ÖAW AUSTRIAN ACADEMY OF SCIENCES

VIENNA INSTITUTE OF DEMOGRAPHY

WORKING PAPERS

01/2024

FERTILITY DECLINE AND AGE-STRUCTURE IN CHINA AND INDIA

THOMAS FENT, STEFAN WRZACZEK, GUSTAV FEICHTINGER AND ANDREAS NOVAK

Vienna Institute of Demography
Austrian Academy of Sciences
Dominikanerbastei 16 | 1010 Vienna, Austria
vid@oeaw.ac.at | www.oeaw.ac.at/vid

VIENNA INSTITUTE OF DEMOGRAPHY

Fertility decline and age structure in China and India: actual vs. optimal development

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

May 6, 2024

ABSTRACT

China and India, two Asian countries that experienced a rapid decline in fertility since the middle of the twentieth century, are the focus of this paper. Although there is no doubt that lower fertility levels have many positive effects on the economy, development and sustainability, little is known about the optimal transition from high to medium or even low levels of fertility. Firstly, implementing policies that have the potential to reduce fertility is costly. Secondly, additional costs arise from adapting the infrastructure to a population that fluctuates quickly not only in terms of size but also with respect to the age structure.

We apply an intertemporal optimisation model that takes the costs and benefits of fertility decline into account. The optimal time path depends on the cost structure, the planning horizon and the initial conditions. In the case of a long planning horizon and high initial fertility, it may even be optimal to reduce fertility temporarily below replacement level in order to slow down population growth at an early stage.

A key finding of our formal investigation is that, under the same plausible parameter settings, the optimal paths for China and India differ substantially. Moreover, our analysis shows that India, where the fertility decline emerged as a consequence of societal and economic developments, followed a path closer to the optimal fertility transition than China, where the fertility decline was state-imposed. The mathematical approach deployed for this analysis provides insights into the optimal long-term development of fertility and allows for policy conclusions to be drawn for other countries that are still in the fertility transition process.

1 INTRODUCTION

India surpassed China as the world's most populous country in 2023, reaching an estimated population size of around 1.429 billion compared to China's 1.426 billion (United Nations and Social Affairs, 2022, mid-year population). This was not only an overtaking in terms of numbers, but the trends also point into opposite

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

[†]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.

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

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

https://doi.org/10.1553/0x003f0d14

directions. The total population in China reached its maximum back in 2021, while India's population continues to grow. In addition, the population in India grew much faster than in most other Asian countries. While the population in China has increased 2.6-fold since the middle of the 20th century, it has quadrupled in India over the same period. As a result, both China and India currently have a population around three times the size of the European Union or four times the size of the USA. The left panel in figure 1 shows the development of the total population in these two countries. The chart illustrates the continuing population growth in India in contrast to China, which already passed it's maximum population size.

Nevertheless, there has been a remarkable decline in fertility in both countries. In China the net reproduction rate (NRR) fell from a peak of 2.84 in 1963 to 0.54 in 2021 and in India it fell from 1.95 in 1964 to 0.93 (see right panel in figure 1). The graphs clearly indicate that the NRR followed a very erratic path in China which is due to fluctuations in enforcement of family policies (Retherford et al., 2005; Piotrowski and Tong, 2016). In contrast, the NRR developed much more evenly in India. It started at a lower level, declined gradually but steadily and is now below replacement level, but well above the very low level reached in China. In both countries, the decline in fertility happened at a much higher speed than in Europe or North America. As a result, not only the population size but also the age structure of the populations changed dramatically during this period. We investigate in this paper how the decline in fertility can be optimised from an economic perspective.

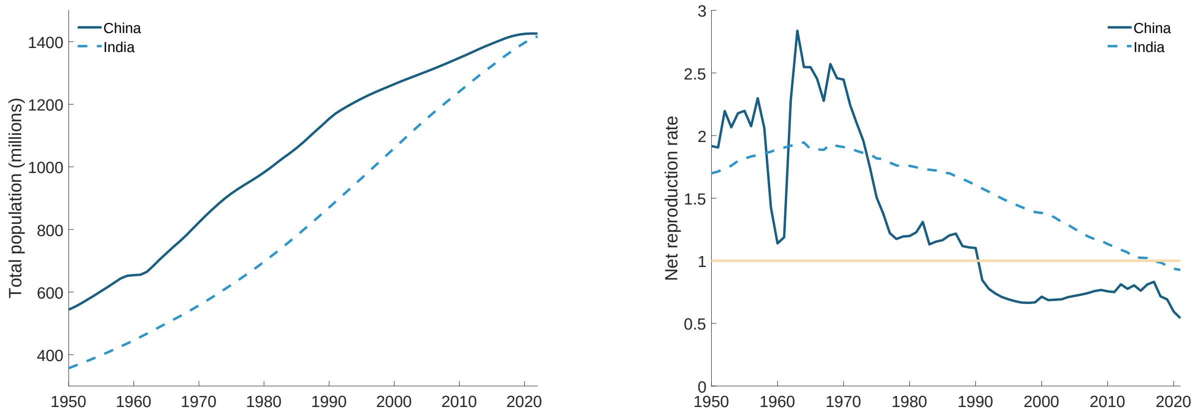

Figure 1: Total population and net reproduction rates in China and India from 1950 to 2021.

Both, the size and the age structure have far-reaching economic, ecologic and geopolitical implications. The size of a population is associated with economic, political and military power. A large population goes hand in hand with a high level of influence in international bodies and negotiating power in international organisations. The change in the ranking in terms of population size could influence regional dynamics. It might impact relationships among China and India—and also other Asian countries not taken explicitly into consideration in this paper—which have a complex relationship with a history of both cooperation and competition. This determines the global influence, both on the economic and political level.

With regard to the economy, a larger population opens up the possibility of producing and consuming more goods, earning, spending and investing more money and hopefully also generating more innovations. In terms of production economics, a larger population is associated with a larger labour force. This can be advantageous in terms of economic growth, productivity, and innovation. Consequently, impressive population growth and, above all, the sheer size of their populations have contributed significantly to the growing economic influence of China and India on the world markets. This development, among others, constitutes the demographic explanation for the increasing economic power of the BRICS[1] countries. The increasing economic strength of these countries is challenging the customary leading role of the G7[2]. However, population size also poses challenges in terms of employment, infrastructure and resource management.

In contrast, the size of a country's population also has an impact on its ecological footprint. A large population puts additional strain on resources like land, water, and energy. Managing these resources sustainably becomes more critical to ensuring the well-being of the population and avoiding potential conflicts. Every inhabitant consumes goods and services and therefore adds to global challenges such as healthcare, food security, climate change and environmental sustainability. Consequently, the actions and policies of populous countries have a significant global impact. A large population brings along social and political challenges, including the need for effective governance, social services, and the development and maintenance of infrastructure. Managing diverse populations and ensuring social cohesion become crucial tasks for political decision-makers. Changes in population size can also have an impact on international migration which, in turn, also affects the population structure in source and destination countries. This could alleviate or aggravate tensions on other countries' labour markets.

Finally, the size of a country's population has an influence on geopolitical issues. Geopolitical power means the ability to assert one's own interests in the face of conflict or opposition. While population size alone does not determine military strength or negotiation power, a larger population could potentially influence military capabilities and strategies. It might impact defence policies and alliances in the region. At present, it is becoming apparent that the previously accepted leadership role of the Western world is no longer guaranteed in the event of international conflicts.

Fertility decline is a complex demographic phenomenon with far reaching societal implications. In this paper we focus on the trends and factors contributing to fertility decline in China and India. These countries with their unique cultural, economic, and political contexts, have experienced significant shifts in population dynamics over the past few decades. Although the decline in fertility took place at nearly the same time in China and India (see right panel in figure 1) there are remarkable differences with respect to the specific causes of the transition from high to low fertility.

China, until recently the most populous country in the world, applied a series of political measures to influence the development of the population. The sudden drop and subsequent recovery in the NRR around 1960 (see right panel in figure 1) was due to China's great famine. Thereafter, the great leap forward, the cultural revolution, the one-child policy, and the economic reform era had a major impact on the development of the population (Piotrowski and Tong, 2016). The later-longer-fewer policy implemented in 1971 imposed a minimum age at marriage for women of 23 years in rural areas and 25 years in urban areas, longer intervals between births and fewer children (Feeney, 1994; Retherford et al., 2005). The widely publicised one-child policy

[1]The abbreviation covers the member states as of 2010; Brazil, Russia, India, China and South Africa, but more countries joined in 2024.

[2]Canada, France, Germany, Italy, Japan, the United Kingdom and the United States

was introduced in 1979 and implemented more seriously in 1980 in response to concerns about overpopulation and too rapid population growth. Due to this policy, urban families were limited to one child each. Moreover, the timing of marriage and childbearing and spacing of children were limited (Piotrowski and Tong, 2016), which resulted in a decline in population growth. This policy was transformed into a two child policy in 2016 and it was further relaxed allowing for up to three children in 2021. As expected, both the later-longer-fewer policy and the one-child policy caused primarily reductions in second- and higher-order births, but did not affect first births as strongly (Feeney and Feng, 1993; Feeney, 1994; Morgan et al., 2009). Similar to these policies, educational expansion also affected urban areas differently than rural areas. Piotrowski and Tong (2016) found that for cohorts born from 1956 to 1960, who began their reproductive years during the cultural revolution but had their late reproductive years during the reform era, both education and urban residence had a negative effect on first birth hazard. The younger cohorts who entered into their reproductive years after the implementation of the one-child policy primarily had a lower average hazard of higher order births. In general, differences between educational groups are significant for all birth cohorts in urban areas, but in rural areas only for the oldest cohorts born from 1945 to 1955 (Piotrowski and Tong, 2016). To summarise, the decline in fertility in China was the result of a series of policies imposed by state authorities rather than a development emerging through societal changes (Feeney, 1994). It resulted not only in the intended slowdown of population growth, but obviously also in rapid population ageing. Among the females, the share of the population aged 65 and older rose from 3.2 per cent in 1950 to 18.9 per cent in 2021. Another not so obvious consequence was a clear shift in the gender ratio. Families who were allowed to have only one child preferred to have a son, which was frequently achieved by sex-selective abortion (Kashyap and Villavicencio, 2016). Even after relaxing these originally very restrictive policies, fertility rates continued to decline due to high costs of living, changing societal norms, and delayed marriage, all influencing individuals' decisions to have fewer children (Piotrowski and Tong, 2016; Wang et al., 2022).

India is not only the world's most populous country, but also among those countries with the most heterogeneous population. The high diversity is the consequence of its multilingual and multi-ethnic society. Due to the heterogeneous society, there are significant regional variations in the fertility development (Dreze and Murthi, 2001; Krishnaji, 2009). In contrast to China, the decline in fertility was not primarily due to state regulatory interventions, but mainly due to social and economic factors. Among them are increasing urbanisation, educational expansion (of women), improved access to family planning services, improved child survival and economic development (Dreze and Murthi, 2001). These, in turn, led to delayed marriage and childbearing and, consequently, to lower completed fertility. Better access to contraceptives and family planning methods, on the other hand, enabled couples to make informed and conscious decisions about their family size (Dreze and Murthi, 2001). Continued economic growth reduced the gender gap in educational attainment, but female labour force participation remained low despite declining fertility (Debnath and Das, 2022). This is because women in India face a severe motherhood penalty. Increases in wealth and income of other household members have the consequence that women become less likely to enter the labour market and more likely to leave it (Sarkar et al., 2019). The lack of compatibility between family and career resulted in a decline in female labour force participation from 2000 to 2020 (International Labour Organization, 2024) despite increases in education (UNESCO Institute for Statistics, 2023). However, since 2020 female labour force participation follows a sharply uprising trend. Despite these positive developments, challenges like socio-economic disparities, cultural norms, and the need for comprehensive healthcare services remain. Similar to China, fertility decline in India also led to an intensified gender bias. For example, Nath (2023) shows that families with two daughters were more likely to have a third child than families with two sons. Gupta and Bhat (1997) found evidence of an increase in sex-selective abortion in India between 1981 and 1991. Aksan (2021) detects a trend towards gender selection at ever lower parities. Generally, the fertility dynamics elapsed rather moderately compared to China. The peak was lower and the decline was more gradual, with the result that fertility in India is now just below replacement level.

As we can see in figure 4 (solid and dashed lines), the initially extraordinarily high fertility levels followed by a rapid decline, temporarily led to an economically advantageous age structure with a high proportion of the working-age population in both countries. If this favourable time interval is used to invest in education, health and infrastructure, this will boost economic growth and development. The resulting increase in economic performance and prosperity is usually addressed as the first demographic dividend (Bloom and Williamson, 1998; Dörflinger and Loichinger, 2024). Since such an advantageous age structure cannot be sustained forever, we postulate that the population has to reach a stationary state in the long term.

To achieve this, the two countries China and India, currently exhibiting fertility below replacement level, need to increase their fertility. We assume that it is generally possible to influence fertility, but at a cost. Policies that reduce time and financial expenditure, such as childcare support and housing policies, have a positive effect on fertility (Baizan et al., 2016; Lui and Cheung, 2021). Fent et al. (2013) showed that the effectiveness of family policy measures depends on the specific social structure. In China, fertility decline was achieved through a very restrictive policy regime. India also tried to influence fertility through a government-imposed population control in the 1950s, but then switched to a more humane approach, assuming that “development is the best contraceptive”.

To investigate the optimal path toward a stationary population, we apply an intertemporal optimisation model that takes into account the costs of adaptation efforts, the costs associated with a deviation from the desired age structure and costs of changes in the total population size. Due to the fact that China starts from a higher initial fertility level, the optimal solutions for China and India are very different, even if we apply the same cost parameters for both countries. We compare the actual temporal development of the NRR, the population size and the age structure in China and India with the optimal solutions under different optimisation approaches. Under plausible parameter settings, our analysis reveals that fertility decline in India was closer to the economically optimal solution than in China over a relatively long period of time. A more detailed description of the model along with a thorough mathematical analysis can be found in (Wrzaczek et al., 2024).

2 THE MODEL

The dynamics of the one-sex (female) population is modelled as a continuous-time dynamical system with time t and age a as independent variables. Mortality and fertility are the crucial rates defining the dynamics of one cohort and the reproduction process at every time t. This is modelled by the well-established McKendrick-von Foerster (partial differential) equation (see Keyfitz and Keyfitz, 1997; Keyfitz and Caswell, 2005, section 7.5.4), which is given by

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

$$P(0,a) = P_0(a) \tag{1b}$$

$$P(t,0) = B(t) = \int_{\underline{a}}^{\overline{a}} f(t,a)R(t)P(t,a), \tag{1c}$$

where $P(t,a)$ denotes the size of the a-year old population at time t. The equation describes the evolution of a cohort born at $t-a$ along the life-cycle (i.e., along both dimensions t and a), which is reduced by the mortality rate $\nu(t,a)$. $P(0,a)$ denotes the exogenously given initial population distribution. The number of newborns, denoted by $B(t)$ and referred to as *births* in the remainder of the paper, equals the newborns of all females at time t aggregated across all fertile age groups. The function $f(t,a)$ determines the fertility distribution over a at t, defined over the fertile ages $a \in [\underline{a}, \overline{a}]$. This fertility distribution is multiplied by the net reproduction rate, $R(t)$, which represents a uniform scaling factor for all ages.

Analogously, the number of deaths at t is defined by

$$D(t) = \int_0^{\omega} \nu(t,a)P(t,a)\,\mathrm{d}a. \tag{2}$$

As a result, the dynamics of the total population $N(t)$, i.e., the sum of persons across all age groups a at time t, can be written as

$$\dot{N}(t) = B(t) - D(t), \quad N(0) = N_0 = \int_0^{\omega} P_0(a)\,\mathrm{d}a, \tag{3}$$

where the initial value N_0 is obtained by summing over all ages at $t = 0$. Here ω denotes the maximum age.

A core assumption of our population dynamics model is the possibility to influence $R(t)$ positively and negatively by investing in fertility adaptation efforts (later on referred to as *adaptation efforts*) $k(t)$. Negative (positive) values of $k(t)$ result in a decrease (increase) of $R(t)$. The effects of the adaptation efforts on the fertility dynamics are assumed to be proportional to the current value of $R(t)$ (following Feichtinger and Wrzaczek, 2024a,b), which yields the following dynamics for $R(t)$ over time

$$\dot{R}(t) = k(t)R(t), \quad R(0) = R_0. \tag{4}$$

The initial net reproduction rate, R_0, can be less or greater than one, which means a decreasing or increasing population in the long term. The latter applies to the recent situation in China and India, which are the two countries that are the focus of this article.

Although the fertility distribution $f(t,a)$ can be observed in the population data of the past, it is difficult, if not impossible, to find reasonable arguments for how it adapts as $R(t)$ changes over time. Analysing the data, however, reveals that the mean age of childbearing did not vary substantially even for drastically changing $R(t)$ in a relatively short period of time. This is due to the fact that in high fertility contexts, women have their first child at a relatively young age, but have more children later on, while in low fertility contexts, women have their first child at an advanced age, but have no or fewer children later in life. Consequently, the age at first birth is subject to major changes, but the mean age at birth varies only slightly. We therefore assume that fertility is concentrated at a unique age μ. Formally this is denoted by

Assumption 1 *Fertility is concentrated at a unique age μ, i.e., $f(t,a) = \delta(a-\mu)$ for $t \geq 0$.*

The function $\delta(x)$ denotes the standard Dirac delta function. The mean age of fertility can be derived from the historical data and differs slightly across China and India. Applying this assumption to (1c), the number of newborns becomes

$$B(t) \quad = \quad P(t,\mu)R(t). \tag{5}$$

We assume that it is desirable for a high-fertility society to achieve a lower level of fertility while pursuing three objectives within a planning horizon of T years. Firstly, and most importantly, $R(t)$ must be equal to replacement level at $t = T$, i.e., $R(T) = 1$. This is an additional (end time) constraint to equation (4) describing the dynamics of $R(t)$. Secondly, the age structure of the population should be as close as possible to a predefined structure $\bar{c}(a)$. Corresponding to the constraint $R(T) = 1$ it makes sense to assume a stationary population structure for $\bar{c}(a)$, but in general the model is open for other age structures. The deviation of the current structure from $\bar{c}(a)$ is costly, which is measured by the quadratic deviation weighted by a constant cost term C_2 aggregated over time. Thus, the total deviation causes the following costs:

$$C_2 \int_0^T \int_0^\omega \left(c(t,a) - \bar{c}(a)\right)^2 \mathrm{d}a\, \mathrm{d}t, \tag{6}$$

where $c(t,a) := \frac{P(t,a)}{N(t)}$ denotes the share of the a-year old females at time t.

Thirdly, changes in the total population size entail costs for adapting the infrastructure (such as e.g. housing, transport, supply of water and electricity, child care facilities, schools and many others). This is captured by the quadratic deviation of the total population at T from the initial population size weighted by a cost term C_3. Hence, the policy maker tries to minimize $C_3\left(N(T) - N_0\right)^2$.

Summing up, three objectives have to be considered: (i) reaching $R(T) = 1$, (ii) minimising the deviation from $\bar{c}(a)$ at every t, and (iii) minimising the deviation of the size of the total population from N_0 at T. Whereas, (i) is a *hard* constraint, which has to be fulfilled no matter about the costs, (ii) and (iii) are *soft* in the sense that only deviation costs should be minimised. Note that (iii) could also be considered for every t, this would be conflicting to (ii), which already captures costly infrastructure adaptations by the deviation from the desired age structure.

The dynamics of the net reproduction rate $R(t)$ can be influenced by adaptation efforts. According to (4), a negative $k(t)$ decreases $R(t)$ in the next (marginally small) time unit, and a positive $k(t)$ implies an increase. Assuming for simplicity that an increase and a decrease are equally difficult to implement and therefore equally costly, we assume quadratic costs with cost parameter C_1 as standard in economics.

As a result, the objective is to minimise the total costs

$$\int_0^T \left[C_1 k^2(t) + C_2 \int_0^\omega \left(c(t,a) - \bar{c}(a)\right)^2 \mathrm{d}a\right] \mathrm{d}t + C_3(N(T) - N_0)^2 \tag{7}$$

by optimal choice of adaptation efforts $k(t)$ subject to the population dynamics and the evolution of $R(t)$, i.e., equations (1)-(5). This is a finite-time age-structured optimal control model, which can be solved with the age-structured Maximum Principle. The optimality conditions are briefly presented in Appendix A. For a more elaborated theoretical discussion of this model we refer to (Wrzaczek et al., 2024). Note that this paper uses the analogous model with a second assumption of concentrated mortality. This assumption follows Feichtinger and Wrzaczek (2024a,b) and allows the derivation of non-existence and existence results of changing signs of adaptation efforts (i.e., increase and decrease of $R(t)$) within the planning period for different assumptions on the cost parameters. I.e., for certain parameters and slightly more restrictive assumptions a theoretical proof of the optimality of under- or overshooting of $R(t)$ (depending on the initial value of R_0) before reaching $R(T) = 1$ can be provided.

3 RESULTS

In the following we present numerical results obtained from the model introduced in section 2. We calibrate the age-structured optimal control model determined by the objective function (7) and the equations (1)-(5) with data from China and India. For the initial state, we take the female age structure, the female life table and the NRR from the year in which the NRR peaked. In China the peak NRR was at 2.84 and occurred in 1963, and India reached a peak at 1.95 in 1964. Therefore, in the following, the time $t = 0$ refers to the year 1963 or 1964, respectively.

The optimal solutions obviously depend not only on the initial conditions of the countries but also on the choice of the cost parameters C_1, C_2 and C_3. A comprehensive discussion of the emerging dynamical systems and their possible outcomes is given in Wrzaczek et al. (2024). An increase in C_1, the costs of the adaptation efforts, leads to a more even approach towards the constraint $R(T) = 1$. An increase in C_2, the costs of deviations from the desired age structure $\bar{c}(a)$, has the effect that the decline in fertility is achieved with only small deviations from $\bar{c}(a)$. In combination with a sufficiently small C_1, this can lead to erratic dynamics with repeated changes in the sign of $k(t)$. Finally, an increase in C_3, the costs of a deviation from the initial population size, has the effect of slowing down population growth more quickly. This can even lead to undershooting behaviour. This means, that the NRR falls below replacement fertility in order to curb population growth as quickly as possible. Due to the end condition $R(T) = 1$, $R(t)$ must increase again in the following. If we compare China and India applying their unique initial conditions but the same numerical parameters, we always get qualitatively different dynamics for these two countries. In the following we discuss a set of solutions that covers a wide range of possible outcomes but is representative for a reasonable and realistic choice of parameters.

3.1 COMPARISON OF THE OPTIMAL SOLUTIONS

In this section we present results obtained for planning horizons of $T = 50$ and $T = 200$ years, respectively. Thus, starting from the time of the highest NRR, the population has to arrive at $R(T) = 1$ while minimising the cost function given in (7). Figure 2 shows the optimal adaptation efforts for a planning horizon of $T = 50$. Both countries start at a high NRR and have to decline to replacement level. Consequently, the adaptation efforts are negative, i.e. $k(t) < 0$. China starts from a higher level than India (see right panel of figure 1) and has to invest more efforts to reduce fertility as we can clearly see from the figure.

While these graphs give the impression that the optimal adaptation efforts remain constant, a closer look reveals significant changes over time. The two graphs in the first row of figure 3 illustrate the optimal time paths of the adaptation efforts for the two countries under consideration separately. From the graphs we can

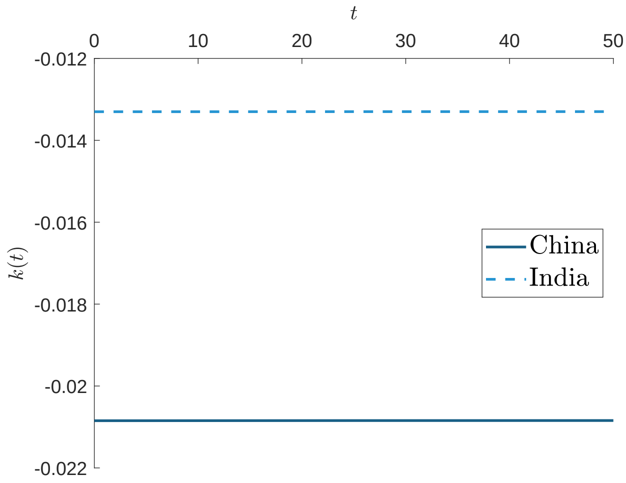

Figure 2: Comparison of optimal adaptation efforts for $T = 50$.

see that in the optimal solution the efforts decrease. Consequently the net reproduction rate $R(t)$ gradually declines to arrive at the target value $R(T) = 1$ at the end of the planning period. As expected, the decline must be faster in China. We can see this in the graph in the lower left panel of figure 3. In both countries $R(t)$ gradually converges towards the end time constraint. Not only does China start from a higher level of fertility, but it's optimum $R(t)$ remains higher than India's throughout the entire planning period. Finally, the graph in the lower right panel shows the relative deviation of the actual NRR from the optimal $R(t)$. Firstly, the chart shows that the NRR was above the optimum in China most of the time, while in India it was mostly below the optimum. This suggests that the decline in fertility happened too fast in China and too slow in India. The graph also indicates that the fertility development in India was closer to the optimum than in China.

Figure 4 shows the broad age structure based on the three age groups $0-19$, $20-65$ and $65+$ over time. The solid lines depict the actual age structure and the shades of purple represent the optimal solution. Again we can see that fertility decline in China was faster than optimal. As a result, the share of the age-group $0-19$ declined too quickly and the share of the age-group $20-65$ rose too much. In contrast, the fertility decline in India was slower than optimal, resulting in a too slow decline of the age-group $0-19$. In the case of a quickly growing population, a high proportion in the youngest age group leads to prolonged rapid population growth. This growth, in turn, results in increased deviations from the target age structure $\bar{c}(a)$ and in increased fluctuations in the total size of the population represented by the cost term $N(T) - N_0$.

Finally, figure 5 shows the optimal age structure over time if age is presented as a continuous variable, i.e. the plot depicts the population density. Despite the smooth decline in fertility, the age structure is subject to marked fluctuations. This is due to irregularities in the initial population, but also to the time lag between a birth of a woman and the time when this woman herself gives birth to a daughter. As a result of this time delay, any change in the number of births at a certain point in time leads to an echo one generation length later. This behaviour is analogous to the findings of Bourgeois-Pichat and Taleb (1970) and Leridon (1983).

In the next step we investigate how the optimal solution changes if we increase the planning horizon from $T = 50$ to $T = 200$. A longer planning horizon means that the cost parameter C_2, which determines the

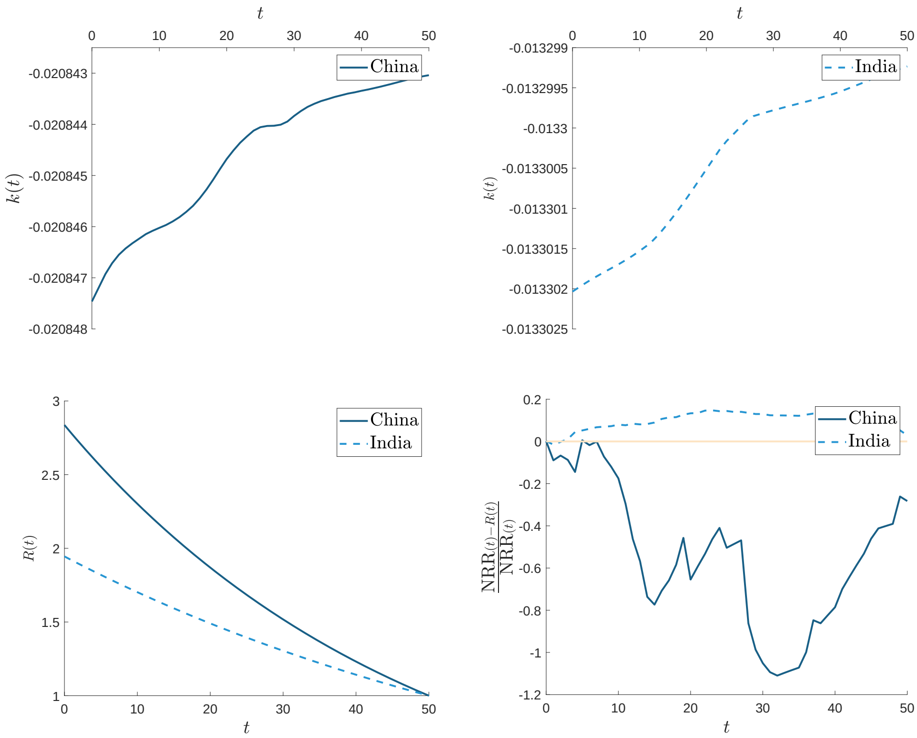

Figure 3: Optimal adaptation efforts, optimal net reproduction rate and relative deviation

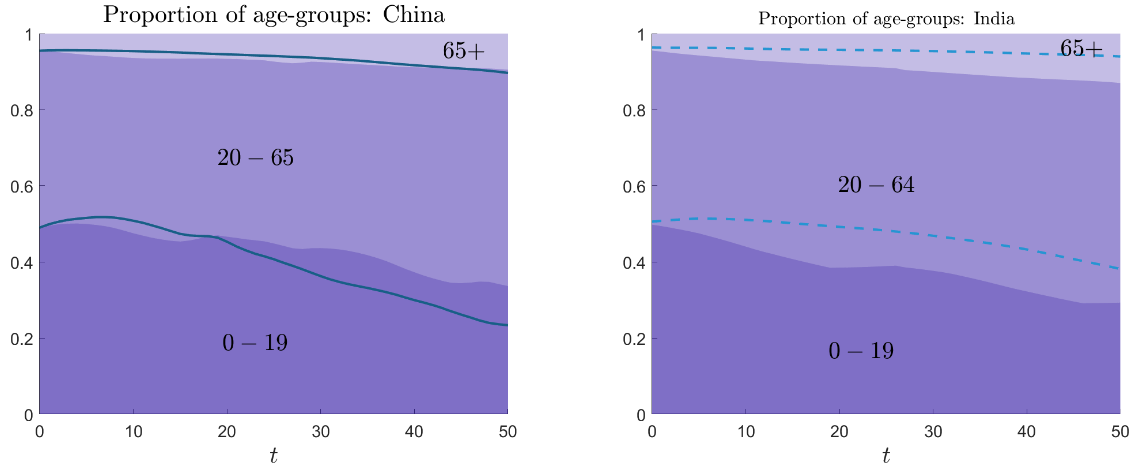

Figure 4: Comparison of the optimal and actual age structure for $T = 50$.

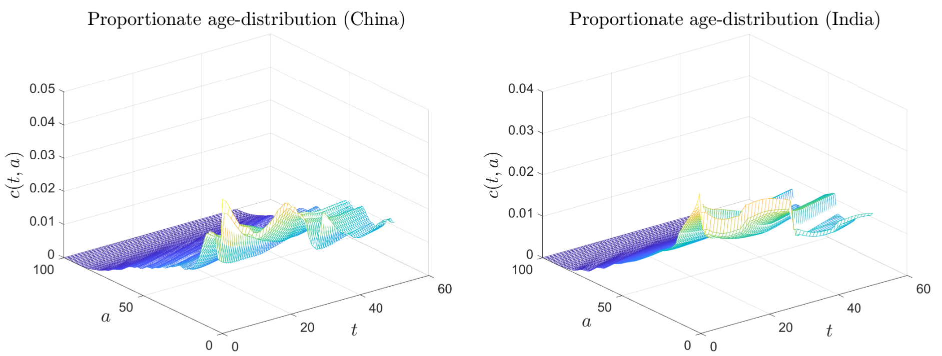

Figure 5: Proportionate age-distributions of the optimal solution for $T = 50$.

costs of deviations from the desired age structure, becomes more important because the deviation from the end constraint is only relevant at the end of the planning horizon. We can see the impact of this effect in figure 6. The first graph shows the optimal adaption efforts over 200 years. Here, too, China has to make greater efforts than India at the beginning due to the higher initial fertility. Nevertheless, both countries initially make much greater efforts than in the previous case with a shorter planning horizon. This is because in the case of a longer planning horizon it is advantageous to slow down population growth early in order to shorten the period during which the age structure deviates substantially from the desired age structure. In contrast to the previous case, the adaption efforts in China change their sign shortly after the halfway point of the planning horizon. The adaptation efforts in India remain negative over the whole planning period—like in the previous case— but also for India the variation is more pronounced than in the case of $T = 50$.

The second graph shows the evolution of the optimal net reproduction rate $R(t)$ over time. China departs from a higher initial fertility, thus, the difference between the actual and the desired age structure, $c(t, a) - \bar{c}(a)$, is greater. In turn, greater efforts are needed to reduce the costs associated with this deviation. This results in undershooting, as we can see in the graph. Thus, fertility falls well below replacement level in order to rapidly slow down population growth. Since the terminal condition $R(T) = 1$ has to be fulfilled, $R(t)$ increases later on to arrive at the desired final value. Unlike the previous case, the two countries change their positions in terms of $R(t)$. China starts at a higher level, but quickly falls below the level of India and also below replacement level. The optimal path in India is more even. Since India starts at a lower level of fertility, adaptation efforts are less pronounced and the net reproduction rate converges monotonically towards its target value. The third graph in this figure again depicts the relative deviation of the actual net reproduction rate from the optimal solution. Since we do not have time series for 200 years, the graph covers only a time span of 50 years. The picture reveals that at the beginning both countries develop close to the optimum but later on in both countries fertility falls below the optimal path. The development in China is more erratic at this stage, and also the deviation from the optimum is larger in China. Towards the end of the planning horizon, however, India moves away from the optimal path, while China moves closer to it.

Figure 7 again compares the broad age structure of the real population (solid and dashed lines) with the optimal solution (purple shades). Notably, the overall variation of the age structure is more pronounced in China than in India. This applies to both the optimum solution and the actual development. Furthermore, in both countries the optimal path exhibits several waves in the shares of the age groups. The causes are again the irregular initial age structure and the time lag between the births of mothers and their daughters. In the case of China, we see that the share of the youngest age group first increased above the optimal time path for a short period but later on it decreased more quickly than in the optimal case. The share of the oldest age group initially increased more slowly than optimal, but later on the actual and optimal share coincided. In the case of India, the proportion of the youngest age group remains above the optimal level for most of the time, but intersects with the optimal path towards the end, while the share of the oldest age group remains below the optimal level throughout the observation period. As the deviations of the youngest and oldest age groups approximately offset each other, the proportion of the middle age group is closer to the optimum for most of the time.

As the last point of this section, figure 8 again shows the optimum continuous age structure, the population

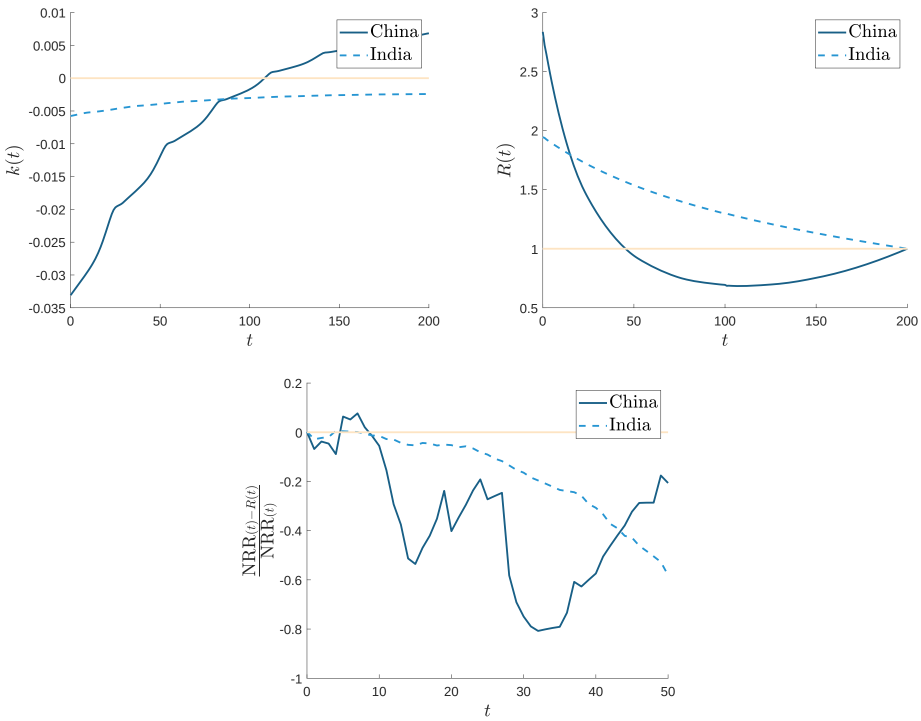

Figure 6: Optimal adaptation efforts, optimal net reproduction rate and relative deviation for $T = 200$.

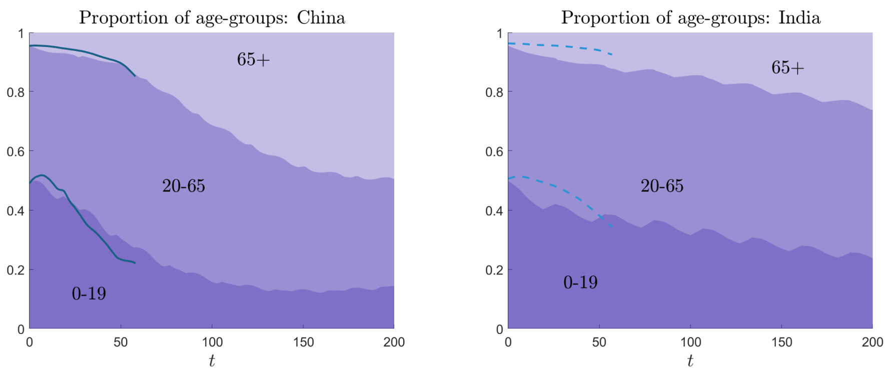

Figure 7: Comparison of the optimal and actual age structure for $T = 200$.

density, over time. Here, too, pronounced waves can be observed both over age and over time.

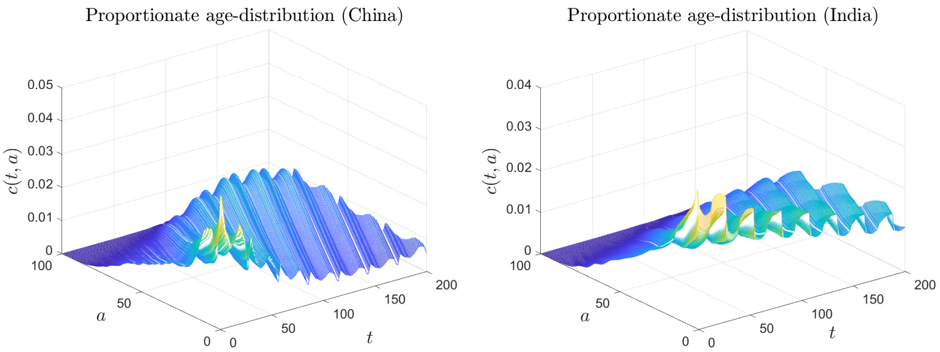

Figure 8: Proportionate age-distributions of the optimal solution for $T = 200$.

3.2 COMPARISON OF THE OPTIMAL SOLUTION WITH ROLLING HORIZON SOLUTION

In this section, we again assume a planning horizon of T years. However, instead of pursuing the optimum solution over the entire period of T years, the optimisation is reactivated after a fixed time interval and a new planning horizon of T years begins. This means that the point in time at which the target value $R(T) = 1$ is reached is shifted into the future each time a new re-optimisation is called up. The idea behind this procedure is the assumption that a new government, once in office, might refocus on long-term goals to be achieved.

Figure 9 illustrates the results obtained with initial values of China, a planning period of $T = 50$ years and a re-optimisation every 5 or 25 years, respectively. The graphs show the adaptation efforts (upper left panel), the net reproduction rate (upper right panel), the total population (lower left panel) and the proportion of three broad age groups (lower right panel). The graphs contain the optimal solution without re-optimisation (black lines), the solutions with re-optimisation every 5 years (red lines), and the solutions with re-optimisation every 25 years (blue lines) and the actual data for China (green lines).

A look at the adaptation efforts reveals that these are reduced with every re-optimisation. This is due to the fact that the time to reach the target value gets extended, therefore, smaller efforts are sufficient to reach the target. The curves representing the NRR or $R(t)$ confirm this. The curves flatten out in order reach the envisaged target later. Consequently, re-optimisation results in a higher total population and a higher share of the youngest age group. Comparing the actual data (green line) with the three different optimisation scenarios reveals that the development in China followed a path closer to the optimal solution without re-optimisation. With due care, we can conclude that the Chinese family planning authorities have once staked out the terrain to reduce fertility and then adhered to the path taken.

Figure 10 shows the same results for a planning horizon of $T = 200$ years. Here too, re-optimisation results in a mitigation of the adaptation efforts and a prolongation of the period required to reach the target value. The differences between the three optimisation scenarios become more pronounced and even a qualitative difference in the time development occurs. We know from the previous section that for a planning horizon of 200 years, the optimal path for China involves undershooting, as shown by the black lines in figure 10. The adaptation

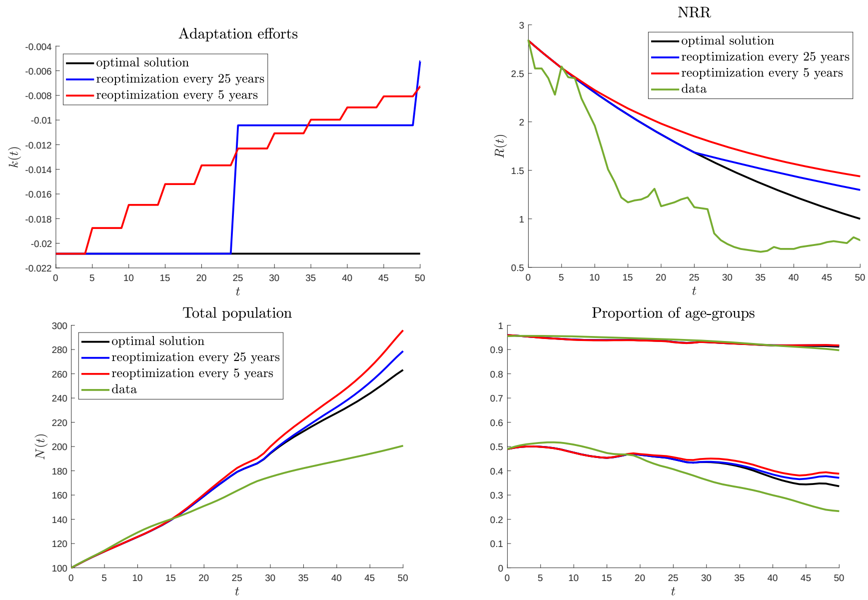

Figure 9: Optimal solution and rolling time horizon of China with $T = 50$.

efforts change their sign, the net reproduction rate falls below one and recovers during the second half of the planning period and the population shows a period of growth followed by a period of shrinkage. In the two scenarios with re-optimisation, however, this undershooting behaviour vanishes. The adaptation efforts remain negative, the net reproduction rate converges monotonically towards one and the total population increases monotonically. Analogous to the results obtained with a shorter planning horizon, the actual data are closer to the optimal solution without re-optimisation, which allows the same conclusions to be drawn.

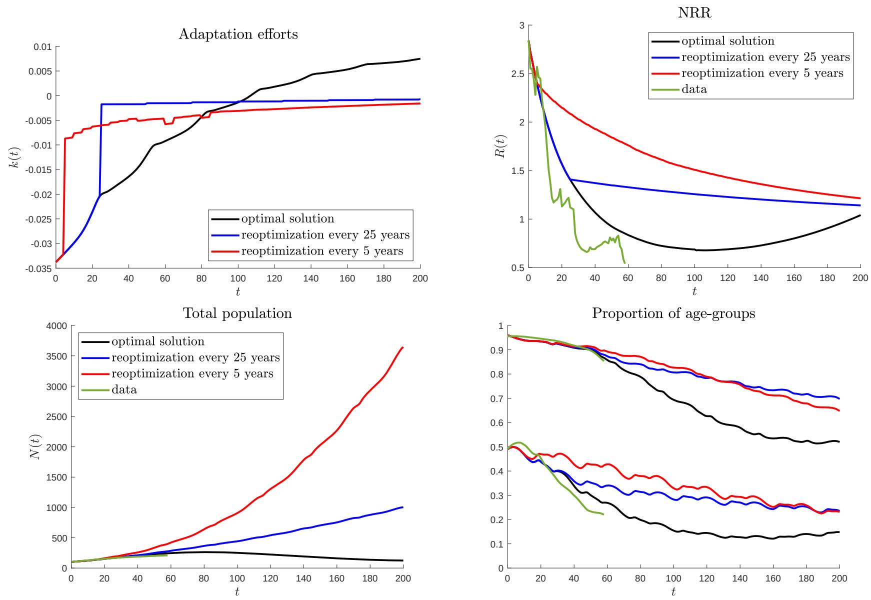

Figure 10: Optimal solution and rolling time horizon of China with $T = 200$.

4 SUMMARY AND CONCLUSIONS

The aim of this paper is to examine the fertility decline in China and India. Both countries had very high fertility, which peaked in the mid-1960s. In China, fertility decline was imposed by the state through a series of strictly enforced policies. The result of these policies, and in particular of the policy changes and adjustments, was a rapid but erratic decline in fertility. This did not only lead to the intended slowdown in population growth, but also appears to have led to rapid population ageing. In India, attempts were made even earlier than in China to curb population growth through state-imposed family planning measures. However, the authorities only pursued this approach on a temporary basis. Later on, the government adopted a more balanced approach, which is conclusively expressed in the statement "Development is the best contraceptive", made by the Minister for Health and Family Planning, Karan Singh, at the World Population Conference in Bucharest in 1974. As a result, India's fertility decline has followed a more gradual path. Apart from the fact that these two countries chose a different approach to reducing fertility, the economic consequences of a decline in fertility are often ignored in the literature. While it is clear that high fertility well above replacement level has several devastating effects in the long term, it is not clear which path to lower fertility levels should be taken.

To fill this gap, we analyse the fertility decline from an economic perspective. For this purpose, we deploy a dynamic model based on the McKendrick-von Foerster (partial differential) equation. We assume that adaptation efforts can influence fertility in both directions. This adaptation can be state-imposed but it can also emerge from societal changes, but in any case it comes at a cost. In addition to the costs of adaptation, we introduce two further cost components. The first one expresses the opportunity costs of deviating from the desired age structure. The second one expresses the costs arising from changes in the size of the total population. The dynamics of the net reproduction rate are influenced by adaptation efforts and constrained by the requirement that the population must stabilise at the end of the planning period, i.e. the net reproduction rate must be equal to one.

The optimal path to replacement level fertility obviously depends on the choice of numerical parameters. Nevertheless, the analysis of our model provides insights into the range of possible optimal solutions. In this paper we present solutions for China and India, for two different planning horizons and finally for rolling horizons. The particular solutions presented are all derived with the same set of numerical parameters. Due to the different initial conditions, i.e. the higher fertility and the resulting younger age structure in China, the optimal solutions for the two countries differ considerably, although the same numerical parameters are used. Our analysis shows that, assuming a 50-year planning horizon, fertility decline has been too rapid in China and too slow in India. Extending the planning horizon to 200 years changes the picture. In this case, fertility declined too quickly in both countries. This change was to be expected, as a longer planning period naturally means more time to achieve the goal. Finally, optimal solutions for China are presented under a rolling horizon. In this case, after a fixed period, the end of the planning horizon is shifted into the future and a new optimal solution is calculated. A comparison of the results with a fixed planning horizon and a rolling horizon with the actual development in China shows that the fertility decline in China was closer to the optimal solution with a fixed planning horizon. This suggests that China consistently pursued the path of fertility decline that it had originally decided upon.

The mathematical approach presented in this paper allows for a wide range of applications. China and India were selected, on one hand, because of their sudden and pronounced drop in fertility. On the other hand, these two countries are convenient because of their low migration rates. This allows us to approximate the population dynamics by leaving migration aside. This in turn enables the derivation of analytical results, which we present in (Wrzaczek et al., 2024) and adopt for this practical application. However, numerical computation of optimal solutions, including migration, is also possible. Moreover, the model cannot only give insights into optimal fertility decline but also provide economically optimal solutions for fertility increases. Furthermore, the end time constraint regarding the net reproduction rate at the end of the planning horizon is not fixed to one but can take any desired target level. This expands the set of countries to which our approach can be applied. The model presented in this paper is therefore very flexible and versatile and can be applied to virtually any country or region where fertility needs to change.

REFERENCES

Aksan, A.-M. (2021). Son preference and the fertility squeeze in india. *Journal of Demographic Economics*, 87(1):67—106.

Baizan, P., Arpino, B., and Delclos, C. (2016). The effect of gender policies on fertility: The moderating role of education and normative context. *European Journal of Population*, 32:1–30.

Bloom, D. E. and Williamson, J. G. (1998). Demographic Transitions and Economic Miracles in Emerging Asia. *The World Bank Economic Review*, 12(3):419–455.

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)*, pages 957–974.

Debnath, A. and Das, S. (2022). Inter-relationship among female labour-force participation, fertility and economic development: Evidences from india. *Economic Affairs*, 67(4):673–680.

Dörflinger, M. and Loichinger, E. (2024). Fertility decline, changes in age structure, and the potential for demographic dividends: A global analysis. *Demographic Research*, 50(9):221–290.

Dreze, J. and Murthi, M. (2001). Fertility, education, and development: Evidence from india. *Population and Development Review*, 27(1):33–63.

Feeney, G. (1994). Fertility decline in East Asia. *Science*, 266(5190):1518–1523.

Feeney, G. and Feng, W. (1993). Parity progression and birth intervals in china: The influence of policy in hastening fertility decline. *Population and Development Review*, 19(1):61–101.

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.

Feichtinger, G. and Wrzaczek, S. (2024a). The optimal momentum of population growth and decline. *Theoretical Population Biology*, 155:51–66.

Feichtinger, G. and Wrzaczek, S. (2024b). The optimal transition to a stationary population for concentrated vitality rates. *Demographic Research*, 50:171–184.

Fent, T., Prskawetz, A., and Aparicio Diaz, B. (2013). Family policies in the context of low fertility and social structure. *Demographic Research*, 29(37):963–998.

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

Grass, D., Caulkins, J. P., Feichtinger, G., Tragler, G., Behrens, D. A., et al. (2008). Optimal control of nonlinear processes. *Berlino: Springer*.

Gupta, M. D. and Bhat, P. N. M. (1997). Fertility decline and increased manifestation of sex bias in india. *Population Studies*, 51(3):307–315.

International Labour Organization (2024). ILO Modelled Estimates and Projections database (ILOEST). ILOSTAT. Accessed February 06, 2024. ilostat.ilo.org/data.

Kashyap, R. and Villavicencio, F. (2016). The dynamics of son preference, technology diffusion, and fertility decline underlying distorted sex ratios at birth: A simulation approach. *Demography*, 053(5):1261—-1281.

Keyfitz, B. L. 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.

Keyfitz, N. and Caswell, H. (2005). *Applied Mathematical Demography*. Springer, New York, third edition.

Krishnaji, N. (2009). Fertility transition in india: Development and diffusion. Development economics working papers, East Asian Bureau of Economic Research.

Leridon, H. (1983). Un taux d'accroissement constant comme objectif démographique et méthode de projection. *Population (french edition)*, pages 343–360.

Lui, L. and Cheung, A. K.-L. (2021). Family policies, social norms and marital fertility decisions: A quasi-experimental study. *International Journal of Social Welfare*, 30(4):396–409.

Morgan, S. P., Zhigang, G., and Hayford, S. R. (2009). China's below-replacement fertility: Recent trends and future prospects. *Population and Development Review*, 35(3):605–629.

Nath, S. (2023). Explaining third birth patterns in india: causal effects of sibling sex composition. *Journal of Population Economics*, 36(4):2169–2203.

Piotrowski, M. and Tong, Y. (2016). Education and fertility decline in China during transitional times: A cohort approach. *Social Science Research*, 55:94–110.

Retherford, R. D., Choe, M. K., Chen, J., Xiru, L., and Hongyan, C. (2005). How far has fertility in China really declined? *Population and Development Review*, 31:57–84.

Sarkar, S., Sahoo, S., and Klasen, S. (2019). Employment transitions of women in india: A panel analysis. *World Development*, 115:291–309.

UNESCO Institute for Statistics (2023). UIS.Stat Bulk Data Download Service. Accessed September 19, 2023.

United Nations, D. o. E. and Social Affairs, P. D. (2022). World population prospects.

Veliov, V. M. (2008). Optimal control of heterogeneous systems: basic theory. *Journal of Mathematical Analysis and Applications*, 346(1):227–242.

Wang, P., Zhan, H. J., Liu, J., and Barrett, P. M. (2022). Does the one-child generation want more than one child at their fertility age? *Family Relations*, 71(2):494–512.

Wrzaczek, S., Feichtinger, G., Fent, T., and Novak, A. (2024). To grow or to fluctuate: Optimal paths to demographic equilibria. IIASA Working Paper. Laxenburg, Austria: WP-24-008.

A OPTIMALITY CONDITIONS

The full model presented in section 2 can be formulated as

$$\min_{k(t)} \quad \int_0^T \left[C_1 k^2(t) + C_2 \int_0^\omega \left(\frac{P(t,a)}{N(t)} - \bar{c}(a) \right)^2 \mathrm{d}a \right] \mathrm{d}t + C_3 (N(T) - N_0)^2 \tag{8a}$$

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

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

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

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

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

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

This is an age-structured optimal control model with concentrated and age-specific state variables. The model can be solved with the extension of the age-structured Maximum Principle for concentrated state and control variables. For details we refer to the appendix of Feichtinger and Wrzaczek (2024a).[3]

In this problem the population $P(t,a)$ represents an age-structured state variable, whereas births $B(t)$ as well as deaths $D(t)$ concentrated ones. Similarly, the birth adjustment effort $k(t)$ is a concentrated control variable. The Hamiltonian of this problem is defined as (time and age arguments are suppressed in the following):

$$\begin{aligned} \mathcal{H} \;=\; & -\frac{1}{\omega} C_1 k^2 - C_2 \left(\frac{P}{N} - \bar{c}(a) \right)^2 \\ & -\xi \nu P + \frac{1}{\omega} \lambda_R k R + \frac{1}{\omega} \lambda_N (B - D) + \eta_B \delta(a-\mu) P R + \eta_D \nu P, \end{aligned} \tag{9}$$

where $\xi(t,a)$ denotes the adjoint variable for the (age-structured) state variable $P(t,a)$, and $\lambda_R(t)$ and $\lambda_N(t)$ those of the (concentrated) state variables $N(t)$ and $D(t)$, respectively. $\eta_B(t)$ and $\eta_C(t)$ indicate the adjoint variables for $B(t)$ and $D(t)$.

[3] Alternatively the age-structured Maximum Principle (see e.g. Feichtinger et al., 2003; Veliov, 2008) can be used, if the concentrated state variables are transformed to age-structured ones.

Taking the first derivative of the Hamiltonian we get the following necessary condition for an inner optimum of $k(t)$ for every $t \in [0, T]$:

$$\frac{\partial \mathcal{H}}{\partial k} = -2\frac{1}{\omega}C_1 k + \frac{1}{\omega}\lambda_R R = 0 \qquad \Longrightarrow \qquad k = \frac{\lambda_R R}{2C_1}. \tag{10}$$

The adjoint variables (representing *dynamic* shadow prices) are obtained by the adjoint system that can be solved together with the transversality conditions. The adjoint system reads

$$\xi_t + \xi_a = \nu\xi + \frac{2C_2}{N}\left(\frac{P}{N} - \bar{c}(a)\right) - \eta_B \delta(a-\mu)R - \eta_D \nu \tag{11a}$$

$$\dot{\lambda}_R = -\lambda_R k - \int_0^\omega \eta_B \delta(a-\mu) P \, \mathrm{d}a \tag{11b}$$

$$\dot{\lambda}_N = -2C_2 \int_0^\omega \left(\frac{P}{N} - \bar{c}(a)\right) \frac{P}{N^2} \, \mathrm{d}a \tag{11c}$$

$$\eta_B = \lambda_N + \xi(t, 0) \tag{11d}$$

$$\eta_D = -\lambda_N. \tag{11e}$$

In contrast to the state variables that are defined by initial conditions (concentrated state variables), or an initial distribution and a boundary condition (age-structured state variables), the transversality conditions depict specific values of the adjoint variables at the end of the time horizon and at the maximum life-time of individuals. For problem (8) the following transversality conditions are obtained

$$\xi(t, \omega) = 0 \tag{12a}$$

$$\xi(T, a) = 0 \tag{12b}$$

$$\lambda_N(T) = -2C_3 \left(N(T) - N_0\right) \tag{12c}$$

$$\lambda_R(T) \qquad \text{no transversality condition.} \tag{12d}$$

For λ_R no transversality condition is available, as the corresponding state variable $R(t)$ is defined by an initial and an end condition. This is usual also in standard optimal control theory (see e.g. Grass et al., 2008).

The numerical solution can be found by the implementation of a gradient based optimisation algorithm, which works with the following basic idea. In this approach an initial guess for the control variables allows the calculation of the state and adjoint variables for this guess. The gradient then defines the direction along which the optimal solution can be approached until the solution converges. A more elaborate description of this approach can be found in Freiberger et al. (2024).

Wittgenstein Centre

FOR DEMOGRAPHY AND GLOBAL HUMAN CAPITAL

IIASA ÖAW universität wien

The Wittgenstein Centre is a collaboration among the Austrian Academy of Sciences (ÖAW), the International Institute for Applied Systems Analysis (IIASA) and the University of Vienna.

www.wittgensteincentre.org