J. Math. Anal. Appl. 405 (2013) 71–89

Contents lists available at SciVerse ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa

Optimal immigration age-patterns in populations of fixed size

C. Simon [a], B. Skritek [b,*], V.M. Veliov [b]

[a] Wittgenstein Center for Demography and Global Human Capital (IIASA, VID/OEAW, WU), Wohllebengasse 12-14, 1040 Vienna, Austria
[b] Institute of Mathematical Methods in Economics, Vienna University of Technology, Argentinierstrasse 8, 1040 Vienna, Austria

ARTICLE INFO

Article history:
Received 30 January 2013
Available online 29 March 2013
Submitted by H. Frankowska

Keywords:
Optimal control
Age-structured system
Maximum principle
Population dynamics

ABSTRACT

Due to a low birth rate many countries need immigration to sustain a given population size. The resulting model consists of specific non-standard age-structured control systems on an infinite horizon. This paper investigates the question of optimal choice of the immigration age-profile within certain bounds. The main qualitative result is that under certain conditions the optimal control (the immigration age-profile) is time-invariant, although a dynamic framework is used. This makes it possible to give a detailed description of the optimal immigration policy and to obtain numerical solutions for the Austrian case. The analysis requires a number of new results obtained in the paper: a new type of transversality condition for the adjoint equation involved in the optimality conditions, generic conditions for well-posedness of the problem and appropriate stability conditions.

1. Introduction

Many countries face low fertility levels combined with an increase in life expectancy especially in older ages. These demographic developments influence the populations' well-being in many ways and lead for example to severe challenges for their social security systems. One possible way to counteract these developments is to steer immigration in an appropriate way.

In this paper, we consider a human population where immigration is allowed, although subjected to restrictions. It is assumed that the intensity of the migration inflow and, to a certain extend, the age-structure of the migrants can be used as control (policy) instruments. The problem we consider is to keep the size of the population constant by choosing appropriate immigration policy which, in addition, optimizes a certain objective function. Of course, the problem is meaningful only if the population would steadily decrease without migration (population with *below-replacement fertility*).

The relevance of this issue is underlined by the United Nations in [12], where the authors investigate whether immigration can be used to hinder a decline or aging of the populations of eight industrialized countries. It is concluded that immigration alone cannot stop the aging of these populations. In [11], the authors determine for a stationary population with below-replacement fertility the optimal age-specific immigration profile that minimizes the dependency ratio while fixing either the population size or the immigration quota. In contrast to [11], in this paper non-stationary populations are investigated leading to the formulation and study of a distributed control problem on an infinite horizon. The population dynamics in this case can be modeled by the McKendrick–von Foerster equation.

From mathematical point of view the considered problem is challenging for three reasons: (i) it has the form of a distributed optimal control problem with state constraints (although rather specific); (ii) the time horizon is infinite and

* Corresponding author.
E-mail addresses: christa.simon@oeaw.ac.at (C. Simon), bernhard.skritek@tuwien.ac.at (B. Skritek), veliov@tuwien.ac.at (V.M. Veliov).

0022-247X/$ – see front matter
http://dx.doi.org/10.1016/j.jmaa.2013.03.061

a theory for infinite-horizon optimal control problems for age-structured systems is missing ([6,8] are exceptions, as well as a few non-sound papers that we do not mention); (iii) we deal with a maximization problem for a non-concave functional, where the existence of a solution and the well-posedness are problematic.

The main contributions of the paper are as follows. We obtain a Pontryagin type maximum principle with a transversality condition in the form of boundedness of the adjoint variable. This is done under suitable stability assumptions which are fulfilled for populations with sufficiently low fertility. Existence of an optimal solution is also proved. The most striking result is that under an additional generic well-posedness condition for a population with time-invariant mortality and fertility the optimal age-density of the migration turns out to be time-invariant and independent of the initial data. This makes it possible to find it by solving the associated steady-state problem, which is an optimal control problem for an ordinary differential equation and was studied in details in [11]. Thanks to this property also qualitative results for the optimal policy are obtained.

As an application we consider the Austrian female population in 2009 and determine from the available data the age-specific mortality and fertility rates and the initial age structure of the population and of the migration. Then we consider the age-profile of the migration as a control (policy) variable, allowing for modifications of the age-profile from 2009. Using the results in this paper we determine numerically the immigration policy that maximizes the aggregate number of workers over time. It turns out that the optimal migration intensity is at its upper bound on a single age-interval and on its lower bound at all other ages.

For the presentation of the problem and the proofs below we use ideas from [8], where the authors investigate the recruitment problem of organizations of fixed size. Like here, a distributed control problem is involved. However, the problem in the present paper is substantially more complicated due to the involvement of births in the boundary condition.

The paper is structured as follows. In Section 2 we present the population dynamics and some auxiliary results. In Section 3 we formulate the optimization problem and in Section 4 we prove necessary optimality conditions. Stationarity, structure and uniqueness of the optimal solution are investigated in Section 5. Finally, in Section 6 we provide numerical illustrations with Austrian data before concluding in Section 7. Most of the proofs are moved to the Appendix.

2. Preliminary statements

Below $t \geq 0$ denotes time, $a \geq 0$ denotes age and $a \mapsto N(t,a) \geq 0$ is the (non-probabilistic) age-density of a population. The mortality and the fertility rate at age a of this population are denoted by $\mu(a)$ and $\varphi(a)$, respectively. The immigration flux (number of immigrants) at time t will be denoted by $R(t) \geq 0$, and the immigration age-density – by $u(t,a)$, $a \in [0,\infty)$ and $t \in [0,\infty)$. That is, u satisfies $u(t,a) \geq 0$, $\int_0^\infty u(t,a)\mathrm{d}a = 1$ and $R(t)u(t,a)$ is the flow of immigrants of age a at time t. Then the evolution of the population is described by the McKendrick–von Foerster equations (see e.g. [14])

$$\left(\frac{\partial}{\partial a} + \frac{\partial}{\partial t}\right) N(t,a) = -\mu(a)\,N(t,a) + R(t)\,u(t,a), \quad t,\ a \geq 0, \tag{1}$$

$$N(0,a) = N_0(a), \quad a \geq 0, \tag{2}$$

$$N(t,0) = \int_0^\infty \varphi(a)\,N(t,a)\mathrm{d}a, \quad t \geq 0, \tag{3}$$

where $N_0(\cdot)$ is the initial population density. The formal meaning of these equations will be given below in this section.

Given an immigration profile, $u(t,a)$, one can always keep the size of the population constant (equal to $M := \int_0^\infty N_0(a)\mathrm{d}a$) by an appropriate choice of the immigration intensity, namely by choosing $R(t)$ in the feedback form

$$R(t) := \int_0^\infty (\mu(a) - \varphi(a))\,N(t,a)\mathrm{d}a. \tag{4}$$

This fact has an obvious demographic meaning and can easily be established by integration of equation (1) with respect to a, provided that N is a differentiable function and the equation is satisfied in the classical sense. This is not necessarily the case (due to a possible inconsistency of the initial and the boundary conditions (2), (3), or due to discontinuities of u; the optimal control for the problem described in the next section is discontinuous indeed). Therefore we give a strict proof below, see Lemma 3.

Let us introduce also the number of births

$$B(t) := \int_0^\infty \varphi(a)\,N(t,a)\mathrm{d}a. \tag{5}$$

Now we pass to a strict formulation of the previous consideration, starting with some basic assumptions (BA).

(BA) The functions $\mu,\ \varphi,\ N_0 : [0,\infty) \to \mathbf{R}$ are non-negative, bounded and Lipschitz continuous; μ satisfies $\mu(a) \geq \mu_0 > 0$ for all sufficiently large a; $\varphi(a)$ and $N_0(a)$ are equal to zero for all sufficiently large a; there is $a_0 \geq 0$ such that $\varphi(a_0) > 0$ and $N_0(a) > 0$ for $a \in [0,a_0]$; N_0 satisfies $\int_0^\infty N_0(a)\mathrm{d}a = M$ with some positive $M < \infty$.

Remark 1. The Lipschitz continuity assumption is made just for technical convenience and can be relaxed. The remaining assumptions about the fertility $\varphi(a)$, the present age-density of the population, $N_0(a)$, and that the population is non-void until some fertile age a_0 are factual. The boundedness assumption for the mortality rate needs some explanation. There is no empirical evidence about boundedness or unboundedness of $\mu(a)$. We can assume equally well that the mortality rate is unbounded close to some maximal age $a = \omega$ in such a way that all the population dies till age ω. An alternative (not less plausible, in our opinion) is that $\mu(a)$ is bounded and large enough after a certain age, say 110 years, so that individuals of age above 130 exist only mathematically. We chose the second option about the mortality rate just for a minor technical convenience.

A few notations are introduced below. For any positive number T we abbreviate

$$D_T := [0, T] \times [0, \infty), \qquad D := [0, \infty) \times [0, \infty).$$

By definition the space $\mathcal{N}$ consists of all functions $N : D \to \mathbf{R}$ which are

(i) measurable and the function $t \mapsto \int_0^\infty |N(t, a)| \mathrm{d}a < \infty$ is finite and locally bounded;
(ii) locally absolutely continuous on almost every line $t - a =$ const. intersected with D (these are the characteristic lines of the differential operator in (1)).

Then for any $N \in \mathcal{N}$ the directional derivative

$$\mathcal{D}\, N(t, a) = \lim_{h \to 0} \frac{N(t + h, a + h) - N(t, a)}{h}, \tag{6}$$

is well-defined for a.e. $(t, a) \in D$.

Let $u : D \to \mathbf{R}$ be an immigration age-profile, that is u is measurable and locally bounded, $u(t, a) \geq 0$ and $\int_0^\infty u(t, a)\mathrm{d}a = 1$. Moreover, let $R : [0, \infty) \to \mathbf{R}$ be also measurable and locally bounded. Then by definition $N \in \mathcal{N}$ is a solution of (1)–(3) if the equations are satisfied almost everywhere with $\left(\frac{\partial}{\partial a} + \frac{\partial}{\partial t}\right) N$ interpreted as $\mathcal{D}N$. Notice that property (i) of the functions from $\mathcal{N}$ implies that the right-hand side of (3) makes sense, and property (ii) implies that the traces $N(0, \cdot)$ and $N(\cdot, 0)$ are (a.e.) well-defined and measurable (see [7] for more details; the above definition of a solution is equivalent to the ones commonly used in the literature, e.g. [1,14]).

Note, that since we consider a human population, negative values for R and N make no sense. However, so far it is not clear whether a solution to the system exists nor that it is non-negative. In the following lemmas, existence, boundedness and non-negativity as well as fixed size of a solution are discussed, both for the original system and R chosen in the feedback form (4).

Lemma 1. *Let u and R be fixed as above. Then system (1)–(3) has a unique solution $N \in \mathcal{N}$ and $N \in L_\infty(D_T)$ for every $T > 0$. The function B is locally bounded.*

The proofs of this and of the next lemmas in this section will be given in the Appendix.

In parallel we consider the system

$$\mathcal{D}N(t, a) = -\mu(a)\, N(t, a) + \int_0^\infty (\mu(s) - \varphi(s))\, N(t, s)\mathrm{d}s\, u(t, a) \tag{7}$$

with side conditions (2) and (3). The meaning of a solution $N \in \mathcal{N}$ is the same as for (1)–(3), regarding the fact that the integral on the right-hand side of (7) is well-defined and finite due to property (i) of $\mathcal{N}$.

Lemma 2. *Let u be fixed as above. Then Eq. (7) with side conditions (2)–(3) has a unique solution $N \in \mathcal{N}$ and N is (essentially) bounded on every subset $D_T \subset D$, $0 < T < \infty$. Moreover, the functions R and B, defined by (4) and (5), are locally Lipschitz continuous.*

Lemma 3. *Let u and R be as above and let N be the unique solution of (1)–(3). Then the population N has a fixed size (that is, $\int_0^\infty N(t, a)\mathrm{d}a = M$) if and only if the function R satisfies (4). In this case N coincides with the unique solution of (7) with side conditions (2)–(3).*

Lemma 4. *Let u be fixed as above. Assume that for the unique solution $N \in \mathcal{N}$ of (7) with side conditions (2)–(3) it holds that $R(t) \geq 0$ for every $t \geq 0$, where R is defined by (4). Then the functions N, B and R are non-negative and bounded, uniformly with respect to u as above for which the assumption $R(t) \geq 0$ is fulfilled.*

3. The optimization problem

The main aim of this paper is to determine optimal age patterns of immigrants in a population of fixed size. The specific optimization problem that we will introduce below arises only for populations that need a positive immigration in order

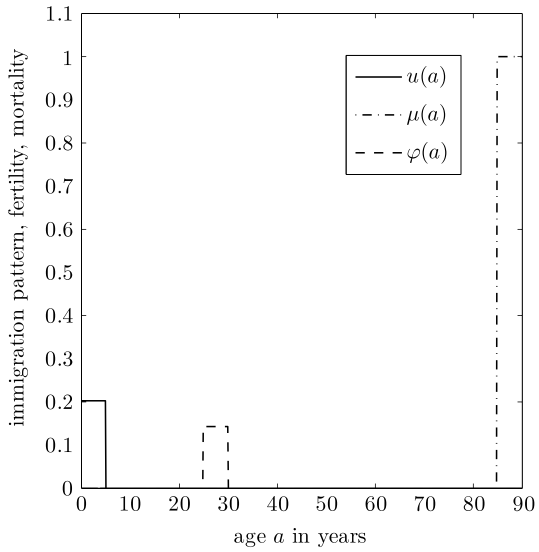

Fig. 1. Functions $\varphi(a)$ (dashed line), $\mu(a)$ (dashed–dotted line) and $u(a)$ (solid line).

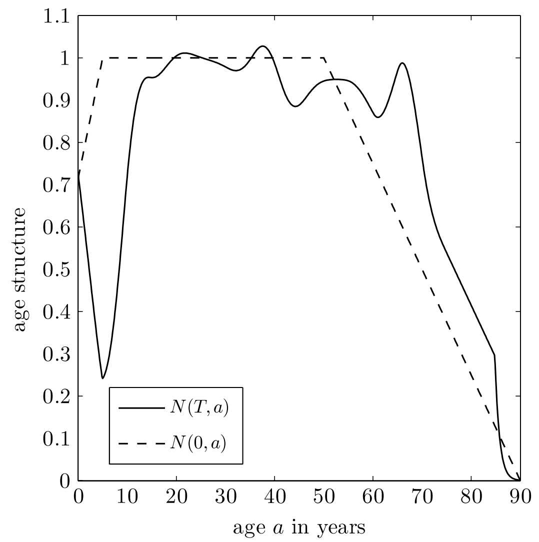

Fig. 2. The initial $N_0(a)$ (dashed line) and the age structure $N(T, a)$ (solid line) at $T = 90$.

to sustain their size, as it is the case for most European countries. Many of these countries had to face below-replacement fertility

$$\int_0^\infty \varphi(a)\, e^{-\int_0^a \mu(\theta)\mathrm{d}\theta} \mathrm{d}a < 1, \tag{8}$$

for a long period. Below replacement fertility implies extinction of the population without immigration and thus immigration is needed to sustain the size. However, it does not imply that the native population will decrease in the short run, therefore negative immigration, i.e. emigration, $R(t) < 0$ may be needed for some t (see [10]). In Figs. 1–3 we provide an illustrative example which shows that the immigration rate $R(t)$ determined by (4) takes negative values for some t several generations after the initial time, although fertility and mortality satisfy condition (8). Fig. 1 presents the fertility, mortality and the immigration profile $u(a)$. The initial population $N_0(a)$ and the population $N(T, a)$ at time $T = 90$ are depicted in Fig. 2.

Since in the present paper we use immigration as a policy instrument and negative immigration is not admissible, we have to eliminate this possibility by introducing assumption (A2), which is stronger than the below-replacement fertility condition. In practice, discrimination of immigrants will not happen based on age only. Nevertheless, the age of applicants

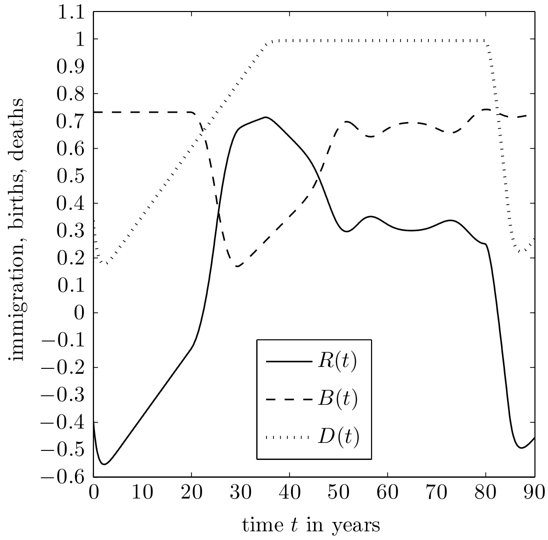

Fig. 3. The number of births $B(t)$ (dashed line), the number of deaths $D(t)$ (dotted line) and the number of immigrants $R(t)$ (solid line) over time. Below-replacement fertility is not enough to ensure $R(t) > 0$.

for a visa is taken into account, for example, by the Australian authorities.[1] There, in the skilled point test, 60 points are needed for a working permit and 30 of those can be gained by being a member of the age group ranging from 25 to 29, while for age 45+ zero points are awarded.

Let $m(a)$ be the present immigration profile, that is, at time $t = 0$, which is historically determined by habits, policies or other factors. Then the present normalized age-density of immigration is given by

$$\hat{u}(a) := \frac{m(a)}{\int_0^\infty m(a)\mathrm{d}a}. \tag{9}$$

Therefore, when using the age-density of the immigration as a control (policy) variable we can implement only slight changes in $\hat{u}(a)$. For this reason we consider control constraints of the form $\underline{u}(a) \le u(t, a) \le \bar{u}(a)$, where the lower and the upper bounds are not much different from the present values $\hat{u}(a)$, say $\underline{u}(a) = (1 - \varepsilon)\hat{u}(a)$ and $\bar{u}(a) = (1 + \varepsilon)\hat{u}(a)$ with some small $\varepsilon > 0$.

The optimization problem we consider in what follows is:

$$\max_{R,u} \int_0^\infty e^{-rt} \left[\int_0^\infty p(a)\, N(t, a)\mathrm{d}a - q\, R(t) \right] \mathrm{d}t, \tag{10}$$

subject to

$$\mathcal{D}N(t, a) = -\mu(a)N(t, a) + R(t)\, u(t, a), \quad (t, a) \in D, \tag{11}$$

$$N(0, a) = N_0(a), \quad a \ge 0, \tag{12}$$

$$N(t, 0) = \int_0^\infty \varphi(a)\, N(t, a)\mathrm{d}a, \quad t \ge 0, \tag{13}$$

$$\int_0^\infty N(t, a)\mathrm{d}a = M, \tag{14}$$

$$\underline{u}(a) \le u(t, a) \le \bar{u}(a), \qquad \int_0^\infty u(t, a)\mathrm{d}a = 1, \tag{15}$$

$$R(t) \ge 0. \tag{16}$$

Here, function $p(a)$ is a weight function that is higher if people of a certain age are more valuable from point of view of the policy maker solving this optimization problem. For example, $p(a)$ could be the function taking the value 1 for ages $a \in [20, 65]$, representing the working ages, and 0 otherwise. The second term penalizes the size of immigration (if $q > 0$).

The intertemporal discount rate is r. The constant q represents the benefits or costs of immigration arising, for example, from possible integration or education expenditures. If $q = 0$ then maximizing the performance value is related to

[1] www.visabureau.com/australia.

minimizing the dependency ratio of the population (considered in a steady state in [11]), that is the fraction of non-workers to workers in a population, which is a measure of how solvent a social security system is.

Additionally to (BA) we make the following assumptions.

(A1) The discount rate r is strictly positive, the function $p : [0, \infty) \to \mathbf{R}$ is measurable and bounded; q is a real number; the functions $\underline{u}, \bar{u} : [0, \infty) \to \mathbf{R}$ are measurable and bounded, and satisfy the relations $0 \le \underline{u}(a) \le \bar{u}(a)$ for all $a \ge 0$, $\bar{u}(a) = 0$ for all sufficiently large a; moreover $\int_0^\infty \underline{u}(a)\mathrm{d}a < 1$ and $\int_0^\infty \bar{u}(a)\mathrm{d}a > 1$.

According to Lemma 3 we can reformulate problem (10)–(16) in the following way:

$$\max_{u\in\mathcal{U}} J(u) := \int_0^\infty e^{-rt}\left[\int_0^\infty [p(a) - q(\mu(a) - \varphi(a))]\, N(t, a)\mathrm{d}a\right]\mathrm{d}t, \tag{17}$$

$$\mathcal{D}N(t, a) = -\mu(a)\, N(t, a) + u(t, a)\int_0^\infty (\mu(s) - \varphi(s))\, N(t, s)\mathrm{d}s, \quad (t, a) \in D, \tag{18}$$

$$N(0, a) = N_0(a), \quad a \ge 0, \tag{19}$$

$$N(t, 0) = \int_0^\infty \varphi(a)\, N(t, a)\mathrm{d}a, \quad t \ge 0. \tag{20}$$

Here, the set of admissible controls $\mathcal{U}$ is defined as

$$\mathcal{U} := \left\{u : D \to \mathbf{R} : \underline{u}(a) \le u(t, a) \le \bar{u}(a), \int_0^\infty u(t, a)\mathrm{d}a = 1\right\}. \tag{21}$$

Due to the last requirement in (A1) the set of admissible controls, $\mathcal{U}$, is nonempty.

The condition for a non-negative immigration rate, $R(t) \ge 0$, is disregarded in the above reformulation. It will be stipulated by the following additional assumption.

(A2) For any $u \in \mathcal{U}$ the immigration intensity R, defined by (4) for the corresponding solution of (18)–(20), is strictly positive for all t.

Even with below-replacement fertility, see (8), it could happen that $R(t) < 0$ for some t. We assume that for the present immigration pattern $\hat{u}(a)$ the resulting immigration size satisfies $\hat{R}(t) \ge R_0 > 0$. Implicitly this property requires that the initial density $N_0(a)$ results from a population which has experienced below-replacement fertility for quite a while before the present time $t = 0$. This is the situation in most of the European countries in the 21st century, for example, as in our case study in Section 6. We assume a bit more, namely that $R(t) > 0$ for any admissible control u, having in mind that all admissible controls are close to $\hat{u}$.

Since $R(t) > 0$, Lemma 4 together with $r > 0$ and the boundedness of p, μ and φ, imply that $J(u)$ is finite for every $u \in \mathcal{U}$ and that $\sup_{u\in\mathcal{U}} J(u)$ is finite. Thanks to this we can use the standard definition of optimality: $u \in \mathcal{U}$ is optimal if $J(u) \ge J(v)$ for every $v \in \mathcal{U}$.

Proposition 1. *Let assumptions* (BA), (A1), (A2) *be fulfilled. Then the optimal control problem* (17)–(20) *has a solution.*

The proof is given in the Appendix. We mention that this proof is not routine since we deal with a problem of maximization of a non-concave functional. Indeed, the mapping $\mathcal{U} \ni u \longrightarrow$ "objective value $J(u)$" is not concave, as argued in [8] even in the substantially simpler case $\varphi = 0$.

4. Necessary optimality conditions

In this section we formulate and prove necessary optimality conditions of Pontryagin's type for problem (17)–(20). The problem at hand has a similar structure as those studied in [5,7,13], with the substantial difference that here the time-horizon is infinite. To our knowledge there are no results in the literature that provide necessary optimality conditions for age-structured optimal problems on infinite horizon, except the ones mentioned in the introduction, which are not applicable for problem (17)–(20).

The Pontryagin-type necessary optimality conditions involve (i) an appropriate *adjoint equation*; (ii) an appropriate *transversality condition* that uniquely determines a solution of the adjoint equation; (iii) a *maximization condition* for each t separately. The word "appropriate" in (i) and (ii) means that the maximization condition in (iii) holds true with the "appropriate" adjoint function.

The appropriate adjoint equation associated with our problem will be shown to have the form

$$\begin{aligned}\mathcal{D}\,\xi(t, a) = {} & (r + \mu(a))\,\xi(t, a) - \varphi(a)\,\xi(t, 0) - (\mu(a) - \varphi(a)) \\ & \times \int_0^\infty \xi(t, \alpha)\, u(t, \alpha)\mathrm{d}\alpha - p(a) + q(\mu(a) - \varphi(a)).\end{aligned} \tag{22}$$

A main challenge is to define an appropriate transversality condition. Following ideas originating in [4] and developed in [3] for ordinary differential systems and also in [8] for a problem which is similar but substantially simpler than (17)–(20), we introduce the "transversality" condition $\|\xi\|_\infty < \infty$. This is justified by Proposition 2 and Theorem 1.

We start with some preliminary results and an additional assumption.

In the following (u, N) denotes an optimal solution of problem (17)–(20) and R and B are the number of immigrants and births given by (4) and (5).

We introduce the notations

$$\rho(a) := r + \mu(a), \qquad \nu(a) := \mu(a) - \varphi(a), \qquad f(a) := p(a) - q\,(\mu(a) - \varphi(a))$$

and the auxiliary variables

$$\lambda(t) := \xi(t,0), \qquad \eta(t) := \int_0^\infty \xi(t,a)\,u(t,a)\mathrm{d}a. \tag{23}$$

Then the adjoint equation becomes

$$\mathcal{D}\,\xi(t,a) = \rho(a)\,\xi(t,a) - \varphi(a)\,\lambda(t) - \nu(a)\,\eta(t) - f(a). \tag{24}$$

Lemma 5. *Let assumptions* (BA), (A1), (A2) *be fulfilled. Then, for any given functions* $\lambda,\ \eta \in L_\infty(0,\infty)$ *Eq.* (24) *has a unique bounded solution on* D *and it is given by the formula*

$$\xi(t,a) = \int_a^\infty e^{-\int_a^s \rho(\theta)\mathrm{d}\theta}\,[\varphi(s)\,\lambda(s+t-a) + \nu(s)\,\eta(s+t-a) + f(s)]\mathrm{d}s, \tag{25}$$

where the integral is absolutely convergent.

Proof. The integral in (25) is absolutely convergent and the function ξ is bounded due to the boundedness of the term in brackets and the inequality

$$\int_a^\infty e^{-\int_a^s \rho(\theta)\mathrm{d}\theta}\mathrm{d}s \le \frac{1}{r}.$$

One can verify by substitution that ξ defined by (25) satisfies (24).

To prove the uniqueness assertion we consider the difference $\Delta\xi(t,a)$ between two bounded solutions which is also bounded. It satisfies the equation

$$\mathcal{D}\,\Delta\xi(t,a) = \rho(a)\,\Delta\xi(t,a).$$

If $\Delta\xi(t,a) \neq 0$ for some t and a, then the function $x(s) := \Delta\xi(t+s, a+s)$, $s \ge 0$, satisfies $\dot{x}(s) = \rho(a+s)\,x(s)$ with $x(0) = \Delta\xi(t,a)$. Due to $\rho(a+s) \ge r > 0$, $x(s)$ is unbounded since $x(0) = \Delta\xi(t,a) \neq 0$. This contradiction completes the proof. □

Similarly as in the proof of Lemma 2 in the Appendix, one can obtain by substituting (25) in (23) that ξ is a bounded solution of (22) if and only if it is the unique bounded solution of (25) with the functions λ and η determined as bounded solutions of the following equations:

$$\lambda(t) = \int_t^\infty e^{-\int_0^{s-t}\rho(\tau)\mathrm{d}\tau}\,[\varphi(s-t)\,\lambda(s) + \nu(s-t)\,\eta(s) + f(s-t)]\mathrm{d}s, \tag{26}$$

$$\eta(t) = \int_t^\infty \int_0^\infty u(t,a)\,e^{-\int_a^{a+s-t}\rho(\tau)\mathrm{d}\tau}\,[\varphi(a+s-t)\,\lambda(s) + \nu(a+s-t)\,\eta(s) + f(a+s-t)]\mathrm{d}a\,\mathrm{d}s. \tag{27}$$

This system can be written as

$$x(t) = \int_t^\infty k(t,s)x(s)\mathrm{d}s + F(t), \tag{28}$$

where $x = (\lambda, \eta)$, $k(t,s) = \big(k_{i,j}(t,s)\big)$ is the matrix

$$k(t,s) = \begin{pmatrix} e^{-\int_0^{s-t}\rho(\tau)\mathrm{d}\tau}\varphi(s-t) & e^{-\int_0^{s-t}\rho(\tau)\mathrm{d}\tau}\nu(s-t) \\ \int_0^\infty u(t,a)\,e^{-\int_a^{a+s-t}\rho(\tau)\mathrm{d}\tau}\varphi(a+s-t)\mathrm{d}a & \int_0^\infty u(t,a)\,e^{-\int_a^{a+s-t}\rho(\tau)\mathrm{d}\tau}\nu(a+s-t)\mathrm{d}a \end{pmatrix}, \tag{29}$$

$$F(t) = \begin{pmatrix} \int_t^\infty e^{-\int_0^{s-t}\rho(\tau)\mathrm{d}\tau} f(s-t)\mathrm{d}s \\ \int_t^\infty \int_0^\infty u(t,a)\,e^{-\int_a^{a+s-t}\rho(\tau)\mathrm{d}\tau} f(a+s-t)\mathrm{d}a\,\mathrm{d}s \end{pmatrix}.$$

A key point in the subsequent analysis is that integral equation (28) has a unique bounded solution. This, however requires an additional assumption about the kernel, which is formulated in terms of the numbers κ introduced below:

$$\kappa_{11} := \int_0^\infty e^{-\int_0^a \rho(\theta)\mathrm{d}\theta}\varphi(a)\mathrm{d}a, \qquad \kappa_{12} := \int_0^\infty e^{-\int_0^a \rho(\theta)\mathrm{d}\theta}|\nu(a)|\mathrm{d}a, \tag{30}$$

$$\kappa_{21} := \max_{a\geq 0}\int_0^\infty e^{-\int_a^{a+\tau} \rho(\theta)\mathrm{d}\theta}\varphi(a+\tau)\mathrm{d}\tau, \qquad \kappa_{22} := \max_{a\geq 0}\int_0^\infty e^{-\int_a^{a+\tau} \rho(\theta)\mathrm{d}\theta}|\nu(a+\tau)|\mathrm{d}\tau. \tag{31}$$

The following condition ensures that the integral operator in (28) is contractive in an appropriate norm:

(A3) The following inequality is fulfilled

$$\frac{1}{2}\left[\kappa_{11}+\kappa_{22}+\sqrt{(\kappa_{11}-\kappa_{22})^2+4\,\kappa_{12}\,\kappa_{21}}\right] < 1. \tag{32}$$

Note, that (A3) imposes a condition solely on the data of the problem. This condition is not only used in the proof of the optimality condition below, but also for the characterization of the optimal solution in the next section (see the proof of Lemma 7). The results in this section are valid under an alternative condition that involves the set of admissible controls and the numbers

$$\bar{\kappa}_{21} := \int_0^\infty u^*(a)\int_0^\infty e^{-\int_a^{a+\tau} \rho(\theta)\mathrm{d}\theta}\varphi(a+\tau)\mathrm{d}\tau\mathrm{d}a,$$
$$\bar{\kappa}_{22} := \int_0^\infty u^*(a)\int_0^\infty e^{-\int_a^{a+\tau} \rho(\theta)\mathrm{d}\theta}|\nu(a+\tau)|\mathrm{d}\tau\mathrm{d}a,$$

where $\hat{u}$ is a reference time-invariant control (see (9) and the explanations there around).

(A3′) For some $\varepsilon > 0$ it holds that

$$\bar{u}(a) \leq (1+\varepsilon)\hat{u}(a), \quad a \geq 0,$$

and the following inequality is fulfilled:

$$\frac{1}{2}\left[\kappa_{11}+(1+\varepsilon)\bar{\kappa}_{22}+\sqrt{(\kappa_{11}-(1+\varepsilon)\bar{\kappa}_{22})^2+4\,(1+\varepsilon)\kappa_{12}\,\bar{\kappa}_{21}}\right] < 1. \tag{33}$$

Remark 2. Assumption (A3) and (A3′) implicitly require that $\kappa_{11} < 1$ which is equivalent to the below-replacement fertility condition if $r = 0$. For $r > 0$ the below-replacement fertility condition is stronger than $\kappa_{11} < 1$. We mention also that $\bar{\kappa}_{21} \leq \kappa_{21}$ and $\bar{\kappa}_{22} \leq \kappa_{22}$, so that (A3′) may be a weaker assumption than (A3) if ε is sufficiently small.

Lemma 6. *Under* (BA) *and* (A1)–(A3)*, system* (26)*,* (27) *has a unique solution in* $L_\infty(0,\infty)$*. The same is true also under* (BA), (A1), (A2), (A3′).

The proof will be given in the Appendix.

As a consequence of the above two lemmas in combination, we obtain the following proposition.

Proposition 2. *Under* (BA) *and* (A1)–(A3)*, the adjoint equation* (22) *has a unique solution in* $L_\infty(D)$*. The same is true also under* (BA), (A1), (A2), (A3′).

Proof. According to Lemma 6 system (26), (27) has a bounded solution. Then according to Lemma 5 Eq. (22) also has a bounded solution, obtained by substitution of the solution (λ,η) of (26), (27) in (24).

For any bounded solution ξ of (22) the functions λ and η defined by (23) are bounded and satisfy (26), (27). Therefore λ and η are uniquely determined (Lemma 6), hence ξ is unique (Lemma 5). □

The next theorem gives a necessary optimality condition of the type of the Pontryagin maximum principle for problem (17)–(20).

Theorem 1. *Let assumptions* (BA), (A1)–(A3) *(or alternatively* (BA), (A1), (A2), (A3′)*) be fulfilled and let* (u, N) *be an optimal solution of problem* (17)–(20)*. Let* ξ *be the unique solution in* $L_\infty(D)$ *of the adjoint equation* (22)*. Then for a.e.* $t \geq 0$ *the optimal control* $u(t,\cdot)$ *maximizes the integral*

$$\int_0^\infty \xi(t,a)v(a)\mathrm{d}a,$$

on the set of measurable functions $v(\cdot)$ *satisfying*

$$\underline{u}(a) \leq v(a) \leq \bar{u}(a), \qquad \int_0^\infty v(a)\mathrm{d}a = 1. \tag{34}$$

Proof. Let J be the optimal objective value and let ξ be the unique bounded solution of the adjoint equation (22) on D (see Proposition 2). Let us fix an arbitrary $\theta > 0$, let $h > 0$ be arbitrary (and presumably small) and $T > 0$ be such that $\theta - h \geq 0$ and $\theta + h \leq T$. Denote $\Theta_h := [\theta - h, \theta + h] \times [0, \infty) \subset D$ and define a "disturbed" control

$$\tilde{u}(t, a) := \begin{cases} u(t, a) & \text{for } (t, a) \notin \Theta_h, \\ v(a) & \text{for } (t, a) \in \Theta_h, \end{cases} \tag{35}$$

where v is any measurable function satisfying (34).

Then $\tilde{u}$ satisfies the control constraints. Let $\tilde{N}$ be the corresponding solution of (18)–(20) and $\tilde{R}$, $\tilde{B}$ be corresponding functions (immigration and birth flows) defined by (4) and (5), while R and B correspond to N. Denote $\Delta J = J(\tilde{u}) - J(u)$, $\Delta u = \tilde{u} - u$, $\Delta N = \tilde{N} - N$, $\Delta R = \tilde{R} - R$, all depending on the chosen h and v. According to (A2), $\tilde{R}$ is non-negative, and all the functions introduced above are bounded (see Lemma 4).

Clearly,

$$\Delta J = \int_0^T e^{-rt} \int_0^\infty f(a)\, \Delta N(t, a) \mathrm{d}a \mathrm{d}t. \tag{36}$$

In order to obtain an expression for ΔN we multiply the equation

$$\mathcal{D} \Delta N(t, a) = -\mu(a)\, \Delta N(t, a) + \int_0^\infty \nu(\alpha)\, [\Delta N(t, s)\, u(t, a) + N(t, s)\, \Delta u(t, a) + \Delta N(t, s) \Delta u(t, a)] \mathrm{d}s \tag{37}$$

resulting from (18) by $e^{-rt}\xi(t, a)$ and integrate on D. Since $D = \{(s, x + s) : s, x \geq 0\} \cup \{(x + s, s) : s, x \geq 0\}$ and the two sets on the right intersect only on a set of measure zero, we may represent

$$\begin{aligned} \int_0^\infty \int_0^\infty \mathcal{D} \Delta N(t, a) e^{-rt} \xi(t, a) \mathrm{d}t \mathrm{d}a = &\int_0^\infty \int_0^\infty e^{-rs} \xi(s, x + s) \frac{\mathrm{d}}{\mathrm{d}s} \Delta N(s, x + s) \mathrm{d}s \mathrm{d}x \\ &+ \int_0^\infty \int_0^\infty e^{-r(x+s)} \xi(x + s, s) \frac{\mathrm{d}}{\mathrm{d}s} \Delta N(x + s, s) \mathrm{d}s \mathrm{d}x. \end{aligned} \tag{38}$$

By integration by parts the first term on the right-hand side gives

$$\int_0^\infty \left[e^{-rs} \xi(s, x + s) \Delta N(s, x + s) \mid_{s=0}^\infty - \int_0^\infty \Delta N(s, x + s) e^{-rs} \left(-r\xi(s, x + s) + \frac{d}{ds} \xi(s, x + s) \right) \mathrm{d}s \right] \mathrm{d}x.$$

The term with $s \to \infty$ is zero because both ξ and N (and therefore ΔN) are bounded, and the term with $s = 0$ is zero because $\Delta N(0, a) = 0$. The second term on the right-hand side of (38) is treated in the same way and combining the two terms we obtain that the right-hand side of (38) is equal to

$$-\int_0^\infty e^{-rt} \xi(t, 0) \Delta N(t, 0) \mathrm{d}t - \int_0^\infty \int_0^\infty e^{-rt} (\mathcal{D}\xi(t, a) - r\xi(t, a)) \Delta N(t, a) \mathrm{d}a \mathrm{d}t. \tag{39}$$

Then taking into account that $N(t, 0) = \int_0^\infty \varphi(a) N(t, a) \mathrm{d}a$ we obtain from (38), by using (11) and rewriting it again in the (t, a)-plane, the equality

$$\begin{aligned} 0 = \int_0^\infty \int_0^\infty e^{-rt} \Bigg[&\xi(t, 0)\varphi(a) + (\mathcal{D}\xi(t, a) - r\xi(t, a) - \mu(a)\xi(t, a)) \Delta N(t, a) \\ &+ \int_0^\infty \nu(s) \xi(t, a) \left(\Delta N(t, s) u(t, a) + N(t, s) \Delta u(t, a) + \Delta N(t, s) \Delta u(t, a) \right) \mathrm{d}s \Bigg] \mathrm{d}a \mathrm{d}t. \end{aligned}$$

Using the adjoint equation (22) we obtain that

$$0 = \int_0^\infty \int_0^\infty e^{-rt} \left[-f(a) \Delta N(t, a) + \int_0^\infty \left(\nu(s) \xi(t, a) N(t, s) \Delta u(t, a) + \nu(s) \xi(t, a) \Delta N(t, s) \Delta u(t, a) \right) \mathrm{d}s \right] \mathrm{d}a \mathrm{d}t.$$

Adding this to (36) we get

$$\begin{aligned} \Delta J &= \int_0^\infty \int_0^\infty \int_0^\infty e^{-rt} \nu(s) \xi(t, a) \Big[N(t, s) \Delta u(t, a) + \Delta N(t, s) \Delta u(t, a) \Big] \mathrm{d}s \mathrm{d}a \mathrm{d}t \\ &= \int_0^\infty \int_0^\infty e^{-rt} R(t) \xi(t, a) \Delta u(t, a) \mathrm{d}a \mathrm{d}t + \int_0^\infty \int_0^\infty \int_0^\infty e^{-rt} \nu(s) \xi(t, a) \Delta N(t, s) \Delta u(t, a) \mathrm{d}s \mathrm{d}a \mathrm{d}t. \end{aligned} \tag{40}$$

Next we shall show that the second term on the right-hand side above is of second order with respect to h (cf. (35)). Note, that $\Delta u(t,a) = 0$ for $t \notin [\theta - h, \theta + h]$, thus we need an estimation of ΔN only on the time-horizon $[0, T]$ with some $T > \theta$, say $T = \theta + 1$.

By solving Eq. (18) along the characteristic lines we obtain the representation

$$N(t,a) = B(t-a)e^{-\int_{0\wedge(a-t)}^{a}\mu(\tau)\mathrm{d}\tau} + \int_{0\wedge(t-a)}^{t} e^{-\int_{a-t+s}^{a}\mu(\tau)\mathrm{d}\tau}R(s)u(s,a-t+s)\mathrm{d}s,$$

where B is extended as $B(t) = N_0(-t)$ for $t < 0$ and $0 \wedge \alpha := \max\{0, \alpha\}$. Due to assumption (BA) we can estimate $e^{-\int_{\alpha}^{a}\mu(\tau)\mathrm{d}\tau} \leq 1$. A similar equality holds for $\tilde{N}(t,a)$ corresponding to the control $\tilde{u}$. Subtracting the two expressions we obtain

$$|\Delta N(t,a)| \leq |\Delta B(t-a)| + \int_{0\wedge(t-a)}^{t} \left|\Delta R(s)\tilde{u}(s,a-t+s) + R(s)\Delta u(s,a-t+s)\right|\mathrm{d}s. \tag{41}$$

From (4), (5) we can estimate $|\Delta R(t)|$ and $|\Delta B(t)|$ by $c_1\int_0^\infty |\Delta N(t,a)|\mathrm{d}a = c_1\|\Delta N(t,\cdot)\|_{L_1}$, where c_1 is a constant depending only on φ and μ. Then it is a matter of routine estimations (taking into account that Δu is non-zero on a set of measure proportional to h) to obtain the inequality

$$\|\Delta N(t,\cdot)\|_1 \leq \int_0^t c_2\|\Delta N(x,\cdot)\|_1 e^{-(t-s)\mu_0}\mathrm{d}x + c_3 h, \quad t \in [0, \theta+1],$$

where c_2 and c_3 are independent of h (although may depend on θ and the data of the problem).

Since $T = \theta + 1$ is finite, Gronwall's lemma gives

$$\|\Delta N(t,\cdot)\|_1 \leq Ch, \quad t \leq T.$$

Now it is straightforward to estimate the last term in (40) by

$$\int_{\theta-h}^{\theta+h} e^{-rt}C\,h\,\|v\|_{L_\infty}\|\xi\|_{L_\infty}\int_0^{\bar{a}} |\Delta u(t,\alpha)|\mathrm{d}\alpha\mathrm{d}t \leq C\,h^2.$$

Using this in the estimation (40) and having in mind that $\Delta J \leq 0$ due to the optimality of u we obtain that

$$\frac{1}{2h}\int_{\theta-h}^{\theta+h} e^{-rt}\int_0^\infty R(t)\,\xi(t,a)\,u(t,a)\mathrm{d}a\mathrm{d}t \geq \frac{1}{2h}\int_{\theta-h}^{\theta+h} e^{-rt}\int_0^\infty R(t)\,\xi(t,a)\,v(a)\mathrm{d}a\mathrm{d}t - \frac{C}{2}\,h.$$

Almost every s is a Lebesgue point of the function $t \to \int_0^\infty R(t)\xi(t,a)u(t,a)\mathrm{d}a$, and $R(t) > 0$. Therefore, we can conclude the proof of the theorem by taking the limit $h \to 0$. □

5. Uniqueness, stationarity, and structure of the optimal immigration pattern

In this section we use Theorem 1 to obtain some qualitative properties of the optimal solution of problem (17)–(20). The most interesting one is that the optimal control u (that is, the optimal immigration profile) is unique and time-invariant: $u(t,a) = u(a)$. This fact is not evident. Its proof is based on stability condition (A3) and on an additional well-posedness condition, which implies also a bang–bang structure of the optimal control $u(a)$.

To prove uniqueness and stationarity of the optimal solution we rewrite the adjoint equation in a feedback form. To do this we introduce the functional $\sigma(\cdot)$:

$$\sigma(g) = \max_{v\in\mathcal{V}} \int_0^\infty g(a)v(a)\,\mathrm{d}a, \quad g \in L_\infty(0,\infty), \tag{42}$$

where $\mathcal{V}$ is the set of functions $v : [0,\infty) \to \mathbf{R}$ satisfying (34). Then using the optimization condition in Theorem 1 we can rewrite the adjoint equation (22) in the feedback form

$$D\xi(t,a) = (r + \mu(a))\xi(t,a) - \varphi(a)\xi(t,0) - (\mu(a) - \varphi(a))\sigma(\xi(t,\cdot)) - f(a). \tag{43}$$

The existence of a solution in $L_\infty(D)$ to this equation follows from the necessity of the maximum principle.

Lemma 7. *If assumption* (A2) *is fulfilled, then Eq.* (43) *has a unique bounded solution.*

The proof is given in the Appendix.

Now we introduce a *regularity assumption* that ensures that the maximization condition in Theorem 1 determines a unique control. As shown in [8] in a simpler version of the problem considered here, without a certain regularity assumption the uniqueness fails and this is due to the non-concavity of the problem. On the other hand the regularity assumption is in a reasonable sense generic and easy to check.

(A4) For all real numbers d_0, d_1 and d_2 it holds that

$$\text{meas}\{a \in [0, \infty] : d_0 + d_1\mu(a) + d_2\varphi(a) - p(a) = 0\} = 0.$$

This assumption requires that μ, φ and p must not be linearly related on a set of positive measure.

Theorem 2. *Let assumptions* (BA), (A1)–(A4) *be fulfilled. Then optimal control problem* (17)–(21) *has a unique optimal control u and it is time-invariant: $u(t, a) = u(a)$.*

Proof. First we shall prove that (43) has a stationary bounded solution $\hat{\xi}(t, a) = \hat{\xi}(a)$. To do this we show that the equation

$$\xi'(a) = (r + \mu(a))\xi(a) - \varphi(a)\xi(0) - (\mu(a) - \varphi(a))\sigma(\xi(\cdot)) - f(a) \tag{44}$$

has a bounded solution. Denote $\lambda = \xi(0)$ and $\eta = \sigma(\xi(\cdot))$, then we can write the solution to the differential equation as

$$\xi(a) = \int_a^\infty e^{-\int_a^s \rho(\theta)\mathrm{d}\theta}[\varphi(s)\lambda + \nu(s)\eta + f(s)]\mathrm{d}s.$$

Using the definition of $\sigma(\xi(\cdot))$, (42), the equations for η and λ are

$$\lambda = \int_0^\infty e^{-\int_0^s \rho(\tau)\mathrm{d}\tau}[\varphi(s)\lambda + \nu(s)\eta + f(s)]\mathrm{d}s$$

$$\eta = \max_{v\in\mathcal{V}} \int_0^\infty v(a) \int_a^\infty e^{-\int_a^s \rho(\tau)\mathrm{d}\tau}[\varphi(s)\lambda + \nu(s)\eta + f(s)]\mathrm{d}s\mathrm{d}a.$$

Denoting the terms independent from λ and η by (b_1, b_2) we can write the equations above as

$$(I - K)\begin{pmatrix}\lambda\\ \eta\end{pmatrix} = \begin{pmatrix}b_1\\ b_2\end{pmatrix}, \tag{45}$$

where I is the 2×2 identity matrix and K is the matrix defined by the right hand side. As in the proof of Lemmas 6 and 7, define a norm on $\mathbf{R}^2$ as $\|(x, y)\| = \max\{|x|, \alpha|y|\}$, $\alpha > 0$. Estimating the operator norm of K in the same way as in the proof of Lemma 6 gives that the norm is smaller or equal to the left hand side in (33). Then assumption (A3) states that the norm is smaller than one, thus $(I - K)$ is invertible and therefore a unique solution to (45) exists. Thus, a bounded solution $\hat{\xi}(a)$ of (44) exists and it is obviously a stationary bounded solution of (43).

According to Lemma 7 the stationary function $\hat{\xi}(a)$ is the unique bounded solution of (43). On the other hand, Theorem 1 claims that for every optimal control u, the adjoint equation (22) has a unique bounded solution $\xi(t, a)$ and for a.e. $t \geq 0$

$$\int_0^\infty \xi(t, a)u(t, a)\mathrm{d}a = \max_{v\in\mathcal{V}} \int_0^\infty \xi(t, a)v(a)\mathrm{d}a = \sigma(\xi(t, \cdot)).$$

Then ξ is a bounded solution also of (43), which implies that $\xi = \hat{\xi}$. The above maximization condition reads now as

$$\int_0^\infty \hat{\xi}(a)u(t, a)\mathrm{d}a = \max_{v\in\mathcal{V}} \int_0^\infty \hat{\xi}(a)v(a)\mathrm{d}a, \tag{46}$$

where $\hat{\xi}$ is the unique bounded solution of (44).

Assumption (A4) obviously implies that the solution $\hat{\xi}$ of (44) cannot be constant on a set of positive measure. Then similarly as in Corollary 5.1 in [8] one can prove that (46) uniquely determines (modulo a set of measure zero) a control $u \in \mathcal{U}$, it is time-invariant and has the following structure: there is a real number l such that

$$u(t, a) = \begin{cases}\underline{u}(a) & \text{if } \hat{\xi}(a) \leq l,\\ \bar{u}(a) & \text{if } \hat{\xi}(a) > l. \quad \square\end{cases} \tag{47}$$

We formulate the last finding in the proof of the above theorem as a corollary.

Corollary 1. *Let $\hat{\xi}$ be the unique bounded solution of* (44)*. Then, there is $l \in \mathbf{R}$ such that the unique optimal control $u(t, a) = \hat{u}(a)$ is determined by* (47)*. This number l is the only one for which the resulting $\hat{u}$ satisfies $\int_0^\infty \hat{u}(a)\mathrm{d}a = 1$.*

Thus the optimal solution is of bang–bang type. A related result is obtained in [11] for a static counterpart of the problem considered in this paper. Since $\xi(a)$ can be interpreted marginally as "shadow price" of an a-year-old individual, the above corollary asserts that there is a critical value l such that it is optimal to encourage as much as possible migration in ages for which the shadow price is higher than l $(u(a) = \bar{u}(a))$ and restrict as much as possible migration in ages for which the shadow price is smaller than l. The remarkable fact here is that the shadow price is time-invariant.

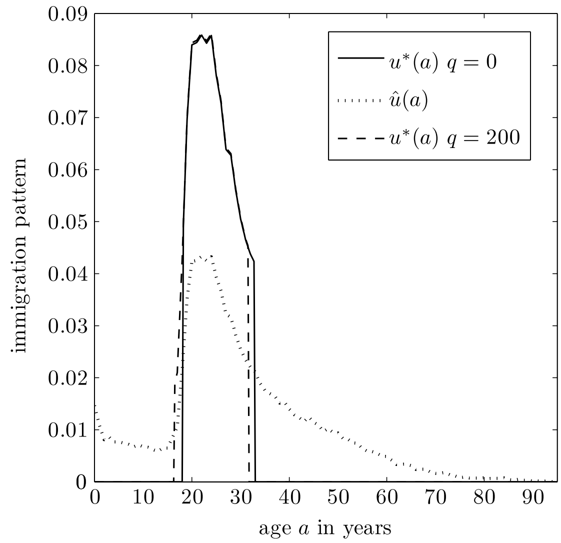

Fig. 4. The actual age density $\hat{u}(a)$ (dotted line) and the optimal immigration density $u^*(a)$ for $q = 0$ (solid) and $q = 200$ (dashed line).

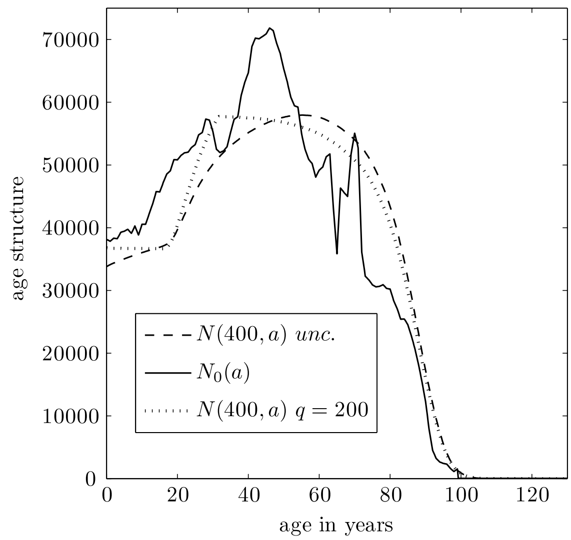

Fig. 5. The initial age structure $N_0(a)$ (solid line) and $N(400, a)$ for the uncontrolled case (dashed) and for the optimal control with $q = 200$ (dotted).

6. A case study: the Austrian population

In this section, we numerically determine the optimal immigration policy of (17)–(20) for the case study of the Austrian population. The numerical results for the optimal time-invariant immigration profile and the population's age structure obtained in this section are based on the analytical results above. In all the numerical calculations below, we specify $p(a)$ in (17) as the characteristic function of the age interval [20, 65]. If we additionally set $q = 0$ the objective function (17) is the discounted and aggregated number of workers over time. It is related to the so-called dependency ratio, which is the ratio of nonworking age population to the working age population. The dependency ratio is an important demographic indicator for the solvency of the social security system of a population. The case of $q > 0$, which is also discussed below, accounts for possible costs for the integration of immigrants.

For the computations, we initialize the age structure of demographic variables referring to Austrian data as of 2009 and interpolate these data piecewise linearly to obtain continuous representations of the vital rates, $\varphi(a)$, $\mu(a)$. As already mentioned in Remark 1, we assume that $\mu(a) = \mu(95)$ for $a \geq 95$. These demographic data together with an intertemporal discount rate of $r = 0.04$ satisfy assumption (A3) with $\kappa_{11} = 0.0737$, $\kappa_{12} = 0.0774$, $\kappa_{21} = 0.1480$, $\kappa_{22} = 0.1506$. For these values the quantity in the left hand side of (32) equals 0.2259 and is therefore well below 1. For the initial age structure $N_0(a)$ we take the annual average numbers of the Austrian female population in 2009, see Fig. 5 (solid line). The normalized immigration age-density of 2009 is denoted by $\hat{u}(a)$, see Fig. 4. We set the lower and upper age-specific limits for

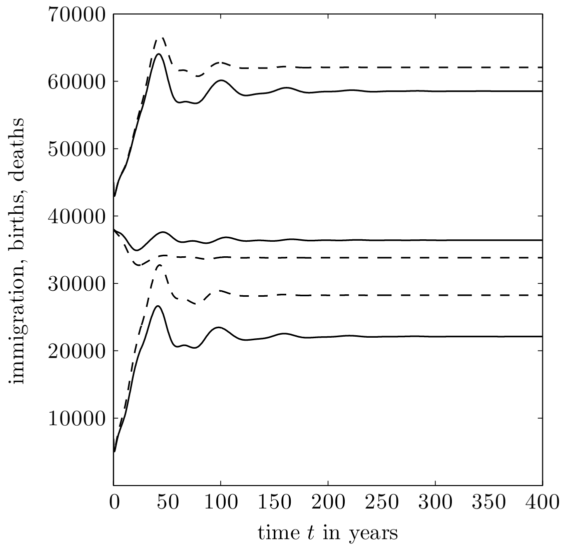

Fig. 6. The evolution of the number of deaths $D(t)$ (upper solid line), the number of births $B(t)$ (middle solid line) and the number of immigrants $R(t)$ (lower solid line) over time for the optimal control and $q = 0$ compared with the uncontrolled case (corresponding dashed lines).

immigration to

$$\underline{u}(a) = 0 \quad \text{and} \quad \overline{u}(a) = 2\hat{u}(a).$$

In the following, we analyze three scenarios: in the uncontrolled case the immigration age density remains the same in the future $u(t, a) = \hat{u}(a)$; then we assume that q in (17) takes the value zero and the immigration age density is chosen optimally $u(t, a) = u^*(a)$; and additionally we set $q = 200$ where again $u(t, a) = u^*(a)$ is chosen optimally. With the last scenario we analyze the effect of immigration costs on the optimal immigration age-pattern, see Fig. 4.

The optimal solution $u(t, a) = u^*(a)$ in the above scenarios is depicted in Fig. 4. As it can be seen in this figure, the optimal age profile of immigrants is at its upper bound from slightly before the lowest working age of $a = 20$ until the mid thirties. It is on its lower bound at any other ages. Note, that increasing the costs of immigration shifts the optimal age pattern to the left.

In Fig. 5 we compare the age structure of the initial population with the stationary population at $t = 400$ which results in the uncontrolled case and when applying the optimal $u^*(a)$ for $q = 200$. The sharp increase of the optimal population $N^*(400, a)$ at the low working ages is due to the annual inflow of immigrants at these ages.

In Fig. 6 we plot the evolution of the number of newborns $B(t)$, the number of deaths $D(t)$ and the recruitment rate $R(t)$ on the time horizon [0, 400], where $D(t) = R(t) + B(t)$. Note, that for the uncontrolled as well as for the controlled immigration, there is a huge increase in the number of immigrants $R(t)$ at the beginning, caused by the high number of deaths as a result of the baby boom, that occurred in Austria in the 50s and 60s.

In Fig. 7 the change of the number of workers and in Fig. 8 the dependency ratio over time are shown. We compare the scenario with $q = 0$ to the case where current age-specific immigration rates would remain the same in the future. Clearly, we can sustain a higher number of workers and simultaneously a lower dependency ratio when applying the optimal immigration pattern $u^*(a)$.

7. Discussion

Two of the key assumptions in this paper pose questions: the stability conditions (A3) (or (A3′)) and the assumption that the discount rate r is strictly positive.

We were not able to prove our main results—the optimality conditions and the stationarity of the optimal immigration age-profile—only assuming below-replacement fertility (8) (which has a clear demographic meaning). A challenging question is whether the "stability" provided by (8) is not enough for the validity of the results. Apparently this question requires more profound analyses of the stability of the involved systems of integral equations.

The question whether the optimality conditions (especially our "transversality" condition for the associated adjoint equation) and the stationarity result can be obtained in the case of no discount (that is, for $r = 0$) seems to be important since discounting is not a common practice in the "evaluation" of demographic processes. Of course, with no discount the notion of optimality should be appropriately defined in the spirit of *overtaking optimality*.

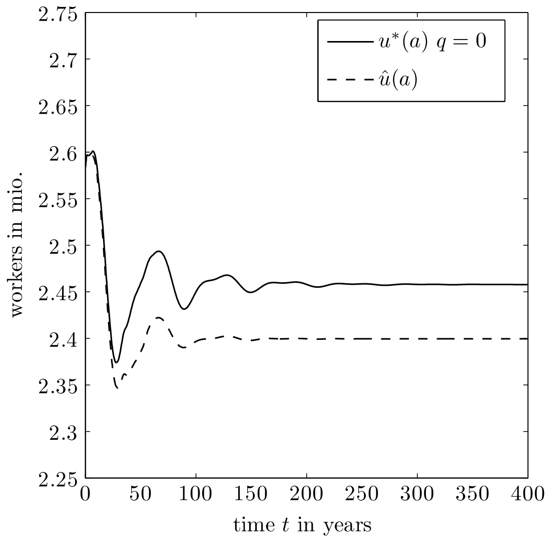

Fig. 7. The evolution of the number of workers over time for $q = 0$ (solid line) and the uncontrolled case (dashed line).

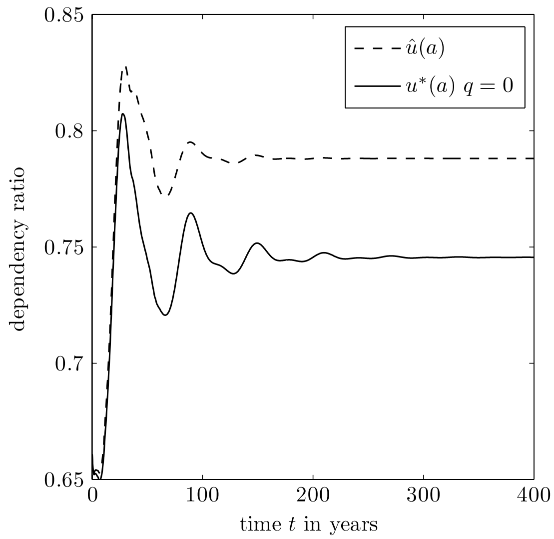

Fig. 8. The evolution of the so-called dependency ratio over time for $q = 0$ (solid line) and the uncontrolled case (dashed line).

Acknowledgments

This research was supported by the Austrian Science Foundation (FWF) under grant No. P20408-G14 (the first author) and No. 1476-N13 (the second and third authors).

Appendix

The proofs are in the order as they appear in the paper. The proof of Lemma 1 is omitted as it is essentially the same as that of Lemma 2, but easier, since R is given and we deal with only one Volterra equation—that for B.

Proof of Lemma 2. Let us start with the uniqueness. Let $N \in \mathcal{N}$ be a solution of (7), (2), (3). Let $R(t)$ and $B(t)$ be defined by (4) and (5), respectively. Both are measurable and locally bounded, according to property (i) of $\mathcal{N}$.

The function N has the following representation, resulting from solving (1) along the characteristic lines:

$$N(t,a) := \begin{cases} e^{-\int_0^a \mu(\tau)\mathrm{d}\tau} B(t-a) + \displaystyle\int_{t-a}^{t} e^{-\int_{a-t+s}^{a} \mu(\tau)\mathrm{d}\tau} R(s)\, u(s, a-t+s)\mathrm{d}s & \text{if } a < t, \\ e^{-\int_{a-t}^{a} \mu(\tau)\mathrm{d}\tau} N_0(a-t) + \displaystyle\int_{0}^{t} e^{-\int_{a-t+s}^{a} \mu(\tau)\mathrm{d}\tau} R(s)\, u(s, a-t+s)\mathrm{d}s & \text{if } a \geq t, \end{cases} \tag{48}$$

on $[0, \infty)$. Inserting this expression for N in (4) and (5) and changing the order of integration in the double integrals we obtain the following system of Volterra equations of the second kind for B and R:

$$\begin{aligned} B(t) &= \int_0^t R(s) \int_0^\infty e^{-\int_a^{a+t-s} \mu(\tau)\mathrm{d}\tau} \varphi(a+t-s)\, u(s,a)\mathrm{d}a\mathrm{d}s \\ &\quad + \int_0^t B(s)\, e^{-\int_0^{t-s} \mu(\tau)\mathrm{d}\tau} \varphi(t-s)\mathrm{d}s + \int_0^\infty e^{-\int_s^{t+s} \mu(\tau)\mathrm{d}\tau}\, \varphi(s+t)\, N_0(s)\mathrm{d}s, \\ R(t) &= \int_0^t R(s) \int_0^\infty e^{-\int_a^{a+t-s} \mu(\tau)\mathrm{d}\tau} \nu(a+t-s)\, u(s,a)\mathrm{d}a\mathrm{d}s \\ &\quad + \int_0^t B(s)\, e^{-\int_0^{t-s} \mu(\tau)\mathrm{d}\tau} \nu(t-s)\mathrm{d}s + \int_0^\infty e^{-\int_s^{t+s} \mu(\tau)\mathrm{d}\tau}\, \nu(s+t)\, N_0(s)\mathrm{d}s, \end{aligned} \tag{49}$$

where $\nu(a) := \mu(a) - \varphi(a)$. Notice that all the four components of the kernel of this system are bounded due to the properties of u and the data. Indeed, take for example the most complicated component

$$\begin{aligned} \left| \int_0^\infty e^{-\int_a^{a+t-s} \mu(\tau)\mathrm{d}\tau} \nu(a+t-s)\, u(s,a)\mathrm{d}a \right| &\le \sup_{a\ge 0} \left\{ e^{-\int_a^{a+t-s} \mu(\tau)\mathrm{d}\tau} |\nu(a+t-s)| \right\} \int_0^\infty u(s,a)\mathrm{d}a \\ &\le \bar{\nu}, \quad 0 \le s \le t < \infty, \end{aligned}$$

where $\bar{\nu} = \sup_{a\ge 0} |\nu(a)| < \infty$.

According to Theorems 5.4 and 5.5 in Chapter 9 of [9] this system has a unique locally bounded solution (B, R), so that B and R are uniquely determined, hence N is also uniquely determined by (48).

On the other hand, from the existence of the locally bounded solution (B, R) we obtain a function N from (48). Due to the local boundedness of B and R and due to $\int_0^\infty N_0(a)\mathrm{d}a = M$, we have that $N \in \mathcal{N}$. It is straightforward to check that N satisfies (7), (2), (3),which proves the existence.

It remains to prove that the functions R and B are locally Lipschitz continuous. We have

$$B(t+h) - B(t) = \int_0^\infty \varphi(a)(N(t+h,a) - N(t+h,a+h))\mathrm{d}a + \int_0^\infty \varphi(a)(N(t+h,a+h) - N(t,a))\mathrm{d}a. \tag{50}$$

Notice that both N and $\mathcal{D}N$ are bounded on every set D_T due to property (i) of $\mathcal{N}$ and (6). Then the second integral is proportional to h because of the absolute continuity of N along the characteristics. For the first integral it holds that

$$\begin{aligned} &\int_0^\infty \varphi(a)(N(t+h,a) - N(t+h,a+h))\mathrm{d}a \\ &= \int_0^h \varphi(a)\, N(t+h,a)\mathrm{d}a + \int_h^\infty \varphi(a)\, N(t+h,a)\mathrm{d}a - \int_0^\infty \varphi(a)\, N(t+h,a+h)\mathrm{d}a \\ &= \int_0^h \varphi(a)\, N(t+h,a)\mathrm{d}a + \int_0^\infty (\varphi(a+h) - \varphi(a))\, N(t+h,a+h)\mathrm{d}a. \end{aligned}$$

The last integral is proportional to h due to the Lipschitz continuity of φ and the boundedness of N. Therefore, the Lipschitz continuity follows. The proof for R is analogous. □

Proof of Lemma 3. The function $t \mapsto M(t) := \int_0^\infty N(t,a)\mathrm{d}a$ is locally Lipschitz (which can be concluded analogously as for R and B in Lemma 2) and thus almost everywhere differentiable. Then the fixed size property is equivalent to $\frac{\mathrm{d}}{\mathrm{d}t}M(t) = 0$ for a.e. t. The latter is equivalent to having the weak derivative of $M(\cdot)$ equal to zero. That is, having

$$\int_0^\infty \Psi(t)\frac{\mathrm{d}}{\mathrm{d}t}M(t)\mathrm{d}t = 0$$

for every $\Psi(t) \in C_0^\infty(0,\infty)$ (the space of all infinitely differentiable function with compact support and $\Psi(0) = 0$).

We have

$$\int_0^\infty \Psi(t)\frac{\mathrm{d}}{\mathrm{d}t}M(t)\mathrm{d}t = [\Psi(t)M(t)]_{t=0}^\infty - \int_0^\infty M(t)\, \frac{\mathrm{d}}{\mathrm{d}t}\Psi(t)\mathrm{d}t.$$

The first term is zero because of the properties of $\Psi(t)$. The second we can rewrite along the characteristic lines of (1) as follows:

$$\begin{aligned} \int_0^\infty \Psi(t)\frac{\mathrm{d}}{\mathrm{d}t}M(t)\mathrm{d}t &= -\int_0^\infty \int_0^\infty N(t,a)\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{d}a\, \Psi(t)\mathrm{d}t \\ &= -\int_0^\infty \int_0^\infty N(s,\tau+s)\frac{\mathrm{d}}{\mathrm{d}s}\Psi(s)\mathrm{d}s\mathrm{d}\tau - \int_0^\infty \int_0^\infty N(\tau+s,s)\frac{\mathrm{d}}{\mathrm{d}s}\Psi(\tau+s)\mathrm{d}s\mathrm{d}\tau. \end{aligned}$$

Integrating again by parts and using (3) we obtain

$$\int_0^\infty \Psi(t)\frac{d}{dt}M(t)dt = -\left[\int_0^\infty N(s,\tau+s)\Psi(s)\right]_0^\infty d\tau + \int_0^\infty \frac{d}{ds}\int_0^\infty N(s,\tau+s)\,\Psi(s)dsd\tau$$
$$-\left[\int_0^\infty N(\tau+s,s)\Psi(\tau+s)\right]_0^\infty d\tau + \int_0^\infty \frac{d}{ds}\int_0^\infty N(\tau+s,s)\,\Psi(\tau+s)dsd\tau$$
$$= \int_0^\infty\int_0^\infty \varphi(a)N(t,a)da\,\Psi(\tau)d\tau$$
$$+\int_0^\infty\int_0^\infty \frac{d}{ds}N(s,\tau+s)\,\Psi(s)dsd\tau + \int_0^\infty\int_0^\infty \frac{d}{ds}N(\tau+s,s)\,\Psi(\tau+s)dsd\tau.$$

Using (1) and rewriting the integral again in the (t,a)-plane we obtain

$$\int_0^\infty \Psi(t)\frac{d}{dt}M(t)dt = \int_0^\infty\int_0^\infty \Psi(t)[-\mu(a)N(t,a)+\varphi(a)N(t,a)+R(t)u(t,a)]dadt.$$

The fact $\int_0^\infty u(t,a)da = 1$ and the arbitrary choice of $\Psi \in C_0^\infty(0,\infty)$ imply that the left hand side is zero if and only if $R(t) = \int_0^\infty[\mu(a)-\varphi(a)]N(t,a)da$.

The last claim of the lemma is evident. □

Proof of Lemma 4. First we shall prove that under the conditions in Lemma 4 we have $B(t) \geq 0$ for all $t \geq 0$. We recall that B is a continuous function due to Lemma 2. Moreover, from (2) and assumption (BA) we have $B(0) = \int_0^\infty \varphi(s)N_0(s)ds > 0$. Denote

$$\theta = \sup\{t \geq 0 : B(s) > 0 \text{ on } [0,t)\}.$$

Assume that θ is finite (otherwise we are done). Then $B(\theta) = 0$. On the other hand we have from (2) that

$$0 = B(\theta) \geq \int_0^\theta B(s)\,e^{-\int_0^{\theta-s}\mu(\tau)d\tau}\varphi(\theta-s)ds + \int_0^\infty e^{-\int_s^{\theta+s}\mu(\tau)d\tau}\,\varphi(s+\theta)\,N_0(s)ds, \tag{51}$$

where we use the assumption $R(t) \geq 0$. Obviously both terms are non-negative. We shall show that at least one of them is strictly positive, which contradicts (51). If $\varphi(a) > 0$ for some $a \in [0,\theta)$, then the first integral in (51) is strictly positive since $B(s) > 0$ on $[0,\theta)$. Alternatively, let $\varphi(a) = 0$ for all $a \in [0,\theta)$. Then $a_0 \geq \theta$ (see assumption (BA)). Take $s = a_0 - \theta$. Then the integrand $e^{-\int_s^{\theta+s}\mu(\tau)d\tau}\,\varphi(s+\theta)\,N_0(s)$ is strictly positive (see (BA)), hence the second integral in (51) is strictly positive, too. The obtained contradiction proves that $B(t) \geq 0$ for all $t \geq 0$. From (48) it follows also that $N(t,a) \geq 0$.

Then the boundedness follows easily:

$$R(t) \leq \int_0^\infty |\mu(a)-\varphi(a)|\,|N(t,a)|da \leq (\bar\mu+\bar\varphi)\int_0^\infty |N(t,a)|da$$
$$= (\bar\mu+\bar\varphi)\int_0^\infty N(t,a)da = (\bar\mu+\bar\varphi)\,M,$$

where $\bar\varphi$ is the upper bound for φ. The same argument proves also boundedness of B. Then the boundedness of $N(t,a)$ follows from (48) and the fact that $u(t,a) = 0$ for all sufficiently large a. □

Proof of Lemma 6. Consider the kernel $k(t,s)$ of the integral equation (28) defined in (29). It defines an operator $K : (L_\infty(0,\infty))^2 \to (L_\infty(0,\infty))^2$. The operator depends on u, so we need the existence of a solution of the integral equation for every admissible u. If the operator norm is smaller than one, a resolvent $L : (L_\infty(0,\infty))^2 \to (L_\infty(0,\infty))^2$ with kernel $l(t,s)$ exists according to Corollary 3.10 and Theorem 3.6 in [9] and can be written as

$$x(t) = F(t) - \int_0^\infty l(t,s)F(s)ds.$$

To show that the norm is smaller than unity, we define for $\alpha > 0$ a new norm on $(L_\infty(0,\infty))^2$ for $\alpha > 0$.

$$\|(x_1,x_2)\| = \max\{\|x_1\|_{L_\infty}, \alpha\|x_2\|_{L_\infty}\}.$$

Take $x \in (L_\infty(0,\infty))^2$ with $\|x\| = 1$, and estimate the norm of $y = Kx$.

$$\|y\| = \max\{y_1, \alpha y_2\}$$
$$= \max\left\{\sup_{t\geq 0}\int_t^\infty [k_{11}(t,s)x_1(s)+k_{12}(t,s)x_2(s)]ds,\ \alpha\sup_{t\geq 0}\int_t^\infty [k_{21}(t,s)x_1(s)+k_{22}(t,s)x_2(s)]ds\right\}$$

$$\leq \max\left\{\sup_{t\geq 0}\int_t^\infty |k_{11}(t,s)|\mathrm{d}s\,\|x_1\|_\infty + \sup_{t\geq 0}\int_0^\infty \frac{1}{\alpha}|k_{12}(t,s)|\mathrm{d}s\,\alpha\,\|x_2\|_\infty,\right.$$
$$\left.\sup_{t\geq 0}\int_t^\infty \alpha|k_{21}(t,s)|\mathrm{d}s\,\|x_1\|_\infty + \sup_{t\geq 0}\int_0^\infty |k_{22}(t,s)|\mathrm{d}s\,\alpha\,\|x_2\|_\infty\right\}. \tag{52}$$

Since $\int_0^\infty u(t,a)\mathrm{d}a = 1$, with κ_{ij} defined in (31), it holds that:

$$\sup_{t\geq 0}\int_0^\infty |k_{ij}(t,s)|\mathrm{d}s \leq \kappa_{ij}, \quad i,j\in\{1,2\}.$$

With this, (52) and $\|x\| = 1$ it can be concluded that

$$\|y\| \leq \max\left\{\kappa_{11} + \frac{1}{\alpha}\kappa_{12}, \alpha\kappa_{21} + \kappa_{22}\right\}.$$

Thus, the right hand side is an estimation for the operator norm of K. The operator norm being smaller than unity is implied by the existence of $\theta_0 < 1$ and $\alpha > 0$ such that

$$\kappa_{11} + \frac{1}{\alpha}\kappa_{12} \leq \theta_0, \tag{53}$$
$$\alpha\kappa_{21} + \kappa_{22} \leq \theta_0. \tag{54}$$

Since the first line is monotonously decreasing and the second increasing in α, for the optimal α, which allows for the smallest possible θ_0, both equations are fulfilled as equality. Therefore, we solve the equation

$$\kappa_{11} + \frac{1}{\alpha}\kappa_{12} = \alpha\kappa_{21} + \kappa_{22}$$

for α and obtain

$$\alpha_{1,2} = \frac{1}{2\kappa_{21}}\left[\kappa_{11} - \kappa_{22} \pm \sqrt{(\kappa_{11}-\kappa_{22})^2 + 4\kappa_{12}\kappa_{21}}\right].$$

We insert the positive solution for α into the second line of (54) to obtain θ_0:

$$\theta_0 = \frac{1}{2}\left[\kappa_{11} + \kappa_{22} + \sqrt{(\kappa_{11}-\kappa_{22})^2 + 4\kappa_{12}\kappa_{21}}\right].$$

The requirement $\theta_0 < 1$ is exactly inequality (33) in condition (A3).

The sufficiency of (A3′) follows because the assumption implies that for all admissible u it holds that $\sup_{t\geq 0}\int_0^\infty k_{ij}(t,s)\mathrm{d}s \leq (1+\varepsilon)\bar{\kappa}_{2j}$ for $j = 1, 2$. System (53)–(54) then reads as

$$\kappa_{11} + \frac{1}{\alpha}\kappa_{12} \leq \theta_0, \tag{55}$$
$$(1+\varepsilon)(\alpha\bar{\kappa}_{21} + \bar{\kappa}_{22}) \leq \theta_0.$$

By following the same steps as above, we obtain that (32) is sufficient for the operator norm of K to be smaller than one. □

Proof of Lemma 7. The proof is similar to the one of Lemma 6. Let us take two bounded solutions, ξ_1 and ξ_2 and denote by $\Delta\xi(t,a)$ the difference between the two. The solutions ξ_i can be written as (cf. (25))

$$\xi_i(t,a) = \int_a^\infty e^{-\int_a^s \rho(\theta)\mathrm{d}\theta}[\varphi(s)\xi_i(s+t-a,0) + \nu(s)\sigma(\xi_i(s+t-a,\cdot)) + f(s)]\mathrm{d}s, \quad i = 1, 2.$$

Let $\Delta\lambda(t) := \xi_1(t,0) - \xi_2(t,0)$ and $\Delta\sigma(t) := \sigma(\xi_1(t,\cdot)) - \sigma(\xi_2(t,\cdot))$, then we obtain the homogeneous system of integral equations

$$\Delta\xi(t,a) = \int_0^\infty e^{-\int_a^{s+a}\rho(\theta)\mathrm{d}\theta}(\varphi(s+a)\Delta\lambda(s+t) + \nu(s+a)\Delta\sigma(\xi(s+t,\cdot)))\mathrm{d}s, \tag{56}$$
$$\Delta\lambda(t) = \int_0^\infty e^{-\int_0^s \rho(\theta)\mathrm{d}\theta}[\varphi(s)\Delta\lambda(s+t) + \nu(s)\Delta\sigma(s+t)]\mathrm{d}s.$$

Denote by $K : (L_\infty(0,\infty))^2 \to (L_\infty(0,\infty))^2$ the integral operator representing the system of integral equations above. The existence of a unique solution to the system is guaranteed if $\|K\| < 1$. To show this, take the norm in the system of equations and use that σ is Lipschitz with Lipschitz constant 1 (cf. Lemma 4.1 in [8]),

$$\|\Delta\xi\|_\infty \le \kappa_{21}\|\Delta\lambda\|_\infty + \kappa_{22}\|\Delta\xi\|_\infty$$
$$\|\Delta\lambda\|_\infty \le \kappa_{11}\|\Delta\lambda\|_\infty + \kappa_{12}\|\Delta\xi\|_\infty.$$

As in the proof of Lemma 6 we define a norm $\|(\Delta\xi, \Delta\lambda)\| := \max\{\|\Delta\xi\|_\infty, a\|\Delta\lambda\|_\infty\}$ for $a > 0$. We choose $a > 0$ in such a way, that the norm of the operator K is minimized. The minimum is exactly the left hand side of (32) and Assumption (A3) guarantees that it is smaller than one. Therefore, a unique solution exists to the homogeneous system, which is obviously $\Delta\xi = 0$. □

Proof of Proposition 1. The proof is a modification of that in [8] while the underlying idea stems from [2]. Denote by $J(u)$ the objective value for an admissible control u. Due to Lemma 4 and $r > 0$ the value $J(u)$ is finite and uniformly bounded with respect to $u \in \mathcal{U}$. Then $J^* = \sup_{u\in\mathcal{U}} J(u)$ is also finite. Pick a maximizing sequence $\{u_k\}$ of admissible controls for which $J(u_k) \ge J^* - \frac{1}{k}$. Denote by N_k the corresponding solution of (18)–(21), and let R_k be defined as in (4). According to Assumption (A2) and Lemma 4 there is a constant C such that $0 \le N_k(t, a) \le C$ and $0 < R_k(t) \le C$ almost everywhere and for all k.

The sequence $\{e^{-rt}N_k\}$ of elements of $L_1(D)$ is weakly relatively compact due to the Dunford–Pettis criterion.

Therefore, there exists a subsequence, which will also be denoted by N_k, such that $e^{-rt}N_k$ converges $L_1(D)$-weakly to some $e^{-rt}N^*$, and N^* is obviously bounded by the same constant C. According to Mazur's lemma there exist a sequence

$$e^{-rt}\tilde{N}_k := \sum_{i=k}^{n_k} p_i^k e^{-rt}N_i, \quad p_i^k \ge 0, \ \sum_{i=k}^{n_k} p_i^k = 1,$$

that (strongly) converges to $e^{-rt}N^*$ in $L_1(D)$. Obviously for every $T > 0$ the sequence $\tilde{N}_k$ converges to N^* in $L_1(D_T)$. With the same weights p_i^k we define

$$\tilde{R}_k(t) := \sum_{i=k}^{n_k} p_i^k R_i(t) = \int_0^\infty (\mu(a) - \varphi(a))\tilde{N}_k(t, a)\mathrm{d}a. \tag{57}$$

Since $R_k(t) > 0$ holds for all $k > 0$ and $t > 0$, this also holds for $\tilde{R}_k$ and we can define

$$\tilde{u}_k(t, a) := \frac{1}{\tilde{R}_k(t)} \sum_{i=k}^{n_k} p_i^k R_i(t)u_i(t, a).$$

Obviously $\tilde{u}_k$ is also an admissible control. Moreover we have that

$$\mathcal{D}\tilde{N}_k = \sum_{i=k}^{n_k} p_i^k(-\mu N_i + R_i u_i) = -\mu\tilde{N}_k + \tilde{R}_k\tilde{u}_k, \tag{58}$$

$$\tilde{N}_k(t, 0) = \sum_{i=k}^{n_k} p_i^k N_i(t, 0) = \sum_{i=k}^{n_k} p_i^k \int_0^\infty \varphi(a)N_i(t, a)\mathrm{d}a = \int_0^\infty \varphi(a)\tilde{N}_k(t, a)\mathrm{d}a, \tag{59}$$

which means that $(\tilde{u}_k, \tilde{N}_k)$ is an admissible control-trajectory pair in problem (17)–(21).

Since $\tilde{N}_k$ converges to N^* in $L_1(D_T)$, we may pass to an almost everywhere converging subsequence that we denote again by $\tilde{N}_k$. Moreover, we may assume (passing again to a subsequence) that $e^{-rt}\tilde{u}_k$ converges to some $e^{-rt}u^*$ weakly in $L_1(D)$. Now we will show that u^* is an admissible control. For every measurable and bounded set $\Gamma \subset [0, \infty)$ it holds that

$$\int_\Gamma \int_0^{\bar{a}} \tilde{u}_k(t, a)\mathrm{d}a\mathrm{d}t \to \int_\Gamma \int_0^{\bar{a}} u^*(t, a)\mathrm{d}a\mathrm{d}t$$

where $\bar{a}$ is such that $\bar{u}(a) = 0$ for $a \ge \bar{a}$, hence also $u^*(t, a) = 0$ (see (A1)). Since $\tilde{u}_k$ are admissible controls the left hand side is equal to meas(Γ), and thus also the right hand side. Since this holds for any measurable and bounded set Γ, this implies that $\bar{u}$ satisfies the integral constraint in (21). The inequality constraints are obviously also satisfied. Therefore, u^* is an admissible control. In the next paragraph we shall prove that the pair N^* solves (18)–(20) with $u = u^*$.

Let us define

$$R^*(t) = \int_0^\infty (\mu(a) - \varphi(a))N^*(t, a)\mathrm{d}a.$$

Due to the pointwise convergence of $\tilde{N}_k$ in D_T we obtain by passing to a limit in (57) that $\tilde{R}_k(t) \longrightarrow R^*(t)$ for a.e. $t \in [0, T)$, and since T is arbitrary this holds for a.e. $t \ge 0$. Moreover, for a.e. t the mapping $[0, T - t] \ni s \longrightarrow \tilde{N}_k(t + s, s)$ is uniformly Lipschitz continuous. From here it easily follows (see [7] for more details) that $N^*(t, 0)$ is well defined for a.e. t

and $\tilde{N}_k(t,0) \longrightarrow N^*(t,0)$. Then by passing to a limit in (59) we obtain that N^* satisfies the boundary condition (20) for a.e. $t \in [0,T]$, hence for all a.e. $t \geq 0$. In a similar way one can prove that N^* satisfies the initial condition (19). In order to show that (18) is also satisfied we take an arbitrary measurable set $\Gamma \subset [0,T]$ integrate with respect to $a \in \Gamma$ the representation (48) of the solution $\tilde{N}_k$ for $u = \tilde{u}_k$. Due to the established properties we can pass on to the limit. Since Γ is arbitrary, we obtain that (u^*, N^*, R^*) satisfy (48) on D_T, hence N^* is a solution of (18) on D for $u = u^*$.

Thus (u^*, N^*) is an admissible control-trajectory pair.

Now let us show that $J(u^*) \geq J^*$. We have

$$\begin{aligned}
J(\tilde{u}_k) &= \int_0^\infty e^{-rt}\left[\int_0^\infty p(a)\tilde{N}_k(t,a)\mathrm{d}a - q\tilde{R}_k(t)\right]\mathrm{d}t \\
&= \int_0^\infty e^{-rt}\left[\int_0^\infty p(a)\sum_{i=k}^{n_k} p_i^k N_i(t,a)\mathrm{d}a - q\sum_{i=k}^{n_k} p_i^k R_i(t)\right]\mathrm{d}t \\
&= \sum_{i=k}^{n_k} p_i^k J(u_i) = \sum_{i=k}^{n_k} p_i^k\left(J^* - \frac{1}{i}\right) \geq J^* - \frac{1}{k}.
\end{aligned}$$

Using this we obtain

$$\begin{aligned}
J^* &\leq \limsup_k \left(J(\tilde{u}_k) + \frac{1}{k}\right) = \limsup_k J(\tilde{u}_k) \\
&= \limsup_k \int_0^\infty e^{-rt}\left[\int_0^\infty p(a)\tilde{N}_k(t,a)\mathrm{d}a - q\tilde{R}_k(t)\right]\mathrm{d}t \\
&= \int_0^\infty e^{-rt}\left[\int_0^\infty p(a)N^*(t,a)\mathrm{d}a - qR^*(t)\right]\mathrm{d}t \\
&= J(u^*). \quad \square
\end{aligned}$$

References

[1] S. Anita, Analysis and Control of Age-Dependent Population Dynamics, Kluwer Academic Publishers, The Netherlands, 2000.
[2] S. Anita, M. Iannelli, M.Y. Kim, E.J. Park, Optimal harvesting for periodic age-dependent population dynamics, SIAM J. Appl. Math. 58 (1998) 1648–1666.
[3] S. Aseev, V. Veliov, Maximum principle for problems with dominating discount, Dyn. Contin. Discrete Impuls. Syst. Ser. B 19 (2012) 43–63.
[4] J.P. Aubin, F.H. Clarke, Shadow prices and duality for a class of optimal control problems, SIAM J. Control Optim. 17 (1979) 567–586.
[5] M. Brokate, Pontryagin's principle for control problems in age-dependent population dynamics, J. Math. Biol. 23 (1985) 75–101.
[6] W.L. Chan, G.B. Zhu, Overtaking optimal control problem of age-dependent populations with infinite horizon, J. Math. Anal. Appl. 150 (1990) 41–53.
[7] G. Feichtinger, G. Tragler, V. Veliov, Optimality conditions for age-structured control systems, J. Math. Anal. Appl. 288 (2003) 47–68.
[8] G. Feichtinger, V. Veliov, On a distributed control problem arising in dynamic optimization of a fixed-size population, SIAM J. Optim. 18 (3) (2007) 980–1003.
[9] G. Gripenberg, S.O. Londen, O. Staffans, Volterra Integral and Functional Equations, Cambridge University Press, Cambridge, UK, 1990.
[10] N. Keyfitz, On the momentum of population growth, Demography 8 (1971) 71–80.
[11] C. Simon, A. Belyakov, G. Feichtinger, Minimizing the dependency ratio in a population with below-replacement fertility through immigration, Theor. Popul. Biol. 82 (3) (2012) 158–169.
[12] United Nations, Replacement Migration: Is It a Solution to Declining and Ageing Populations? UN Population Division, Department of Economic and Social Affairs, New York, 2001. ST/ESA/SER.A/206.
[13] V. Veliov, Optimal control of heterogeneous systems: basic theory, J. Math. Anal. Appl. 346 (2008) 227–242.
[14] G. Webb, Theory of Nonlinear Age-Dependent Population Dynamics, Marcel Dekker, New York, 1985.