International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Address: IIASA, Schlossplatz 1, A-2361 Laxenburg, Austria

Email: repository@iiasa.ac.at

Working Paper **WP-24-008**

To grow or to fluctuate: Optimal paths to demographic equilibria

Stefan Wrzaczek (wrzaczek@iiasa.ac.at)
Gustav Feichtinger (gustav@eos.tuwien.ac.at)
Thomas Fent (thomas.fent@oeaw.ac.at)
Andreas Novak (andreas.novak@univie.ac.at)

Approved by

Michael Kuhn
Program Director, Economic Frontiers Program
2 May 2024

Table of contents

ZVR 524808900

Disclaimer, funding acknowledgment, and copyright information:
IIASA Working Papers report on research carried out at IIASA and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the institute, its National Member Organizations, or other organizations supporting the work.

The authors gratefully acknowledge funding from IIASA and the National Member Organizations that support the institute (The Austrian Academy of Sciences; The Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES); The National Natural Science Foundation of China (NSFC); The Academy of Scientific Research and Technology (ASRT), Egypt; The Finnish Committee for IIASA; The Association for the Advancement of IIASA, Germany; The Technology Information, Forecasting and Assessment Council (TIFAC), India; The Iran National Science Foundation (INSF); The Israel Committee for IIASA; The Japan Committee for IIASA; The National Research Foundation of Korea (NRF); The Research Council of Norway (RCN); The Russian Academy of Sciences (RAS); Ministry of Education, Science, Research and Sport, Slovakia; The National Research Foundation (NRF), Sub-Saharan Africa Regional Member Organization (SSARMO); The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS); The Ukrainian Academy of Sciences; The Research Councils of the UK; The National Academy of Sciences (NAS), USA; The Vietnam Academy of Science and Technology (VAST).

The authors gratefully acknowledge funding from the Austrian Science Fund (FWF) for the research project 'Life-cycle behaviour in the face of large shocks to health' (No. P 30665-G27).

To grow or to fluctuate: Optimal paths to demographic equilibria

Stefan Wrzaczek*, Gustav Feichtinger†, Thomas Fent‡, Andreas Novak§

May 2, 2024

Abstract

The principal goal of this paper is to establish a firm fundament of population policy. It is shown how intertemporal optimisation theory can be used to fulfill this important task. In particular, this will be illustrated by calculating the optimal trade-off between the further growth (or shrinking) of a population and the fluctuations of its age-structure generated by the decline (or the increase) of the fertility.

While the system dynamics of the age-structured optimal control model considered in this paper is described by the McKendrick-von Foerster partial differential equation, its objective functional is given by the discounted stream of the adaptation costs of the net reproduction rate (NRR) and the aforementioned trade-off. For the sake of obtaining analytic results the model is reduced to concentrated vitality rates. Its essential results for growing and declining populations are that under-/over shooting of the NRR is optimal for a short time horizon, whereas a fluctuating NRR takes over if the time horizon is extended. Numerical simulations for a stylized population structure show how the change in the NRR carries over to the total population and age-groups along time.

1 Introduction

The main target of population policy is to reach a desirable demographic equilibrium in an efficient manner. Three issues are essential to understand the logic behind this approach. The first ingredient is the *objective* the politician has in mind. During the sixties and the seventies of the past century, *zero population growth* was one predominant goal. To reach this goal, certain *instruments* must be available to the decision-maker, i.e., to the politician. This means that, as a second point, she has to select the values of these decision variables. Thirdly, and this step is essential in the logical framework, the decision instruments have to be connected with the objective. The way in which the former influence the latter is described by a suitable mathematical model. For the dynamic focus of this task, dynamic (i.e., intertemporal) optimization methods are an appropriate way to cope with these population policy problems. The rest of this paper illustrates the proposed approach in a particular case originating in a general discussion some fifty years ago as briefly outlined in the following. In the seventies General W.H. Draper, the representative of the US in the Population Commission of the UN,

*International Institute for Applied Systems Analysis (IIASA), Schlossplatz 1, Laxenburg, Austria; and Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/OeAW, University of Vienna), Austria. Email: wrzaczek@iiasa.ac.at.

†Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/OeAW, University of Vienna), Austria; and Institute of Statistics and Mathematical Methods in Economics (Research unit VADOR), Vienna University of Technology, Vienna, Austria. Email: gustav@eos.tuwien.ac.at

‡Austrian Academy of Sciences (OeAW), Vienna Institute of Demography (VID), Vienna, Austria; and Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/OeAW, University of Vienna), Austria. Email: thomas.fent@oeaw.ac.at.

§University of Vienna, Vienna, Austria. Email: andreas.novak@univie.ac.at

formulated the utopian objective of reaching zero population growth for his country by the end of the 20^{th} century, i.e., within 30 years. Soon after Draper's keen stipulation, Bourgeois-Pichat and Taleb [1] showed that his objective was virtually unreachable. Taking the female part of the fastly growing Mexican population, the two authors illustrated that the fertility rate would exhibit drastic fluctuations. Their calculations underline the fact that the goal of a uniform decrease of the natural growth rate within a few decades is quite unrealistic.

In a follow-up paper, another prominent French demographer, Leridon [19], stresses that fertility is the key variable to be influenced rather than the natural population growth rate. Taking the French female population, Leridon contrasted Draper's target by a more reasonable objective, namely a final stable population which is reached by a gradually declining fertility. The take-away of both papers consists in the fact that complex problems usually have no simple solutions. General Draper was not aware of how population dynamics really worked, particularly on the inertia of age-structure. In the early seventies, Frejka [14] studied the way of populations to their stationary equilibria for different ferility reduction patterns for all world regions. Essential for all this work is the usage of the classical cohort component method for population projections. The approach proposed in this paper is a different, mathematically deeper method, namely the McKendrick-von Foerster partial differential equation to formulate the dynamics of an age-structured population (see Keyfitz and Keyfitz [17]). The reason for this application is a paradigmatic change from *descriptive* alternative paths to equilibria to a *normative* approach proposed in this paper. It should be noted that Masnick [20] in a review of the Frejka compendium can be considered in a certain sense as precursor of the optimization approach if he writes as follows:

> *Age structure .. is not given (by Frejka) the attention it deserves. When ... we are shown the unevenness in world age structure that would result from an immediate fertility decline, contrasted with the smooth transition that accompanies a gradual fertility reduction, the tradeoff between excessive growth and excessive fluctuations in age structure becomes apparent. Nowhere, however, does Frejka really begin to discuss the implications of growth vs, instability.*
>
> *The subtitle of the book is "Alternative Paths to Equilibrium," but are his alternatives well chosen? There may be some advantages to be gained if every country approached replacement fertility on a direct and unwavering path with as little delay as possible. But to fail to consider the less rational paths, or to fail to point out the penalties of wavering in the process of moving in the equilibrium direction, or to fail to make clear the consequences of a too rapid approach to equilibrium, which is likely to be associated with overshooting the target of replacement fertility, is to fail to consider the alternatives that are likely to become reality.*

The remainder of the paper is structured as follows. Section 2 sets out the model and its basic assumptions, which can be solved with the set of optimality conditions presented in section 3. Analytic results are presented together with numerical illustrations in section 4. The paper closes with final remarks in section 5.

2 The model and assumptions

As it is frequently assumed in the theoretical literature on population dynamics, we consider a one-sex (female) population, where population size is denoted by $P(t,a)$ with t and a as time and age, respectively. The dynamic evolution of one cohort is described by the well-known McKendrick-von Foerster partial differential equation (see Keyfitz and Keyfitz [17] for details):[1]

$$P_t(t,a) + P_a(t,a) \quad = \quad -p(t,a)P(t,a). \tag{1}$$

[1] Note that subscripts denote partial derivatives, e.g., $P_t(t,a)$ denotes the partial derivative w.r.t. time t, etc.

This equation models the reduction of population size along the life-cycle of a cohort born at $t-a$ by the mortality rate $p(t,a)$, which naturally depends on age as well as on time to capture changes in health and life-expectancies along time as observed everywhere in the world and at any time in history. The maximum attainable age an individual is able to reach is marked as ω. The condition $P(t,a) = P(t-a,0)e^{\int_0^\omega p(t-a+s,s)\,\mathrm{d}s}$, translating to $\int_0^\omega p(t-a+s,s)\,\mathrm{d}s = +\infty$, is sufficient that the complete $(t-a)$-cohort has been extincted at the age of ω.

The initial (i.e., at $t=0$) age-structured population size is given by the exogenous initial condition $P(0,a) = P_0(a)$. The number of newborns at t, on the other hand, is endogenously defined by

$$B(t) \;=\; \int_0^\omega f(t,a)P(t,a)\,\mathrm{d}a, \tag{2}$$

where $f(t,a)$ denotes the fertility rate at t and age a. As $B(t)$ initiates a new cohort $P(t,0) = B(t)$ holds for all t. Analogously, the number of deaths $D(t)$ at time t are defined as

$$D(t) \;=\; \int_0^\omega p(t,a)P(t,a)\,\mathrm{d}a. \tag{3}$$

Summing up $P(t,a)$ across ages at every t defines the total population size $N(t)$ at t. The dynamics is naturally determined by the difference between births $B(t)$ and deaths $D(t)$, i.e.,

$$\dot{N}(t) = B(t) - D(t), \tag{4}$$

where the initial population size $N(0) = N_0$ is given by the initial condition of $P(0,a)$:

$$N_0 \;=\; \int_0^\omega P_0(a)\,\mathrm{d}a. \tag{5}$$

System (1)-(3) (together with (4)) defines a non-local partial differential equation modeling the evolution of the population over time. To get some crispy analytic results we follow Feichtinger and Wrzaczek [8, 9] and assume *concentrated* vitality rates as follows:

(A1) Mortality: Mortality is realized only at the maximum age ω, i.e., $p(t,a) = \delta(a-\omega)$ for all $t \geq 0$.

(A2) Fertility: The population reproduces at a unique age μ, i.e., $f(t,a) = R(t)\delta(a-\mu)$ for all $t \geq 0$.

Here $\delta(x)$ denotes the Dirac delta function, and $R(t)$ the *net reproduction rate* (NRR) at t, i.e., the mean number of girls born by a female. We, furthermore, assume that the fixed ages of death and giving births (equal to ω and μ in (A1) and (A2)) do not change over time.

The crucial assumption of this paper is, that the NRR is not an exogenous parameter (or exogenous function over time), but can be influenced, i.e., increased or decreased, by the policy maker. This critical assumption already considered in a number of theoretical works on population dynamics (see Coale [3], Feichtinger and Wrzaczek [8, 9]), does certainly apply to the real world though the extent of as well as the cost of adaptation efforts are not explored exhaustively to the best of our knowledge.

In the following we assume that the policy maker can invest into NRR-adaptation efforts $k(t)$ that change the NRR in the following way over time:[2]

$$\dot{R}(t) \;=\; k(t)R(t), \quad R(0) = R_0, \tag{6}$$

where the initial NRR is exogenously given by $R(0) = R_0$. This can be greater or smaller than replacement level 1.

[2] Note that (6) could also be formulated as $\dot{R}(t) = k(t)$. However, as argued also in Feichtinger and Wrzaczek [9] such a specification would lead to very similar optimal results, but needs an additional constraint on the state variable $R(t)$ to guarantee non-negative values.

Having the population dynamics (1)-(6) in mind, the policy maker tries to match the following three goals over a fixed time horizon T by choosing the appropriate (i.e., optimal) adaptation efforts along time:

(i) In accordance with relevant parts of the theoretical demographic literature the NRR has to be adapted (time period between $t = 0$ and $t = T$) such that it ends at replacement level $NRR = 1$ at T, i.e., $R(T) = 1$. If and how often $R(t)$ crosses the replacement level (i.e., if the NRR shows wave-like behavior) is not taken into account by this point.

(ii) Knowing that changes in population shares (i.e., age-composition of the population) are problematic from several points of view (e.g., sustainability of retirement system) the policy maker tries to drive the age-composition of the population defined as $c(t, a) := \frac{P(t,a)}{N(t)}$ ($N(t)$ representing the total population at t) to the age-composition of a stable population $\bar{c}(a)$ at replacement level as fast as possible. According to (A1), mortality is realized at the maximum age, implying that $\bar{c}(a) = \frac{1}{\omega}$. All deviations are costly (with increasing marginal costs) and aggregated over time and age. Thus, deviation costs can be defined as $C_2 \int_0^\omega (c(t, a) - \bar{c}(a))^2 \,\mathrm{d}a$, where C_2 is the corresponding cost parameter.

(iii) Similar to (ii) the change of the size of the population is a major pressing problem. However, the policy maker considers a deviation from some intended target population $\bar{N}$ only at T.[3] The structure is analogous as before, i.e., a deviation causes convex costs $C_3(N(T) - \bar{N})^2$ with C_3 as cost parameter. For the ease of exposition we fix $\bar{N}$ at the initial population size N_0.

Adaptation efforts $k(t)$ are certainly costly, both if the NRR has to be increased and if it has to be decreased. In accordance with the large economics literature we assume convex costs, leaving the policy maker with adaptation costs of $C_1k^2(t)$ at t (using C_1 as cost parameter). We, moreover, make the simplifying assumption that positive or negative efforts are equally expensive, which certainly can be viewed critically in real world applications. However, different costs would not change the theoretical conclusions of the optimal solution, which is the central contribution of this paper.

Combining (i)-(iii) the policy maker ends up with the following objective function:

$$\mathcal{V}(R_0) := \int_0^T \left[C_1k^2(t) + C_2 \int_0^\omega (c(t, a) - \bar{c}(a))^2 \,\mathrm{d}a \right] \mathrm{d}t + C_3(N(T) - N_0)^2. \tag{7}$$

The resulting full model analyzed in the following section reads

$$\min_{k(t)} \quad \mathcal{V}(R_0) \tag{8a}$$

$$\text{s.t.} \quad P_t(t, a) + P_a(t, a) = 0 \tag{8b}$$

$$P(0, a) = P_0(a), P(t, 0) = P(t, \mu)R(t) \tag{8c}$$

$$\dot{R}(t) = k(t)R(t), \quad R(0) = R_0, R(T) = 1 \tag{8d}$$

$$\dot{N}(t) = B(t) - D(t), \quad N(0) = N_0 = \int_0^\omega P_0(a)\,\mathrm{d}a \tag{8e}$$

$$B(t) = \int_0^\omega \delta(a - \mu)P(t, a)R(t)\,\mathrm{d}a = P(t, \mu)R(t) \tag{8f}$$

$$D(t) = \int_0^\omega \delta(a - \omega)P(t, a)\,\mathrm{d}a = P(t, \omega). \tag{8g}$$

[3]Note, that considering the difference at every t would give a very similar objective as compared to (ii). The concentration on T adds another different qualitative channel of the optimal solution, as will be shown in section 4.

3 Optimality conditions

The full model (8) is a finite time horizon age-structured optimal control model.[4] As the model includes age-structured as well as concentrated state variables (i.e., variables that evolve along time only), the problem is non-standard implying that an application of the standard age-structured Maximum Principle (see Feichtinger et al. [5]) requires a reformulation of the concentrated state variables as age-structured ones. Alternatively, Feichtinger and Wrzaczek [9] provide an extension of the age-structured Maximum Principle explicitly allowing for concentrated state variables.

Using the extension of the age-structured Maximum Principle, the Hamiltonian reads (suppressing t and a, and the arguments of the Hamiltonian)

$$\begin{aligned}\mathcal{H} &= -\frac{1}{\omega}C_1k^2 - C_2\left(\frac{P}{N}-\bar{c}(a)\right)^2 \\ &\quad +\xi\cdot 0 + \frac{1}{\omega}\lambda_R kR + \frac{1}{\omega}\lambda_N\left(B-D\right) + \eta_B\delta(a-\mu)PR + \eta_D\delta(a-\omega)P, \end{aligned} \tag{9}$$

where $\xi(t,a)$ denotes the adjoint variable of the age-structured population, and $\lambda_R(t)$ and $\lambda_N(t)$ the adjoint variables of the NRR and the total population. The adjoint variables of births and deaths are denoted by $\eta_B(t)$ and $\eta_D(t)$ respectively. Note that the direction along which the adjoint variables evolve carries over from the corresponding state variables. While ξ depends on time and age, λ_R, λ_N, η_B, and η_D depend on time only. Analogous to standard optimal control theory, the adjoint variables can be interpreted as dynamic shadow prices which give the (marginal) increase of the objective value for a (marginal) increase of the corresponding state variable. In other words, an adjoint variable evaluates the discounted effect of one (marginal) unit of a state variable over the remaining time horizon according to the objective function of the underlying optimal control model.

By derivation of the Hamiltonian we obtain the first order condition for optimal adaptation efforts at any time $t\in[0,T]$:

$$\frac{\partial\mathcal{H}}{\partial k} = -2\frac{1}{\omega}C_1k + \frac{1}{\omega}\lambda_R R = 0, \tag{10}$$

which balances marginal costs (first term) with marginal effect on the NRR (second term). By reformulation, the optimal adaptation efforts can be expressed as a function of the discounted future effect of a marginal increase of the NRR multiplied by the current NRR, weighted by the marginal adaptation costs:

$$k = \frac{\lambda_R R}{2C_1}. \tag{11}$$

Taking the derivative with respect to time immediately gives an analytical expression for the evolution of the optimal adaptation efforts:

$$\dot{k} = -\frac{RP(t,\mu)}{2C_1}\left(\lambda_N + \xi(t,0)\right). \tag{12}$$

This is referred to as *Euler equation* (among others) in theoretical economics and crucial in the proof of the central results of this paper. Also for the curvature of k an analytic expression can be formulated (analogously to Feichtinger and Wrzaczek [8]), as sum of the growth rates of the NRR, the fertile population and the marginal cost of population change.

[4]Note that the special structure of (8), i.e., the constant value of $P(t,a)$ along the life-cycle of a cohort, allows for a reformulation as an optimal control model with time-lag (dynamics of total population) and delay (second cost term in objective function). For details see Freiberger et al. [12].

Moving forward with the analysis, the adjoint variables are defined by the adjoint equations

$$\xi_t + \xi_a = \frac{2C_2}{N}\left(\frac{P}{N} - \bar{c}(a)\right) - \eta_B \delta(a-\mu)R - \eta_D \delta(a-\omega) \tag{13a}$$

$$\dot{\lambda}_R = -\lambda_R k - \int_0^\omega \eta_B \delta(a-\mu) P \,\mathrm{d}a \tag{13b}$$

$$\dot{\lambda}_N = -2C_2 \int_0^\omega \left(\frac{P}{N} - \bar{c}(a)\right) \frac{P}{N^2} \,\mathrm{d}a \tag{13c}$$

$$\eta_B = \lambda_N + \xi(t,0) \tag{13d}$$

$$\eta_D = -\lambda_N, \tag{13e}$$

and the corresponding transversality conditions

$$\xi(t,\omega) = 0 \tag{14a}$$

$$\xi(T,a) = 0 \tag{14b}$$

$$\lambda_N(T) = -2C_3\left(N(T) - N_0\right) \tag{14c}$$

$$\lambda_R(T) \quad \text{no transversality condition.} \tag{14d}$$

Here, (14a)-(14c) are standard and can be obtained by derivative of a salvage value function. (14d), however, differs due to fixed end value of the NRR at replacement level, i.e., $R(T) = 1$. By use of the assumption on the vitality rates (assumption (A1) and (A2)), the adjoint equations can be simplified as follows:

$$\xi_t + \xi_a = \begin{cases} \frac{2C_2}{N}\left(\frac{P}{N} - \bar{c}(a)\right) - \left(\lambda_N + \xi(t,0)\right) R & \text{for } a = \mu \\ \frac{2C_2}{N}\left(\frac{P}{N} - \bar{c}(a)\right) + \lambda_N & \text{for } a = \omega \\ \frac{2C_2}{N}\left(\frac{P}{N} - \bar{c}(a)\right) & \text{else} \end{cases} \tag{15a}$$

$$\dot{\lambda}_R = -\lambda_R k - P(t,\mu)\left(\lambda_N + \xi(t,0)\right) \tag{15b}$$

$$\dot{\lambda}_N = -\frac{2C_2}{N} \int_0^\omega \left(\frac{P}{N} - \bar{c}(a)\right) c \,\mathrm{d}a. \tag{15c}$$

The three-part representation of (15a) allows a backward solution of ξ and is the basis for the proofs of the theoretical results of this section. Further, note that the term $\xi(t,0)$ in (15a) adds the value of progeny to the adjoint variable. This term is a generalization of Fisher's reproductive value (see Fisher [11]) and occurs generically in age-structured optimal control models with endogenous fertility as shown by Wrzaczek et al. [23] and Feichtinger et al. [7].

Backward solution of (15a) for cohorts that reach the end of life before T gives (using (14a)) an explicit expression of $\xi(t,a)$:

$$\xi(t,a) = \underbrace{\Delta(t,a,\omega)}_{=:(a)} + \underbrace{\left(\lambda_N(t-a+\mu) + \xi(t-a+\mu,0)\right) R(t-a+\mu)}_{=:(b)} - \underbrace{\lambda_N(t-a+\omega)}_{=:(c)} \tag{16}$$

with

$$\Delta(t,a,\omega) := \int_a^\omega -\frac{2C_2}{N}\left(\frac{P}{N} - \bar{c}(s)\right) \mathrm{d}s. \tag{17}$$

The interpretation of the three terms (a)-(c) is as follows. Term (a) aggregates the marginal costs of deviation from target age-composition $\bar{c}(a)$ along the remaining life of the individual. The last argument in the function Δ denotes the age at which the cohort exits the model. For all cohorts that reach their maximum lifetime before T (i.e., born no later than $T-\omega$), this argument is ω. For later born cohorts this argument equals $T-(t-a)$. Term (b) depicts the marginal costs of (the expected number of) newborns of the individuals, composed by contribution to total population at time of birth and the reproductive value along the life-cycle of the newborn.

Term (c) represents the reduction of marginal costs (or benefit)[5] to the total population at the time of death. This term corresponds to the third part of the objective function, which is to end up with a total population that is as close as possible to the initial size.

Due to the assumptions on the concentrated vitality rates (assumptions (A1) and (A2)) the explicit expression (16) has a clear structure. This structure is illustrated in Figure 1. Aggregated costs of deviation from the target age-composition add at every t and a. Marginal costs/benefits to the total population enter at age ω, marked in blue, or eventually at time T at age $T-(t-a)$ marked in red. The reproductive part at age μ as colored in green.

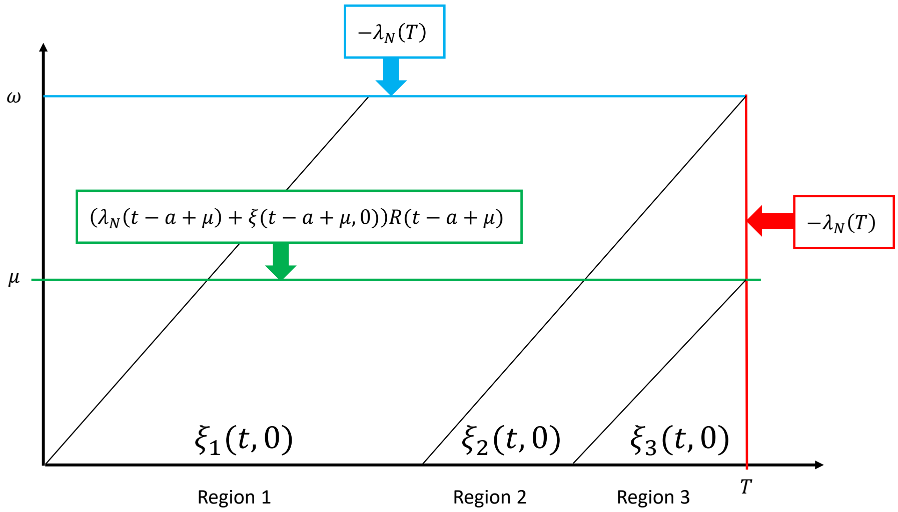


Figure 1: Illustration of structure of the concentrated vitality rates and how they enter $\xi(t,a)$.

Going back to the Euler equation (12) it is obvious that $\xi(t,0)$ is key. Here we exploit the assumption of the concentrated vitality rates and specify the explicit expression of $\xi(t,0)$ by dividing the time horizon into three regions, as visualized in Figure 1:

- Cohorts in region 1 $(t+\omega<T)$ experience the whole life-cycle before the end of the planning period,

$$\xi_1(t,0)=\Delta(t,0,\omega)+\left(\lambda_N(t+\mu)+\xi(t+\mu,0)\right)R(t+\mu)-\lambda_N(t+\omega). \tag{18a}$$

- Cohorts in region 2 $(t+\mu<T<t+\omega)$ reach their reproductive age within the planning period, but live beyond T,

$$\xi_2(t,0)=\Delta(t,0,T-t)+\left(\lambda_N(t+\mu)+\xi(t+\mu,0)\right)R(t+\mu). \tag{18b}$$

- Cohorts in region 3 $(t+\mu>T)$ neither reach reproduction not the end of their life before T,

$$\xi_3(t,0)=\Delta(t,0,T-t). \tag{18c}$$

Note that there are cohorts that are born at the transition from region 1 to 2 $(t+\omega=T)$ and from region 2 to 3 $(t+\mu=T)$. However, as time (and age) are continuous variables both cohorts have measure zero and can therefore be neglected.

The conditions and properties of this section will be exploited in the following section to derive and characterize properties of the optimal solution along time.

[5] The adjoint λ_N can be interpreted as marginal costs if the sign is negative and as benefit otherwise.

4 Results

In the following we will investigate analytically *if*, *how many*, and *under which conditions*, the optimal solution uses under-/overshooting or fluctuating behavior along the optimal path of adaptation efforts and/or the corresponding NRR towards reaching $R(T) = 1$ under minimal deviations of the target age-composition at every t and of the total population at T. For this reason we first look at the polar cases of the cost parameters $C_2 = 0$ (subsection 4.1) and $C_3 = 0$ (subsection 4.2). This allows to exploit the special structure of the positive cost term in the objective function. If and how the specific solution properties carry over to the general case will be investigated afterwards (subsection 4.3).

4.1 Over- and undershooting ($C_2 = 0$)

In the case of $C_2 = 0$, the policy maker is not concerned about the age-composition of the population in relation to the target age-composition $\bar{c}(a)$, but only about the total population size at the end of the planning period. The objective function reduces to

$$\min_{u(t)} \int_0^T C_1 k^2(t)\,\mathrm{d}t + C_3(N(T) - N_0)^2, \tag{19}$$

which means that a target population size (in our case equal to N_0), which technically refers to a salvage value function, should be met as close as possible under minimizing adaptation costs over time. Thus, the policy maker is not concerned about the structure of the population at any point in time, but only on the final size of the total population. The following theorem exploits the explicit representation of $\xi(t, 0)$ in the three regions as illustrated in Figure 1 and provides insight into the qualitative solution structure over time.

Theorem 1 *Let us assume that $C_1 \geq 0$ is sufficiently small, $C_2 = 0$ and $C_3 > 0$. If the initial population distribution is stable regarding $R_0 \neq 1$, the optimal adaptation efforts $k(t)$ change the sign once. The sign of $\dot{k}(t)$ does not change. Due to continuity that holds also for sufficiently small C_2.*[6]

Proof. From (15c) we can conclude that $\lambda_N(t) = -2C_3\,(N(T) - N_0)$ is a constant (therefore denoted by λ_N^0) for $C_2 = 0$. Then we have the following $\lambda_N(t) + \xi(t, 0)$ for the three regions

$$\begin{aligned}
\lambda_N(t) + \xi_1(t, 0) &= \left(\lambda_N^0 + \xi(t + \mu, 0)\right) R(t + \mu) && \text{(20a)}\\
\lambda_N(t) + \xi_2(t, 0) &= \lambda_N^0 + \left(\lambda_N^0 + \xi(t + \mu, 0)\right) R(t + \mu) && \text{(20b)}\\
\lambda_N(t) + \xi_3(t, 0) &= \lambda_N^0. && \text{(20c)}
\end{aligned}$$

The sign of these three expressions defines the sign of $\dot{k}$ by the Euler equation (12), where we also note that $R(t) > 0$, $P(t, \mu) > 0$ holds for all t:

- Region 3: The sign of $\dot{k}$ does not change as $\lambda_N(t) + \xi_1(t, 0)$ is constant at λ_N^0.
- Region 2: Here we distinguish two cases.
  First consider that $\xi(t+\mu, 0)$ lies in region 3. Then the above expression (20b) reduces to $\lambda_N(t)+\xi_2(t, 0) = \lambda_N^0\,(1 + R(t + \mu))$, which implies the same sign.

[6] Note that this conclusion strongly depends on the implicit assumption of a (locally) continuous value function of (8) with respect to C_2. Certainly, we are aware that this is not necessarily the case. (see the advanced literature on Skiba points and surfaces). However, it can be expected that this interesting phenomenon (which has not been explored in age-structured models so far extensively) does not apply for marginally small C_2, as this is not a typically causal mechanism.

Second, consider the opposite, that $\xi(t+\mu,0)$ lies still in region 2. Then iterative usage of the above expressions yields

$$\begin{aligned}
\lambda_N(t)+\xi_2(t,0) &= \lambda_N^0+\left(\lambda_N^0+\xi_2(t+\mu,0)\right)R(t+\mu)\\
&= \lambda_N^0+\left(\lambda_N^0+\left(\lambda_N^0+\xi(t+2\mu,0)\right)R(t+2\mu)\right)R(t+\mu)\\
&= \lambda_N^0+\lambda_N^0R(t+\mu)+\lambda_N^0R(t+2\mu)R(t+\mu)+\xi(t+2\mu,0)R(t+2\mu)R(t+\mu)
\end{aligned} \tag{21}$$

for which we get the same sign if $\xi(t+2\mu,0)$ is in region 3. Otherwise, follow the same manipulation until (after j iterations) the $\xi(t+j\mu,0)$ is in region 3 and the same conclusion applies.

- Region 1: The sign does not change because of the same argument as for region 2.

The second assertion therefore follows immediately.

For the first assertion we consider a sufficiently small $C_1>0$ which, together with $C_2=0$, implies that $C_3(N(T)-N_0)^2$ remains in the objective function. As the NRR can be adapted almost without costs (recall $C_1>0$ small) the optimal solution implies $R(T)=1$ (end constraint for the NRR) and almost $N(T)=N_0$.[7] If we start from $R_0>1$ (analogously for $R_0<1$) the latter can only be reached if $k(t)<0$ for $t\in[0,\tau)$ and $k(t)>0$ for $t\in(\tau,T]$. According to the first assertion $\dot{k}(t)>0$ and the $R(t)$ trajectory crosses 1 only once on the open interval $(0,T)$. ∎

Theorem 1 allows for a few important conclusions. First note that the result gives sufficient conditions (i.e., parameter constellation and stable population) for one wave in the optimal adaptation efforts over time. This means that for $R_0>1$ (for $R_0<1$ an analogous argument applies) the policy maker starts to reduce the NRR even below replacement level by $k(t)$ that is negative for the first part of the time horizon. As $k(t)$ strictly increases over the entire time horizon and $R(t)$ has to end up at replacement level, $k(t)$ crosses 0 and becomes positive afterwards. Due to the unique sign of $\dot{k}(t)$ this *undershooting* occurs only once. The intuition is that the NRR has to be reduced even below replacement level in order to end up with a total population that is (almost) equal to N_0. The latter objective could certainly be also reached by several changes in the sign of k (i.e., NRR crossing 1 several times), however, even for a small cost parameter C_1 this would be more costly than the solution explained above. The change of the sign of $k(t)$ certainly carries over to the NRR in terms of taking values above or below 1. It is not possible to conclude undershooting or even a wave for certain ages or age-groups directly from Theorem 1.

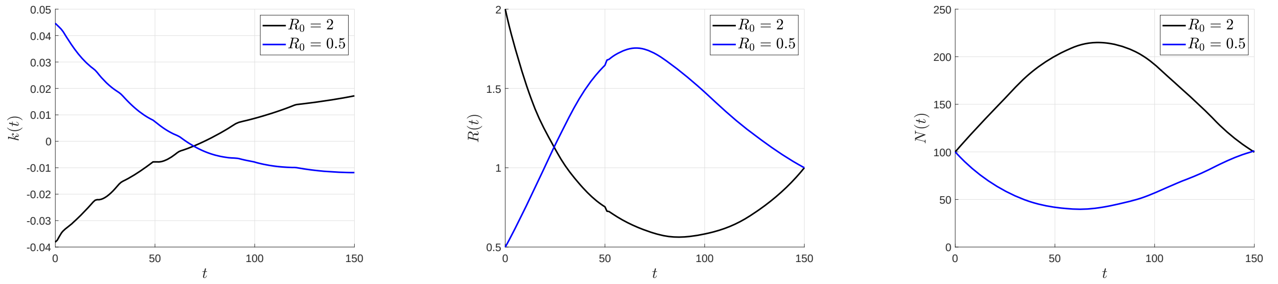
To illustrate the statement of the Theorem we use a numerical solution of model (8) with stylized parameters and the assumed cost parameter specifications. A stable initial population distribution with $R_0\neq 0$ implies $P(0,a)=P(0,0)e^{-\frac{\ln R_0}{\mu}a}$.[8] Figure 2 plots the optimal solution for a maximum individual life-time of 100 years over a time horizon of 150 years. The colors of the solution trajectories correspond to different initial values of the NRR. Black ones mark the solution with $R_0=2$, blue ones to $R_0=0.5$. Panel 2a confirms both assertions of the Theorem. For $R_0=2$ the NRR is too high initially and $k(t)$ crosses zero once and increases over the entire time horizon (assertion 2). Due to the negative values of $k(t)$ the NRR decreases steeply at the beginning of the time horizon and crosses replacement already quite soon (panel 2b). Below replacement level the NRR starts to increase (when $k(t)$ changes its sign) to end up at $R(T)=1$. The total population follows the evolution of the NRR. It increases initially, reaches a peak, but starts to decrease until it eventually ends up very close to N_0. The interpretation for $R_0=0.5$ works accordingly.

In a nutshell, the policy maker faces very low adaptation costs in order to reach the terminal condition $R(T)=1$ and to deviate from N_0 only marginally. For $R_0=2$ ($R_0=0.5$) she therefore undershoots (overshoots)

[7] To be precise: For any small $\epsilon>0$ we can find a small $\bar{C}_1$ such that $|N(T)-N_0|\leq\epsilon$ holds for any $C_1\leq\bar{C}_1$.

[8] By $P(0,a)=R_0P(0,a+\mu)$ we have $P(0,a+\mu)=\frac{1}{R_0}P(0,a)=e^{\mu r_0}P(0,a)$. From $e^{\mu r_0}=\frac{1}{R_0}$ we get $r_0=-\frac{\ln R_0}{\mu}$ and therefore, $P(0,a)=P(0,0)e^{-\frac{\ln R_0}{\mu}a}$.

the NRR below (above) replacement level to decrease (increase) the population size in the second part of the time horizon to offset the initial increase (decrease) in the population.



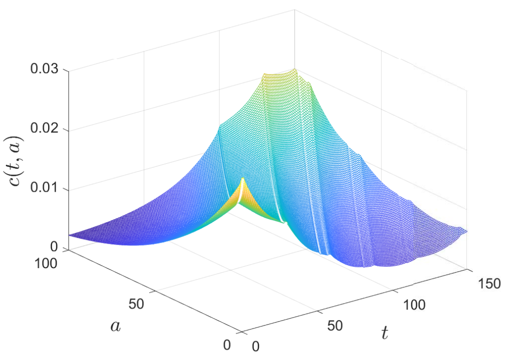
(a) Adaptation of fertility $k(t)$

(b) NRR(t)

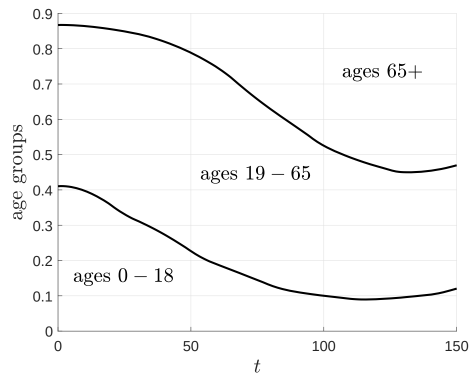
(c) Total population $N(t)$

Figure 2: Optimal policy over time for $C_2 = 0$ (case Theorem 1).

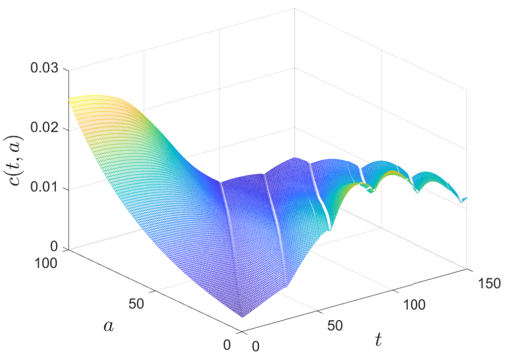
Figure 3 shows the corresponding age-compositions of the population (panels 3a and 3c) and relation between aggregated age-groups (panels 3b and 3d) over time. Considering the case $R_0 = 2$ (first line of the figure) shows the waving behavior in the population age-composition as well as in age-groups though the NRR and the total population undershoots once as discussed above. This effect is due to the time-lag in the fertility, which means that babies born at t reach their fertile age at age μ in our model. Moreover, the size of a cohort cannot be changed by the policy maker once it is born. It follows the mortality process, which is realized only at the maximum age according to our simplifying assumption (A1). Model (8) only allows for a direct adaptation of the newborns at any point in time, which carries over to the reproduction process μ years later. Together with the initial stable population structure, which means that together with assumption (A1) the initial population $N_0(a)$ decreases in a for $R_0 > 0$, this implies the observed waves.



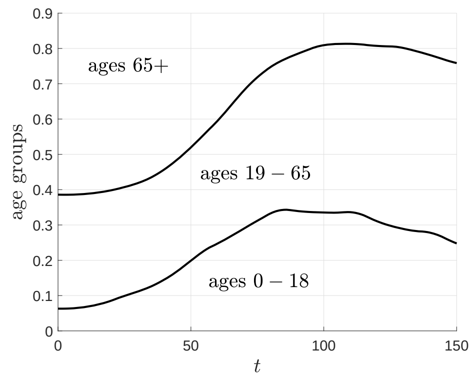
(a) Age-composition of the population for $R_0 =$ 2.



(b) Age-groups for $R_0 = 2$.



(c) Age-composition of population for $R_0 = 0.5$.



(d) Age-groups for $R_0 = 0.5$.

Figure 3: Population densities and age-groups over time for $C_2 = 0$ (case Theorem 1).

The results of Figures 2 and 3 are quite robust to the parameters C_1 and C_3. Note that not the absolute parameter values but only their relation impacts the objective function, as the multiplication with a positive parameter does change the value $\mathcal{V}(R_0)$ but not the optimal solution trajectories. So, fixing C_3 the result remains unchanged (see again footnote 7) as soon as C_1 is small enough such that the adaptation costs are dominated by the costs of mismatching the final total population size N_0.

4.2 Waves in adaptation and NRR ($C_3 = 0$)

As a next step we consider the opposite of what has been assumed in Theorem 1. Assume that the policy maker is not concerned about the development of the total population, but only about the age-distribution at every t. The objective function therefore reads

$$\int_0^T \left[C_1 k^2(t) + C_2 \int_0^\omega \left(c(t,a) - \bar{c}(a)\right)^2 \mathrm{d}a \right] \mathrm{d}t, \tag{22}$$

and includes costs incurred at every t, but no salvage value as in the previous section. The following theorem characterizes the qualitative structure of the optimal solution reflecting this different parameter setup. The proof approaches the problem along a different line and uses a direct argument that shows that it is optimal to change the sign of the adaptation efforts several times during the planning period.

Theorem 2 *Let us assume that $C_1 \geq 0$ is sufficiently small, $C_2 > 0$, $C_3 = 0$, and $T > \mu$ holds. If the initial population distribution is stable, the DM changes the sign of the adaptation efforts more than once during the planning priod. Likewise the NRR crosses the replacement level several times.*

Proof. Let us consider the case $R_0 < 1$ (for $R_0 > 1$ the proof work analogous) and let us further consider a marginally small C_1. Then the dynamic optimization reduces to an (almost) static optimization problem, i.e., the policy maker minimizes the cost term $C_2 \int_0^\omega \left(\frac{P(t,a)}{N(t)} - \bar{c}(a)\right)^2 \mathrm{d}a$ at every t as the NRR can be adapted almost for free. Thus, in a marginally short time interval $(t, t+\delta_t)$ the policy maker increases R (starting from $R_0 < 1$), such that $\int_0^\omega \left(\frac{P(t,a)}{N(t)} - \bar{c}(a)\right)^2 \mathrm{d}a$ is minimal at t in the following way:

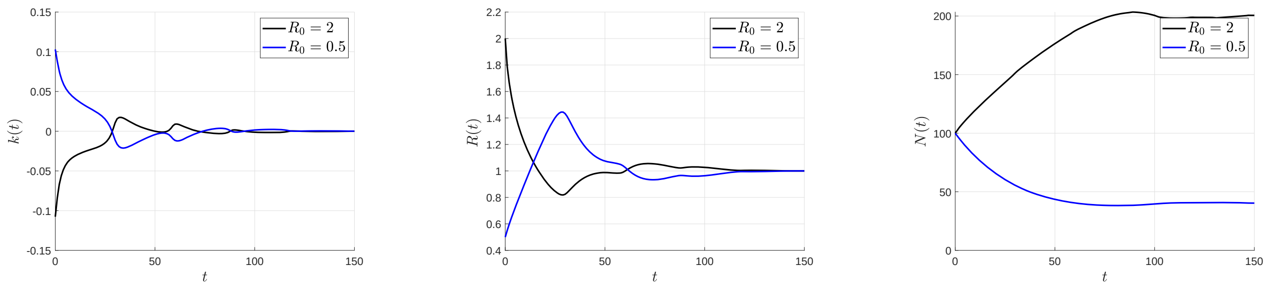
A stable population with $R_0 < 1$ implies that (recall $\bar{c}(a) = \frac{1}{\omega}$) the $P_a(0,a) > 0$. Therefore it is obvious that $\frac{P(0,0)}{N(0)} < \frac{1}{\omega} < \frac{P(0,\omega)}{N(t)}$ (recall also that $\int_0^\omega \frac{P(t,a)}{N(t)} \mathrm{d}a = 1$). Thus the policy maker will immediately, i.e., in a marginally small time interval $(0, \delta_t)$, increase the NRR (i.e., positive $k(t)$) such that $\frac{P(\delta_t,0)}{N(\delta_t)} = \frac{R(\delta_t)P(\delta_t,\mu)}{N(\delta_t)} = \frac{1}{\omega}$. As a result, the objective function (i.e., the costs) will decrease and $N(t)$ has decreased as $P(0,\omega)$ (with fraction $\frac{P(0,\omega)}{N(0)} > \frac{1}{\omega}$) have left and $P(0,0)$ with fraction ($\frac{P(0,0)}{N(0)} < \frac{1}{\omega}$) have entered the system. In the next marginally small time interval the argument is analogous.

The same pattern is followed during the time interval $[0, \mu)$. At $t = \mu$ the first cohort that has been 'optimized' by the policy maker reaches the fertile age. This cohort is now larger then the cohort immediately before, i.e., $P(\mu,\mu) > P(\mu, \mu + \delta_t)$. As a result, the NRR is too high and the policy maker has to reduce the NRR implying a negative $k(t)$.

This pattern is followed during the entire planning period. By continuity with respect to the cost parameters, the result holds also for small C_1. ■

Theorem 2 neither extends nor contrasts Theorem 1. It corresponds to another cost part of the objective function, i.e., the deviation from a stable age-composition of the population, and shows that this gives an incentive to the policy maker for changing several times between an increasing and a decreasing NRR. Certainly several switches in the NRR adaptation strategies towards ending up with the NRR at replacement level are costly and can analytically be proven only exploiting $C_1 = 0$ which extends to sufficiently small $C_1 > 0$. However, it reveals the relative age-composition of the population as key factor for several adaptation waves. The policy maker does not care about the final size, but considers every period separately.

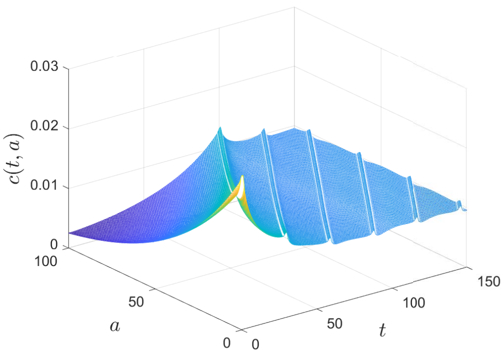
Figures 4 and 5 show the solution of the analogous numerical setup with different cost parameters. Panels 4a and 4b of Figure 4 immediately illustrate the result of Theorem 2, i.e., fluctuating behavior during the planning period until $R(T)$ reaches replacement level. Interestingly, the downward and upward branches of the waves do evolve at the same pace, which is induced by the initial stable population structure (i.e., $N_0(a)$ decreases in a for $R_0 > 1$, and opposite for $R_0 < 1$). Moreover, the figures show that the waves are dampening over time, which is an indication that the deviation from the target age-composition (which is $\bar{c}(a) = \frac{1}{\omega}$ for all a) becomes smaller. Another interesting observation is that the length of the waves decreases over time (for both $R_0 = 2$ and $R_2 = 0.5$). As before, this is also implied by the stable age-composition of the population at $t = 0$, which implies monotony of N_0, as well as by the fact that smaller deviations can be 'corrected' by smaller adaptations. During the process of a changing NRR this monotony breaks, which gives the policy maker rise to change the adaptation efforts earlier and earlier as time elapses. The total population (panel 4c) reflects that the policy maker does not consider meeting N_0 at T at all, i.e., $C_3 = 0$. $R_0 = 2$ ($R_0 = 0.5$) implies that the total population approximately doubles in the long run. This clearly carries over from the initial value of the NRR, and the fact that the period lengths of being above or below replacement level (see argument on the length of the waves above) decrease over time.



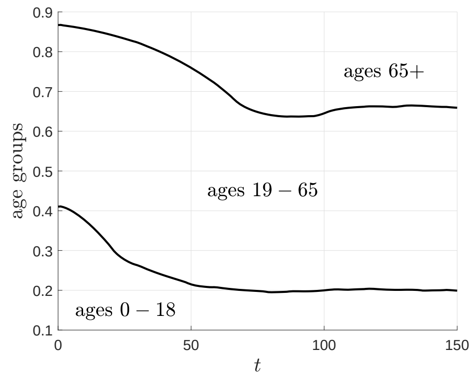
(a) Adaptation of fertility $k(t)$ (b) NRR(t) (c) Total population $N(t)$

Figure 4: Optimal policy over time for $C_3 = 0$ (case Theorem 2).

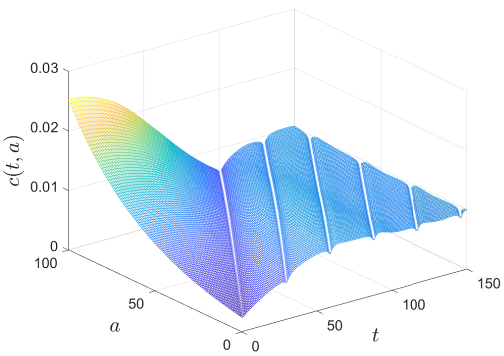
Figure 5 adds the corresponding population densities and the relations between aggregated age-groups to $R_0 = 2$ and $R_0 = 0.5$. The left panels show relative quick convergence to the target age-composition. The aggregated age-group size follow likewise. Both type of figures also show waves, which dampen over time.



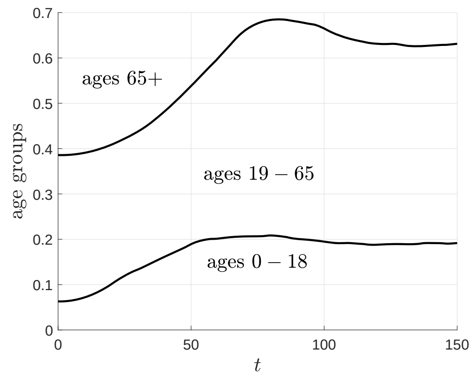
(a) Age-composition of the population for $R_0 =$ 2.



(b) Age-groups for $R_0 = 2$.



(c) Age-composition of the population for $R_0 =$ 0.5.



(d) Age-groups for $R_0 = 0.5$.

Figure 5: Population densities and age-groups over time for $C_3 = 0$ (case Theorem 2)..

Note that the small bumps in the figures for the population densities are numerical artefacts that emerge from the time-lag of the problem in combination with the step-size of the discretization of the numerical implementation. Reducing the step-size would make these bumps smaller and smaller, but not change the overall qualitative structure of the solution.

4.3 Complete picture

The proofs of Theorems 1 and 2 heavily exploited the basic assumptions $C_2 = 0$ and $C_3 = 0$, respectively. It is at hand that this this strategy will fail in the general case where both parameters, C_2 and C_3, are positive. However, using the results and assumptions of theorems, another route can be followed. Recalling the general objective function

$$\int_0^T \left[C_1 k^2(t) + C_2 \int_0^\omega (c(t,a) - \bar{c}(a))^2 \,\mathrm{d}a \right] \mathrm{d}t + C_3 (N(T) - N_0)^2, \tag{23}$$

we observe a different structure (regarding the cost parameters) as compared to the reduced forms (19) and (22) used for Theorems 1 and 2 and with respect to the length of the planning period. While C_2 effects costs at every point in time, C_3 is in place only at $t = T$. I.e., if the time horizon T is varied from a conceptual point of view (from the perspective of the policy maker it certainly remains fixed), C_2 changes the impact compared to C_3, which remains as a constant for every specific T. The following theorem makes use of this property and shows that the qualitative behavior of Theorems 1 and 2 are in place for different lengths of the planning period.

Theorem 3 *Let us assume that $C_1 \geq 0$ is sufficiently small. If the initial population distribution is stable, the following two assertions hold for the full model (8) with $C_2 > 0$ and $C_3 > 0$:*

1. *If T is sufficiently small, the sign of $\dot{k}(t)$ does not change, and the sign of $k(t)$ changes once during the planning period: Theorem 1 holds.*
2. *If T is sufficiently high, $k(t)$ changes the sign more than once during the planning period: Theorem 2 holds.*

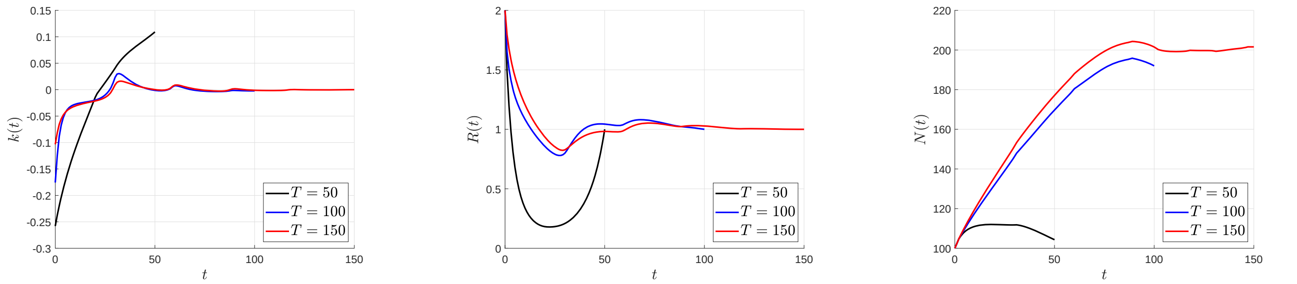
Proof. For the first assertion we consider the cost term for deviations from $\bar{c}(a)$ for all a. Due to definition of $N(t)$ (see (4) and (5)) we obtain $\left(\frac{P(t,a)}{N(t)} - \bar{c}(a)\right)^2 \leq \max\left\{\frac{(\omega-1)^2}{\omega^2}, \frac{1}{\omega^2}\right\} \leq 1$. For the entire planning period this implies $C_2 \int_0^T \int_0^\omega \left(\frac{P(t,a)}{N(t)} - \bar{c}(a)\right)^2 \mathrm{d}a\,\mathrm{d}t \leq C_2 \omega T$. For a sufficiently short planning period T, this term is therefore negligibly small and the assertion can be proven analogously to Theorem 1.

The second assertion will be proven by contradiction, i.e., let us assume that $k(t)$ changes the sign once at $\hat{t}$. W.l.o.g. we assume $R_0 > 1$ (for $R_0 < 1$ the analogous arguments are in place). As a result the planning period divides into two periods, $[0, \hat{t})$ and $(\hat{t}, T]$. As above we consider the second cost term and get $C_2 \int_0^T \int_0^\omega \left(\frac{P(t,a)}{N(t)} - \bar{c}(a)\right)^2 \mathrm{d}a\,\mathrm{d}t \geq C_2 \int_0^{\hat{t}} \int_0^\omega \left(\frac{P(t,a)}{N(t)} - \bar{c}(a)\right)^2 \mathrm{d}a\,\mathrm{d}t$. As for the first period $\frac{P(t,a)}{N(t)} \neq \bar{c}(a)$ except one specific a, this cost term tends to infinity for $T \to \infty$. Therefore, by an analogous argument as for the first assertion, this cost term dominates the third one, which proves the assertion. ∎

As mentioned before, the qualitative structure of the optimal solution of (8) is unclear if the cost parameters are not restricted somehow. However, Theorem 3 provides a criterion for how the results of the previous theorems spill over to the general case (at least for small adaptation costs) depending on the length of the time horizon. The proof uses again a different approach, i.e., the relation of the deviation cost parameters C_2 and C_3 that changes as T is varied. For small T the policy maker sticks to over-/undershooting behavior as she is mainly interested in the size of the total population at T. As T becomes large enough the age-composition along time becomes important and fluctuating adaptations take over.

The transition from over-/undershooting to fluctuating behavior takes place for varying T and can either be smooth or by a qualitatively disruptive change in the strategy. A smooth transition takes place if the importance of C_2 smoothly wipes out C_3. A qualitative disruptive transition, if it exists, is caused by the time-lagged nonlinearity in the model, i.e., $P(t,\mu)R(t)$ in (8c). The nonlinearity of the latter product can affect the competing cost factors in the objective function, which is a known source for Skiba surfaces (see e.g., Rowthorn and Maciejowski [21] Caulkins et al. [2], Grass et al. [13]). However, the detection and structured analysis of Skiba surfaces in age-structured optimal control models is a very advanced topic that has hardly been addressed in the literature so far and is an interesting and important topic for future research.

For the illustration of this property in Figure 6 we focus on $R_0 = 2$ only and vary the length of the time horizon. For $T = 50$ (black curves) the policy maker follows undershooting behavior with $k(t)$ changing the sign only once. The only slight difference to Theorem 1 is that the total population size does end up slightly, but unambiguously, above N_0. This is due to a positive C_1, which has not been chosen marginally small in this scenario. A further reduction of T would bring $N(T)$ as close to N_0 as required. Here the same argument of the relative size of the parameters in response to the length of the planning period applies. For $T = 100$ (blue lines) the situation now differs and the age-composition of the population becomes more prominent in the objective function. Now the policy maker follows waving behavior in $k(t)$ and $R(t)$, which now implies ending at a considerably high $N(T)$. For $T = 150$ (red lines) the situation is similar to $T = 100$ but extends over a longer time period. The waves are less pronounced as also strong waves produce higher costs by deviating from the target age-composition.



(a) Adaptation of fertility $k(t)$ (b) NRR(t) (c) Total population $N(t)$

Figure 6: Optimal policy over time for different time horizons (case Theorem 3).

5 Conclusions

The paper considers a theoretical model on the optimal adaption of the net reproduction rate over a fixed time horizon. Until ending with an NRR at replacement level, the decision maker chooses adaptation efforts to minimize the deviation from a given age-composition of the population and the deviation from the initial population at the terminal time. Using age-structured optimal control theory it is shown that over- or undershooting (depending on whether $R_0 > 1$ or $R_0 < 1$) of the NRR along the planning period is optimal if the deviation from the initial population is considered only, whereas in the opposite case of only including the age-composition at every t, fluctuations in the adaptation efforts are optimal. Finally, it is shown that these two polar-case results carry over to the general case depending on the length of the time horizon. From a mathematical point of view, a proof of fluctuating behavior is new in the literature of age-structured optimal control theory. For a similar result we refer to Feichtinger et al. [6] or Wrzaczek and Kort [22]. However, in these models fluctuations in the control variables are optimal because of the anticipation to a non-continuous change (i.e., a jump) of a model parameter (i.e., productivity of production capital). An important open question for future research is therefore a mathematical rigorous definition of (necessary and sufficient) conditions or model classes for which fluctuations of the control variables are optimal.

One critical point of the proposed model if it is applied to concrete data is the correct choice of the three cost parameters in the objective function.[9] The problem here is twofold:

(i) Although a number of possible efforts (in the form of incentives rather than in the form of forcing) to increase or decrease the population are well known (including costs), the efficacy, i.e., the effect on the NRR, remains unknown. As the control variable in the above model is measured in relative change units per time unit, the unknown effect carries over to a very imprecise idea of resulting cost.[10] However, the qualitative structure of the optimal solution for a very short or a sufficiently long time horizon is known as implied by Theorem 3. (ii) Similarly it is difficult to measure the costs for deviation from a stable age-composition as well as from the initial population size (or more general from an arbitrary population size) at T. Of course, this costs correspond to adaptations of the infra structure, the social system, etc. (see the introduction of Fent et al. [10] for a brief overview), a profound estimation is difficult. Because of (i) and (ii) we therefore recommend running the model for different cost parameters to develop a feeling for the form of the optimal solution depending on different cost combinations if the model is applied to realistic demographic data.

Another critical point in our model is the assumption of concentrated vitality rates. Differently to the uncertainty of the cost parameters, the concentrated vitality rates are only necessary for the formal proofs which are the focus of this paper. Both, the fertility distribution as well as the mortality rate can be included in a numerical study of the general problem. Based on the analysis in Fent et al. [10] (same model structure as here with realistic mortality data) and Feichtinger and Wrzaczek [9] (realistic fertility and mortality data but a slightly different objective function), we are optimistic that the theoretical results remains true for realistic mortality data. Finally, we would like to stress once more, that the main focus of the paper was to show that over- and undershooting as well as fluctuations of the NRR in the transition to a NRR at replacement level can be optimal for a rational policy maker.

In addition to the concentrated vitality rates, the theoretical model presented in this paper is also quite stylized regarding other dimensions and allows for several extensions appealing for demography such as, for instance, (i) including exogenous or endogenous migration which can hardly be neglected for a substantial number of countries, (ii) addressing health related public measures in combination with a relaxation of the assumption of the concentrated mortality age, or (iii) considering different cost/benefit terms in the objective function reflecting the policy makers evaluation of the population development and its age-composition. A resulting generalized model then could be applied to specific countries to derive and evaluate how to respond best to different demographic scenarios along time.

References

[1] Bourgeois-Pichat, J., Taleb, S.-A. (1970) Un taux d'accroissement nul pour les pays en voie de developpement en l'an 2000. Reve ou realite? Population (French Edition) 25, 957-974.

[9] Note that the three parameters can be reduced to two if one of these is chosen as numeraire parameter. As this can easily be done by dividing the entire objective function by the numeraire parameter, the optimal solution remains unchanged. One only has to take this into account if the value of the objective function is reported in money.

[10] Note also that our stylized model assumes that increasing and decreasing efforts of the NRR are equally costly. Needless to say that this assumption is another artificial one, it does not affect the main theoretical results of the paper as this parameter is not critical in any the proofs. On the other hand, although the model could easily be generalized to allow for different costs for positive or negative adaptation efforts, the analysis will complicate without additional insight in the solution or the qualitative structure of the model.

[2] Caulkins, J.P., Grass, D., Feichtinger, G., Hartl, R.F., Kort, P.M., Prskawetz, A., Seidl, A., Wrzaczek, S. (2021) The optimal lockdown intensity of COVID-19. Journal of Mathematical Economics 93, 102489.

[3] Coale, A.J. (1972) Growth and Structure of Human Populations. A Mathematical Investigation. Princeton University Press.

[4] Feichtinger, G. (1979) Demographische Analyse und populationsdynamische Modelle. Grundzüge der Bevölkerungsmathematik. Springer, Wien.

[5] Feichtinger, G., Tragler, G., Veliov, V.M. (2003) Optimality conditions for age-structured control systems. Journal of Mathematical Analysis and Applications 288, 47–68.

[6] Feichtinger, G., Hartl, R.F., Kort, P.M., Veliov, V.M. (2006) Anticipation effects of technological progress on capital accumulation: a vintage capital approach. Journal of Economic Theory 126, 143-164.

[7] Feichtinger, G., Kuhn, D. R., Prskawetz, A., Wrzaczek, S. (2011) The reproductive value as part of the shadow price of population. Demographic Research 24, 709–718.

[8] Feichtinger, G., Wrzaczek, S. (2024) The optimal transition to a stationary population for concentrated vitality rates. Demographic Research 50, 171-184.

[9] Feichtinger, G., Wrzaczek, S. (2024) The optimal momentum of population growth and decline. Theoretical Population Biology 155, 51-66.

[10] Fent, T., Wrzaczek, S., Feichtinger, G., Novak, A. (2024) Fertility decline and age-structure in China and India. VID Working Paper 01/2024, Vienna, Austria.

[11] Fisher, R.A. (1930) The Genetical Theory of Natural Selection. Dover Publications, New York.

[12] Freiberger, M., Kuhn, M., Prskawetz, A., Sanchez-Romero, M., Wrzaczek, S. (2024) Optimization in age-structured dynamic economic models. IIASA Working Paper. Laxenburg, Austria: WP-24-004.

[13] Grass, D., Wrzaczek, S., Caulkins, J.P., Feichtinger, R.F., Hartl, R.F., Kort, P.M., Kuhn, M., Prskawetz, A., Sanchez-Romero, M., Seidl, A. (2024) Riding the waves from epidemic to endemic: Viral mutations, immunological change and policy responses. Theoretical Population Biology 156, 46-65.

[14] Frejka, T. (1973) The Future of Population Growth: Alternative Paths to Equilibrium. Wiley, New York.

[15] Grass, D., Caulkins, J. P., Feichtinger, G., Tragler, G., Behrens, D.A. (2003) Optimal Control of Nonlinear Processes: With Applications in Drugs, Corruption and Terror. Springer.

[16] Keyfitz, N. (1971) On the momentum of population growth. Demography 8 (1), 71–80, Publisher: Springer.

[17] Keyfitz, N., Keyfitz, B. (1997) The McKendrick partial differential equation and its uses in epidemiology and population study. Mathematical Computation Modelling 26, 1-9.

[18] Keyfitz, N., Caswell, H. (2005) Applied mathematical demography, 3rd ed. In: Statistics for biology and health. Springer, New York, NY.

[19] Leridon, H. (1983) Un taux d'accroissement constant comme objectif démographique et méthode de projection. Population (French Edition) 38 (2), 343-360.

[20] Masnick, G. S. (1974) Review of: The Future of Population Growth: Alternative Paths to Equilibrium by Tomas Frejka. Demography 11 (2), 337-341.

[21] Rowthorn, R., Maciejowski, J. (2020) A cost-benefit analysis of the COVID-19 disease. Oxford Review of Economic Policy 36, 38-55.

[22] Wrzaczek, S., Kort, P. (2012) Anticipation in innovative investment under oligopolistic competition. Automatica 48, 2812-2823.

[23] Wrzaczek, S., Kuhn, M., Prskawetz, A., Feichtinger, G. (2010) The reproductive value in distributed optimal control models. Theoretical Population Biology 77, 164-170.