Demographic Transition in the Ramsey Model: Do Country-Specific Features Matter?*

Ulla Lehmijoki
University of Helsinki and HECER

ISSN 1459-3696, ISBN 952-10-1542-X
610: 2004

November 2, 2004

Abstract

The paper modifies the Ramsey model to take demographic transition into account. The non-linear discount factor problem is solved in virtual time. The model may have multiple steady states. Family planning programs may be important in solving indeterminacy in the model. The transitional dynamics of the model show that economic growth fluctuates along with demographic growth. Country-specific features of transition determine the intensity of the fluctuation.

JEL Classification: O41, O11, J10.

Keywords: demographic transition, economic growth, neo-classical growth models.

---

[*] I would like to thank Pertti Haaparanta, Matti Pohjola, Pekka Ilmakunnas, Tapio Palokangas, Seppo Honkapohja, Alexia Fuernkranz-Prskawetz , Carl Johan Dalgaard, and Norman Sedgley for helpful comments. I am also indebted to Yrjö Leino from CSC, the Finish IT center for science for calculating and drawing the bifurcation diagram of the paper, and to Heikki Ruskeepää for his advice in MATHEMATICA© −calculations.

1 Introduction

Current theoretical models on demographic transition suggest that transition occurred due to a rising rate of return to human capital (Becker *et al.* 1990), or due to an increase in the price of a mother's time (Galor and Weil 1996), or because technical progress motivated to substitute child quality for child quantity (Galor and Weil 2000, Lucas 2002, Galor 2004).

On the other hand, current growth empirics mainly rely on the Ramsey model (Ramsey 1928) which ignores demographic transition in assuming that the population growth rate is constant. This assumption would not be so problematic if the transition everywhere had followed the same pattern so that all countries were parallely affected. But the data on demographic transition in Figure 1 show that the features of transition greatly varied from country to country and symmetry in its economic effects is not to be expected. On the contrary, the fact that demographic transition in some countries has been of a different magnitude implies that economic consequences have been of different dimensions as well.

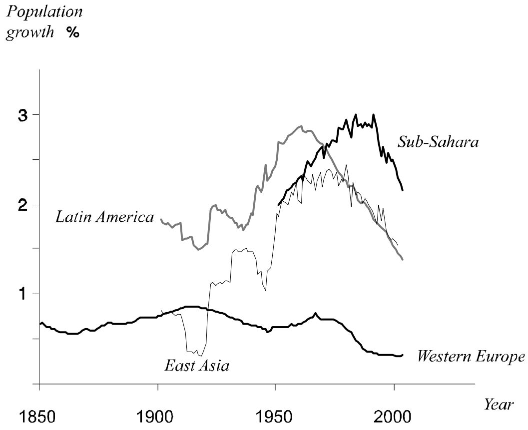

Figure 1: Demographic transition in selected groups. Source: Maddison 2003.

In this paper we want to discover the role of the country-specific features of demographic transition in the growth performance of countries. We introduce the transition into the Ramsey model by assuming that the population growth rate is not constant but a function of per capita income such that population growth first increases and then decreases. This simple assumption is in line with the data (Lucas 2002) and with those microfoundations in which increases in income are accompanied by increases in the price of time, and the dominance of the income effect changes to the dominance of the price-of-time effect so that the demand for normal goods like children first increases and then decreases (Becker 1982). The explanations provided by Galor and Weil (1996 and 2000), Lucas (2002), and Galor (2004) lean essentially on the role of technical progress but even these models predict that the correlation between income and population growth is first positive and then negative.

An explanation for *differences* in demographic transitions was suggested by Watkins (1990) who argued that diffusion of technology and information has been important. At the onset of demographic transition some countries were close to the technical frontier but some far behind. Income and technology advanced slowly in the former but were available “on a tray” in the latter (Williamson 1998). Therefore, demographic transition also proceeded at an accelerated rate in the adopting countries as is suggested by Figure 1.

We concentrate on three country-specific features in demographic transition: on the intensity of population growth, on its sensitivity to income, and on the level of income from which on population growth keeps decreasing. We find that if demographic transition takes an aggravated form the model has multiple steady states and a poverty trap. The model also predicts that, during the transitional period, economic growth fluctuates and this fluctuation is stronger the more prominent the demographic transition is. Fluctuations imply that convergence of incomes fails. Therefore, cross-country growth statistics should be reconsidered to make them compatible with demographic transition.

The mechanism of the model is the following: consumers choose between consumption and accumulation in the knowledge that the latter leads to some predictable changes in population growth. Therefore, consumers also choose that population growth rate which maximizes their utility in the long-run. Compared to the fertility decisions on a day-to-day basis (e.g., Palivos 1995), the long-run optimization keeps the model in one sector and provides easy access to the transitional dynamics of the model. The argument is that demographic transition, as the name implies, is a transitional phenomenon which goes back one or two hundred years. Hence, the empirics can be best understood from a transitional perspective.

The outline of the paper is the following: Chapter 2 introduces the modified Ramsey and solves it in virtual time (Uzawa 1968). Chapter 3 discusses how the dynamics are related to country-specific features and what was the role of family planning programs in solving indeterminacy of the model. A calibrated model is provided. The main analysis deals with the competitive version but the central planner's version is given in Appendix. Chapter 4 closes the paper.

2 The Ramsey Model Modified

2.1 The Economy and the Population

Consider an economy with capital $K(t)$ and labor $L(t)$ so that per capita capital is $k(t) = K(t)/L(t)$. Assume that the per capita production function $y(t) = f[k(t)]$ satisfies $f' > 0$, $f'' < 0$ and $\lim_{k\to 0} f'(k) = \infty$ and $\lim_{k\to\infty} f'(k) = 0$. Per capita capital accumulates according to

$$\dot{k}(t) = f[k(t)] - c(t) - (\delta + n) k(t), \tag{1}$$

in which $c(t)$, δ and n are per capita consumption, depreciations, and the population growth rate respectively. The economy maximizes $U = \int_0^\infty u[c(t)] L(t) e^{-\rho t} dt$, i.e., utility is derived both on per capita consumption and on the number of people. For $L(0) = 1$ and $L(t) = e^{nt}$ the integrand takes the familiar expression $u[c(t)] e^{-(\rho - n)t}$. This is the standard Ramsey model that can be considered as a central planner's problem or as a problem of a decentralized competitive economy. In the latter n should be considered as the growth of family size which is equal to population growth because households are identical. In the text, we concentrate on the competitive model; the planner's solution is given in Appendix A.

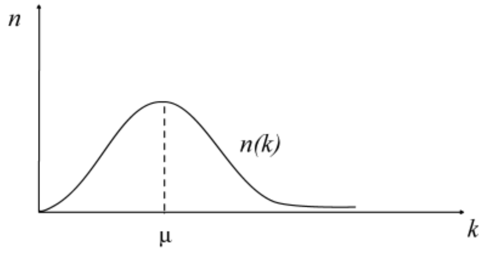

Figure 2: The population function.

We now modify the model by assuming that the population growth rate is a function of per capita income y. Further, because y is a monotonous in terms of k it is convenient to write population growth as a function of k.[1] *The population function* $n = n\left[k\left(t\right)\right]$ then becomes

$$\begin{aligned} n'\left[k\left(t\right)\right] > 0 &\Leftrightarrow k\left(t\right) < \mu, \\ n'\left[k\left(t\right)\right] = 0 &\Leftrightarrow k\left(t\right) = \mu, \\ n'\left[k\left(t\right)\right] < 0 &\Leftrightarrow k\left(t\right) > \mu. \end{aligned} \tag{2}$$

The capital stock $k(t) = \mu$ is the stock from which the number of children keeps decreasing (income $y = f(\mu)$ respectively). Further, we assume $\lim_{k\to 0}\left\{n'\left[k\left(t\right)\right]\right\} < \infty$, $\lim_{k\to\infty}\left\{n'\left[k\left(t\right)\right]\right\} = 0$. Defined in this way, the population function $n = n\left[k\left(t\right)\right]$ is in line with the data and with the microfoundations discussed above. Figure 2 illustrates. The size of population at time t becomes $L\left(t\right) = e^{\int_0^t n[k(\tau)]d\tau}$ and the expressions of U can now be replaced by

$$U = \int_0^\infty u\left[c\left(t\right)\right] \cdot \exp\left\{-\int_0^t \left\{\rho - n\left[k\left(\tau\right)\right]\right\} d\tau\right\} dt. \tag{3}$$

In (1) the effective depreciation $\left(\delta + n\right) k\left(t\right)$ becomes $\left[\delta + n(k\left(t\right))\right] k\left(t\right)$. We assume $\rho > n(k)$ for all k.

Equations (3) - (1) define an infinite horizon discount problem in which the discount rate is variable (see Uzawa 1968). To solve the problem we move from unit steps in natural time t to those in virtual time Δ by defining

$$\Delta\left(t\right) = \int_0^t \left\{\rho - n\left[k\left(\tau\right)\right]\right\} d\tau,$$

which gives $\frac{d\Delta(t)}{dt} = \rho - n\left[k\left(t\right)\right]$. The problem can be rewritten in terms of $\Delta\left(t\right)$:

$$U = \int_0^\infty \frac{u\left[c\left(t\right)\right]}{\rho - n\left[k\left(t\right)\right]} e^{-\Delta(t)} d\Delta(t), \tag{4}$$

$$\frac{dk(t)}{d\Delta(t)} = \frac{f\left[k\left(t\right)\right] - c\left(t\right) - \left(\delta + n\left[k\left(t\right)\right]\right) k\left(t\right)}{\rho - n\left[k\left(t\right)\right]}. \tag{5}$$

In the virtual time the discount factor is constant and the problem can be solved by standard methods (Benveniste and Scheinkman 1982).[2] The cur-

---

[1] Solow (1956) suggested the formula $n = n(k)$ but did not interprete in terms of demographic transition.

[2] We abandon time and functional indicies if possible. Recall, however, that $n = n(k)$.

rent value Hamiltonian and the necessary conditions become $H(k,c,\lambda) = \frac{1}{\rho-n}\{u+\lambda(\Delta)[f-c-(\delta+n)k]\}$, and $\partial H/\partial c = 0$, and :

$$\begin{aligned} \frac{d\lambda(\Delta)}{d\Delta} &= -\frac{\partial H(k,c,\lambda)}{\partial k} + \lambda(\Delta), && (6) \\ && (7) \\ \lim_{\Delta\to\infty}\left\{\lambda(\Delta)\cdot e^{-\Delta}\cdot k\right\} &= 0, \end{aligned}$$

together with (5). Condition (6) reverts back to natural time by writing $\dot{\lambda} = \frac{d\lambda}{d\Delta}\frac{d\Delta}{dt} = (\rho-n)\left\{-\frac{\partial H(k,c,\lambda)}{\partial k}+\lambda\right\}$. The condition $\partial H/\partial c = 0$ implies $u' = \lambda$. We eliminate λ in the usual way. After some algebra the differential equation for consumption becomes

$$\frac{\dot{c}}{c} = \frac{-u'}{u''\cdot c}\left\{f' - (\delta+\rho) - n'\cdot k + \frac{n'}{u'}H(k,c)\right\}, \tag{8}$$

in which $H(k,c) = \frac{1}{\rho-n}\{u+u'[f-c-(\delta+n)k]\}$ refers to *optimized Hamiltonian* derived by elimination of λ. The Euler equation of the model is:

$$f' - \delta = -\frac{u''c}{u'}\cdot\frac{\dot{c}}{c} + \rho + n'\cdot k - \frac{n'}{u'}H(k,c).$$

The Euler equation says that an investment is profitable if its (net) marginal product covers the loss of utility. This loss of utility consists, in addition to the ordinary terms, (elasticity of intertemporal substitution and time preference) of terms $n'\cdot k$ and $\frac{n'}{u'}H(k,c)$. The term $n'\cdot k$ says that because investment changes per capita capital, the population growth rate changes and a changed number of new people must be provided with new capital. Note that if $n'(k) < 0$, this factor alleviates the productivity requirement. But a changed number of new people also consume. The optimized Hamiltonian refers to the total utility derived by a person $H(k,c)/u'$; a change in population growth changes the total flow of utils in the future.

2.2 The Solution

Equation (8) is easier to handle if we adopt the *CIES* utility function $u(c) = \frac{c^{1-\theta}}{1-\theta}, \theta > 0, \theta \neq 1$, in which $\frac{-u'(c)}{u''(c)c} = \frac{1}{\theta}$. Hall (1988) suggests that high values for θ are empirically most plausible. Therefore, through the analysis we assume $\theta > 1$ but nothing essential is changed if the reverse assumption is adopted. Then the optimized Hamiltonian is $H(k,c) = \frac{1}{(\rho - n)}\left\{\frac{c^{1-\theta}}{(1-\theta)} + c^{-\theta}\left[f - c - (\delta + n)\, k\right]\right\}$ and the differential equations for consumption are

$$\frac{\dot{c}}{c} = \frac{1}{\theta}\left[f' - (\delta + \rho) - n' \cdot k + \frac{n'}{c^{-\theta}} H(k,c)\right], \tag{9}$$

The $\dot{k} = 0$ and $\dot{c} = 0-$lines in the $k - c-$ space are given by

$$\dot{k} = 0 \Rightarrow c = f - (\delta + n)\, k. \tag{10}$$

$$\dot{c} = 0 \Rightarrow c = \frac{\theta - 1}{\theta}\{[f' - (\delta + \rho)]\,(\frac{\rho - n}{n'}) + [f - (\delta + \rho)\, k]\}. \tag{11}$$

The $\dot{k} = 0-$line runs from the origin and intersects the $k-$axis at $\tilde{k}$ where $f(\tilde{k})/\tilde{k} = \delta + n(\tilde{k})$. Even if $f(k)$ is concave the $\dot{k} = 0-$line has non-concave areas because $n = n(k)$.[3]

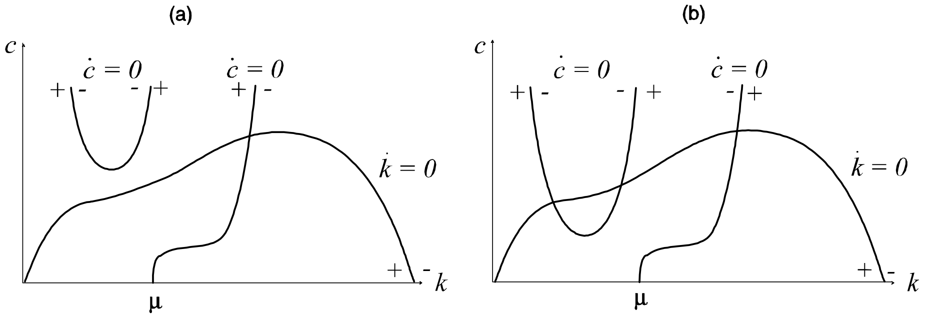

Figure 3: The phase diagrams.

To capture the shape of the $\dot{c} = 0-$line we concentrate on its limit behavior. In addition to the constant $\frac{\theta-1}{\theta} > 0$ the line consists of three expressions. First, the expression $f - (\delta + \rho)\, k$ is positive for $k < \breve{k}$ where $f(\breve{k})/\breve{k} = (\delta + \rho)$. This expression has no effect on the limit behavior but affects the shape of the $\dot{c} = 0-$line in the vicinity of the horizontal axis. Second, $f'(k) - (\delta + \rho)$ approaches $+\infty$ as k goes to zero, intersects the $k-$axis from above at $\hat{k}$ where $f'(\hat{k}) = (\delta + \rho)$ and approaches $-(\delta + \rho)$ as k goes to infinity. Third, to the assumptions above the expression $\frac{\rho - n}{n'}$ approaches a finite positive number as

---

[3] It is in principle possible that the isocline cuts the k-axis for $k < \tilde{k}$ due to a strong demographic transition. This, however, would imply that population grows at a high rate even if consumption is zero — a situation impossible in real life.

k goes to zero. Further, it approaches $+\infty$ as $k \to \mu$ from the left but $-\infty$ as $k \to \mu$ from the right, and it has a point of discontinuity at $k = \mu$. To determine the behavior of $[f' - (\delta + \rho)]\left(\frac{\rho - n}{n'}\right)$ close to μ we make the following assumption:

Assumption 1. *Demographic transition peaks at* $k = \mu$ *so that* $\mu > \hat{k}$ *where* $\hat{k}$ *is given by* $f'(\hat{k}) = (\delta + \rho)$.

Assumption 1 says that population growth peaks at a relatively low level of per capita capital (income) and it is justified by the fact that everywhere demographic transition has occurred at the beginning of industrialization and development.[4] Therefore, $f'(k = \mu) - (\delta + \rho) > 0$ and $\lim_{k \uparrow \mu} \left\{[f' - (\delta + \rho)]\left(\frac{\rho - n}{n'}\right)\right\} = +\infty$ and $\lim_{k \downarrow \mu} \{\cdot\} = -\infty$. Further, because n' goes (from negative) to zero as k goes to infinity we have $\lim_{k \to \infty} [f' - (\delta + \rho)]\left(\frac{\rho - n}{n'}\right) = +\infty$. By definition $\hat{k} < \breve{k} < \tilde{k}$.

To summarize, the limit behavior of the $\dot{c} = 0-$line is

$$\begin{aligned}
\lim_{k \to 0} (\dot{c} = 0) &= +\infty, \\
\lim_{k \uparrow \mu} (\dot{c} = 0) &= +\infty, \quad \lim_{k \downarrow \mu} (\dot{c} = 0) = -\infty, \\
\lim_{k \to \infty} (\dot{c} = 0) &= +\infty.
\end{aligned}$$

This limit behavior implies that $\dot{c} = 0-$line takes a $U-$shaped graph for $k < \mu$, but swings from $-\infty$ to $+\infty$ for $k > \mu$. For $k = \tilde{k}$ the $\dot{k} = 0-$line hits the $k-$axis but the $\dot{c} = 0-$line is positive and the model has at least one interior steady state.

The phase diagram depicted in Figure 3 shows that two generic cases arise. The $U-$part of the $\dot{c} = 0-$line can lie so high that the number of interior steady states is one (panel a). Alternatively, the $U-$part lies low and the number of interior steady states is three (panel b).[5] Local stability analysis shows that the outermost steady states (the single steady state in panel a) are saddle points with stable paths running from southwest and northeast while the steady state

---

[4] For discussion of concrete numbers, see page 11.

[5] The non-generic tangent case is not analyzed. Because of non-concavities, additional steady states can not be excluded a priori. Parametric calculations below show that cases in Figure 3 are typical. We concentrate on these cases.

between them is an unstable focus or node. We assume the former; the analysis of the latter is not much different.[6]

Now compare Figures 2, 3, and 4. In each steady state k^* population growth holds constant $n = n(k^*)$. However, it is apparent that the low-income steady state (low k^*) is located on the increasing part of the population function, i.e., left of μ, whereas the high-income steady state is located right of μ. Therefore, the economy which is led to the low-income steady state never reaches the peak of its demographic transition. on the other hand it is not possibly to *a priory* conclude in which steady state population growth is higher; in principle it is possible that income stagnates at such a low level that demographic transition never really gets started.

In case of three steady states the saddle paths can adopt several shapes. At least two alternatives are present: path B towards the high-income steady state can emanate out of the unstable focus as depicted in Figure 4 or it can run from the origin as depicted in Figure 5.[7] In the latter case the high-income steady state is reachable from all initial states but in the former the capital stock must be at least k_l initially, i.e., the model has a poverty trap.

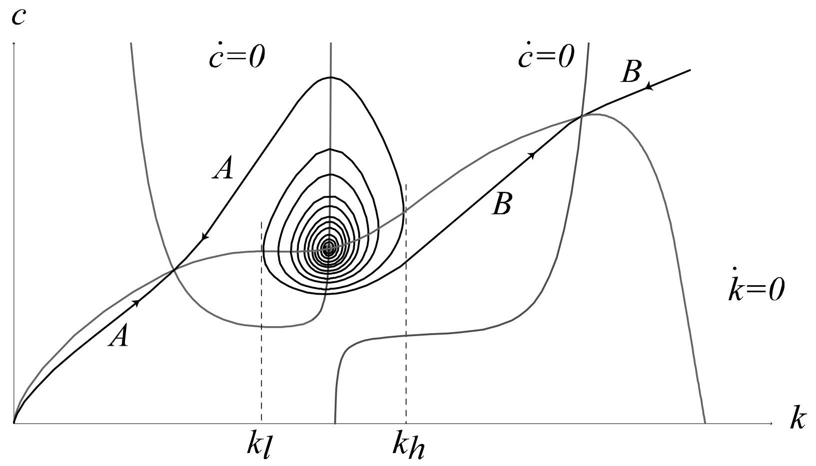

Figure 4: Stable saddle paths A an B, the spiraling case. Capital stock k_l (k_h) is the lowest (highest) initial stock from which the high-income (low-income) steady state can be reached.

---

[6] Palivos (1995) analyzes the case of an unstable node in his two-sector model.

[7] Essential parts in Figures 4 and 5 are parametrically drawn by applying parameters as reported in Table 1. Mathematica 4.02 files to draw the original figures are available from the author.

If several paths for some initial state $k(0)$ are available, and if households are unable to predict which of them gets realized, they are unable to make their decisions. Therefore, the model is indeterminate for $k(0) \in [k_l, k_h]$ in Figure 4 and for $k(0) < k_h$ in Figure 5. A way out of indeterminacy was suggested by Matsuyama (1991) who argued that if consumers adopt similar expectations and behave accordingly their expectations become self-fulfilling. Now consider a developing country which implements a family planning program in order to reduce birth rates. These programs usually apply concrete measures that increase information and availability of contraceptives but they also try to make small families more attractive by suggesting that they are "modern" or "families of the future". This may shape people's beliefs about the expected behavior of their neighbors and relatives. They may start to believe that the small family alternative is the most likely in the future and calculate that social services and education policies will be formulated to benefit the majority and, finally, they may choose to become part of the majority. Indeterminacy is solved and path B becomes optimal for an individual family. A well formulated program may behave like a self-fulfilling prophecy; it may shape people's reproductive behavior to a much higher extent than can deduced from its concrete measures.

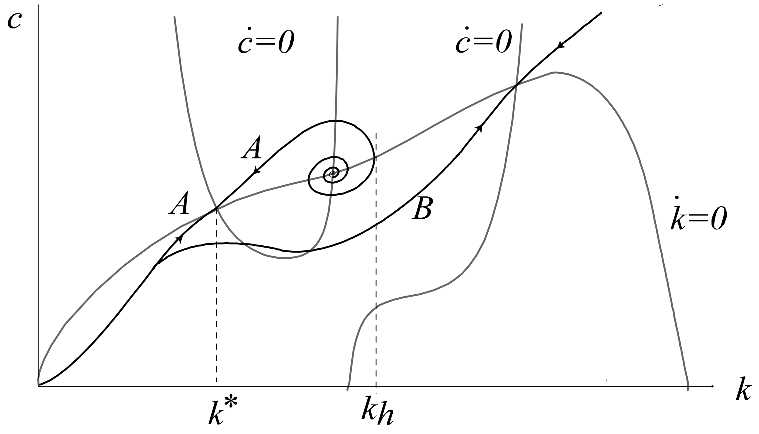

Figure 5: Stable saddle paths A an B, saddle B from origin.

$\alpha = 0.7$	The share of broad capital
$\rho = 0.045$	Time preference factor
$\theta = 3$	The negative of the elasticity of marginal utility
$\delta = 0.05$	The rate of depreciation
$10 < \sigma < 120$	The (inv. of) elasticity of pop. gr. to p.c. capital (income)
$150 < \mu < 778.5 = \hat{k}$	The peak stock of per capita capital
$0.01 < \eta < 0.045$	The peak population growth rate
$y = k^{\alpha}$	Cobb-Douglas production function
$u(c) = \frac{c^{1-\theta}}{1-\theta}$	CIES utility function
$n(k) = \eta e^{-\frac{1}{2}(\frac{k-\mu}{\sigma})^2}$	Population function

Table 1: The functional forms and the values of the parameters.

3 Do Country-Specific Features Matter?

Panels *a* Figure 3 and Figures 4 and 5 refer to three alternative solutions of the model. In this chapter we try to discover whether country-specific features can discriminate between these solutions, i.e., whether we can identify the features of transition that give birth to each of them. For this purpose we introduce a calibrated version of the model.[8] Several functional formulas satisfy Equation (2), among them the logistic formula which, however, fails the requirement that demographic transition ultimately levels-off, i.e., $\lim_{k\to\infty} \{n'[k(t)]\} = 0$. In this paper we suggest the formula

$$n(k) = \eta \cdot \exp\left\{-\frac{1}{2}\left(\frac{k-\mu}{\sigma}\right)^2\right\},$$

in which η is the (peak) population growth rate[9], μ is the peak stock of per capita capital, and $1/\sigma$ controls elasticity in terms of capital (income); low values for $1/\sigma$ refer to low elasticity.

We use the Cobb-Douglas production function and parameters close to those of Barro and Sala-i-Martin (1995). To evaluate the limits for parameters η, μ, and δ note that the data on the peak population growth rate η are readily available from demographic statistics and it ranges from approximately 0.01 to 0.04 (see also Figure 1). To find limits for σ, write $L(t) = L(0) \cdot \exp\left\{\int_0^t \eta e^{-\frac{1}{2}(\frac{k(\tau)-\mu}{\sigma})^2} d\tau\right\}$ in which $\exp\left\{\int_0^t \eta e^{-\frac{1}{2}(\frac{k(\tau)-\mu}{\sigma})^2} d\tau\right\}$ is the population

---

[8] Matsuyama (1991) has analyzed this question in a constant discount rate model by using the global bifurcation technique.

[9] For $k = \mu$ we have $n(k) = \eta$. Note, however, that for any k high η referes to high population growth rate.

multiplier that shows by how many fold population grows during the transition. Empirical estimates on multiplier are between 2.5 and 20 (see Livi-Bacci 1997) which gives limits $10 < \sigma < 120$. To find limits for μ note that *Assumption 1* requires $f'(k = \mu) - (\delta + \rho) > 0$. By applying values $\delta = 0.05$, $\rho = 0.045$ and $\alpha = 0,7$ we derive $\mu < 778.5$. The Cobb-Douglas formula implies that the per capita income produced by the per capita capital $k = \mu = 778.5$ is 106. The data provided by Maddison (2003) show that the highest per capita incomes during the peak of demographic transition have been approximately 3000 and the lowest approximately 1000 international 1990 (Geary-Khamis) dollars. Therefore, by applying multiplier 30 to move between the model and 1990 dollars we derive the lowest limit for $k = \mu \approx 150$. The parameters are summarized in Table 1.

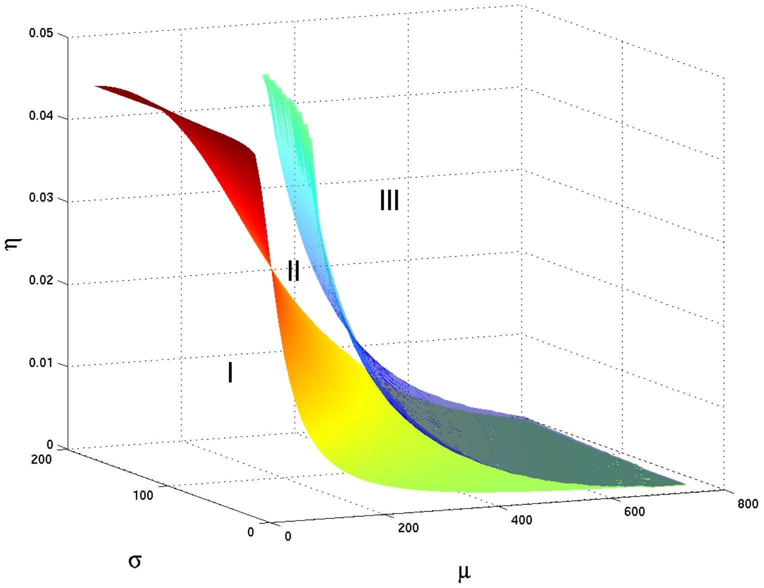

Figure 6: Effect of the parameters in the calibrated model. Area I: single steady state. Area II: three steady states, the south-western saddle path B starts from the origin. Area III: three steady states, the south-western saddle path B emanates spirally from steady state 2. The figure was calculated and drawn by Yrjö Leino from CSC.

Figure 6 shows the combined effects of parameters η, μ, and σ. The two surfaces divide the space into three areas I, II, and III which refer to panel a in Figure 3 (single steady state), to Figure 5 (path B from origin), and to Figure 4 (path B spirals) respectively. Consider first area III in which η, μ, and $1/\sigma$ are all high. For intuition note that every unit investment must be divided between capital deepening and capital widening. Therefore

- high value for η means that population grows at a high rate and the burden of capital widening is high for all k,
- high $1/\sigma$ refers to high elasticity. Every increase in capital stock is accompanied by a large increase in population growth. Therefore, the *marginal* burden of capital widening is high,
- high μ means that population growth peaks for large values of capital and every newcomer must be provided with a large stock. Further, because of diminishing returns, the capital widening may be excessive and the economy may stagnate into the low-income steady state.

Next consider countries in area II with still relatively high values for η, μ, and $1/\sigma$. If we assume that indeterminacy is solved in favor of path B as described above, then countries in area II proceed towards the high-income steady state. Equation (1) gives the off-steady state growth rate for per capita capital as

$$\gamma_k = \frac{\dot{k}}{k} = \frac{f}{k} - \frac{c}{k} - (\delta + n).$$

Barro and Sala-i-Martin (1995) show that for constant population growth rate, $\dot{\gamma}_k = d\left(\frac{f}{k} - \frac{c}{k}\right)/dt < 0$ and because $y = f(k)$ the growth rate of per capita income also decreases. In our model $n = n(k)$ and $\dot{n} = n'(k)\dot{k}$ and a monotonic decrease is not implied. The transitional dynamics in Figure 7 (heavy line) show that the economic growth rate actually greatly varies in area II. Figure 7 also predicts that economic growth and capital accumulation maximizes during the transition peak because it is *optimal* to pass the peak as soon as possible. Apparently, this result is not very realistic. It is due to the assumption that the supply of labor is inelastic and the dependency burden is constant. In the real world, the dependency burden varies and is heaviest when population growth is at its highest (Williamson 1998). For example in 1965 the dependency rate in Eastern Asia was 0.76 (per one adult of working age) whereas this rate currently has decreased to 0.46 (United Nations 2003). Typical changes in the dependency burden tend to postpone the period of maximal economic growth from that predicted in the model.

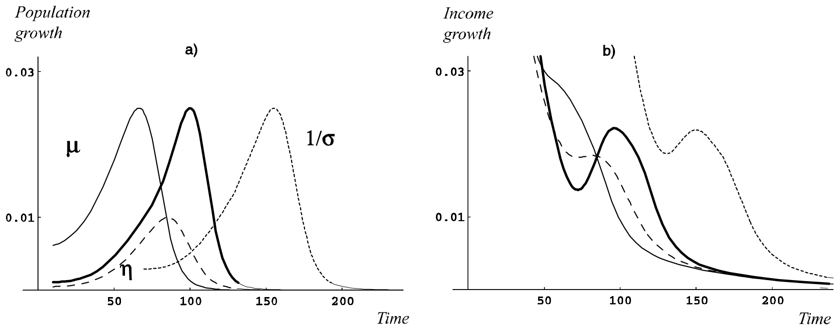

Figure 7: The time paths for population growth rate and the growth rate of per capita income. The original parameters on are II are $\eta = 0.025$, $\mu = 250$, and $\sigma = 100$ (heavy line). The changed parameters are $\eta = 0.01$, $\mu = 250$, and $\sigma = 120$.

Finally, to compare area I with area II we further decrease η, μ, and$1/\sigma$, *in turn*, so that the new combination of parameters lies on area I. Panel *a* in Figure 7 shows that a low value of η (naturally) makes the time path of population growth flatter. Further, if μ is low, population growth peaks early but the effect of low $1/\sigma$ is in the opposite direction. Panel *b* gives analogous changes in economic growth showing that a decrease in η, μ, or in$1/\sigma$ decreases the amplitude of fluctuations. Especially, a decrease in μ almost eliminates it. Therefore, panel *b* predicts that the effect of demographic transition on economic growth is rather neglible in area I if compared to its effect in area II.

To summarize, the calibrated model implies that economies which experience precipitous and exceedingly drastic demographic transitions in the sense that η, μ, and$1/\sigma$ *all* take remarkably high values (area III) are in danger of being caught in a poverty trap. However, because demographic numbers everywhere are decreasing, it is likely that most countries have evaded the trap. Countries in areas II (assuming that indeterminacy is solved in favor of path B) and I both proceed towards the high-income steady but in the former economic growth fluctuates much more than in the latter.

4 Discussion

The modified Ramsey model helps us to understand the role of country-specific features in the growth performance of countries. To make some preliminary contemplations take the extreme cases, Western Europe and Sub-Saharan Africa, as depicted in Figure 1. In Western Europe and Sub-Saharan Africa population growth peaked in 1913 and 1991 with peak population growth rates of 0.86% and $2,99\%$ and the peak-year per capita incomes of 3458 and 1522 (1990 international Geary-Khamis) dollars respectively (Maddison 2003). From 1850 to 1913 (in 63 years) per capita income in Western Europe increased by 120% but population growth increased by only 0.19 percentage points whereas in Sub-Sahara income increased from 1950 to 1991 (in 41 years) only by 64% but population growth increased by 0.99 percentage points showing that the income sensitivity of population growth was much higher in Sub-Sahara. An explanation may be in diffusion. As modern technology (new production methods but also pesticides and drugs) entered Africa a rapid decrease in mortality elevated´population growth to high levels. But at the time of the demographic peak per capita income was remarkably low making the burden of capital widening easier. This may have offered some compensation and helped the Sub-Saharan countries to endure the otherwise unbearable demographic growth rates.

Countries in Eastern Asia followed the same pattern but with an earlier peak in demographic growth. In Latin America development was exceptional in that at the time of the population peak per capita income was almost identical to that in Western Europe (3337$ in 1964). On the other hand, population growth reached the same rates as in other developing countries. The special features of Latin America — the European origin of the white population, early onset industrialization which then faded, and a large disparity between social groups — may have triggered such a development (Chesnais 1992). Whatever the explanation, the model indicates that the combined effect of high income and high population growth may have made the economic effects of demographic transition especially pronounced in Latin America.

The modified Ramsey model also has a bearing on cross-country growth empirics. First, note that discrepancies at the onset of industrialization led to the post-war situation in which developed countries proceeded towards the end of demographic transition whereas some developing countries bypassed the peak of transition and some others just arrived the transition. Therefore, to discover convergence one should find the countries which, during the research period, were in the same phase of their demographic transitions (Sala-i-Martin 1996). Lehmijoki (2003) applied the regression tree technique to find the number and the members of such clubs and found convergence in three of four. Further, understanding how countries proceed in their demographic careers and how the country-specific features change helps us to gain a better understanding of convergence outlooks. The model predicts that most countries bypass the demographic peak and proceed towards the high-income equilibrium so that, at least from a demographic point of view, optimistic rather than pessimistic expectations in terms of convergence seem most appropriate.

A Appendix: Central Planner's Solution

The central planner chooses $c(t)$ to maximize (3) subject to (1). If several saddle paths are available for some initial state $k(0)$, the planner chooses the path which maximizes the value of the program. For a constant discount rate problem, along any trajectory leading to a steady state the value of the program equals the value of optimized Hamiltonian evaluated at time zero and divided by the discount rate (Skiba 1978). The result generalizes to virtual time (discount rate unity). The proof and the discussion below utilize Tahvonen and Salo (1996).

Proposition 1 *Along any stable saddle path, the value of the program is* $H[k\left(0\right), c(0)]$, *in which* $c(0)$ *lies on that path.*

Proof. *The current value Hamiltonian* $H\left(k, c, \lambda\right) = H = \frac{1}{\rho - n}\left(u + \lambda\dot{k}\right)$ *and the conditions* $\frac{\partial H}{\partial c} = 0$, $\dot{\lambda} = \left(\rho - n\right)\left(\frac{-\partial H}{\partial k} + \lambda\right)$ *and* $\dot{k} = \left(\rho - n\right)\frac{\partial H}{\partial \lambda}$ *imply* $\frac{dH}{dt} = \frac{\partial H}{\partial c}\dot{c} + \frac{\partial H}{\partial k}\dot{k} + \frac{\partial H}{\partial \lambda}\dot{\lambda} = \frac{\partial H}{\partial \lambda}\left(\rho - n\right)\lambda = \lambda\dot{k}$. *Then*

$$\begin{aligned} -\frac{d\left(e^{-\Delta(t)}H\right)}{dt} &= -e^{-\Delta(t)}\left[\frac{dH}{dt} - \left(\rho - n\right)H\right] \\ &= -e^{-\Delta(t)}\left[\lambda\dot{k} - \left(\rho - n\right)H\right] \\ &= u \cdot e^{-\Delta(t)}. \end{aligned}$$

Recall that $e^{-\Delta(t)} = e^{-\int_0^t\{\rho - n[k(\tau)]\}d\tau}$ *and* $e^{-\Delta(0)} = 1$. *Then*

$$\begin{aligned} &\int_0^\infty u \cdot e^{-\Delta(t)}dt = -\int_0^\infty \left[e^{-\Delta(t)}\frac{dH}{dt}\right]dt \\ = \quad &H[k\left(0\right), c(0), \lambda\left(0\right)] - \lim_{t\to\infty} e^{-\int_0^t\{\rho - n[k(\tau)]\}d\tau}H[k\left(t\right), c(t), \lambda\left(t\right)]. \end{aligned}$$

Along any path leading to a steady state $H[k\left(t\right), c(t), \lambda\left(t\right)]$ *tends to be constant and* $\lim_{t\to\infty} e^{-\int_0^t\{\rho - n[k(\tau)]\}d\tau}H[k\left(t\right), c(t), \lambda\left(t\right)] = 0$. *Thus* $\int_0^\infty u[c(t)]e^{-\int_0^t\{\rho - n[k(\tau)]\}d\tau}dt$ $= H[k\left(0\right), c(0), \lambda\left(0\right)]$. *On a saddle path* $\lambda\left(0\right) = u'\left[c\left(0\right)\right]$ *so that* $H[k\left(0\right), c(0), \lambda\left(0\right)] = H[k\left(0\right), c(0)]$. ■

We apply *Proposition 1* to the case in which saddle B spirals out of the focus as depicted in Figure 4. Let k_l (k_h) be the lowest (highest) capital stock from which the high-income (low-income) steady state is reachable. The problem is to choose between two alternative saddle paths for initial capital $k_l < k(0) < k_h$ so that the value of the program is maximized. We utilize the approach suggested by Tahvonen and Salo (1996) which was based on two properties of the optimized Hamiltonian $H\left(k,c\right) = \frac{1}{\rho-n}\left(u + u' \cdot \dot{k}\right)$:

$$\Pr operty\ 1: \ \frac{\partial H(k,c)}{\partial c} = \left[u' + u''\dot{k} - u'\right]\frac{1}{\rho - n} = \frac{u''}{\rho - n}\dot{k}.$$

Each optimal path satisfies

$$\frac{dc}{dk} = \frac{\dot{c}}{\dot{k}} = \frac{-\frac{u'}{u''}\left\{\frac{-n'H(k,c)}{u'} - \left[f' - (\delta + \rho) - n' \cdot k\right]\right\}}{\dot{k}}.$$

Along any optimal path, $c = c(k)$. Then

$$\begin{aligned}
\Pr operty\ 2 \quad : \quad \frac{dH[k,c(k)]}{dk} &= \frac{\partial H[k,c(k)]}{\partial k} + \frac{\partial H[k,c(k)]}{\partial c}\frac{\dot{c}}{\dot{k}} \\
&= \frac{n'}{(\rho - n)^2}\left(u + u'\ \dot{k}\right) + \frac{u'}{\rho - n}\left[f' - (\delta + n) - n' \cdot k\right] - \frac{u''\dot{k}}{\rho - n}\frac{\dot{c}}{\dot{k}} \\
&= \frac{n'}{\rho - n}H(k,c) + \frac{u'}{\rho - n}\left[f' - (\delta + n) - n' \cdot k\right] - \frac{u''\dot{c}}{\rho - n} \\
&= u' > 0.
\end{aligned}$$

Property 1 is available to compare two paths lying on the same side of the $\dot{k} = 0-$line. Assume that $k(0) = k_l$. Denote the initial consumption chosen on path A and B by c_l^A and c_l^B, respectively. Then $H(k_l, c_l^A)$ and $H(k_l, c_l^B)$ are the values of the program if path A or B is chosen respectively. Note that $c_l^A > c_l^B$. Point (k_l, c_l^B) lies on the $\dot{k} = 0-$line but (k_l, c_l^A) above it implying $H(k_l, c_l^A) > H(k_l, c_l^B)$ and for $k(0) = k_l$ the value of the program is maximized on path A. By an analogous argument, for $k(0) = k_h$ the value of the program is maximized on path B.

Property 2 can be used to compare two paths as k changes. Because $u'' < 0$, the increase of $H[k, c(k)]$ as a function of k is faster the lower the value of $c(k)$ is. We show that it is never optimal to move along the spiral: Assume that for some $k(0) \in (k_l, k_h)$ path A *is* optimal. Path A can be reached by choosing one of several initial consumptions (Figure 4). Assume that the lowest possible initial consumption is chosen. To reach the steady state it is first necessary to move along A by $k(0) - k_h$ and then by $k_h - k(0)$ (Figure 4). The former (latter) increases (decreases) the value of the program. Because the former lies below the latter (has lower values for c) the value of the program increases. Therefore, for those initial capital stocks for which path A is optimal, it is always best to choose the highest possible consumption initially. By an analogous argument, *if* B is optimal, the lowest possible consumption should be chosen.

We compare paths A and B for initial values $k(0) \in (k_l, k_h)$. Because for all $k(0) \in (k_l, k_h)$ the best value of $c(k)$ is lower on B than on A (Figure 4), $H[k, c(k)]$ increases faster along B than along A as k increases. Because $H(k_l, c_l^A) > H(k_l, c_l^B)$ but $H(k_h, c_h^A) < H(k_h, c_h^B)$ and because $H[k, c(k)]$ is continuous in k, there exists a unique $k_m \in (k_l, k_h)$ so that $H(k_m, c_m^A) = H(k_m, c_m^B)$. For $k(0) = k_m$ the planner is indifferent regarding A and B. For all $k(0) < k_m$ it is optimal to choose A but for all $k(0) > k_m$ path B is optimal.

Consider the case depicted in Figure 5. For $k(0) \leq k^*$ path A lies above B and they both lie below the $\dot{k} = 0-$line and *Property 1* implies $H(k, c^A) < H(k, c^B)$. For $k^* < k(0) < k_h$, path B further lies below A and *Property 2* implies that the value of the program increases faster along B as $k(0)$ increases. For $k(0) \geq k_h$ only B is available. Thus, path B is globally optimal.

References

BARRO, R. J., AND X. SALA-I-MARTIN (1995): *Economic Growth.* McGraw-Hill, New York.

BECKER, G. S. (1982): *A Treatise on the Family.* Harvard University Press Massachusetts.

BECKER, G. S., K. M. MURPHY, AND R. TAMURA (1990): "Human Capital, Fertility, and Economic Growth," *Journal of Political Economy*, 98(5), S12–S37.

BENVENISTE, L., AND J. SCHEINKMAN (1982): "Duality Theory for Dynamic Optimization Models of Economics: The Continuous Time Case," *Journal of Economic Theory*, 27, 1–19.

Chesnais, J.-C. (1992): *The Demographic Transition: Stages, Patterns, and Economic Implication. A Longitudinal Study of Sixty-Seven Countries Covering the Period 1720-1984*. Clarendon Press, Oxford.

Galor, O., and D. N. Weil (1996): "The Gender Gap, Fertility, and Growth," *American Economic Review*, 86(3), 374–385.

——— (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and beyond," *American Economic Review*, 90(4), 806–826.

Hall, R. E. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96(2), 339–357.

Livi-Bacci, M. (1997): *A Concise History of World Population.* Blackwell Publishers, Oxford, U.K.

Lucas, R. E. J. (2002): *Lectures on Economic Growth.* Harvard University Press, Cambridge, Massachusetts.

Maddison, A. (2003): *The World Economy, Historical Statistic CD-ROM.* OECD, Paris.

Matsuyama, K. (1991): "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics*, 106(2), 617–650.

Palivos, T. (1995): "Endogenous Fertility, Multiple Growth Paths, and Economic Convergence," *Journal of Economic Dynamics and Control*, 19, 1489–1510.

Ramsey, F. P. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 38, 543–559.

Sala-i-Martin, X. X. (1996): "The Classical Approach to Convergence Analysis," *Economic Journal*, 106, 1019–1036.

Skiba, A. (1978): "Optimal Growth with a Convex-Concave Production Function," *Econometrica*, 46(3), 527–539.

Solow, R. M. (1956): "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics*, 70, 65–94.

TAHVONEN, O., AND S. SALO (1996): "Nonconvexities in Optimal Pollution Accumulation," *Journal of Enviromental Economics and Management*, 31, 160–177.

United Nations (2000): *World Population Prospects. The 2000 Revision, Vol I: Comprehensive Tables.* New York.

UZAWA, H. (1968): "Time Preference, the Consumption Function, and Optimum Asset Holdings," in (Wolfe 1968), chap. 21.

WATKINS, S. C. (1990): "From Local to National Communities: The Transformation of Demographic Regimes in Western Europe, 1870-1960," *Population and Development Review*, 16(2), 241–272.

WILLIAMSON, J. G. (1998): "Growth, Distribution, and Demography: Some Lessons from History," *Explorations in Economic History*, 35, 241–271.

WOLFE, J. (ed.) (1968): *Value, Capital, and Growth.* Aldine, Chicago.