : [0, \infty) \to [k_0, \hat{k}^\eta)$ and $F^{-1} : [k_0, \hat{k}^\eta) \to [0, \infty)$. Calculating the derivative leads to

$$\frac{\partial \tau}{\partial y_1} = \frac{1}{vF_1(\tau, k_0)} > 0.$$

3.3. **Transition densities.** Knowing the density of τ, it is possible to calculate the conditional density of K_{i+1} given $K_i = k_0$ and $V = v$ by using density transformation. Taking expectation with respect to the c.d.f. $H(\cdot; \eta)$ of the random variable V leads to the transition density.

$$p(k_1, k_0; \eta) = \lambda \int_{\frac{k_1}{\hat{k}^\eta}}^1 \frac{e^{-\lambda F^{-1}\left(\frac{k_1}{v}, k_0.\eta\right)}}{v\, F_1\left(F^{-1}\left(\frac{k_1}{v}, k_0; \eta\right), k_0; \eta\right)} \frac{1}{\eta} g_Z\left(\frac{1-v}{\eta}\right) \mathrm{d}v \tag{3.6}$$

The integration boundary follows from the domain of F^{-1} together with the fact that $0 \leq V \leq 1$.

From this result one can also derive the related conditional distribution function

$$(3.7)\qquad \begin{aligned} P(k_1, k_0; \eta) &= \mathbb{P}\left(K_i = k_1 | K_{i-1} = k_0\right) \\ &= \int_{\frac{k_1}{\hat{x}\eta}}^{1} \left[1 - e^{-\lambda F^{-1}\left(\frac{k_1}{v}, k_0.\eta\right)}\right] \frac{1}{\eta} g_Z\left(\frac{1-v}{\eta}\right) \mathrm{d}v, \end{aligned}$$

which may be used in order to simulate realizations of the process K_i.

4. Aiming at the Trapping Probability

Based on the optimal dynamics of $DP(\delta + c(\eta))$ and using waiting times and proportional losses as described above, we aim at "optimal" retention rates η. One possible objective consists in minimizing the probability that the capital reaches or falls below the trapping point $\tilde{x}$ - the smallest possible Skiba point (which results from setting $\eta = 1$). If this happens, then there is no chance for escaping from the lower stable equilibrium in the long run, because already the deterministic dynamics leads to decreasing capital, and all jumps decrease the capital further. We refer to this probability as the trapping probability in the following. Such an approach was suggested in Kovacevic and Pflug (2011) for a simpler dynamics with a trapping point, without using an underlying optimal control problem and without the possibility to change the retention rate η after a catastrophic event.

In such a setup, one searches for a (point-wise) minimal trapping probability $Q(k)$, defined on $(\tilde{k}, \hat{k}]$, where $\tilde{k} = \tilde{k}^1$ is the smallest Skiba point and $\hat{k} = \hat{k}^1$- the largest upper equilibrium. This is the probability of eventually reaching $\tilde{k}$ or any point below at some point in time, after starting with a capital of $k > \tilde{x}$. Again it is assumed that fraction η of non insured capital can be readjusted after each catastrophic event, hence $\eta = \eta(k)$ can be considered as a function of the starting capital. The function Q fulfills the functional equation

$$(4.1)\qquad Q(k) = \min_{\eta} \left[\int_0^{\tilde{k}} p(y, k; \eta)\, dy + \int_{\tilde{k}}^{\hat{k}^{\eta}} Q(y) p(y, k; \eta)\, dy\right]$$

and the function defined by the argmin, $\eta(k)$, describes the optimal fraction of uninsured capital for each starting capital x. Basically the trapping probability equals the probability of falling below the poverty line immediately after the next jump plus the expectation (with respect to the transition densities) of the trapping probabilities for capital values above the poverty line. We assume that $Q(x)$ is a bounded function $0 \leq Q(\cdot) \leq 1$ and denote the set of such functions by $\mathcal{B}$. Because of these bounds $Q(y)$ is integrable w.r.t. any probability density $p(y, k; \eta)$, hence the right hand side of (4.1) is well defined. This integrability property also ensures a bounded minimum. With T we denote the operator, defined by the right hand side, which is a mapping

$\mathcal{B} \to \mathcal{B}$, because p is a probability density for all possible values of η. Any solution of (4.1) is a fixed point of T.

It should be noted that $Q(k) \equiv 1$ is always a (trivial) solution of (4.1), which shows the existence of a fixed point. However, basically we seek for a nontrivial solution that is smaller than one for at least some amounts of capital. Unfortunately classical contraction arguments (e.g. applications of the Banach fixed point theorems) can not be applied, because they lead to unique fixed points.

However let $(\mathcal{B}, \leq)$ denote the vector space $\mathcal{B}$ together with the point-wise partial order, i.e. $Q_1 \leq Q_2$ when $Q_1(k) \leq Q_2(k)$ for all $x \in (\tilde{k}, \hat{k}]$. Then we can show the following:

Proposition 1. *The operator T has a smallest and a largest fixed point, Q_* and Q^*in $(\mathcal{B}, \leq)$, which can be obtained by*

$$Q_* = \sup \{Q \in \mathcal{B} : TQ \geq Q\} \tag{4.2}$$

$$Q^* = \inf \{Q \in \mathcal{B} : TQ \leq Q\} . \tag{4.3}$$

Here the infimum has to be understood in the point-wise sense, induced by the point-wise order $\leq$.

Proof. The operator T is monotone: assume $Q_1 \leq Q_2$ and define

$$\eta_2(k) = \arg\min_{\eta} \left[\int_0^{\tilde{k}} p(y, k; \eta)\, dy + \int_{\tilde{k}}^{\hat{k}^{\eta}} Q_2(y) p(x, k; \eta)\, dy \right],$$

then we have for any x

$$\begin{aligned}
(TQ_1)(k) &= \min_{\eta} \left[\int_0^{\tilde{k}} p(y, k; \eta)\, dy + \int_{\tilde{k}}^{\hat{k}^{\eta}} Q_1(y) p(y, k; \eta)\, dy \right] \\
&\leq \int_0^{\tilde{k}} p(y, k; \eta_2(x))\, dy + \int_{\tilde{k}}^{\hat{k}^{\eta_2(x)}} Q_1(y) p(y, k; \eta_2(x))\, dy \\
&\leq \int_0^{\tilde{k}} p(y, k; \eta_2(x))\, dy + \int_{\tilde{k}}^{\hat{k}^{\eta_2(x)}} Q_2(y) p(x, y; \eta_2(x))\, dy \\
&= \min_{\eta} \left[\int_0^{\tilde{k}} p(y, k; \eta)\, dy + \int_{\tilde{k}}^{\hat{k}^{\eta}} Q_2(y) p(x, k; \eta)\, dy \right] = (TQ_2)(k).
\end{aligned}$$

Now $(\mathcal{B}, \leq)$ is a complete lattice with smallest element $Q(x) \equiv 0$ and largest element $Q(x) \equiv 1$ and we can apply the Knaster-Tarski theorem, see Tarski (1955), in order to show the existence of a largest and a smallest fixed point together with properties (4.2)-(4.3). The smallest fixed point is found by starting with $Q^0(x) \equiv 0$ and applying the operator T until some stopping criterion is fulfilled. □

Based on the proposition, it is clear that the smallest trapping function Q is given by the smallest fixed point of (4.2) and can be obtained by (4.2). It is also a simple fact that the constant function $Q(x) = 1$

is in $\mathcal{B}$ and fulfills the functional equation. So this constant function is the largest element of $\mathcal{B}$ and hence also the largest fixed point of the operator T.

While Proposition (1) may lead to useful algorithms in case of infinite upper equilibria $\hat{k} = +\infty$, unfortunately it is not applicable when $\hat{x}$ is finite as for our optimal dynamics F. It can even be shown that in this case $Q(k) \equiv 1$ is the only solution (and fixed point) of (4.1).

Proposition 2. *If $\hat{k} < +\infty$ and the support of the random variable $\hat{k}V$ contains the value $\tilde{k}$ then $Q(k) = 1$ for any $k \in (\tilde{k}, \hat{k}]$.*

Proof. Higher starting capital leads to a smaller ruin probability. Therefore we have

$$\begin{aligned}
Q(k) &\geq Q(\hat{k}) \\
&= \min_{\eta} \left[\int_0^{\tilde{k}} p(y, \hat{k}; \eta)\, dy + \int_{\tilde{k}}^{\hat{k}\eta} Q(y) p(y, \hat{k}; \eta)\, dy \right] \\
&= \int_0^{\tilde{k}} p(y, \hat{k}; \eta^*)\, dy + \int_{\tilde{k}}^{\hat{k}^*} Q(y) p(y, \hat{k}; \eta^*)\, dy \\
&\geq \int_0^{\tilde{k}} p(y, \hat{k}; \eta^*)\, dy + Q(\hat{k}) \int_{\tilde{k}}^{\hat{k}^*} p(y, \hat{k}; \eta^*)\, dy.
\end{aligned}$$

Here η^* is the optimal solution of the second line.

This leads to

$$Q(\hat{k}) \geq P + (1 - P) Q(\hat{k})$$

or - after reordering and dividing by P - to

$$Q(\hat{k}) \geq 1$$

and hence

$$Q(k) \geq 1.$$

Here $P > 0$ is ensured by the condition on the support of $\hat{k}V$, which ensures that regions below or at $\tilde{k}$ can be reached with positive probability after starting at $\hat{k}$. Because Q is a probability it is possible to conclude

$$Q(k) = 1.$$

□

With a sure transition below the Skiba point in the long run, it is not meaningful to use the trapping probability as a measure of success in a meaningful way. An alternative would be to aim at expected first passage times (trapping times), i.e. the expectation of the random variable $\inf \left\{ t : K(t) \leq \tilde{k} \right\}$. Unfortunately first passage times are difficult to treat and in particular to optimize. For arbitrary processes it is already difficult to calculate first passage times using e.g. recursive algorithms of Laplace transforms (see Nyberg et al. 2016)). Applying such results in a situation where already the transition densities can be computed only in a numerically costly way and an important parameter of the process should change over time in an optimal way is not tractable with reasonable computational effort.

5. Optimizing the Expected Capital After Jumps

Observe now that starting with larger capital always must be better than starting with lower capital if decisions on the retention rate can be taken after catastrophic jumps. This is true whichever objective should be optimized (as long as larger capital is counted as better than smaller capital).

Fact 3. *Consider two processes K_i and K_i' as described in section* (3), *and controlled by processes with retention rates η_i, η_i' chosen by a decision maker. If $K_0 \geq K_0'$ (i.e. process K starts at the higher capital level), then setting $\eta_i = \eta_i'$ for all $i \in \mathbb{N}_0$ implies*

$$K_i \geq K_i' \text{ for all } i \in \mathbb{N}_0.$$

Proof. Assume that after jump j the capital of the first process is not smaller than the capital of the second process, i.e. $K_j \geq K_j'$. Then

$$F(t - T_j, K_j; \eta') \geq F(t - T_j, K_j'; \eta')$$

until the next jump, because of continuity: $F(t - T_j, K_j; \eta') < F(t - T_j, K_j'; \eta')$ can only happen if

$$F(t_1 - T_j, K_j; \eta') = F(t_1 - T_j, K_j'; \eta')$$

at some t_1, but this would imply

$$F(t - T_j, K_j; \eta') = F(t - t_1, F(t_1 - T_j, K_j; \eta'); \eta') = F(t - T_j, K_j'; \eta')$$

for $t \geq t_1$until the next jump. So

$$K(T_{j+1}^-) \geq K'(T_{j+1}^-)$$

and hence

$$K(T_{j+1}) = V_{j+1} K(T_{t+1}^-) \geq V_{j+1} K'(T_{j+1}^-) = K'(T_{j+1}).$$

Therefore, because $K_0 \geq K_0'$ by assumption, the assertion $K_i \geq K_i'$ follows by induction. □

In consequence, a choice of $\eta(K_i)$ maximizing K_{i+1} given K_j in each step would maximize the trapping time. This is not possible, because of the random jumps and event times. However, in the following we will analyze a strategy that maximizes the discounted expectation $\mathbb{E}\left[e^{-\rho\tau} K_{j+1}\right]$ when K_j is known, which is motivated by discussed fact. The expectation can be replaced by other relevant acceptability measures in order to take into account the risk dimension of the problem in a better way. Moreover, it is possible to extend this approach to a fully dynamic decision problem. We leave such extensions for future research and stick to the myopic formulation in the present work.

The decision problem can be formulated as the bi-level problem

$$\begin{aligned} \max_{\eta} \quad & \mathbb{E}_{\tau,V}\left[e^{-\rho\tau}V(\eta)F(\tau,K;\eta)\right] \\ & s.t. F(\cdot,\cdot;\eta) \text{ is the optimal dynamics of } DP(\delta+c(\eta)) \\ & \quad \text{for retention rate } \eta, \text{ see (3.5).} \end{aligned}$$

Because already $DP(\delta+c(\eta))$ can be solved only numerically, the same is true for the overall bi-level problem. Because solutions of $DP(\delta+c(\eta))$ are unique for given parametrization, in a numerical setup the task simplifies considerably: it is possible to calculate approximations of the function F already in advance, such that the bi-level problem reduces to an optimization problem without constraints.

5.1. **Interpolating the Function** F**.** If the function values of F are calculated on a grid $\mathcal{T}\times\mathcal{K}\times\mathcal{E}$, where $\mathcal{T}$ contains (finitely many) points in time, $\mathcal{K}$ contains values for the start capital and $\mathcal{E}$ contains possible values for η, this information can be used to interpolate the function F over the relevant domain - at least if $\mathcal{T},\mathcal{K}$ and $\mathcal{E}$ are chosen sufficiently fine. It should be ensured that $\mathcal{E}$ contains values in the interval $[0,1]$. Moreover K should include the value zero and sufficiently many possible capital values up to a level that contains the largest upper equilibrium $\hat{x}$.

It would be a very slow approach to calculate $F(t,K;\eta)$ for such a large grid by fully solving $DP(\delta+c(\eta))$ for all start values K and retention rates η over a large time range. Therefore we start with the derivative

$$f(K,\eta)=F_1(t,K;\eta). \tag{5.1}$$

Note that it does depend on t only via K (as the closed loop control u also depends on K). We calculate an estimate $\hat{f}$ by applying the NMPC procedure (see the appendix) with given η and start value K over a small time horizon $[0,\Delta]$ with step size δt and plug the resulting optimal control value u_0 (i.e. the optimal control at time $t=0$, calculated by the NMPC-procedure, see the appendix) into the right hand side of equation (2.3) for $t=0$. This leads to an estimate $\hat{f}(K,\eta)$ for all pairs $(K,\eta)\in\mathcal{K}\times\mathcal{E}$. In further calculations we use the function $\hat{f}_I(K,\eta)$, which takes values $\hat{f}(K,\eta)$ for $(K,\eta)\in\mathcal{K}\times\mathcal{E}$ and else interpolates by cubic splines (if (K,η) is at least in the range of $\mathcal{K}\times\mathcal{E}$). For our standard parametrization, the interpolating function $\hat{f}_I(K,\eta)$ is shown in Figure 5.1 Using the estimated derivatives then gives a very convenient way to calculate the equilibrium points of the deterministic dynamics for any relevant start value K. These can be found by searching for stationary points K^s, i.e. by finding the solutions of the

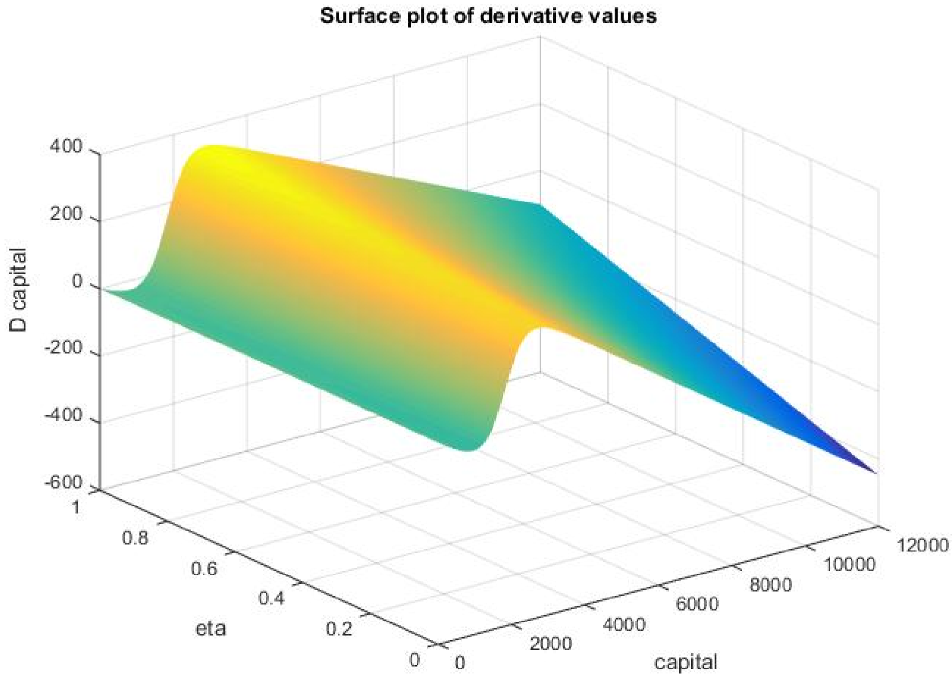

FIGURE 5.1. The interpolating function $\hat{f}_I(K, \eta)$ of derivatives for the standard parametrization

equations

$$\hat{f}_I(K^s, \eta) = 0.$$

separately for all $\eta \in \mathcal{E}$. By interpolation we find estimates $\tilde{k_I}^\eta, \hat{k_I}^\eta$ for the unstable middle and the stable upper equilibrium given the retention rate for any $\eta \in [0, 1]$. The results for our standard parametrization were already shown in Figure 3.2.

In a next step, the interpolated time derivatives $\hat{f}_i$ are used to reconstruct the paths of the deterministic dynamics for any $\eta \in [0, 1]$, i.e. the function F. Let now $\varepsilon > 0$ be a small real number. We define a function

$$\tilde{F}(t, \eta) = \begin{cases} \tilde{k}_I^\eta + \varepsilon & t = 0 \\ \tilde{F}(t_-, \eta) + \hat{f}_I(\tilde{F}(t_-, \eta), \eta)\,(t - t_-) & \text{for any other } t \in T \\ & \text{ and } t_- = \max\{d \in \mathcal{T} : d < t\} \end{cases}$$

on $\mathcal{T} \times \mathcal{E}$ and its interpolated version $\tilde{F}_I(t, \eta)$. Figure 5.2 shows the interpolated function $\tilde{F}_I(t, \eta)$ for our standard parametrization.

Finally we get an estimate $\hat{F}(t, K; \eta)$ for the function $F(t, K; \eta)$ which describes the optimal dynamics of the modified deterministic problem with start capital K and retention rate η. Because of the autonomous nature of the control problem, the trajectory for any start

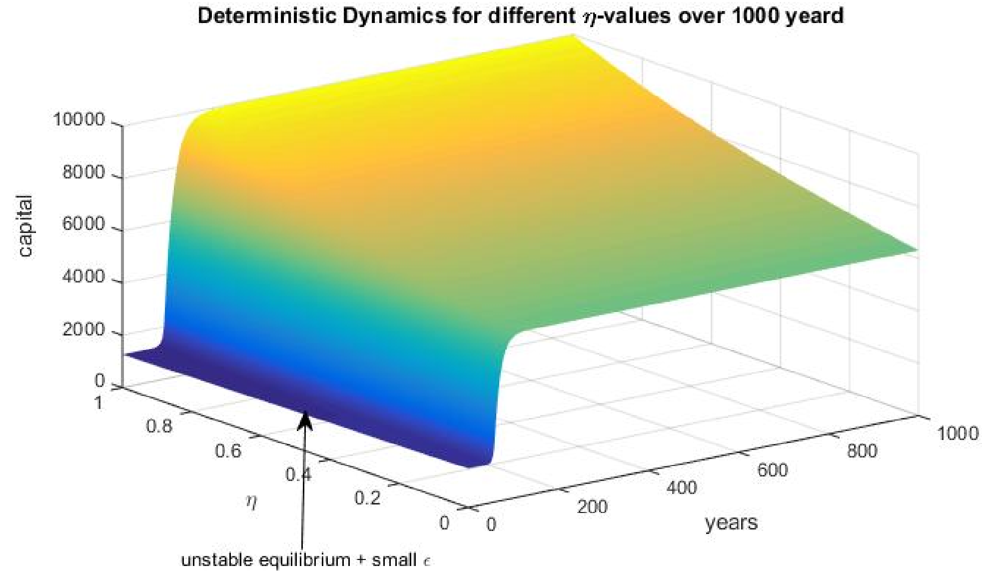

FIGURE 5.2. The interpolating function $\hat{F}_I(t, \eta)$ of deterministic trajectories with retention rate η and start capital $\tilde{k}_I^\eta$.

value $K \geq \tilde{k}_I^\eta + \varepsilon$ can be reconstructed from the trajectory with start value $\tilde{k}_I^\eta + \varepsilon$ using the relation

$$\hat{F}(t, K, \eta) = \hat{F}_I(\hat{F}_I^{-1}(K, \eta) + t, \eta),$$

where $\hat{F}_I^{-1}(\cdot, \eta)$ denotes the inverse function of F_I with respect to the first argument and with η fixed.

In this way it can be avoided to apply the full NMPC procedure for each possible starting capital over the full planning horizon, which is considerably faster. Due to some random test instances, the loss in accuracy compared to the full NMPC procedure turned out to be very minor. On the other hand, the substantial gain in speed makes it possible to use the deterministic control problem for simulation and optimization of expected capital values after the jumps, as described below.

5.2. **Simulation of capital values and optimization of the remaining capital.** If the aim is to maximize the expected value of remaining capital after the next catastrophic event, different approaches are possible to calculate the expectation. One may plug the interpolating functions $\hat{f}_I, \hat{F}_I$ into equation (3.6) for F_1, F and use the resulting interpolating estimate of the transition density p in order to calculate the relevant expectations dependent on η as integrals. However already calculation of the estimated densities affords integration, so the calculations become numerically very involved, especially when the final aim is to optimize over the resulting expectations.

Therefore in this study the expectations for any relevant retention rate η are calculated by simple Monte-Carlo simulation: the expectation for given start capital K_0 and retention rate η, i.e.

$$\mu(K_0,\eta)=\int_0^\infty\int_0^{\hat{x}^\eta}\lambda e^{-(\rho+\lambda)t}yp(y,K_0;\eta)\,dy\,dt,$$

is estimated by the mean

$$\hat{\mu}(K_0,\eta)=\frac{1}{n}\sum_{i=1}^{n}e^{-\rho\tau_i}V_i\hat{F}_I(\tau_i,K_0,\eta).$$

Here V_i,τ_i are (independent) pseudo-random sequences, obtained from the distribution functions of fractional losses and waiting times, H and G_τ. We used the simple inversion method, see Press et al. (2007, pp. 27).

For given start capital K_0 it is then possible to calculate $\hat{\mu}(K_0,\eta)$ for $\eta\in E$ and to use a spline-interpolated (in the second argument) version $\hat{\mu}_I(K_0,\eta)$ for finding the optimal retention rate

$$\hat{\eta}^*(K_0)=\arg\min_\eta\{\hat{\mu}_I(K_0,\eta):\ 0\le\eta\le 1\}\ \text{ for } K_0\in\mathcal{K}\ .$$

Assuming a singular optimizer in the interval $[0,1]$, golden section search (see e.g. Press (2007), Section 10.2) was used to find $\eta^*(K_0)$ in an efficient way.

Figure 5.3 demonstrates the approach and the resulting optimal strategy for our standard example. The upper part gives an overview for the whole range of start capital values, while the lower part analyzes the situation for smaller start capital, in particular start capital near the trapping point $\tilde{k}$. In both pictures, the colors indicate the approximate size of the expected present value $\hat{\mu}(K_0,\eta)$ of capital after the next catastrophic event. Lighter colors are related to higher values according to the color codes shown at the right margin. In addition the contour lines show points (K_0,η) with constant values $\hat{\mu}(K_0,\eta)=c$.

The violet regions have special interpretations. The violet region at the right side of the upper picture contains combinations of start capital K_0 and retention rate η where the stable upper equilibrium $\hat{k}^\eta$ lies below the start capital. This leads to a decreasing deterministic dynamics and therefore can not be optimal when compared with η chosen such that the upper equilibrium equals the starting capital. The violet region to the left (of both pictures) is related to points (K_0,η) where $K_0\le\tilde{k}^\eta$, i.e. the start capital is below the trapping point of the deterministic dynamics with retention rate η. Here again already the deterministic dynamics leads to a decrease in capital, and η cannot be optimal (e.g. compared with η chosen such that the start capital lies already below the resulting trapping point).

The red line shows (in both pictures) the optimal retention rate $\eta^*(K_0)$ for each relevant start capital. One can see (lower picture) that

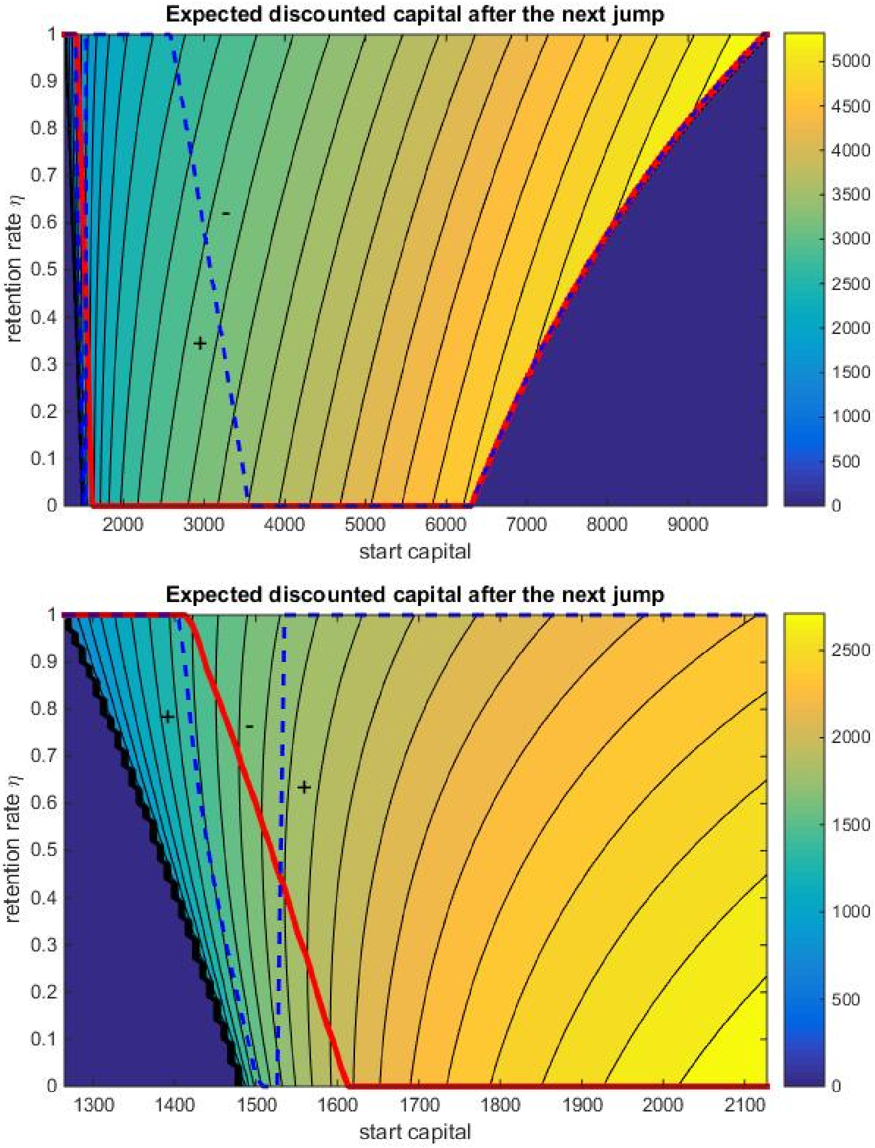

FIGURE 5.3. Contour plots of expected discounted capital (EPV), i.e. $\hat{\mu}(K_0, \eta)$ after the next catastrophic event dependent on start capital K_0 and retention rate η. Lighter regions are related to higher values of EPV. The red line indicates the optimal retention rate $\eta^*(K_0)$. The dashed blue line shows combinations of K_0 and η where the EPV equals K_0 and separates regions with $EPV > K_0$ (indicated by +) from regions with $EPV < K_0$ (indicated by $-$).
upper picture: full range of possible start capital, **lower picture**: start capital in the range [1264.7, 2127.5].

near the trapping point $\tilde{k}$, in a region where the contour lines are strict monotone decreasing, it is optimal to require no insurance at all ($\eta^* = 1$). The costs of insurance lead to a braked deterministic dynamics, which is especially dangerous near the unstable equilibria. Stated in a different way, near the lower boundary any insurance premium costs more in terms of growth than what the related claims payments could replace in average.

There is also a region, where the contour lines are not unequivocally curved in one direction. Here the optimal $\eta^*(K_0)$ decreases almost linearly from $\eta = 1$ to $\eta = 0$, i.e. from no insurance to full insurance. In the (largest) third regions (see the upper picture) the contour lines are curved to the right, hence here it is optimal to choose the retention rate as low as possible without starting below the upper equilibrium (and therefore inducing negative growth).

Finally, the dashed blue line separates (in both pictures) regions where the expected discounted capital $\hat{\mu}(K_0, \eta)$ after the next jump is below (indicated by +), respectively above (indicated by +) the start capital K_0. It can be seen that near the lower bound the strategy of denying any insurance leads to a positive effect. When the optimal retention rate decreases, we pass through a region where the optimal amount of insurance leads to a discounted expectation smaller than the start capital, which means that it is hard to pass through this region when repeating the selection of optimal retention rates after several successive events. In some sense this region separates the poor from the wealthy: With still decreasing optimal retention rate the blue line is passed again and the optimal strategy leads to the expectation of increased discounted capital. Finally, for very high capital values (upper picture) the optimal strategy again leads to the expectation of decreased discounted capital due to the effects of a finite stable upper equilibrium in the deterministic dynamics and the negative dynamics above this boundary.

6. Conclusions/Discussion

Given the recent rise in frequency of climate related disasters, severely affecting countries and regions, the issue of recovering lost asset by an insurance scheme has become an important issue. As we show there are mechanisms after disaster shocks that enhance the likelihood of falling into a poverty trap, even with insurance - if the retention rate is chosen too small. Though we start with a stylized deterministic dynamic model, with possibly generating multiple equilibria paths, the deterministic dynamics is then overlayed by random dynamics where catastrophic events happen at random points of time. The number of catastrophic events follows a homogeneous Poisson process and the proportional size of the disasters are modeled by a beta distribution. Our approach represents a bi-level decision model which is hard to compute analytically. Based on the NMPC procedure, we therefore apply a new algorithm that helps to compute numerical results. Even if a fraction of capital loss is insured and an optimal insurance premium, including possibly an appropriate risk loading, can be computed, falling into a poverty trap is still feasible. The expected discounted capital after the next catastrophic event, if a certain fraction of capital is insured, is computed in dependence of the (changing) initial size of capital. As also shown insurance against disaster shocks close to the cliff might not pay-off, thus other policies are needed in this case, see Mittnik et al. (2019). However, for larger start capital insurance is a valuable strategy for reducing the probability of falling off the cliff. This feature that the optimal insurance premium of insuring a certain fraction of assets may be also a helpful device to compute state dependent credit cost and to assess risk premia and creditworthiness of borrowers when a sequence of shocks at uncertain times and of uncertain size is expected.

Yet, further research is needed. In particular the underlying economic control problem can be enhanced. So far, only capital is taken into consideration. However, certain types of severe disasters may also have an impact on the work force, which per se cannot be insured. Insurance of the capital stock might be even more attractive in such a setup because it can partially compensate for the decreasing workforce and may mitigate the effects on production[8]. In order to develop the general algorithmic approach, in the present paper we have restricted the analysis to expected values, which neglects risk. Therefore the effect of optimizing risk sensitive functionals instead of the expected value has to be analyzed next. An important step forward will be the introduction of the funds generated by the premium payments into the deterministic model as an additional state variable. This allows to optimize the risk loading (which was assumed as given in the present paper) in addition to the insured fraction of capital. In such a manner it will be possible to explore risk loading drivers in its interaction with macroeconomic effects. Finally, despite the fact that one falls off the cliff for sure as shown, it is possible to analyze the time until falling below the cliff. Optimizing the acceptability of this first passage time will lead back from the myopic optimization to a dynamic approach for the upper level problem.

[8]The authors thank an anonymous referee for pointing out this possibility.

.

References

[1] Arrow, K. (1962). The Economic Implications of Learning by Doing. Review of Economic Studies, Vol. 29: 155-73.

[2] Azariadis, C. (2001). The Theory of Poverty Traps: What Have We Learned?, mimeo, UCLA.

[3] Azariadis, C. and Drazen (1990). Threshold Externalities in Economic Development, Quarterly Journal of Economics, vol. 105, 2: 501-526.

[4] Azariadis, C. and J. Stachurski (2004). Poverty Traps, forthcoming Handbook of Economic Growth, ed. by P. Aghion and S. Durlauf, Elsevier Publisher

[5] Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century, The Quarterly Journal of Economics, 121(3):823–866.

[6] Barro, R. J. and J. F. Ursua (2008). Macroeconomic crises since 1870, Working Paper 13940, National Bureau of Economic Research.

[7] Burke, E. J., R. H. J. Perry, S. J. Brown (2010). An extreme value analysis of UK drought and projections of change in the future, Journal of Hydrology, 388 (2010) 131–143

[8] Devroye, Luc (1986). Non-Uniform Random Variate Generation, SpringerLink. New York: Springer-Verlag, . http://www.eirene.de/Devroye.pdf.

[9] Burke, M. S. M. Hsiang and E. Miguel (2011). Global non-linear effect of temperature on economic production, Nature volume 527, pp. 235–239

[10] Gabaix, X. (2011). Disasterization: A simple way to fix the asset pricing properties of macroeconomic models, The American Economic Review, 101(3):406–409.

[11] Greiner, A., W. Semmler and G. Gong (2005). The Forces of Economic Growth: A Time Series Perspective, Princeton, Princeton University Press.

[12] Grüne, L. and W. Semmler (2004). Using Dynamic Programming with Adaptive Grid Scheme for Optimal Control Problems in Economics, Journal of Economic Dynamics and Control, vol. 28: 2427-2456.

[13] Gruene, L., W. Semmler, and M. Stieler (2015). Using Nonlinear Model Predictive Control for Solving Dynamic Decisions Problems in Economics, Journal of Economic Dynamic and Control, 60, 112–133.

[14] Gumbel, E. (1935). Les valeurs extremes des distributions statistiques, Annales de l'institut Henri Poincare, 5(2):115–158.

[15] Gumbel, E. (1958). Statistics of Extremes, Columbia University Press.

[16] Hochrainer, S., Mechler, R., and Pflug, G. (2009). Assessing current and future impacts of climate-related extreme events - The case of Bangladesh, Interim Report IR-09-030, International Institute for Applied Systems Analysis.

[17] Hochrainer-Stigler, S., Mechler, R., Pflug, G., and K. Williges (2014). Funding public adaptation to climate-related disasters. estimates for a global fund, Global Environmental Change, 25:87–96.

[18] Independent Evaluation of the Asian Development Bank (2015). Global Increase in Climate-Related Disasters

[19] Kovacevic, R. M. and G. C. Pflug (2011). Does insurance help to escape the poverty trap? - a ruin theoretic approach, Journal of Risk and Insurance, 78(4):1003–1028.

[20] Kremer, M., A. Onatski and J. Stock (2001). Searching for Prosperity, NBER Working Paper Series, Working Paper 8020.

[21] Mittnik, S., W. Semmler and A. Haider (2019), Climate Disaster Risks — Empirics and a Multi-Phase Dynamic Model", IMF working paper WP/19/145

[22] Myrdal, G. (1957). Economic Theory and Under-Developed Regions, London:Duckworth
[23] Lucas, R.E. (1988). The Mechanism of Economic Development, Journal of Monetary Economics, 22: 3-42.
[24] Mikosch, T.. (2009). Non-Life Insurance Mathematics - An Introduction with the Poisson Process, Second Edition, Berlin-Heidelberg, Springer Verlag
[25] Nurske, R. (1953). Problems of Capital Formation n Underdeveloped Countries, Oxford: Oxford University Press.
[26] Nyberg, M. et al. (2016). A simple method to calculate first-passage time densities with arbitrary initial conditions, New Journal of Physics 18, 063019.
[27] Press, W.H. et al. (2007). Numerical Recipes: The Art of Scientific Computing, 3rd edition, New York, Cambridge University Press.
[28] Quah, D.T. (1996). Twin-Peaks: Growth and Convergence in Models of Distribution Dynamics, The Economic Journal 106, S.1045-1055
[29] Rietz, T. A. (1988). The equity risk premium a solution, Journal of Monetary Economics, 22(1):117–131.
[30] Romer, P.M. (1990). Endogenous Technological Change, Journal of Political Economy, Vol. 98: S71-102.
[31] Rosenstein-Rodan,P. (1943). Problems of Industrialization of Eastern and Southeastern Europe, Economic Journal Vol.53, p.202-211
[32] Rosenstein-Rodan, P. (1961). Notes on the Theory of the Big Push, in H.S. Ellis and H.C. Wallich, editors, Economic Development in Latin America, New York: Macmillian
[33] Semmler, W. and Ofori, M. (2007). On poverty traps, thresholds and take-offs - Structural Change and Economic Dynamics, 18(1):1–26.
[34] Singer, H.W. (1949), Economic Progress in Underdevelped Countries, Social Research, vol. 15, no. 2
[35] Skiba, A.K. (1978). Optimal Growth with a Convex-Concave Production Function, Econometrica 46, no. 3: 527-539.
[36] Tarski, A. (1955). A lattice-theoretical fixpoint theorem and its applications, Pacific Journal of Mathematics 5, no 2, 285-309.
[37] Uzawa, H. (1965). Optimum Technical Change in an Aggregative Model of Economic Growth, International Economic Review, Vol. 6: 18-31
[38] Yumashev, D. Hope, K. Schaefer, K. Riemann-Campe, F. Iglesias-Suarez, E. Jafarov, E.J. Burke, P. J. Young, Y. Elshorbany, and G. Whiteman (2019). Climate policy implications of nonlinear decline of Arctic land permafrost and other cryosphere elements, Nature Communications.Appendix: Numerical Solution Procedure

Appendix: Numerical Solution Procedure (NMPC)

For the numerical solution of the deterministic model presented in section 2, and used further in the next sections, we do not apply here the dynamic programming (DP) approach as presented in (Grüne and Semmler 2004) and as used in the original paper of Semmler and Ofori (2007). Though DP method also can find the global solution to an optimal growth model with multiple equilibria by using a fine grid for the control as well state variables but its numerical effort typically grows exponentially with the dimension of the state variable. Thus, even for moderate state dimensions it may be impossible to compute a solution with reasonable accuracy.

Instead computing the solution at each grid point as DP does we here use a procedure that is easier to implement. We are using what is called nonlinear model predictive control (NMPC) as proposed in Grüne and Pannek (2012) and Grüne et al. (2015). Instead of computing the optimal solution and value function for all possible initial states, NMPC only computes single (approximate) optimal trajectories at a time. To describe the NMPC procedure we can write the optimal decision problem as:

$$\text{maximize} \quad \int_0^\infty e^{-\rho t}\ell(x(t),u(t))dt, \tag{6.1}$$

where $x(t)$ satisfies

$$\dot{x}(t) = g(x(t),u(t)), \quad x(0) = x_0 \tag{6.2}$$

By discretizing this problem in time, we obtain an approximate discrete time problem of the form

$$\text{maximize} \quad \sum_{i=0}^{\infty} \beta^i \ell(x_i,u_i), \tag{6.3}$$

where the maximization is now performed over a sequence u_i of control values and the sequence x_i that satisfies $x_{i+1} = \Phi(h, x_i, u_i)$. Hereby $h > 0$ is the discretization time step. For details and references where the error of this discretization is analyzed we refer to Grüne et al. (2015).

The procedure of NMPC consists in replacing the maximization of the infinite horizon functional (3) by the iterative maximization of finite horizon functionals

$$\text{maximize} \quad \sum_{k=0}^{N} \beta^i \ell(x_{k,i},u_{k,i}), \tag{6.4}$$

for a truncated finite horizon $N \in \mathbb{N}$ with $x_{k+1,i} = \Phi(h, x_{k,i}, u_{k,i})$. Hereby the index i indicates the number of iterations. Note that neither β nor ℓ nor Φ changes when passing from (6.3) to (6.4). The procedure works by moving ahead with a receding horizon.

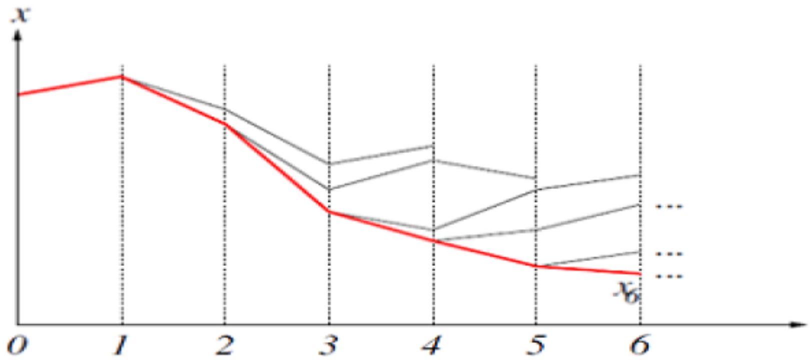

FIGURE 6.1. Receding horizon solution

The decision problem (6.4) is solved numerically by converting it into a static nonlinear program and solving it by efficient NLP solvers, see Grüne and Pannek (2012). In our simulations, we have used a modification of NMPC, as developed by Grüne and Pannek (2012), in their routine `nmpc.m,` available from `www.nmpc-book.com`, which uses MATLAB's `fmincon` NLP solver in order to solve the static optimization problem. Our modification employs a discounted variant of the NMPC MATLAB version, see Grüne et al. (2015).

Given an initial value x_0, an approximate solution of the system (6.1)-(6.2) can be obtained by iteratively solving (6.4) such that for i=1,2,3, that solves for the initial value initial value $x_{0,i} := x_i$ the resulting optimal control sequence by $u^*_{k,i}$, but uses only the first control $u_i := u^*_{0,i}$ and iterates forward the dynamics $x_{i+1} := \Phi(h, x_i, u_i)$ by employing only the first control. Thus, the algorithm yields a trajectory x_i, $i = 1, 2, 3, \ldots$ whose control sequence u_i consists of all the first elements $u^*_{0,i}$ of the optimal control sequences of the finite horizon problem (6.4). Under appropriate assumptions on the problem, it can be shown that the solution (x_i, u_i),which depends on the choice of N in (6.4), converges to the optimal solution of (6.1) as $N \to \infty$, see Grüne et al. (2015).

Figure A1 illustrates the working of the algorithm. The upper black line represents the solution at the step i=1 with the decision horizon N=4. This is iterated forward 6 times, thus we have i=1....6. The lower red line is the outer envelop of the piecewise solutions using the horizon N=4 multiple times, in our case 6 times. The figure A1 shows the solution for 6 iterations.

While the algorithm can be used to solve for optimal trajectories of x and u, it can also be applied for estimating time derivatives $\dot{x}(t)$: This is achieved by plugging the optimal decision u_0 into the differential equation (6.2). Using this estimate, avoids tedious recalculation of trajectories throughout the present paper.

The main requirement in these assumptions is the existence of an optimal equilibrium for the infinite horizon problem (6.1)-(6.2). If this equilibrium is known, it can be used as an additional constraint in (6.4), in order to improve the convergence properties. In our solution of the model in section 2, and further on, we did not use the terminal condition to solve the model but moved forward with a receding horizon to find the (approximate optimal) trajectories. Thus, without a priory knowledge of this equilibrium this convergence can also be ensured. Though the proofs in earlier work were undertaken for an undiscounted NMPC procedure, however, the main proofs carry over to the discounted case, details of which can be found in Grüne et al. (2015).