DEMOGRAPHIC RESEARCH
A peer-reviewed, open-access journal of population sciences

DEMOGRAPHIC RESEARCH

VOLUME 50, ARTICLE 6, PAGES 171–184
PUBLISHED 19 JANUARY 2024
http://www.demographic-research.org/Volumes/Vol50/6/
DOI: 10.4054/DemRes.2024.50.6

Formal Relationship

The optimal transition to a stationary population for concentrated vitality rates

Gustav Feichtinger

Stefan Wrzaczek

Contents

Demographic Research: Volume 50, Article 6
Formal Relationship

The optimal transition to a stationary population for concentrated vitality rates

Gustav Feichtinger[1]
Stefan Wrzaczek[2]

Abstract

BACKGROUND
Several countries nowadays and in the past face a birth rates below replacement level. To what extent should the fertility of this shrinking population be increased during a given planning period such that it approaches stationarity at the end as close as possible? Both immediate adaptation to the replacement level as well as delaying it to the end of the planning period are suboptimal.

METHODS
Distributed parameter optimal control theory provides an appropriate tool to ascertain the efficient intertemporal trade-off between costly birth control and zero population growth.

RESULTS
It turns out that the optimal adaptation rate of the net reproduction rate (NRR) balances between unacceptable adjustment costs for fertility and huge deviations of the terminal age composition from the desired stationary one. The optimal adaptation rate is monotonically increasing with a curvature that depends on the growth rates of the NRR, the fertile population, and the value of newborns.

CONTRIBUTION
The paper analytically characterizes the shape of the transition to a stationary population in an optimal way.

---

[1] Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/OeAW, University of Vienna); and Institute of Statistics and Mathematical Methods in Economics (Research unit VADOR), Vienna University of Technology, Vienna, Austria.

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

1. Relationship

Keyfitz (2009) asks how much a rapidly growing population continues to increase when the net reproduction rate (NRR) falls immediately to bare replacement. This momentum of demographic growth arises because the age distribution of such a population is favorable to increase. However, since nowadays the fertility of many industrialized countries is at a subreplacement level and their populations are shrinking, the problem is opposite: To avoid a more extensive population decrease, how must the NRR be increased on its way to stationarity? Due to the inertia of age-structures, even an instantaneous increase in fertility to replacement level cannot prevent an additional population decrease. Since the application of population policy measures needs time, we assume a planning period in which the fertility can be gradually adapted to reach a NRR equal to 1 at its end as well as a stationary age-structure as close as possible. However, the meaning of "gradual adaptation" remains an open question. Does it refer to a uniform increase or a higher change at the beginning of the planning interval or at its end? The answer leads us to optimality.

Consider a one-sex age-structured population (females) that is closed with respect to migration. In what follows, we deal with population policy questions, fertility being the only control instrument.

Although birth control is affected by socioeconomic, cultural, political, and other measures, in the present context we assume that births can be adapted directly. Assuming a continuous deterministic framework and denoting the NRR at time t as $R(t)$, its development over time is described by the ordinary differential equation:

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

with the adaptation rate $k(t)$ and initial NRR of R_0. Extending Coale (1972, chapter 4), who considers populations with a constant negative rate, $k(t)$ denotes a variable, which can be chosen by a central planner. Negative values of $k(t)$ decrease reproduction, whereas positive ones lead to an increasing NRR.

Two simplifying assumptions allow to obtain analytical results in the following intertemporal optimization procedure: firstly, a unique age of females, μ, at which they give birth to their girls,[3] and secondly, rectangular mortality, meaning that each individual survives to a maximal age of ω years. The population density $P(t,a)$ (aged a at time t) evolves according to the McKendrick–von Foerster partial differential equation (see

---

[3] Note that this assumption made by Coale (1972, chapter 4) and his followers is crucial for the analysis of populations whose fertility declines at a constant rate.

Keyfitz and Kezfitz 1997)

$$P_t(t,a) + P_a(t,a) \quad = \quad 0, \tag{2}$$

where the zero at the right-hand side of (2) results from the rectangular mortality assumption. $P(0,a) = P_0(a)$ (for $a \in [0,\omega]$) denotes the initial age distribution, and $P(t,0) = B(t) = P(t,\mu)$ the number of newborns at t.

To simplify the analysis we assume further that negative and positive adaptation rates $k(t)$ are equally costly with decreasing marginal efficiency reflected by convex (i.e., in our case quadratic) costs $C_1 k^2(t)$, where C_1 is a model parameter. The decision maker has two objectives over a fixed-time horizon T. The first intention is to end up with a NRR of 1 as a hard constraint in the sense that $R(T) = 1$ has to be met no matter what the costs.

Due to the inertia of age structures, the population at end time T is not stationary. Thus, we additionally demand a terminal age-structure as close as possible to the stationary one (as implied by the life table) as a second goal. Note that this is a soft constraint.

Thus, we look for the optimal trade-off between (quadratic) costs for birth control and a target $R(T) = 1$, as well as the (quadratic) deviation from the uniform distribution $\bar{c}(a)$ (due to the rectangular survival assumption):

$$\mathcal{V}(R_0) \quad := \quad \int_0^T C_1 k(t)^2 \mathrm{d}t + C_2 \int_0^\omega \left(c(T,a) - \bar{c}(a)\right)^2 \mathrm{d}a, \tag{3}$$

where $c(T,a) := \frac{P(T,a)}{\int_0^\omega P(T,a)\mathrm{d}a}$ (for $a \in [0,\omega]$).

In mathematical terms $\mathcal{V}(R_0)$ denotes the objective function, $C_1 k^2(t)$ the objective functional, and $C_2(c(T,a) - l(a))^2$ the salvage value function, where C_2 measures the unit cost of the mismatch of the final stationary age distribution. Together with the differential equations (1) and (2) as constraints, the decision maker faces the following finite time optimal control problem:

$$
\begin{aligned}
\min_{k(t)} \quad & \mathcal{V}(R_0) && \text{(4a)}\\
\text{s.t.} \quad & P_t(t,a) + P_a(t,a) = 0, \quad P(0,a) = P_0(a) && \text{(4b)}\\
& P(t,0) = B(t) && \text{(4c)}\\
& \dot{R}(t) = k(t)R(t), \quad R(0) = R_0, R(T) = 1 && \text{(4d)}\\
& \dot{N}(t) = B(t) - D(t), \quad N(0) = N_0 && \text{(4e)}\\
& B(t) = \int_0^\omega \delta(a-\mu)P(t,a)R(t)\mathrm{d}a && \text{(4f)}\\
& D(t) = \int_0^\omega \delta(a-\omega)P(t,a)\mathrm{d}a, && \text{(4g)}
\end{aligned}
$$

where $N(t)$ and $D(t)$ define the total population and deaths at t that are necessary to fit $c(t,a)$ in the form of a standard optimal control model. $\delta(x)$ denotes the Dirac delta function (introduced for mathematical convenience to apply the Maximum Principle).

To reach the terminal goal $R(T) = 1$, according to (4d) it is intuitively clear that the adaptation rate $k(t)$ has to be positive. The question remains how to distribute these efforts for a fertility increase over time to guarantee an optimal trade-off between adaptation costs and a penalty for deviating from the stationary age-structure.

Relation - shape of adaptation. Assume that an optimal solution for (4) exists and let $k^*(t)$ denote the optimal efforts to adapt the NRR. If fertility and mortality are concentrated at μ and ω, respectively, $k^*(t)$ is positive and strictly decreasing if $R_0 < 1$ over the entire time horizon, that is,

$$
\begin{matrix} k^*(t) > 0 \\ \dot{k}^*(t) < 0 \end{matrix} \qquad \text{for } R_0 < 1, t \in [0,T]. \tag{5}
$$

The curvature can be expressed as the product of the slope of $k^*(t)$ and the sum of the growth rates of the NRR, the fertile population, and the value of newborns, that is,

$$
\ddot{k}(t) \quad = \quad \left(\frac{\dot{R}(t)}{R(t)} + \frac{P_t(t,\mu)}{P(t,\mu)} + \frac{\xi_t(t,0)}{\bar{\lambda}_N + \xi(t,0)} \right) \dot{k}(t), \tag{6}
$$

with

$$\frac{\dot{R}(t)}{R(t)} > 0, \quad \frac{P_t(t,\mu)}{P(t,\mu)} \gtreqless 0, \quad \frac{\xi_t(t,0)}{\bar{\lambda}_N + \xi(t,0)} < 0. \tag{7}$$

If the growth rate of the NRR is smaller than the sum of the growth rates of the fertile population and the value of a newborn, $k^*(t)$ is convex (locally, i.e., pointwise) at $t \in [0,T]$ and concave otherwise.

The decomposition of the curvature as a product of the slope and the sum of growth rates (as explicitly written in (6)) has the following intuition. The curvature denotes (from a mathematical point of view) the slope of $k^*(t)$. In that respect the three mentioned growth rates are disentangling three effects driving the population dynamics and the (value of the) deviation from $\bar{c}(a)$ at T. Let us discuss the effects one after the other. The NRR directly relates to the number of newborns at t (as product with the fertile population) and additionally must fulfill the hard constraint to end up at 1 at T. The fertile population is mirroring the effect of the NRR. The value of a newborn corresponds to the marginal effect to the deviation from $\bar{c}(a)$ at T, which is a soft constraint in the sense that deviations are allowed but (in terms of the objective function) costly.[4]

The reason why the curvature of $k(t)$ can change the sign lies in the time-lag of problem (4), which is due to the time that elapses from the birth date of a newborn to reaching the fertile age itself, which is μ years. To build an intuition on the parameter dependence (note that μ, C_1, and C_2 are the only parameters of model (4)) of the relation let us consider the hypothetical situation of $\mu = 0$. In this case a change in the growth rate of the NRR has an immediate effect on the growth rate of the fertile population (i.e., the time-lag is zero), yielding $\frac{P_t(t,\mu)}{P(t,\mu)} > 0$ for $R_0 < 1$. Sufficiently high-cost parameters C_1 and C_2 imply that $\bar{c}(a)$ (uniform for rectangular mortality) is approached as flat as possible, yielding a nearly uniform final age-distribution (because of parameter C_2) and without any fluctuations during the transitional period (because of C_1). The latter one is possible as the time-lag is zero. Continuity of the optimal control model (4) implies that convex $k(t)$ (for $R_0 < 1$) is more likely for younger fertility ages and high C_1.

---

[4] Note that meeting $c(T,a) = \bar{c}(a)$ (for all $a \in [0,\omega]$) as a hard constraint would probably not be possible within a finite time horizon due to the time-lagged effect of the NRR on the density of the population.

2. Proof

Model (4) is an age-structured optimal control model (see Feichtinger et al. 2003) integrating age-structured state variables ($P(t,a)$) as well as concentrated state variables ($R(t)$ and $N(t)$).[5]

The Hamiltonian reads

$$\begin{aligned}\mathcal{H} \;&=\; -\frac{1}{\omega}C_1k^2(t)+\xi(t,a)\cdot 0+\frac{1}{\omega}\lambda_R(t)k(t)R(t)+\frac{1}{\omega}\lambda_N(t)(B(t)-D(t))\\ &\quad+\eta_B(t)\delta_\mu P(t,a)R(t)+\eta_D(t)\delta_\omega P(t,a), \qquad (8)\end{aligned}$$

with $\xi(t,a)$, $\lambda_R(t)$, and $\lambda_N(t)$ denoting the adjoint variables (dynamic shadow price) of (distributed and concentrated) state variables $P(t,a)$, $R(t)$, and $N(t)$, respectively. $\eta_B(t)$ and $\eta_D(t)$ are the adjoint variables of the aggregated state variables $B(t)$ and $D(t)$. Derivation yields the following first order condition:[6]

$$\omega\cdot\mathcal{H}_k \;=\; -2C_1k(t)+\lambda_R(t)R(t)=0 \qquad\Longrightarrow\qquad k^*(t)=\frac{\lambda_R(t)R(t)}{2C_1}, \qquad (9)$$

where the asterisk denotes the optimal value of $k(t)$.

Following the Maximum Principle, the shadow prices (referred to as adjoint or costate variables in optimal control theory) evolve according to the following dynamics:

$$\begin{aligned}\xi_t(t,a)+\xi_a(t,a) &= -\eta_B(t)\delta_\mu R(t)-\eta_D(t)\delta_\omega\\ \dot\lambda_R(t) &= -\lambda_R(t)k(t)-\int_0^\omega \eta_B(t)\delta_\mu P(t,a)\mathrm{d}a\\ \dot\lambda_N(t) &= 0\\ \eta_B(t) &= \xi(t,0)+\lambda_N(t)\\ \eta_D(t) &= -\lambda_N(t) \qquad (10)\end{aligned}$$

---

[5] For the age-structured Maximum Principle, including concentrated state variables (which is not presented in standard literature) see Feichtinger and Wrzaczek (2023), where also a sketch of a proof based on the variational principle of needle variations is presented.

[6] Since $k(t)$ is not bounded $\mathcal{H}_k = 0$ applies.

together with the transversality conditions

$$\begin{aligned}
\xi(t,\omega) &= 0\\
\xi(T,a) &= -\frac{2C_2}{N(T)}\int_0^\omega \left(c(T,a)-\bar{c}(a)\right)\mathrm{d}a\\
\lambda_N(T) &= \frac{2C_2}{N(T)}\int_0^\omega c(T,a)\left(c(T,a)-\bar{c}(a)\right)\mathrm{d}a\\
\lambda_R(T) & \quad \text{no transversality condition,}
\end{aligned} \tag{11}$$

where we define $\bar{\lambda}_N := \frac{2C_2}{N(T)}\int_0^\omega c(T,a)\left(c(T,a)-\bar{c}(a)\right)\mathrm{d}a$ and $\bar{\xi} := -\frac{2C_2}{N(T)}\int_0^\omega \left(c(T,a)-\bar{c}(a)\right)\mathrm{d}a$. Using $\dot{\lambda}_N(t)=0$, the transversality conditions as well as the definition of the Dirac delta function, we simplify the first two equations of (10) and obtain

$$\begin{aligned}
\xi_t(t,a)+\xi_a(t,a) &= \begin{cases} -\left(\xi(t,0)+\bar{\lambda}_N\right)R(t) & \text{for } a=\mu\\ \bar{\lambda}_N & \text{for } a=\omega\\ 0 & \text{else}\end{cases}\\
\dot{\lambda}_R(t) &= -\lambda_R(t)k(t)-\left(\xi(t,0)+\bar{\lambda}_N\right)P(t,\mu).
\end{aligned} \tag{12}$$

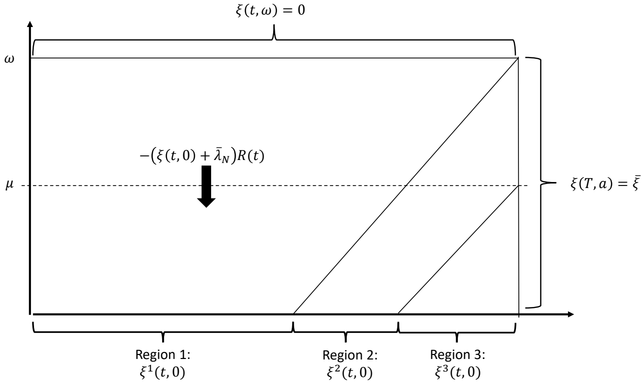
While $\bar{\lambda}_N$ is a constant according to the final age-structure of the population, $\xi(t,0)$ can be understood as the marginal value (i.e., the dynamic shadow price) of a newborn at time t that develops over time. By backward integration and using the definition of the Dirac delta function, we arrive at different expressions for $\xi(t,0)$ within (i) individuals born before $T-\omega$ (referred to as region 1), (ii) individuals born thereafter but before $T-\mu$ (region 2), and finally (iii) individuals born after $T-\mu$ (region 3).[7] Figure 1 gives an overview of the different regions within the Lexis diagram (showing time on the abscissa and age on the ordinate). Individuals still alive at T are weighted by $\bar{\xi}$ (from $\xi(T,a)=\bar{\xi}$), all others (dying before T) by 0 (from $\xi(t,\omega)=0$). During the individual life, $\xi(t,a)$ remains constant except at childbearing age $a=\mu$ (crossing the dashed line in Figure 1), where the expected reproduction enters as the sum of the effect to the total population and the generalized Fisher's reproductive value (see Fisher 1930 and Wrzaczek et al. 2010 or Feichtinger et al. 2011 for an extensive discussion).

Application of this derivation yields the following values for $\xi(t,0)$ within the three

---

[7] Note that the crossing from region 1 to 2 and from region 2 to 3 are have zero measure.

Figure 1: **Derivation of the (shadow) value of newborns within the Lexis diagram**



regions (superscript referring to the corresponding region) as

$$\begin{aligned} \xi^1(t,0) &= \xi(t+\mu,0)R(t+\mu) + \bar{\lambda}_N R(t+\mu) - \bar{\lambda}_N \\ \xi^2(t,0) &= \xi(t+\mu,0)R(t+\mu) + \bar{\lambda}_N R(t+\mu) + \bar{\xi} \\ \xi^3(t,0) &= \bar{\xi}. \end{aligned} \tag{13}$$

Solving $\dot{\lambda}_R(t)$ on the other hand gives

$$\lambda_R(t) = \lambda_R(T) e^{\int_t^T k(s')\mathrm{d}s'} + \int_t^T e^{\int_t^{s'} k(s'')\mathrm{d}s''} \left(\bar{\lambda}_N + \xi(s',0)\right) P(s',\mu)\mathrm{d}s', \tag{14}$$

which shows that the sign of $\lambda_R(t)$ is determined by the signs of $\bar{\lambda}_N + \xi(s',0)$ (noting that population is nonnegative).

To analyze the evolution of the adaptation efforts to the NRR over time, we take the time derivative of (9) and obtain the so-called Euler equation

$$\dot{k}(t) = -\frac{R(t)P(t,\mu)}{2C_1} \left(\bar{\lambda}_N + \xi(t,0)\right), \tag{15}$$

where we already used $\lambda_N(t) = \bar{\lambda}_N$. The curvature of $k(t)$ is obtained by taking the second derivative, for which we obtain

$$\ddot{k}(t) \quad = \quad \left(\underbrace{\frac{\dot{R}(t)}{R(t)}}_{=:(a)} + \underbrace{\frac{P_t(t,\mu)}{P(t,\mu)}}_{=:(b)} + \underbrace{\frac{\xi_t(t,0)}{\bar{\lambda}_N + \xi(t,0)}}_{=:(c)} \right) \dot{k}(t), \tag{16}$$

which is the sum of the growth rates of (a) the NRR, (b) the fertile population, and (c) the value of a newborn times $\dot{k}(t)$. (See Appendix A-1 for the complete derivation.)

For analyzing signs we start with adding $\bar{\lambda}_N$ to (13) and obtain

$$\begin{aligned} \bar{\lambda}_N + \xi^1(t,0) &= \left(\xi(t+\mu,0) + \bar{\lambda}_N\right) R(t+\mu) \\ \bar{\lambda}_N + \xi^2(t,0) &= \left(\xi(t+\mu,0) + \bar{\lambda}_N\right) R(t+\mu) + \left(\bar{\xi} + \bar{\lambda}_N\right) \\ \bar{\lambda}_N + \xi^3(t,0) &= \left(\bar{\xi} + \bar{\lambda}_N\right). \end{aligned} \tag{17}$$

Thus $\bar{\lambda}_N + \xi^3(t,0)$ and $\lambda_R(t)$ do not change the sign within $[T-\mu, T]$. For smaller t ($t \in [T-\omega, T-\mu)$) we need to ensure that a possible different sign of $\bar{\lambda}_N$ does not change the sign of $\bar{\lambda}_N + \xi^3(t,0)$ and (as a further unlikely consequence) the sign of $\lambda_R(t)$. This is because $\bar{\lambda}_N$ will be dominated by $\bar{\xi}$ if $c(t,a)$ is sufficiently close to $\bar{c}(a)$, which is guaranteed by a sufficiently high T (an assumption which is by no means necessary). The analogous consideration carries over to $t \in [0, T-\omega)$. The explicit solution

$$R(T) \quad = \quad R_0 e^{\int_0^T k(t) \mathrm{d}t} = 1 \tag{18}$$

(together with continuity of the adjoint variable) further ensures that $\lambda_R(T)$ has the same sign as $\bar{\lambda}_N + \xi^3(t,0)$.

Finally, we have to put things together starting from $R_0 < 1$:

- Sign of $k(t)$: According to (9) the sign of $k(t)$ is determined only by $\lambda_R(t)$, which does not change the sign. Due to the explicit solution (18), optimal adaptation efforts are positive over the entire time horizon - that is, $t \in [0, T]$.
- Monotony (sign of $\dot{k}(t)$): According to the Euler equation (15), the term $\bar{\lambda}_N + \xi(t,0)$ determines the sign of $\dot{k}(t)$. Since this term does not change the sign during the time horizon (for a sufficiently high T, as discussed in the previous paragraph), $k(t)$ decreases over the entire time horizon (negative $\dot{k}(t)$) - that is, $t \in [0, T]$.
- Curvature (sign of $\ddot{k}(t)$): According to (16) the curvature has the sign of $\dot{k}(t)$ (which is negative for $R_0 < 1$) times the sum of (a), (b), and (c). The sign (a)

> is positive for all t due to the positive sign of $k(t)$ for all t. For (c) we observe that $\bar{\lambda}_N + \xi(t, 0)$ is the aggregation of future terms having a positive sign. Therefore, the derivative with respect to time is negative. The sign (b) is ambiguous. The sum of these three terms therefore need not be constant over the time horizon. However, if (a) is smaller than the sum of (b) and (c), $k(t)$ is convex at t (pointwise) and concave otherwise.

These three points prove the assertion of the relation.

3. History and related results

In the highly stylized model we proposed, birth control causes convex costs. Whereas in practice changes in reproductive behavior happen in a socioeconomic and cultural framework (family planning, education of women, etc.), the proposed approach assumes that fertility can be directly influenced. For some discussion regarding the dependence of fertility on policy measures, see Sobotka et al. (2019) and Gauthier (2007). The issue of influencing fertility may be difficult to study in isolation. It depends on the labor market, welfare system, economic development, education, and gender equality but also culture and norms. Moreover, the assumption that negative and positive efforts to adapt fertility are equally costly is highly questionable. Negative $k(t)$ - that is, decreasing reproduction - requires different measures than positive adaptation efforts, such as supporting women by establishing childcare facilities. Note that these heroic simplifications allow if-then-statements of the following kind: If fertility can be adapted in a certain manner, then the target of replacement and approximate stationarity are reached at the terminal time in an optimal way.

To motivate why finding the optimal transition to stationary population is not a trivial question, consider two polar cases: (i) increase of the NRR immediately (i.e., as fast as possible) and (ii) increase of the NRR just before the end of the planning period. In case (i), an immediate increase of the NRR to replacement level, implies huge population waves in the birth trajectory and the age composition of the population, which is in contrast to stationarity at T (for T not too large). To dampen these waves in the following manner, it would be necessary to counteract by decreasing and increasing the NRR several times alternately by costly adaptation efforts. In case (ii), a too-slow increase just at the end does not create huge waves but implies that it is impossible to reach stationaritity of the population due to the time-lag of newborns until they reach the fertile age. Optimal adaptation efforts, therefore, have to balance both effects - that is, obtaining huge waves versus reaching stationarity at low acceptable costs - by a gradual change of the NRR with the properties shown in the relation on the shape of the adaptation.

Clearly, all these population dynamics could be carried out by the usual component method of demographic projection - that is, by numerical simulations. The great advantage of the analytical approach used in the present context is, that we are able to get insights on which properties of the fundamentals are responsible for the resulting optimal paths. For this fact, we mention equation (6) for the curvature of the optimal fertility adaptation rate.

4. Applications

Most Eastern–European countries complain about a declining population. Besides heavy outmigration, subreplacement fertility is the second cause for this development. Both aspects are negatively assessed by responsible politicians. A personal communication of the former director of the Vienna Institute of Demography, Wolfgang Lutz, might be seen as typical in this context. He was asked by the president of Serbia what can be done to fight against depopulation in his country. In 2021, Serbia had a NRR of 0.73 and a negative natural rate of population growth (i.e., shrinking rate) of 0.97%. The government of this Balkan republic was much concerned about the future prospects. The purpose of the model in this note is to investigate how pronatalist measures for given planning intervals should be set most efficiently.

This raises the question whether countries with shrinking populations could benefit from our analysis. It is a fact that emigration is the essential reason for depopulation. Actual population policy coping with shrinking should be clearly a mix between migration and birth control measures. Neglecting migration aspects and restricting to fertility planning, our model calculations are not appropriate to deliver insights for specific population policy advice. Nevertheless, we are able to obtain information as to what extent quantity and timing of fertility must change to approach stationarity of the population in an efficient way within a given time horizon.

Let us finally comment on an additional aspect of the transition to stationarity. Lotka's ergodicity result shows among other things that a given population subject to constant age-specific vitality rates exhibits temporary fluctuations in its size and age distribution. While the mathematics of those ups and downs has been studied in chapter 3 of the seminal book by Coale (1972), fertility-dependent fluctuations of age pyramids have found interest already earlier; see, for example, Bourgeois-Pichat and Taleb (1970) and Bourgeois-Pichat (1971). In an extension the trade-off between temporary fluctuations and the terminal population size will be considered.

Our analytic result is obviously due to the assumption of concentrated vitality rates. However, extending (4) by unimodal distributions of the fertility and mortality rates will most probably (i.e., it cannot be shown analytically, but rather be backed up with numerical simulations) not change the assertion of the relation on the shape of the adaptation. The discussion (on effect of the time–lag) at the end of Section 1 then corresponds to the mean age of the fertility distribution. For bimodal distributions as observed, for instance, in some South American countries, we are skeptical since that also implies that the dynamics of $\xi(t,a)$ (see (10)) changes the sign more than once. Already in a hypothetical scenario of a fertility distribution with two isolated ages of fertility, the assertion of the relation is violated because of a more complicated structure of the adjoint variables that prevents several conclusions within the proof.

We conclude by pointing out the key of this paper, which is the combination of a formal demographic relationship with optimal control theory. Several problems in theoretical demography can be formulated in terms of optimization, which is the path that leads to our core result.

References

Bourgeois-Pichat, J. (1971). *La démographie*. Collection Idées. Paris: Gallimard.

Bourgeois-Pichat, J. and Taleb, S.-A. (1970). Un taux d'accroissement nul pour les pays en voie de développement en l'an 2000. Rêve ou réalité? *Population* (French Edition) 25(5): 957. doi:10.2307/1530303.

Coale, A.J. (1972). *Growth and structure of human populations. A Mathematical Investigation*. Princeton, NJ: Princeton University Press.

Feichtinger, G., Kuhn, M., Prskawetz, A., and Wrzaczek, S. (2011). The reproductive value as part of the shadow price of population. *Demographic Research* 24: 709–718. doi:10.4054/DemRes.2011.24.28.

Feichtinger, G., Tragler, G., and Veliov, V.M. (2003). Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications* 288(1): 47–68. doi:10.1016/j.jmaa.2003.07.001.

Feichtinger, G. and Wrzaczek, S. (2024). The optimal momentum of population growth and decline. *Theoretical Population Biology* 155: 51-66. doi:10.1016/j.tpb.2023.12.002.

Fisher, R.A. (1930). *The genetical theory of natural selection*. Oxford: Clarendon Press. doi:10.5962/bhl.title.27468.

Gauthier, A.H. (2007). The impact of family policies on fertility in industrialized countries: A review of the literature. *Population Research and Policy Review* 26(3): 323–346. doi:10.1007/s11113-007-9033-x.

Keyfitz, B. and Keyfitz, N. (1997). The McKendrick partial differential equation and its uses in epidemiology and population study. *Mathematical and Computer Modelling* 26(6): 1–9. doi:10.1016/S0895-7177(97)00165-9.

Keyfitz, N. (1971). On the momentum of population growth. *Demography* 8(1): 71–80. doi:10.2307/2060339.

Sobotka, T., Matysiak, A., and Brzozowska, Z. (2019). Policy responses to low fertility: How effective are they? UNFPA working paper series, Working paper No. 1.

Wrzaczek, S., Kuhn, M., Prskawetz, A., and Feichtinger, G. (2010). The reproductive value in distributed optimal control models. *Theoretical Population Biology* 77(3): 164–170. doi:10.1016/j.tpb.2010.01.003.

Appendix

1. Derivation of (16)

We start by manipulating the Euler equation (15)

$$\begin{aligned} \dot{k}(t) &= -\frac{R(t)P(t,\mu)}{2C_1}\left(\bar{\lambda}_N + \xi(t,0)\right) \\ &= -\frac{1}{2C_1}\left(R(t)P(t,\mu)\left(\bar{\lambda}_N + \xi(t,0)\right)\right). \end{aligned} \tag{19}$$

By taking the derivative we get

$$\begin{aligned} \ddot{k}(t) &= -\frac{1}{2C_1}\Big(\dot{R}(t)P(t,\mu)\left(\bar{\lambda}_N + \xi(t,0)\right) \\ &\quad + R(t)P_t(t,\mu)\left(\bar{\lambda}_N + \xi(t,0)\right) + R(t)P(t,\mu)\xi_t(t,0)\Big) \\ &= -\frac{1}{2C_1}\Big(\frac{\dot{R}(t)}{R(t)}R(t)P(t,\mu)\left(\bar{\lambda}_N + \xi(t,0)\right) \\ &\quad + \frac{P_t(t,\mu)}{P(t,\mu)}R(t)P(t,\mu)\left(\bar{\lambda}_N + \xi(t,0)\right) \\ &\quad + \frac{\xi_t(t,0)}{\bar{\lambda}_N + \xi(t,0)}R(t)P(t,\mu)\left(\bar{\lambda}_N + \xi(t,0)\right)\Big) \\ &= -\frac{R(t)P(t,\mu)}{2C_1}\left(\bar{\lambda}_N + \xi(t,0)\right)\left(\frac{\dot{R}(t)}{R(t)} + \frac{P_t(t,\mu)}{P(t,\mu)} + \frac{\xi_t(t,0)}{\bar{\lambda}_N + \xi(t,0)}\right) \\ &= \left(\frac{\dot{R}(t)}{R(t)} + \frac{P_t(t,\mu)}{P(t,\mu)} + \frac{\xi_t(t,0)}{\bar{\lambda}_N + \xi(t,0)}\right)\dot{k}(t). \end{aligned} \tag{20}$$