Minimizing the Dependency Ratio in a Population with Below-Replacement Fertility through Immigration

C. Simon[b,*], A. Belyakov[a], G. Feichtinger[a,b]

[a] *Institute of Mathematical Methods in Economics, Vienna University of Technology, Argentinierstraße 8, 1040 Vienna, Austria, +43 (0)1 58801 10541*
[b] *Vienna Institute of Demography, Austrian Academy of Sciences, Wohllebengasse 12-14, 1040 Vienna, Austria, +43 (0)1 515 81 7702*

Abstract

Many industrialized countries face fertility rates below replacement level, combined with declining mortality especially in older ages. Consequently, the populations of these countries have started to age. One important indicator of age structures is the dependency ratio which is the ratio of the nonworking age population to the working age population. In this work we find the age-specific immigration profile that minimizes the dependency ratio in a stationary population with below-replacement fertility. It is assumed that the number of immigrants per age is limited. We consider two alternative policies. In the first one, we fix the total number of people who annually immigrate to a country. In the second one, we prescribe the size of the receiving country's population. For both cases we provide numerical results for the optimal immigration profile, for the resulting age structure of the population, as well as for the dependency ratio.

Keywords: optimal control theory, stationary population, immigration, below-replacement fertility

*Corresponding author

Email addresses: `christa.simon@oeaw.ac.at` (C. Simon), `anton.belyakov@tuwien.ac.at` (A. Belyakov), `gustav.feichtinger@tuwien.ac.at` (G. Feichtinger)

1. Introduction

In many industrialized countries fertility rates are below-replacement level. Additionally, these countries face a mortality decline, in particular at ages after retirement. Since fertility decline is very often the dominating effect, the population of these countries would decline without immigration. Moreover, the age structure of these countries' population is changing, showing a growth in the number of elderly people and a declining number of young people.

One important indicator of age structures is the so-called dependency ratio, which is the ratio of persons of nonworking age to persons of working age, usually the 20 to 65-year-olds. A low dependency ratio is desirable because it indicates that there are proportionally more adults of working age who can support the young and the elderly of the population. This in turn is advantageous for the countries' health-care system and pension schemes. A downfall of the relative number of working people in a population also has negative impacts on the growth path of the economy. A possible way to counter the risks of these demographic changes is to step up immigration.

Similar to (Arthur and Espenshade, 1988; Mitra, 1990; Schmertmann, 1992; Wu and Li, 2003), in this work we consider a population where we assume that immigration, fertility, and mortality rates are constant and fertility is below-replacement level. These studies already have shown that such populations eventually converge to stationary populations. Following (Schmertmann, 1992) from now on we will denote this kind of population as SI, meaning *stationary through immigration*. Below-replacement level fertility and mortality rates indicate that without immigration the population would converge to zero. In our model we assume that the age-specific fertility rates of immigrants equal those of the natives. Following (Schmertmann, 2011) we do not account for emigration.

In this work, we aim to find the age-specific immigration profile that minimizes the dependency ratio in a stationary population. We do so by applying optimal control theory which is a rather new methodology in demographic research, see for example (Feichtinger and Veliov, 2007). We formulate an optimal control problem where the age-specific immigration profile is the control variable and the age structure of the population is the state variable.

A similar question to the one posed here is asked in (United Nations, 2001) where the authors determine whether migration of a country can be used to hinder a decline or aging of its population. They refer to this as

replacement migration. They examine the situation of eight industrialized countries during the time period from 1995 to 2050.

In (Schmertmann, 2011) the question is raised how age-targeted immigration policy can be used to increase the relative number of working people in a population. There, the total number of annual immigrants is fixed and the problem is reduced to a static optimization problem. What is shown is that the highest relative number of workers can be achieved if all immigrants arrive at one single age under the assumption that at each age an arbitrarily high number of immigrants can be recruited. Schmertmann's paper leaves open the question of what the optimal age-specific immigration profile would look like if not all immigrants are admitted at one single age. This issue, among others things, is tackled below.

From a mathematical point of view, a similar linear optimal control problem to the one proposed here is considered in (Dawid et al., 2009). The authors determine the optimal recruitment policy of a stationary learned society, i.e. a professional and hierarchical organization, that minimizes the average age of the organization for a fixed number of recruits. (Feichtinger and Veliov, 2007) extended their study to the transitory case. Remarkably, the optimal recruitment is the same as in the stationary case. That is why we also start with the stationary case.

In the following, we consider two alternative policies in order to investigate their impact on the optimal immigration profile.

Policy 1: We fix the total number of people who annually immigrate to a country.

Policy 2: We prescribe the (stationary) total size of the receiving country's population.

Moreover, we assume that there are age-specific upper bounds for immigration. The optimal immigration profile for both policies exhibits a *bang-bang* pattern, meaning that the solution jumps from one age-specific bound to the other and takes no values in between. We prove that for the optimal profile under Policy 1 it is a fact that besides immigration at young and middle ages, immigration takes place also in the vicinity of the maximum attainable age. Such counterintuitive old-age immigration does not happen under Policy 2. We show that under reasonable assumptions about the vital rates and the age-specific immigration bounds, the optimal immigration profile under Pol-

icy 2 is such that it is optimal to allow maximum immigration on not more than two separate age intervals before the retirement age.

The optimal control approach enables us to determine the marginal value of an immigrant at a certain age in terms of the dependency ratio, cf. (Wrzaczek et al., 2010), by interpretation of the so-called adjoint variable, cf. (Grass et al., 2008), whose clear meaning will be defined in Section 3. As a consequence we are able to decide what age-specific immigration profile is optimal for minimizing the dependency ratio. Moreover, the impact of an a-year-old immigrant on the dependency ratio can be represented as a sum of two components. The first component, which is referred to as the *direct* effect, accounts for a woman's expected life time inside and outside the work force. The second component, known as the *indirect* effect accounts for the effect on the dependency ratio contributed by her expected number of descendants. Clearly, when an immigrant arrives towards the end of childbearing age she will have less children than a younger woman and therefore she will be less of a burden for the dependency ratio of the resulting stationary population. However, the expected remaining time in the working population is then also reduced, meaning that she will be dependent for a relatively longer time.

The rest of the paper is organized as follows. In Section 2 we state the problem. The optimal age-specific immigration profile for a fixed annual number of immigrants is characterized in Section 3. There, we also present numerical results for the case study of the Austrian population based on demographic data from 2008. In Section 4 we consider the total stationary population size as fixed and provide also some numerical results for the optimal immigration profile and the dependency ratio. In Section 5 the effect of an additional woman of a certain age on the dependency ratio is explained. It is shown how this effect can be separated in two parts. In Sections 6 and 7 we discuss the obtained results and indicate points of future work. In Appendix A we formulate Pontryagin's Maximum Principle (Alekseev et al., 1987). In Appendix B and Appendix D we apply the Maximum Principle to obtain necessary conditions for the optimal solution. The proof of the counterintuitive result that old-age immigration is optimal is given in Appendix C.

2. Model Description and Definitions

In the following, α and β denote the lower and upper age limits determining the working age population and ω is the maximum attainable age of an individual. We aim to minimize the dependency ratio given as

$$D(M(\cdot)) := \frac{\int_0^\alpha N(a)\,\mathrm{d}a + \int_\beta^\omega N(a)\,\mathrm{d}a}{\int_\alpha^\beta N(a)\,\mathrm{d}a}, \quad 0 < \alpha < \beta < \omega,$$

by choosing the age distribution of immigrants $M(\cdot)$. With $D(M(\cdot))$ we mean the dependency ratio that results when realizing the immigration profile $M(\cdot)$ and $N(a)$ denotes the number of resulting females in the population of age a.

We come up with the following optimal control problem:

$$\min_{M(a)} D, \tag{2.1}$$

subject to

$$\dot{N}(a) = -\mu(a)N(a) + M(a), \tag{2.2}$$

$$N(0) = \int_0^\omega f(a)N(a)\,\mathrm{d}a, \tag{2.3}$$

$$0 \le M(a) \le \bar{M}(a). \tag{2.4}$$

Here, the age a is considered as a continuous variable and $\dot{N}(a)$ denotes the derivative of $N(a)$ with respect to a. The immigration age profile $M(\cdot)$ is referred to as *control*, since it is the controlable input to the optimization problem. The population structure is determined by the so - called state variable of the problem, $N(a)$, which is the number of people of age a. In contrast to the control, the state variable cannot be directly influenced. By $f(a)$ and $\mu(a)$, we denote age-specific fertility and mortality rates which do not change with time and are continuous functions in a. Additionally, we assume that $\int_0^\omega \mu(a)\,\mathrm{d}a = +\infty$, cf. (Anita, 2000), which ensures that $N(\omega) = 0$ holds. With $\bar{M}(a)$ we denote the age-specific immigration bounds which are assumed to be continuous.[1]

---

[1]From a mathematical point of view the reason for imposing these age-specific bounds is the applicability of Pontryagin's Maximum Principle. However, more practically spoken these bounds are justifiable because they may reflect the fact that age is not the only factor that should be taken into account when determining the optimal immigration policy and also the number of potential immigrants of a certain age is limited.

Note, that the dynamics (2.2) describing the age structure of the population only holds for a stationary population.

In this work, we consider two alternative policies:

Policy 1. We prescribe the total number of immigrants M_{tot}

$$M_{\text{tot}} = \int_0^\omega M(a) \, \mathrm{d}a. \tag{2.5}$$

Policy 2. We prescribe the stationary population size N_{tot}

$$N_{\text{tot}} = \int_0^\omega N(a) \, \mathrm{d}a. \tag{2.6}$$

These policies represent constraints on the number of immigrants and the total population size.

We define $l(a) := \mathrm{e}^{-\int_0^a \mu(x) \, \mathrm{d}x}$ which is the probability that a female survives at least a years.

We recall the *reproductive value* of an a-year-old female, introduced in (Fisher, 1930) (see also (Keyfitz, 1977)), which is the expected number of future daughters of an individual from her current age onward, given that she has survived to this age as

$$v(a) = \int_a^\omega \frac{l(x)}{l(a)} f(x) \, \mathrm{d}x.$$

Accordingly, the *population's net reproduction rate* in a stationary population is the average number of daughters a female will have,

$$R = \int_0^\omega l(a) f(a) \, \mathrm{d}a.$$

The support of $f(\cdot)$ is a subset $[a_{\min}, a_{\max}] \subset [0, \omega]$, where $a_{\min}$ and $a_{\max}$ denote the youngest and oldest age of childbearing, respectively. Presumably, fertility $f(\cdot)$ is below-replacement, which means it is not high enough to replace the current population. This property of below-replacement fertility (and mortality) can be expressed in terms of the population's reproduction rate, meaning that $R < 1$ must hold.

Note, that the control $M(\cdot)$ enters linearly the problem. This property of the optimal control problem is responsible for the bang-bang behavior of the solution obtained below.

3. The optimal immigration profile for a fixed number of immigrants

In this section we analyze problem (2.1)–(2.5) by making use of optimal control theory. Our aim is to find the optimal immigration profile $M^*(\cdot)$ that minimizes D.

In order to determine the optimal immigration profile we derive necessary conditions to characterize the optimal solution. Therefore, we need to introduce constants λ_1, λ_2, called the Lagrange multipliers, and another notion from optimal control theory, the *adjoint variable ξ*.

The adjoint variable and its interpretation as shadow price

As the name implies, the adjoint variable is related to another variable: the state variable N. It is the derivative of the so-called value function, i.e. the objective function evaluated at the optimal solution, with respect to the state variable. Therefore, in economic applications of optimal control theory, the adjoint variable is interpreted as shadow price of the state variable. In line with this interpretation, here $\xi(a)$ gives the shadow price of an individual of age a. As can be seen below, for this particular optimal control problem considered in here the shadow price is a part of the effect of adding an additional immigrant of age a.

The term shadow price is commonly used in capital theory, cf. (Dorfman, 1969; Léonard and Long, 1992). There, it is interpreted as the highest hypothetical — therefore also called shadow — price a rational decision-maker would be willing to pay for owning an additional unit of the corresponding state variable at time a measured by the discounted (extra) future profit. Note, that the shadow price is not a real market price and therefore can also have a negative value. See also (Grass et al., 2008) for a more detailed discussion of the economic interpretation of the maximum principle.

The Lagrange multipliers

We also introduce the constants λ_1, λ_2 and refer to them as *Lagrange multipliers*. The Lagrange multiplier λ_1 reflects the marginal worth of an increase in the annual flow of newborns. The constant λ_2 gives the marginal change in the dependency ratio when adding an additional immigrant per year.

Note, that for a given age interval $[\alpha, \beta] \subseteq [0, \omega]$ the function $\mathbb{I}_{[\alpha,\beta]}(\cdot)$ is defined as

$$\mathbb{I}_{[\alpha,\beta]}(a) = \begin{cases} 1 & \text{if } a \in [\alpha, \beta], \\ 0 & \text{otherwise.} \end{cases}$$

Then the optimal immigration profile $M^*(\cdot)$ and the corresponding population structure $N^*(\cdot)$ can be characterized as:

Theorem 1. *If $(N^*(\cdot), M^*(\cdot))$ is an optimal solution of problem (2.1)–(2.5), then there are Lagrange multipliers $\lambda_1, \lambda_2 \in \mathbb{R}$ and an adjoint variable $\xi(\cdot)$, such that:*

(i) the continuous function $\xi(\cdot)$ on $[0, \omega]$ satisfies

$$\dot{\xi}(a) = \mu(a)\xi(a) - \lambda_1 f(a) - \frac{(D(M^*(\cdot)) + 1)^2}{N_{\text{tot}}(M^*(\cdot))} \mathbb{I}_{[\alpha,\beta]}(a) + \frac{(D(M^*(\cdot)) + 1)}{N_{\text{tot}}(M^*(\cdot))},$$

$$\xi(0) = \lambda_1, \quad \xi(\omega) = 0, \tag{3.7}$$

(ii) and the following maximum principle holds for almost every $a \in (0, \omega)$, i.e. besides of isolated points:

$$(\xi(a) - \lambda_2)M^*(a) = \max_{0 \le M \le \bar{M}(a)} (\xi(a) - \lambda_2)M. \tag{3.8}$$

Proof 1. *For the proof of Theorem 1 see Appendix B.*

Theorem 1 provides necessary conditions (3.7),(3.8) for the solution of problem (2.1)–(2.5), meaning that they constitute requirements that the optimal solution has to fulfill. The existence of an optimal solution follows from a general argument.

From (3.8) it can immediately be concluded that the optimal control is of bang-bang type, jumping from one boundary to the other. Therefore, function $\xi(\cdot) - \lambda_2$ is usually referred to as *switching function* because the change of its sign determines the ages a at which the optimal control switches from one boundary to the other in consequence of (3.8):

$$M^*(a) = \begin{cases} \bar{M}(a) & \text{if } \xi(a) > \lambda_2, \\ \text{not determined} & \text{if } \xi(a) = \lambda_2, \\ 0 & \text{if } \xi(a) < \lambda_2. \end{cases} \tag{3.9}$$

8

We assume that equality $\xi(a) = \lambda_2$ happens only in isolated points, so that the values $M^*(a)$ at these points have no effect on the dependency ratio. This assumption holds for fertility and morality rates that are not linearly related, which can be concluded by additionally requiring that the derivative of the switching function, $\xi(\cdot) - \lambda_2$, is 0 on an interval $[\underline{a}, \overline{a}]$ which would be a violation of the assumption that the switching function is 0 only in isolated points [2].

To obtain the optimal immigration profile it remains to determine $\xi(\cdot)$ and λ_2. The right hand side of the differential equation (3.7) is discontinuous at ages $a = \alpha$ and $a = \beta$ and therefore the solution ξ has two kinks at each of these ages.

We note that (3.7) is a boundary value problem for a linear differential equation. By using the Cauchy formula for the solution of an ordinary differential equation (3.7) we obtain the solution

$$\xi(a) = \lambda_1 v(a) + \frac{(D(M^*(\cdot)) + 1)}{N_{\text{tot}}(M^*(\cdot))} \Big((D(M^*(\cdot)) + 1) e_{[\alpha,\beta]}(a) - e_{[0,\omega]}(a) \Big).$$

(3.10)

Using the boundary condition $\xi(0) = \lambda_1$ and noting that $R = v(0)$ we obtain that

$$\lambda_1 = \frac{\frac{(D(M^*(\cdot))+1)}{N_{\text{tot}}(M^*(\cdot))} \Big((D(M^*(\cdot)) + 1) e_{[\alpha,\beta]}(0) - e_{[0,\omega]}(0) \Big)}{1 - R}$$

(3.11)

holds. Here, $e_{[0,\omega]}(a) = \int_a^\omega \frac{l(x)}{l(a)} \, dx$ is the life expectancy in $[0,\omega]$ at age a. Similarly, $e_{[\alpha,\beta]}(a) = \int_a^\omega \frac{l(x)}{l(a)} \mathbb{I}_{[\alpha,\beta]}(x) \, dx$ is the working life expectancy of an a-year-old, reflecting the expected number of years an a-year-old would spend working. Clearly, $e_{[\alpha,\beta]}(a) = 0$ for $a \geq \beta$. With (3.9) and expressions (3.10)–(3.11) we are now able to obtain the optimal immigration profile $M^*(\cdot)$, where the Lagrange multiplier λ_2 has to be determined in such a way, that (2.5) holds for the resulting solution.
In Appendix C it is shown, that under the additional condition that the contribution of an additional a-year-old immigrant to the number of workers

---

[2]More precisely, this assumption is fulfilled for fertility and mortality rates such that for any $d \in \mathbb{R}$ meas$\{a \in \Omega : \lambda_2 \mu(a) - \lambda_1 f(a) = d\} = 0$., i.e. this equality only holds on a set of measure zero, where the measure is meant in the sense of Lebesque.

in the resulting SI population (measured in years) must not be proportional to its contribution to the whole population for almost every age $a \in [0, \omega]$, the optimal immigration profile $M^*(\cdot)$ is such that arbitrarily close to the maximum age ω there are ages where immigration is optimal. An individual's contribution consists of her own expected years lived in the host country and the analog contribution of all her future descendants. Since we aim to minimize the relative number of dependent people in the population, the fact that immigration at the end of the life horizon is optimal seems to be counterintuitive. This counterintuitive property of the optimal solution is due to the age-specific immigration bounds, (2.4), that are introduced in this model. If they are removed, as done in a static set up in (Schmertmann, 2011), this effect probably cannot be observed anymore. We also overcome this counterintuitive result in Section 4 by considering Policy 2, where we fix the size of the stationary population.
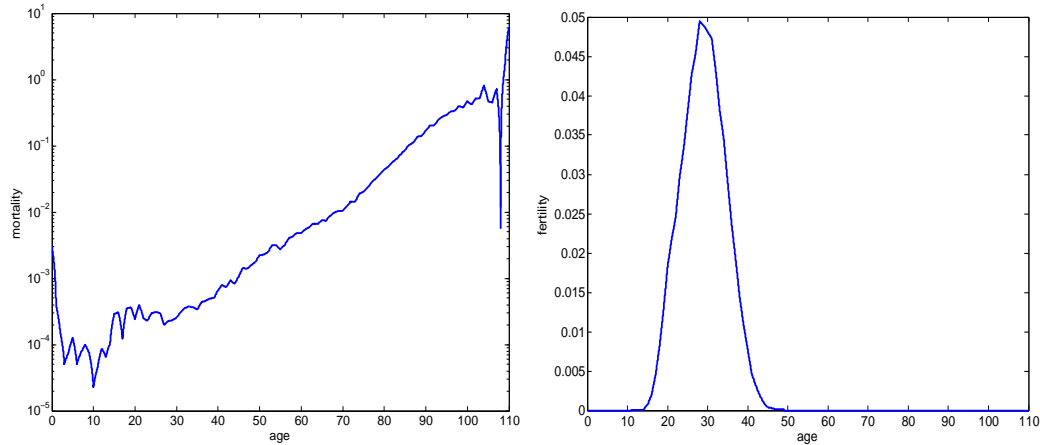
3.1. A case study: the Austrian case

Figure 3.1: Female mortality rate $\mu(a)$ (left, logarithmic scale) and fertility rate $f(a)$ (right); Austria 2008.

The numerical results for the optimal immigration profile and the dependency ratio obtained in this section are based on the analytical derivations above. In the following, we will assume that $\alpha = 20$, $\beta = 65$, and $\omega = 110$. For the computations we initialize the age structure of demographic variables referring to Austrian data as of 2008, cf. Figure 3.1, and interpolate these

10

data piecewise linearly to obtain continuous representations of the vital rates. The actual age-specific immigration numbers of 2008 are denoted by $M_{\text{act}}(a)$.

Scenario 1. We set

$$\bar{M}(a) = 2M_{\text{act}}(a), \quad \forall a \in [0, \omega],$$

which corresponds to a doubling of the number of immigrants at all ages compared to the 2008 level. For M_{tot} we prescribe a total volume of 80000 females.

The resulting age profile that fulfills the maximization condition (3.8) is

$$M^*(a) = \begin{cases} \bar{M}(a) & \text{if } a \in [11, 49] \cup [82, 110], \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

This can be concluded from the values of the adjoint variable $\xi(a)$ at ages a depicted in Figure 3.2a. The solid line in Figure 3.2a corresponds to the λ_2-level. Consequently, for ages where $\xi(a)$ has values larger than λ_2 immigration is at its upper bound and for ages where $\xi(a)$ is smaller than λ_2 the optimal immigration profile is zero. The adjoint variable $\xi(\cdot)$ exhibits two kinks at ages $\alpha = 20$ and $\beta = 65$, due to the discontinuity of the right hand side of the differential equation (3.7). For a detailed explanation of the shape of the adjoint variable as a function of a see Section 5. The resulting optimal immigration profile $M^*(a)$ is illustrated in Figure 3.2c. In Figure 3.2e, the age structure of the optimal SI population is depicted. As typical for a closed stationary population, the age structure of an SI population exhibits a flat line at young ages due to the low mortality at these ages. The resulting minimal dependency ratio is 75.14%[3], which corresponds to about 75 dependents per 100 workers. The resulting total SI (stationary through immigration) population size is 13.0 million females.

Scenario 2. We also performed the calculations with $M_{\text{tot}} = 50000$ which is close to the actual total number of (female) immigrants for Austria in 2008. The age-specific upper bound was set to $\bar{M}(a) = 20M_{\text{act}}(a)$ which corresponds to a high supply of immigrants at all ages. From the switching function depicted in Figure 3.2b, we can conclude that the optimal

---

[3]Typically, the dependency ratio is expressed as percentage, i.e. $100 \times D$.

immigration profile reads

$$M^*(a) = \begin{cases} \bar{M}(a) & \text{if } a \in [33, 36] \cup [109, 110], \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

See also Figure 3.2d. The resulting minimal dependency ratio is 72.24%. This corresponds to a share of 58.1% workers in the population. The resulting total size of the female SI population is 4.1 million.

Figure 3.2f represents the age structure of the optimal SI population. What is striking is that it is optimal to let people immigrate at the end of the age interval, although they are part of the economically dependent population. This can be explained by the fact that (2.5) has to be fulfilled and the age-specific bounds hold.

4. The optimal immigration profile for a fixed population size

We slightly change the model and instead of fixing the volume of immigrants (Policy 1), we require that the number of people in the population equals a prescribed value (Policy 2), i.e. we consider problem (2.1)–(2.4) with the additional constraint (2.6). Theorem 4 below states necessary conditions for the optimal solution.

Theorem 2. *If $(\tilde{N}^*(\cdot), \tilde{M}^*(\cdot))$ is an optimal solution of problem (2.1)–(2.4) and (2.6), then there are Lagrange multipliers $\tilde{\lambda}_1, \tilde{\lambda}_2$, and an adjoint variable $\tilde{\xi}(\cdot)$ such that:*
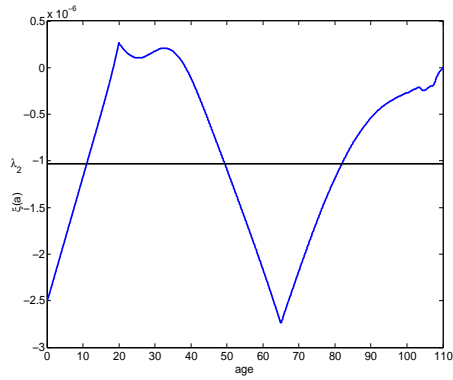
i) the continuous function $\tilde{\xi}(\cdot)$ on $[0, \omega]$ satisfies

$$\dot{\tilde{\xi}}(a) = \mu(a)\tilde{\xi}(a) - \tilde{\lambda}_1 f(a) - \mathbb{I}_{[\alpha,\beta]}(a) + \tilde{\lambda}_2,$$
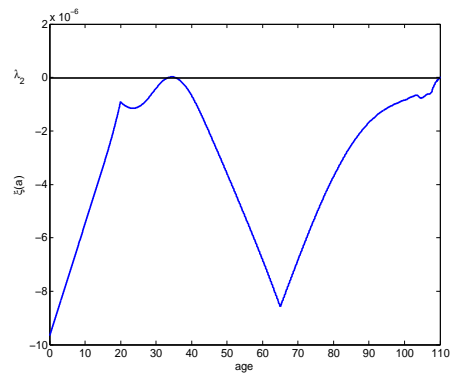$$\tilde{\xi}(0) = \tilde{\lambda}_1, \quad \tilde{\xi}(\omega) = 0, \tag{4.14}$$

ii) and the maximum principle holds for almost every $a \in (0, \omega)$

$$\tilde{\xi}(a)\tilde{M}^*(a) = \max_{0 \le M \le \bar{M}(a)} \tilde{\xi}(a)M. \tag{4.15}$$
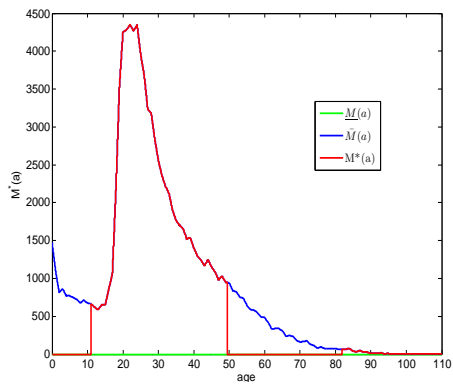
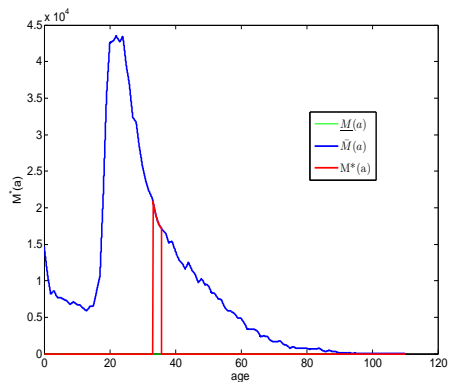**Proof 2.** *For the proof see Appendix D.*

12

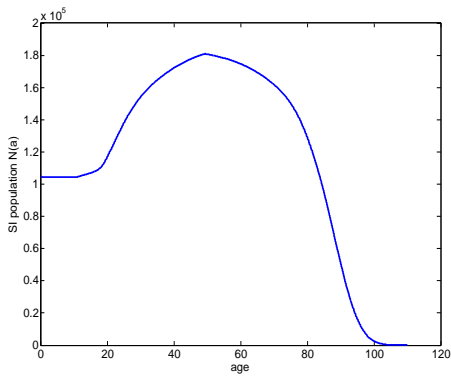(a) Scenario 1: The adjoint variable $\xi(\cdot)$.

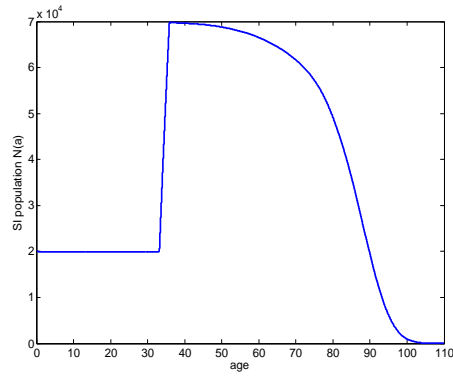(b) Scenario 2: The adjoint variable $\xi(\cdot)$.

(c) Scenario 1: The optimal immigration profile $M^*(\cdot)$.

(d) Scenario 2: The optimal immigration profile $M^*(\cdot)$.

(e) Scenario 1: The age structure of the SI population $N^*(\cdot)$.

(f) Scenario 2: The age structure of the SI population $N^*(\cdot)$.

Figure 3.2: **Policy 1:** The adjoint variable and optimal solution of problem (2.1)–(2.4) under Scenario 1 to the left and under Scenario 2 to the right. The black line indicates the switching line.

13

The optimal immigration profile is again of bang-bang type,

$$\tilde{M}^*(a) = \begin{cases} \bar{M}(a) & \text{if } \tilde{\xi}(a) > 0, \\ \text{not determined} & \text{if } \tilde{\xi}(a) = 0, \\ 0 & \text{if } \tilde{\xi}(a) < 0, \end{cases} \qquad (4.16)$$

and it remains to determine $\tilde{\xi}(\cdot)$. Similar calculations as in Section 3 give

$$\tilde{\xi}(a) = \tilde{\lambda}_1 v(a) + e_{[\alpha,\beta]}(a) - \tilde{\lambda}_2 e_{[0,\omega]}(a), \qquad (4.17)$$

where, using the boundary condition $\xi(0) = \tilde{\lambda}_1$, we obtain

$$\tilde{\lambda}_1 = \frac{e_{[\alpha,\beta]}(0) - \tilde{\lambda}_2 e_{[0,\omega]}(0)}{1 - R}. \qquad (4.18)$$

In order to determine the optimal solution $(\tilde{N}^*(\cdot), \tilde{M}^*(\cdot))$, the Lagrange multiplier $\tilde{\lambda}_2$ in (4.14) has to be chosen in such a way that condition (2.6) is fulfilled. Therefore, the value of $\tilde{\lambda}_2$ depends on the choice of the prescribed value N_{tot}. Note, that $\tilde{\xi}(\cdot)$ is independent of the optimal solution $(\tilde{N}^*(\cdot), \tilde{M}^*(\cdot))$ and can therefore be calculated separately for each $\tilde{\lambda}_2$.

4.1. Analytical study of the optimal immigration profile

In the following, we derive general results for the optimal immigration profile for given age-specific fertility $f(a)$ and mortality $\mu(a)$ rates. We show that the optimal immigration profile attains its upper bound $\bar{M}(a)$ on no more than two separate intervals.

Since the change of the sign of the adjoint variable $\tilde{\xi}(a)$ determines the switches of the optimal solution from one limit to the other, we count how many times the switching function (4.17) can cross its switching level $\tilde{\xi}(a) = 0$. To estimate this number from above we count how many times the derivative in (4.14) can change its sign at level $\tilde{\xi}(a) = 0$ from positive to negative

$$\dot{\tilde{\xi}}(a)\Big|_{\tilde{\xi}=0} = -\tilde{\lambda}_1 f(a) - \mathbb{I}_{[\alpha,\beta]}(a) + \tilde{\lambda}_2 = 0. \qquad (4.19)$$

Assumption 1. *The upper limit $\bar{M}(a)$ is such that if $M(a) \equiv \bar{M}(a)$, then $\int_0^\omega N(a)\,\mathrm{d}a > N_{\text{tot}}$.*

Assumption 2. *If $M(a) \equiv 0$, then $\int_0^\omega N(a)\,\mathrm{d}a = 0$.*

14

That means that below-replacement fertility (and mortality) holds.

Corollary 1. *There should be at least one interval with $\xi(a) > 0$.*

Otherwise the optimality condition (4.16) requires $M(a) = 0$ for almost every a. This, however, leads to the contradiction between Assumption 2 and $N_{\text{tot}} > 0$ in (2.6). $\square$

Proposition 1. $\tilde{\lambda}_2 \in [0, 1]$.

Indeed, $\tilde{\lambda}_2 < 0$ leads to $\tilde{\lambda}_1 > 0$ in (4.18) and both lead to $\dot{\tilde{\xi}}(a)\Big|_{\tilde{\xi}=0} < 0$ in (4.19) for all $a \in [0, \omega)$ so that $\xi(a) > 0$ on $a \in [0, \omega)$ which contradicts Assumption 1. If $\tilde{\lambda}_2 > 1$ then $\tilde{\lambda}_1 < 0$ because $e_{[\alpha,\beta]}(0) < e_{[0,\omega]}(0)$, thus the derivativein (4.19) has the following property $\dot{\tilde{\xi}}(a)\Big|_{\tilde{\xi}=0} = -\tilde{\lambda}_1 f(a) - \mathbb{I}_{[\alpha,\beta]}(a) + \tilde{\lambda}_2 > -1 + \tilde{\lambda}_2 > 0$ for all $a \in [0, \omega]$, since $\min\{f(a)\} = 0$. But to satisfy terminal condition $\tilde{\xi}(\omega) = 0$, for the adjoint variable it should hold, that $\tilde{\xi}(a) < 0$ for $a \in [0, \omega)$. That contradicts Assumption 2 and $N_{\text{tot}} > 0$ in (2.6). $\square$

Proposition 2.

a) $\tilde{\xi}(a) < 0$ *if* $\tilde{\lambda}_2 > 0$ *for all* $a \in [\beta, \omega)$,

b) $\tilde{\xi}(a) = 0$ *if* $\tilde{\lambda}_2 = 0$ *for all* $a \in [\beta, \omega]$.

Indeed, since $e_{[\alpha,\beta]}(a) = 0$ and $v(a) = 0$ holds for all $a \in [\beta, \omega]$ it follows from (4.17) and Proposition 1 that $\tilde{\xi}(a) = -\tilde{\lambda}_2 e_{[0,\omega]}(a) \leq 0$, $a \in [\beta, \omega]$. Thus, **b)** is obvious and **a)** follows from the inequality $e_{[0,\omega]}(a) > 0$ for all $a \in [0, \omega)$ provided that $l(a) > 0$ for all $a \in [0, \omega)$. $\square$

Assumption 3. *The fertility schedule $f(a)$ is single peaked with support to the left from β and to the right from 0, i.e. $a_{min} < \alpha < a_{max} \leq \beta$.*

Let us denote the maximal fertility age as $a_{fmax} = \arg\max(f(a))$.

Lemma 1. *The maximal number of separate intervals on which $\tilde{\xi}(a) > 0$ is two and denoting these intervals Γ_1 and Γ_2 so that for all $a_1 \in \Gamma_1$ and $a_2 \in \Gamma_2$ the inequality $a_1 < a_2$ holds, we have (see Fig. 4.3):*
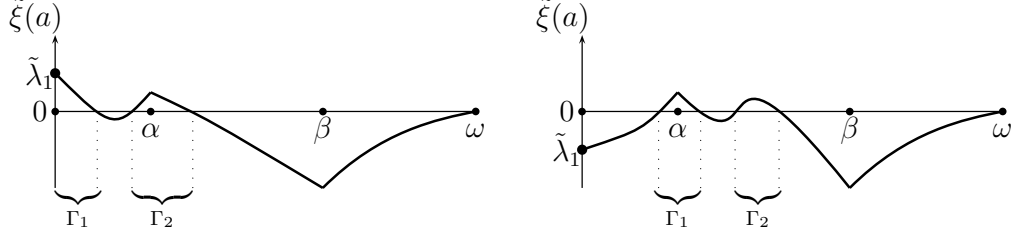
Figure 4.3: The adjoint variable $\tilde{\xi}(\cdot)$ determining the optimal immigration for two separate age intervals Γ_1 and Γ_2 in two cases: **a)** $\tilde{\lambda}_1 > 0$ (left) and **b)** $\tilde{\lambda}_1 < 0$ (right).

a) *if $a_{fmax} < \alpha$ then $\alpha \in \Gamma_2$,*

b) *if $a_{fmax} > \alpha$ then $\alpha \in \Gamma_1$.*

It follows from Proposition 2 that $\Gamma_1, \Gamma_2 \subset [0, \beta]$.

The derivative (4.19) can change its sign from plus to minus only at $a = \alpha$ because of the jump of the function $\mathbb{I}_{[\alpha,\beta]}(a)$ or/and at $a = a_0$, where a_0 is such a root of the equation $\dot{\tilde{\xi}}(a_0)\Big|_{\tilde{\xi}=0} = 0$ that $\ddot{\tilde{\xi}}(a_0)\Big|_{\tilde{\xi}=0} = -\tilde{\lambda}_1 \dot{f}(a_0) < 0$. It follows from Proposition 1 and Assumption 3 that equation $\dot{\tilde{\xi}}(a_0)\Big|_{\tilde{\xi}=0} = 0$ cannot have more than two roots all located either in $[0, \alpha)$ or in $[\alpha, \beta]$ depending on the sign of $\tilde{\lambda}_1$.

If $\tilde{\lambda}_1 > 0$ then equation $\dot{\tilde{\xi}}(a_0)\Big|_{\tilde{\xi}=0} = 0$ can only have roots in $[0, \alpha)$, where a_0 is the first root, if any, of the equation $-\tilde{\lambda}_1 f(a) + \tilde{\lambda}_2 = 0$.

If $\tilde{\lambda}_1 < 0$ then equation $\dot{\tilde{\xi}}(a_0)\Big|_{\tilde{\xi}=0} = 0$ can have roots only in $[\alpha, \beta]$ so a_0 is the second root, if any, of the equation $-\tilde{\lambda}_1 f(a) - 1 + \tilde{\lambda}_2 = 0$, which can happen only when $a_{fmax} > \alpha$.

Thus, it follows from the continuity of the function $\tilde{\xi}(a)$ that it can be positive on not more than two separate intervals. It is also easy to see graphically in Fig. 4.3 that if the function $\tilde{\xi}(a)$ is positive on two separate intervals $\Gamma_1, \Gamma_2 \subset [0, \beta]$, these intervals must contain both points a_0 and α where derivative (4.19) changes its sign, so that $a_0 \in \Gamma_1$, $\alpha \in \Gamma_2$ when $\tilde{\lambda}_1 > 0$ and $\alpha \in \Gamma_1$, $a_0 \in \Gamma_2$ when $\tilde{\lambda}_1 < 0$. $\square$

16

4.2. A case study: the Austrian case

For the calculations we initialize again the fertility and mortality profiles with Austrian data as of 2008, cf. Figure 3.1. The Lagrange multiplier $\tilde{\lambda}_2$ is calculated such that condition (2.6) is fulfilled by the optimal solution. For the total population size we prescribe the resulting sizes from Section 3, i.e. $N_{\text{tot}} = 13.0$ million and $N_{\text{tot}} = 4.1$ million, respectively.

Scenario 1. Therefore, by setting $N_{\text{tot}} = 13.0$ million and $\bar{M}(a) = 2M_{act}(a)$, we achieve a corresponding dependency ratio $D = 74.73\%$ which is slightly smaller than the one we obtain above and the resulting volume of immigrants is 72000. The corresponding optimal immigration profile reads as

$$\tilde{M}^*(a) = \begin{cases} \bar{M}(a) & \text{if } a \in [9, 41], \\ 0 & \text{otherwise,} \end{cases} \tag{4.20}$$

which is determined according to (4.16). Figure 4.4a shows the corresponding adjoint variable $\tilde{\xi}(\cdot)$ and Figure 4.4c the optimal immigration profile. The optimal age structure is depicted in Figure 4.4e.

Scenario 2. We also calculate the optimal immigration profile for $N_{\text{tot}} = 4.1$ million females and $\bar{M}(a) = 20M_{act}(a)$,

$$\tilde{M}^*(a) = \begin{cases} \bar{M}(a) & \text{if } a \in [33, 36], \\ 0 & \text{otherwise.} \end{cases} \tag{4.21}$$

Figure 4.4b shows the adjoint variable $\tilde{\xi}(\cdot)$ and Figure 4.4d the optimal immigration profile. The resulting optimal population age structure is shown in Figure 4.4f. For these parameter values we achieve a corresponding dependency ratio $D = 72.24\%$ and the resulting volume of immigrants is 50000. We observe that although the optimal immigration profiles differ, we obtain the same numerical results for the dependency ratio D and the total number of immigrants, for problem (2.1)–(2.5) and problem (2.1)–(2.6) for these numerical values. This is because the upper bound $\bar{M}(a)$ is zero for ages $a > 95$.

From the switching functions in Figure 4.4a and Figure 4.4b we also see that it is not optimal that people immigrate towards the end of the life cycle.

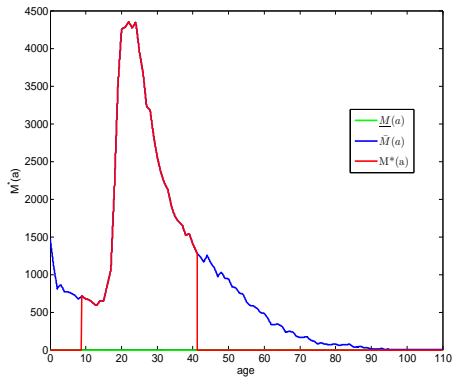## 5. Direct and indirect effect of an additional individual

The adjoint variables $\xi(a)$ and $\tilde{\xi}(a)$ may also be interpreted as shadow prices of $N(a)$ and $\tilde{N}(a)$, meaning that they reflect the decrease of the depen-
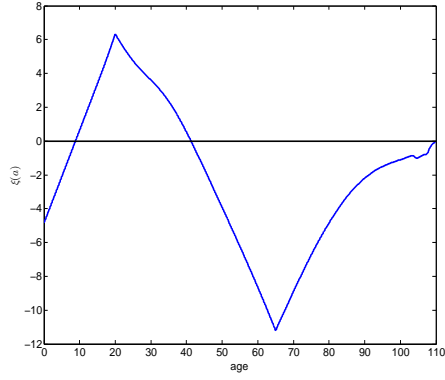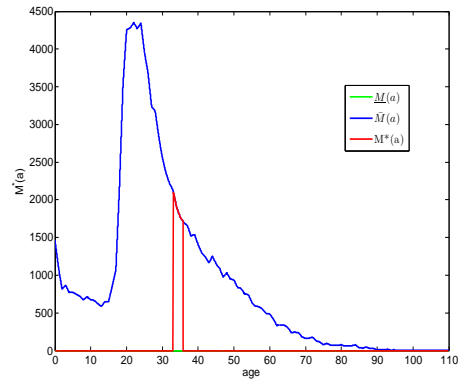
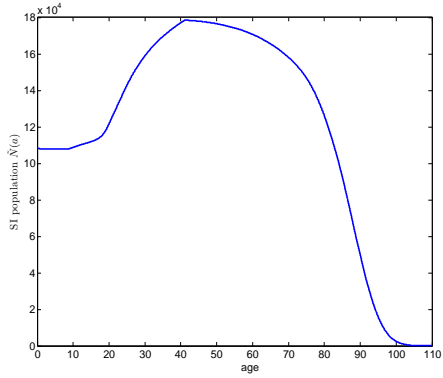(a) Scenario 1: The adjoint variable $\tilde{\xi}(\cdot)$
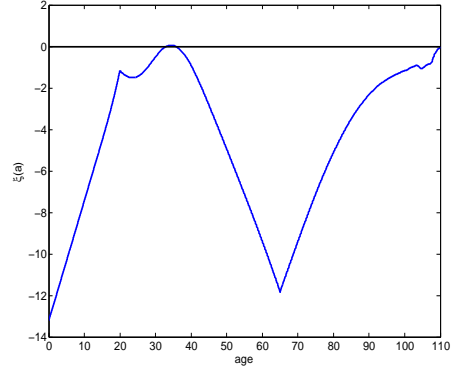
(b) Scenario 2: The adjoint variable $\tilde{\xi}(\cdot)$

(c) Scenario 1: The optimal immigration profile $\tilde{M}^*(\cdot)$.

(d) Scenario 2: The optimal immigration profile $\tilde{M}^*(\cdot)$.

(e) Scenario 1: The age structure of the SI population $\tilde{N}^*(\cdot)$.
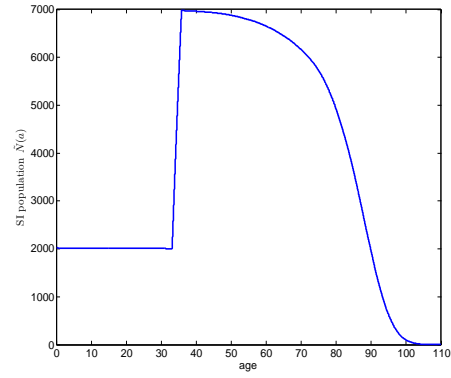
(f) Scenario 2: The age structure of the SI population $\tilde{N}^*(\cdot)$.

Figure 4.4: **Policy 2:** The adjoint variable and optimal solution of problem (2.1)–(2.4) under Scenario 1 to the left and under Scenario 2 to the right. The black line indicates the zero line.

18

dency ratio, when the optimal age structure of the population is marginally increased at age a, roughly speaking, when the population is increased by one a-year-old. A positive value of the adjoint variable means a decrease in the dependency ratio.

Note, that in this dynamic set up, a change of the (optimal) age structure at one particular age, also affects the age structure at other ages.

This shadow price, see Equation (3.10) and (4.17), consists of two parts

$$\xi(a) = \xi^d(a) + \lambda_1 v(a). \tag{5.22}$$

As pointed out in a more general setting in (Wrzaczek et al., 2010), the *direct* effect $\xi^d(a)$ determines the marginal value of an individual of age a given by her participation in the labor force. The direct effect accounts positively for her expected remaining years in $[\alpha, \beta]$ and negatively for her remaining life expectancy in $[0, \alpha]$ (for $a \leq \alpha$) and $[\beta, \omega]$.

The *indirect* effect of an a-year-old, $\lambda_1 v(a)$, is her reproductive value, i.e. the number of expected future daughters, weighted by the shadow price of newborns, λ_1, since $\xi(0) = \lambda_1$. Therefore, the indirect effect can be interpreted as the value of expected future births of an a-year-old in units of the dependency ratio. This is a generalization of the interpretation of the reproductive value, cf. (Fisher, 1930; Wrzaczek et al., 2010). Note, that the indirect effect can also be negative, namely when an additional newborn is negatively valued for the population. The corresponding interpretation holds for $\tilde{\lambda}_1$ in (4.17).

The Lagrange multiplier λ_2 may also be interpreted as the marginal effect on the dependency ratio when changing the total number of immigrants M_{tot}. Similarly, $\tilde{\lambda}_2$ measures the effect of a marginal change of the prescribed population size N_{tot} on the dependency ratio.

In Figure 5.5 we plotted the direct and indirect effect of an additional a-year-old separately. We consider again the Austrian case for problem (2.1)–(2.5), where we set $M_{\text{tot}} = 50000$ and $\bar{M}(a) = 20 M_{act}(a) \ \forall a$. The dotted line corresponds to the weighted reproductive value, representing the indirect effect. The dashed line corresponds to the direct effect. The sum of these two lines, by definition, exhibits $\xi(\cdot)$, which is depicted by the solid line. As it can be seen in Figure 5.5, the indirect effect reduces the absolute value of the adjoint variable $\xi(\cdot)$ in early ages, preventing these ages to be optimal. Furthermore, this effect is zero for ages older than the maximum age of childbearing. Therefore, after this age the direct effect and the adjoint variable coincide.

19

We also see from equation (3.10) that the direct effect always increases until age 20. This is due to the fact, that the remaining life expectancy decreases, implying a higher value of this individual in units of the dependency ratio and also because the ratio between number of person-years lived in the working ages, $\int_\alpha^\beta l(x)\,\mathrm{d}x$, and the individual's probability to survive until age a, $l(a)$, increases with a.

Moreover, we see in Figure 5.5, that the direct effect reaches its maximum at age 20, since these individuals spend their whole working life in the receiving country, and then falls monotonically until age 65. However, the sharp increase in the indirect effect between ages $[20, 40]$ shifts the optimal age away from 20 and further to the right.

The increase of the direct effect after age 65 is due to the fact that the remaining life expectancy in $[0, \omega]$, which is the only term left in equation (3.10), is decreasing with age, and therefore the burden induces by these females on the dependency ratio is reduced.

Moreover, for the particular optimal control problem considered in here it holds that under Policy 1, $\xi(a) - \lambda_2$, and under Policy 2, $\tilde{\xi}(a)$, give the decrease in the dependency ratio when changing the optimal age structure of immigrant inflows. So, under Policy 2 the shadow price is only a part of the total effect of an additional immigrant.

6. Discussion

The aim of the present paper is to determine the age-specific immigration policy that minimizes the dependency ratio in a population with below-replacement fertility assuming that the vital rates remain constant over time. We apply optimal control theory which is a rather new approach in demographic research. We assume that there are age-specific bounds that constrain the immigration profile from above. We consider two alternative policies. In the first one, we prescribe the total number of immigrants. In the second one, we fix the total population size while the rest of the model remains the same. Since the immigration profile enters the problem linearly, the solution exhibits a bang-bang behavior, which depends on th sign of the so-called switching function. The shape of the switching function with varying age a is determined by the adjoint variable.

In the model with a fixed total number of immigrants, it is shown that in the optimal solution there are ages in the vicinity of the maximum attainable age where immigration occurs. When we fix the total population size of the
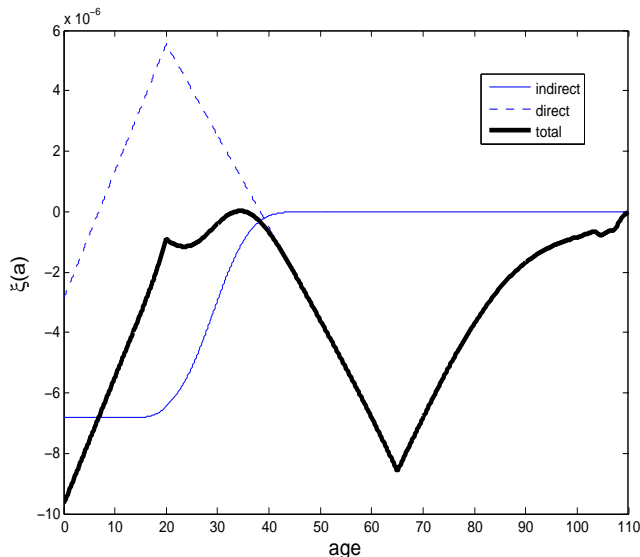
Figure 5.5: The direct (dashed line), indirect (solid blue line) and total (solid black line) effect of an additional a-year-old.

receiving country, the optimal solution is that immigration happens at not more than two separate age intervals and in ages younger than the retirement age. We present numerical results for a case study of the Austrian population based on demographic data from 2008 which underline our theoretical findings.

Moreover, by analyzing the shape of the switching function or, equivalently, the adjoint variable, and interpreting it as a shadow price, we determine the marginal value of an a-year-old individual in terms of the objective function.

7. Extensions

In future work, we also aim to study the transitory case, where we consider time varying fertility, mortality and immigration rates. Similar as in (Feichtinger and Veliov, 2007), the resulting problem is a distributed control problem, which is formalized on infinite horizon. The state dynamics is a first order partial differential equation, which is of McKendrick-type (Keyfitz, 1977; Keyfitz and Keyfitz, 1997). Although, the similarity in the structure of the problem indicates that as in (Feichtinger and Veliov, 2007)

it holds that for stationary data, i.e. fertility and mortality rates, the optimal solution is also stationary, this result does not follow immediately and needs some deeper mathematical involvement. Also in the transitory case, similar analysis of the adjoint variables which again can be interpreted as shadow prices can be carried out, cf. (Wrzaczek et al., 2010). Therefore, optimality conditions for this distributed parameter control model have to be derived in order to obtain necessary conditions for the optimal solutions. These optimality conditions obtain partial differential equations, which have to be solved numerically.

8. Acknowledgements

Appendix A. Pontryagin's Maximum Principle

We shall formulate Pontryagin's Maximum Principle in the form obtained in (Alekseev et al., 1987). Let us consider the problem

$$
\mathcal{L}_0(x(\cdot), u(\cdot), t_0, t_1) = \int_{t_0}^{t_1} f_0(t, x(t), u(t)) \, \mathrm{d}t
$$
$$
+ \psi_0(t_0, x(t_0), t_1, x(t_1)) \to \inf, \qquad (A.1)
$$
$$
\frac{\mathrm{d}x}{\mathrm{d}t} = \varphi(t, x(t), u(t)), \quad u(t) \in \mathfrak{U}, \qquad (A.2)
$$
$$
\mathcal{L}_i(x(\cdot), u(\cdot), t_0, t_1) = \int_{t_0}^{t_1} f_i(t, x(t), u(t)) \, \mathrm{d}t
$$
$$
+ \psi_i(t_0, x(t_0), t_1, x(t_1)) \lessgtr 0, \qquad (A.3)
$$

where $i = 1, 2, \ldots, m$. We introduce the so-called Pontryagin function

$$
H(t, x, u, p) = p \, \varphi(t, x, u) - \sum_{i=0}^{m} \lambda_i \, f_i(t, x, u), \qquad (A.4)
$$

where function $p(t)$ is called the adjoint and λ_i are the Lagrange multipliers.
In Alekseev et al. (1987), p. 218 the following theorem is proven.

Theorem 3 (Pontryagin Maximum Principle). *Let G be an open set in the space $\mathbf{R} \times \mathbf{R}^n$, let W be an open set in the space $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R} \times \mathbf{R}^n$ and let $\mathfrak{U}$ be an arbitrary topological space. Let the functions f_i: $G \times \mathfrak{U} \to \mathbf{R}$, $i = 0, 1, \ldots, m$, φ: $G \times \mathfrak{U} \to \mathbf{R}^n$, and their partial derivatives with respect to x be continuous in $G \times \mathfrak{U}$, and let the functions ψ_i, $i = 1, \ldots, m$ are continuously differentiable in W.*

If $(x^*(\cdot), u^*(\cdot), t_0^*, t_1^*,)$ is an optimal process for the problem (A.1)–(A.3), then there are Lagrange multipliers

$$\lambda_0 \geq 0, \quad \lambda = (\lambda_1, \ldots, \lambda_m),$$

not all zero, and an adjoint variable $p(\cdot)$ such that:

a) *the adjoint equation*

$$
\begin{aligned}
\frac{\mathrm{d}p}{\mathrm{d}t} &= -p(t)\frac{\partial \varphi}{\partial x}(t, x^*(t), u^*(t)) + \sum_{i=0}^{m} \lambda_i \frac{\partial f_i}{\partial x}(t, x^*(t), u^*(t)) \\
&= -\frac{\partial H}{\partial x}(t, x^*(t), u^*(t), p(t)),
\end{aligned}
\tag{A.5}
$$

along with the transversality conditions

$$p(t_1^*) = -\sum_{i=0}^{m} \lambda_i \frac{\partial \psi_i}{\partial x_1}(t_0^*, x_0^*, t_1^*, x_1^*) \tag{A.6}$$

$$p(t_0^*) = \sum_{i=0}^{m} \lambda_i \frac{\partial \psi_i}{\partial x_0}(t_0^*, x_0^*, t_1^*, x_1^*) \tag{A.7}$$

$$H^*(t_1^*) = -\sum_{i=0}^{m} \lambda_i \frac{\partial \psi_i}{\partial t_1}(t_0^*, x_0^*, t_1^*, x_1^*) \tag{A.8}$$

$$H^*(t_0^*) = \sum_{i=0}^{m} \lambda_i \frac{\partial \psi_i}{\partial t_0}(t_0^*, x_0^*, t_1^*, x_1^*) \tag{A.9}$$

the maximum principle in Hamiltonian (Pontryagin) form

$$H^*(t) \equiv H(t, x^*(t), u^*(t), p(t)) \equiv \max_{v \in \mathfrak{U}} H(t, x^*(t), v, p(t)) \tag{A.10}$$

the function $H^*(t)$ being continuous on the closed interval $[t_0^*, t_1^*]$.

b) *the condition of concordance of signs holds:*

$$\lambda_i \gtrless 0; \tag{A.11}$$

23

c) the conditions of complementary slackness hold:

$$\lambda_i \mathscr{L}_i(x^*(\cdot), u^*(\cdot), t_0^*, t_1^*) = 0, \quad i = 1, 2, \ldots, m \qquad (A.12)$$

(inequalities (A.11) mean that $\lambda_i \geq 0$ if $\mathscr{L}_i \leq 0$ in condition (A.3), $\lambda_i \leq 0$ if $\mathscr{L}_i \geq 0$, and λ_i may have an arbitrary sign if $\mathscr{L}_i = 0$).

Appendix A.1. Interpretation of the adjoint variable

One reason why the Maximum Principle is very often applied to economic problems, is the interpretation of the adjoint variable as shadow price. If an optimal control u^* is implemented and the corresponding optimal evolution of the state is x^*, then an infinitesimal *external* change of the state variable Δx^* at any time t would change the optimal performance by $p(t)\Delta x^*$. This results from the heuristic proof of the Maximum Principle via the Hamilton-Jacobi-Bellman equation, cf. (Léonard and Long, 1992), where one can see that for the adjoint variable $p(t)$ it holds that

$$p(t) := \frac{\partial V^*}{\partial x}(t, x(t)).$$

Function $V^* : [0, T] \times \mathbb{R}^n \to \mathbb{R}$ is called *value function*:

$$V^*(t, x(t)) = \inf_{u \in \mathfrak{U}} \int_t^{t_1} f_0(\tau, x(\tau), u(\tau)) \, \mathrm{d}\tau.$$

It gives the optimal objective value beginning at t in state $x(t)$, cf. (Dorfman, 1969; Léonard and Long, 1992).

Appendix A.2. Example of a Fishery Model

To illustrate the application of the above theorem, let us turn to an idealized nonlinear dynamic optimization problem which is a variation of the fishery model stated in Clark (1976). The state equation is

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \varphi(x) - u \qquad (A.13)$$

with given initial condition

$$x(0) = x_0. \qquad (A.14)$$

It is assumed that $\varphi(x)$ is continuously differentiable in x. Now we suppose that the revenue obtained from harvesting is a nonlinear function $R(u)$, with

$R(u)$ being a smooth, convex, nonnegative function of harvesting $u \in [0, \bar{u}]$. For simplicity we neglect the costs of harvesting. The objective function is then

$$\mathcal{L}_0(x(\cdot), u(\cdot)) = \int_0^T e^{-\rho t} R(u(t)) \, \mathrm{d}t \to \max_{u(\cdot) \in [0, \bar{u}]} \qquad (A.15)$$

subject to the state equation (A.13) and the integral constraint due to a fishing quota $\psi_1 > 0$

$$\mathcal{L}_1(x(\cdot), u(\cdot)) = \int_0^T u(t) \, \mathrm{d}t - \psi_1 \leq 0, \qquad (A.16)$$

where $\rho > 0$ is the discount rate. We assume that the specifications of the problem are such that for any feasible control u no extinction of the species is possible at any time $t \in [0, T]$.

In terms of problem (A.1)–(A.3) we have $f_0(t, u(t)) = -e^{-\rho t} R(u(t))$ and $f_1 = u(t)$. The Pontryagin function (A.4) of this problem is

$$H = p(\varphi(x) - u) + \lambda_0 e^{-\rho t} R(u) - \lambda_1 u. \qquad (A.17)$$

Thus, the necessary optimality conditions (A.5)–(A.12) are as follows. The adjoint equation (A.5) along with transversality condition (A.6) is

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -p \frac{\partial \varphi}{\partial x}(x), \quad p(T) = 0. \qquad (A.18)$$

The maximum principle (A.10)

$$\max_{u \in [0, \bar{u}]} \left(p(\varphi(x) - u) + \lambda_0 e^{-\rho t} R(u) - \lambda_1 u \right) \qquad (A.19)$$

where $\lambda_0 \geq 0$ and $\lambda_1 \geq 0$ provided that $\lambda_1 \left(\int_0^T u(t) \, \mathrm{d}t - \psi_1 \right) = 0$. If $\lambda_0 \neq 0$, like in this example, then it is set $\lambda_0 = 1$.

Appendix B. Proof of Theorem 1

In order to prove the existence of Langrange multipliers $\lambda_1, \lambda_2 \in \mathbb{R}$ and an adjoint variable $\xi(\cdot)$, such that the optimal solution $N^*(\cdot), M^*(\cdot))$ can be characterized by conditions $(3.7) - (3.8)$, we state the optimal control problem in the form of (A.1)–(A.3) so that the maximum principle in Appendix A is applicable.

Therefore, in addition to $N(a)$ we introduce the auxiliary state variables $X(a)$, $Y(a)$, being continuous functions of a. The corresponding state equations read

$$\dot{X}(a) = N(a), \quad X(0) = 0, \tag{B.1}$$

$$\dot{Y}(a) = \mathbb{I}_{[\alpha,\beta]}(a)N(a), \quad Y(0) = 0. \tag{B.2}$$

Equivalently, it holds that

$$X(a) = \int_0^a N(\tau)\,\mathrm{d}\tau \quad \text{and} \quad Y(a) = \int_\alpha^a \mathbb{I}_{[\alpha,\beta]}(\tau)N(\tau)\,\mathrm{d}\tau.$$

In this way we can express the objective function (2.1) by evaluating functions $X(\cdot)$ and $Y(\cdot)$ at the terminal value ω. Therefore, solving problem (2.1)–(2.4) with the additional constraint (2.5) is equivalent to solving

$$\min_{M(a)} \frac{X(\omega)}{Y(\omega)}, \tag{B.3}$$

subject to

$$\dot{N}(a) = -\mu(a)N(a) + M(a), \tag{2.2}$$

$$\dot{X}(a) = N(a), \quad X(0) = 0, \tag{B.1}$$

$$\dot{Y}(a) = \mathbb{I}_{[\alpha,\beta]}(a)N(a), \quad Y(0) = 0, \tag{B.2}$$

$$N(0) = \int_0^\omega f(a)N(a)\,\mathrm{d}a, \tag{2.3}$$

$$0 \le M(a) \le \bar{M}(a), \tag{2.4}$$

$$M_{\mathrm{tot}} = \int_0^\omega M(a)\,\mathrm{d}a. \tag{2.5}$$

Introducing Lagrange multipliers λ_1, λ_2, and the adjoint variables $\xi(\cdot)$, $\zeta(\cdot)$ and $\eta(\cdot)$ we define Pontryagin's function as

$$H(a, N, X, Y, M, \xi, \zeta, \eta) =$$
$$\xi\left(-\mu(a)N + M\right) + \zeta N + \eta \mathbb{I}_{[\alpha,\beta]}(a)N$$
$$- \lambda_0 \frac{X(\omega)}{Y(\omega)} - \lambda_1 f(a)N - \lambda_2 M. \tag{B.4}$$

The conditions stated in the maximum principle (Appendix A) provide necessary conditions for the optimal solution (N^*, X^*, Y^*, M^*) of problem (B.3)–(2.5) which can be summarized by the following expressions.

The maximum principle according to (A.10) reads

$$(\xi(a) - \lambda_2)M^*(a) = \max_{0 \leq M \leq \bar{M}(a)} H = \max_{0 \leq M \leq \bar{M}(a)} (\xi(a) - \lambda_2)M, \qquad \text{(B.5)}$$

and the adjoint equation (A.5) along with transversality condition (A.6) can be reduced to

$$\dot{\xi}(a) = \mu(a)\xi(a) - \lambda_1 f(a) - \frac{X(\omega)}{Y^2(\omega)}\mathbb{I}_{[\alpha,\beta]}(a) + \frac{1}{Y(\omega)}, \quad \xi(0) = \lambda_1, \quad \xi(\omega) = 0,$$
$$\text{(B.6)}$$

where $\eta = \frac{X(\omega)}{Y^2(\omega)} = \frac{(D+1)^2}{N_{\text{tot}}}$ and $\zeta = \frac{1}{Y(\omega)} = \frac{D+1}{N_{\text{tot}}}$. $\square$

Appendix C. Optimal immigration arbitrary close to the maximum attainable age

The precise formulation of this result reads as follows.

Theorem 4. *Let $M(\cdot)$ be an arbitrary immigration profile which fulfills (2.4),(2.5) and additionally $M(a) < \bar{M}(a)$ for $a \in [\omega - \delta, \omega]$ and some $\delta > 0$. Then there is an immigration profile $\tilde{M}(\cdot)$ which satisfies (2.4),(2.5) such that*
$$D(\tilde{M}) < D(M).$$

For the proof of Theorem 4 we consider the maximization problem

$$\max_{M(a)} J(M(\cdot)), \qquad \text{(C.1)}$$

subject to

$$M_{\text{tot}} = \int_0^\omega M(a)\,\mathrm{d}a, \qquad \text{(C.2)}$$

$$0 \leq M(a) \leq \bar{M}(a). \qquad \text{(C.3)}$$

It holds that $J(M(\cdot)) = \frac{\int_0^\omega F(a)M(a)\,da}{\int_0^\omega G(a)M(a)\,da}$, where functions $F(a)$ and $G(a)$ are obtained by the Cauchy formula for equations (2.2), (2.3). This problem is equivalent to the minimization problem (2.1)–(2.5). Note, that $J(M(\cdot)) = 1 - D(M(\cdot))$. We now determine $F(a)$ and $G(a)$:

$$N(a) = l(a)N(0) + \int_0^a \frac{l(a)}{l(s)} M(s)\,ds, \tag{C.4}$$

$$N(0) = \frac{1}{1-R} \int_0^\omega f(a) \int_0^a \frac{l(a)}{l(s)} M(s)\,ds\,da, \tag{C.5}$$

$$= \frac{1}{1-R} \int_0^\omega M(s)v(s)\,ds. \tag{C.6}$$

Taking the integral of equation (C.4) over the interval $[0, \omega]$ yields

$$\int_0^\omega N(a)\,da = \int_0^\omega \frac{l(a)}{1-R}\,da \int_0^\omega M(s)v(s)\,ds + \int_0^\omega \int_0^a \frac{l(a)}{l(s)} M(s)\,ds\,da,$$

$$= \frac{e_{[0,\omega]}(0)}{1-R} \int_0^\omega M(s)v(s)\,ds + \int_0^\omega e_{[0,\omega]}(s)M(s)\,ds,$$

$$= \int_0^\omega \left(\frac{e_{[0,\omega]}(0)}{1-R} v(s) + e_{[0,\omega]}(s) \right) M(s)\,ds,$$

$$= \int_0^\omega G(s)M(s)\,ds,$$

where $G(s) := \frac{e_{[0,\omega]}(0)}{1-R} v(s) + e_{[0,\omega]}(s)$. Analogously, we obtain

$$\int_\alpha^\beta N(a)\,da = \int_0^\omega F(s)M(s)\,ds,$$

where $F(s) := \frac{e_{[\alpha,\beta]}(0)}{1-R} v(s) + e_{[\alpha,\beta]}(s)$. Function $F(s)$ can be interpreted as an s-year-old immigrant's effect on the age structure of the population. The first term is then the contribution of all of her future native-born descendants to the age group $[\alpha, \beta]$, (counting children, grandchildren and so forth) in the resulting SI population and the second term may be viewed as the immigrant's own effect by being within the working age. The analogous interpretation holds for the above function $G(s)$ for the age interval $\in [0, \omega]$.

Furthermore, we assume that the following assumption holds:

28

Regularity Assumption 1. *For any $c > 0$ it holds that*

$$F(a) \neq c\, G(a),$$

almost everywhere in $[0, \omega]$.

This assumption means that an immigrant's effect on the working population is not proportional to its effect on the overall population.

For the proof of Theorem 4 we need the following Lemma:

Lemma 2. *For any immigration profile $M(\cdot)$ satisfying (2.5), (2.4) there exists a set $\Gamma \subset [0, \omega]$, $meas(\Gamma) > 0$ such that $M(a) > 0$ for $a \in \Gamma$ and*

$$\frac{F(a)}{G(a)} < J(M(\cdot)), \quad \forall a \in \Gamma,$$

holds.

Assume that

$$\frac{F(a)}{G(a)} \geq \frac{\int_0^\omega F(s)M(s)\,\mathrm{d}s}{\int_0^\omega G(s)M(s)\,\mathrm{d}s}, \quad \forall a \in \{s : M(s) > 0\} =: \Gamma^0, \ meas(\Gamma^0) > 0.$$

Because of the regularity assumption the strict inequality

$$F(a)\int_0^\omega G(a)M(a)\,\mathrm{d}a > G(a)\int_0^\omega F(a)M(a)\,\mathrm{d}a,$$

holds on a subset $\Gamma \subset \Gamma^0$ of positive measure. Multiplying both sides by $M(a)$ and integrating on $[0, \omega]$ we obtain

$$\int_0^\omega F(a)M(a)\,\mathrm{d}a \int_0^\omega G(a)M(a)\,\mathrm{d}a > \int_0^\omega G(a)M(a)\,\mathrm{d}a \int_0^\omega F(a)M(a)\,\mathrm{d}a,$$

which gives a contradiction. $\square$

Proof of Theorem 4: Let Γ be the set from Lemma 2, and let $b \in \Gamma$ be a Lebesgue point. Recall that almost every point of Γ is such. Let us define an immigration profile $\tilde{M}(\cdot)$

$$\tilde{M}(a) := \begin{cases} M(a) & a \notin [b - \delta, b] \cup [\omega - \delta, \omega], \\ M(a) - h & a \in [b - \delta, b], \\ M(a) + h & a \in [\omega - \delta, \omega], \end{cases}$$

where $M(a) > 0$ and $0 < h \leq \bar{M}(a) - M(a)$ holds. The corresponding objective value reads as

$$J(\tilde{M}(\cdot)) = \frac{\int_0^\omega F(a)M(a)\,\mathrm{d}a - h\int_{b-\delta}^b F(a)\,\mathrm{d}a + h\int_{\omega-\delta}^\omega F(a)\,\mathrm{d}a}{\int_0^\omega G(a)M(a)\,\mathrm{d}a - h\int_{b-\delta}^b G(a)\,\mathrm{d}a + h\int_{\omega-\delta}^\omega G(a)\,\mathrm{d}a}.$$

We define
$$H(\delta) := h\int_{x-\delta}^x F(a)\,\mathrm{d}a, \quad x = b, \omega,$$

where, by transformation of the independent variable, $H(\delta) = h\int_0^\delta F(x-t)\,\mathrm{d}t$ holds. By Taylor expansion around 0 we obtain

$$H(\delta; x) = h(H(0; x) + \delta H'(0; x) + \delta^2 H''(0; x) + o(\delta^2)),$$
$$= h\delta F(x) + h\delta^2 F'(x) + ho(\delta^2).$$

By $o(\delta^2)$ we mean, that F'' grows slower than δ^2. The same approach is used for G. Therefore, by neglecting all terms but the linear one in δ,

$$J(\tilde{M}(\cdot)) = \frac{\int_0^\omega F(a)M(a)\,\mathrm{d}a - \delta h F(b) + \delta h F(\omega)}{\int_0^\omega G(a)M(a)\,\mathrm{d}a - \delta h G(b) + \delta h G(\omega)}.$$

Note, that $G(\omega) = F(\omega) = 0$ and therefore it holds that

$$J(\tilde{M}(\cdot)) - J(M(\cdot)) > 0$$
$$\Leftrightarrow \quad \frac{\int_0^\omega F(a)M(a)\,\mathrm{d}a - \delta h F(b)}{\int_0^\omega G(a)M(a)\,\mathrm{d}a - \delta h G(b)} > \frac{\int_0^\omega F(a)M(a)\,\mathrm{d}a}{\int_0^\omega G(a)M(a)\,\mathrm{d}a}$$
$$\Leftrightarrow \quad -F(b)\int_0^\omega G(a)M(a)\,\mathrm{d}a > -G(b)\int_0^\omega F(a)M(a)\,\mathrm{d}a$$
$$\Leftrightarrow \quad \frac{F(b)}{G(b)} < \frac{\int_0^\omega F(a)M(a)\,\mathrm{d}a}{\int_0^\omega G(a)M(a)\,\mathrm{d}a}$$

which is fulfilled by the choice of $b \in \Gamma$ as was proven in Lemma 2.

Since problem (2.1)–(2.5) and problem (B.3)–(2.4) and therefore $J(\tilde{M}(\cdot)) > J(M(\cdot))$ and $D(\tilde{M}(\cdot)) < D(M(\cdot))$ are equivalent we have thus proven Theorem 4. $\square$

Appendix D. Proof of Theorem 2

We consider problem (2.1)–(2.4) with the additional constraint (2.6). Note, that minimizing the dependency ratio D in a population with fixed size is equivalent to maximizing the number of working people

$$\max_{M(a)} \int_0^\omega \mathbb{I}_{[\alpha,\beta]}(a) N(a) \, \mathrm{d}a. \tag{D.1}$$

Again, we define the Pontryagin function as

$$H(a, N, X, Y, M, \tilde{\xi}) =$$
$$\tilde{\xi}(-\mu(a)N + M) + \mathbb{I}_{[\alpha,\beta]}(a)N - \tilde{\lambda}_1 f(a)N - \tilde{\lambda}_2 N, \tag{D.2}$$

and aim to apply Pontryagin's maximum principle presented in Appendix A. The optimality conditions for (N^*, M^*) can be formulated by the following expressions

$$\tilde{\xi}(a)M^*(a) = \max_{0 \leq M \leq \tilde{M}(a)} H = \max_{0 \leq M \leq \tilde{M}(a)} \tilde{\xi}(a)M(a), \tag{D.3}$$

$$\dot{\tilde{\xi}}(a) = \mu(a)\tilde{\xi}(a) - \tilde{\lambda}_1 f(a) + \mathbb{I}_{[\alpha,\beta]}(a) + \tilde{\lambda}_2, \quad \tilde{\xi}(0) = \tilde{\lambda}_1, \quad \tilde{\xi}(\omega) = 0, \tag{D.4}$$

where $\tilde{\lambda}_1$ should be calculated in such a way that (2.6) is satisfied for the resulting optimal solution. $\square$

References

Alekseev, V., Tikhomirov, V., Fomin, S., 1987. Optimal Control. Consultants Bureau, New York.

Anita, S., 2000. Analysis and Control of Age-Dependent Population Dynamics. Springer Netherlands.

Arthur, W., Espenshade, T., 1988. Immigration policy and immigrants' ages. Population and Development Review 14(2), 315–326.

Clark, C., 1976. Mathematical Bioeconomics. John Wiley & Sons.

Dawid, H., Feichtinger, G., Goldstein, J., Veliov, V., 2009. Keeping a learned society young. Demographic Research 20, 541–558.

Dorfman, R., 1969. An economic interpretation of optimal control theory. The American Economic Review 59 (5), 817–831.

Feichtinger, G., Veliov, V., 2007. On a distributed control problem arising in dynamic optimization of a fixed-size population. SIAM Journal on Optimization 18(3), 980–1003.

Fisher, R., 1930. The Genetical Theory of Natural Selection. Dover Publications, New York.

Grass, D., Caulkins, J., Feichtinger, G., Tragler, G., Behrens, D., 2008. Optimal Control of Nonlinear Processes. Springer Berlin Heidelberg.

Keyfitz, N., 1977. Applied Mathematical Demography. John Wiley & Sons.

Keyfitz, N., Keyfitz, B., 1997. The McKendrick partial differential equation and its uses in epidemiology and population study. Math. Comput. Modelling Vol. 26(6), l–9.

Léonard, D., Long, N., 1992. Optimal Control Theory and Static Optimization in Economics. Camebridge University Press.

Mitra, S., 1990. Immigration, below-replacement fertility, and long-term national population trends. Demography 27(1), 121–129.

Schmertmann, C., 1992. Immigrants' ages and the structure of stationary populations with below-replacement fertility. Demography 29(4), 595–612.

Schmertmann, C., 2011. Stationary populations with below-replacement fertility. Working Paper.

Statistics Austria, 2010a. Einjährige Fertilitätsziffern seit 1961, Data download on 4 July 2011.
URL http://www.statistik.at/web_en/statistics/population

Statistics Austria, 2010b. Segment V6C Außenwanderung nach Herkunfts-/ Zielland, Alter, Staatsangehörigkeit, Geschlecht und NUTS 3-Einheiten 2002 bis 2010, Data download on 4 July 2011.
URL http://www.statistik.at/web_de/services/datenbank_isis/index.html

United Nations, 2001. Replacement Migration: Is it a Solution to Declining and Ageing Populations? New York: UN Population Division, Department of Economic and Social Affairs; ST/ESA/SER.A/206.

University of California, Berkeley and Max Planck Institute for Demographic Research, 2010 Human Mortality Database, Data download on 4 July 2011. URL `http://www.mortality.org`

Wrzaczek, S., Kuhn, M., Prskawetz, A., Feichtinger, G., 2010. The reproductive value in distributed optimal control models. Theoretical Population Biology 77(3), 164–170.

Wu, Z., Li, N., 2003. Immigration and the dependency ratio of a host population. Mathematical Population Studies 10(3), 21–39.