No. 43

UNIVERSITY OF COLOGNE

WORKING PAPER SERIES IN ECONOMICS

HISTORY-DEPENDENT INDIVIDUAL BEHAVIOR, POLARIZATION, AND PARETO-IMPROVING ACTIVATING WELFARE

PETER FUNK

Department of Economics
University of Cologne
Albertus-Magnus-Platz
D-50923 Köln
Germany

http://www.wiso.uni-koeln.de

History-Dependent Individual Behavior, Polarization, and Pareto-Improving Activating Welfare

Peter Funk

University of Cologne*

Preliminary version, July 2009

Abstract

This paper assumes that human capital not only generates market incomes but is a direct source of utility as well. In an otherwise standard framework it is shown that the interaction between human capital and effort in raising human capital and in generating utility naturally leads to history-dependent optimal individual behavior. Depending on the initial distribution of skills, this history-dependence divides each group of otherwise identical households into two perpetually separated groups: one rich and educated, the other poor and uneducated. If the rich have a common interest in the education of the poor (for instance financing public goods), such polarized equilibria are typically Pareto-inefficient. While unconditional transfers only reduce the incentives of the uneducated to accumulate skills, it is shown that there exist activating tax-transfer systems that Pareto-dominate any non-redistributing tax-system and involve a negative marginal income tax on household income below a certain threshold.

"Give a man a fish and you feed him for a day. Teach a man to fish and you feed him for a lifetime." Chinese proverb.

1 Introduction

Human capital as a direct source of satisfaction Most people value their education, their personal abilities, their knowledge and skills, and their physical and intellectual fitness beyond their role in generating incomes, in particular when these assets have been actively acquired with past effort. While standard models of human capital recognize the crucial role of human capital as an asset

*Staatswissenschaftliches Seminar, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Germany. Tel.: +49/221/4704496, Fax: +49/221/4705143, Email: funk@wiso.uni-koeln.de

yielding a return through higher market earning capacity, they generally abstract from its role as a direct source of utility.

Despite the promotion by such illustrious advocates as Gary Becker and Amartya Sen (see below), the non-market or the direct effect on utility of human capital remains largely neglected in the economic literature. This is certainly not because economic researchers don't value their knowledge and skill for their own sake or are unaware of the higher wages their human capital may earn outside academics. The reason is more likely the general perception that the non-economic returns of education are not an economic issue and are better taken care off by other social sciences. As Becker (Becker [1993]) already noted, such a perception is obviously ill justified since ignoring human capital as a direct source of satisfaction or more generally, as a source of non-economic returns, quantitatively underestimates the total return to the investment into human capital and therefore underestimates the motivation to invest.

The crucial assumption of the present paper is that human capital (education, skill, fitness or knowledge) is a direct source of pleasure and satisfaction. It explicitly includes the state variable "human capital" as an argument in the instantaneous utility function of the individual household intertemporal optimization problem. At the same time it retains the standard assumptions of this problem: Human capital generates income, effort raises the stock of human capital and reduces (instantaneous) utility.

Path-dependent individual behavior Ignoring the direct utility of human capital not only underestimates the total return to the investment into human capital. More importantly it also alters the qualitative nature of the interaction between human capital, the investment into human capital (effort), and the economic return to human capital (income): This paper shows that the interaction between skill and effort in raising the future stock of human capital and in generating utility naturally leads to history-dependent optimal individual behavior.

The possibility of history dependent behavior requires that the incentive to exert effort rises with rising skill. This requirement involves both the accumulation and the utility of human capital. Concerning the former it will be assumed that the effectiveness of effort (in raising skill) is increasing in the skill level so far attained: Effort is the more effective, the more one already knows (Equivalently, one could assume that effort is the more pleasant the more one knows). As for the latter, the rising effectiveness of effort in generating skill should not be neutralized by too strongly decreasing marginal utility of skill. This second part of the requirement will follow from what I will call the "Maslow condition". The condition is reminiscent of Maslow's "Hierarchy of needs" (Maslow [1943,1954]) which features in many introductory text book on the psychology of motivation. Maslow posited a hierarchy of human needs based on two groupings: deficiency needs and growth needs. In Maslow's hierarchy, deficiency needs must be met before an individual can draw significant satisfaction from growth needs. Adapting this idea to a formal framework with strictly monotone (non-satiable) preferences, the present paper assumes that the marginal (direct) utility does not tend to infinity when human capital tends to zero (Inada condition violated) and that it does not fall with rising human capital, while the marginal utility of consumption tends to infinity when consumption tends to zero (Inada condition satisfied) and falls fast with increasing consumption (the absolute value of the elasticity of the marginal utility of consumption is not smaller than one). Thus, the standard Inada condition is met for small consumption but not for small levels of human capital.

Under this basic assumptions, the history dependence of optimal individual behavior occurs whenever at low consumption the marginal (instantaneous) utility of consumption is sufficiently small compared to the marginal (instantaneous) disutility of effort, while the marginal utility of human capital rises sufficiently fast with rising consumption. There then exists a threshold skill level, below which the effort necessary to lift a household beyond the threshold is unattractively high. Any household initially below this level chooses a path towards increasing passivity and sustained poverty. In contrast, households with initial human capital above the threshold – motivated both by economic and non-economic rewards – chose a path toward sustained activity, high skill and income.

Polarization and Pareto-Efficiency Path dependent individual behavior generates a source of polarization: If within a group of otherwise identical households not all are initially positioned at the same side of the threshold, then this group will eventually become completely polarized into two types of identical households, one uneducated and poor and one educated and rich.

In the absence of external effects any intertemporal equilibrium in the threshold economy, even an extremely polarized one, is Pareto-efficient. Of course, the poor and uneducated would be better off being rich and educated, but *given* their initial human capital below threshold they choose a path towards even more passivity because it is optimal for them to do so! The path-dependence is not a result of coordination failures between households but simply of individual optimization.

Pareto-optimal or not, the extreme inequality coming along with the polarization would call for massive redistribution from rich to poor in most existing industrial countries. In part, the predominance of redistributing tax-transfer systems may originate in the balance of political power that *forces* the rich to finance transfers to the poor. In part, these redistributing systems arise because actually there do exist external effects inducing the rich to *voluntarily* vote for such systems.

Irrespective of which of these reasons explains why existing fiscal systems redistribute, an important aspect of redistribution in an economy with endogenous human capital accumulation must be the question how the tax-transfer scheme affects the individual incentives to accumulate human capital. The present paper addresses this question in a framework with external effects, such that a polarized market equilibrium will in general not be Pareto-efficient. Positive externalities from the action, the human capital or the income of the poor affect the utility of the rich and educated. While in reality there may be numbers of such externalities, this paper concentrates on one that is most common in standard economics: It assumes the existence of a public good.[1] The larger the number of financially strong households that can participate in financing the public good, the better for everybody. The rich and educated benefit from the education of the so far uneducated simply because this either reduces their own contribution or raises the amount of the public good. Whether or not polarization is Pareto-inefficient now depends on how much it would cost the rich to motivate the poor towards a path of learning. In the present setting the rich have strictly positive willingness to pay to contribute to a welfare system if this guarantees the education of the poor beyond the threshold sufficiently fast. If this willingness to pay is high enough to compensate the poor's' disutility of the effort necessary to raise their education, then polarization is Pareto-inefficient.

All essential issues of the present paper could be discussed in a framework without externalities (no public good, no altruism or paternalism) and a government that wants to maximize a utilitarian social welfare function. It is clear that such a government would want to redistribute a polarized market outcome (otherwise the ratio of the poor's to the rich's marginal utility would grow without bound and redistribution would raise social welfare). In this alternative framework too the rich and educated benefit from the education of the so far uneducated, now because this limits the period during which they have to support the poor.

Simple transfers As has been explained the only reason to exert effort for households with low skill is the economic motive. A simple unconditional transfer, by raising consumption and thus reducing the marginal utility of consumption, while keeping constant the high disutility of effort at low skill, thus only reduces the incentives for such households that have the potential to overcome the cost of effort. In fact, it will be shown that such a transfer turns any active low-skill household into an inactive household. Even if at low consumption the marginal utility of consumption is always sufficiently high to motivate low-skill households to activity and learning (no threshold), a simple (non-contingent) transfer introduces a threshold.

Suppose that, starting at a Pareto-inefficient polarized equilibrium, a hypothetical social planner or a real government wants to implement a Pareto-improving tax-transfer system. Such a system has to

[1] If one agrees with the assertion, that most rich and educated people dislike abolute poverty and the complete lack of education of their neighboor, one may still disagree on the motives for this aversion. More direct externalities than those assumed here, also leading to the Pareto-inefficiency of polarization are altruism (the rich's utility is raised by raising the utility of the poor) or paternalism (the rich's utility is raised by raising income or human capital of the poor, or by raising effort and education of the poor).

win the approval of the poor and at the same time motivate them to raise their effort such as to increase their human capital. If simple transfers were the only available policy, this double aim of feeding and activating would establish an unsolvable dilemma for the government, a version of the "Samaritan's Dilemma": An unconditional transfer alleviates the material misery of the poor and unskilled but at the same time reduces their material motivation to exert effort and thus tends to perpetuate low skill and low earnings. Even worse, those low skill households that would have emancipated themselves in the absence of the transfer will now be induced to passivity. Simple transfers can only mitigate current material misery by perpetuating the passivity of the poor and thus reduce the utility of the rich financing the transfer.

Activating Pareto-Improving Welfare and negative marginal income tax Since education cannot be bought by money,[2] unconditioned transfers raise the utility of the unskilled, by generating sufficient consumption in a more comfortable way, without raising their education. They raise the passivity of the poor and thus do not provide any incentive to the rich to finance the transfer. Starting from a polarized economy, a welfare system able to bring about a Pareto-improvement has to guarantee

1. **Activation**: The welfare system has to provide the incentive for the unskilled to exert enough effort such as to raise their human capital above the threshold.

2. **Voluntary participation by the unskilled**: Activation could be forced by a punishment for *not* exerting effort, but this would reduce the utility of the activated household. Thus, activation has to be sweetened by a transfer, which should be high enough to compensate the beneficiaries for their additional effort. This necessitates a transfer financed by the rest of the society.

3. **Voluntary participation by the contributor**: While the transfer should be high enough to activate the unskilled, it should be low enough to be worth financing by the contributor. In addition the contributors have to be compensated by a sufficient increase of the public good (or reduction of their own tax bill) in the sufficiently near future.

Unconditional transfers obviously satisfy (2) but fail to activate and therefore also fail to satisfy (3). Only a conditional transfer can at the same time satisfy satisfying (1) to (3). To activate, a transfer has to be conditioned either on sufficient effort or on the fruits of effort – human capital. At the same time an activating transfer that raises the utility of the poor has to be sufficiently large to compensate for the disutility of the additional effort. The present paper considers tax-transfer systems that condition current transfers on current income alone. Can such transfers systems Pareto-improve a

[2] This also is the reason why the introduction of perfect credit-markets does not eliminate polarization.

polarized equilibrium? It will be shown that this is generally possible. Furthermore the corresponding tax-system is shown to involve negative marginal income tax.

Literature

This paper brings together three issues and argues that they are most naturally related: Human capital in the instantaneous utility function, the possibility of threshold dynamics of individual behavior caused by the additional interaction between state and control variables, and the activating welfare the threshold dynamics calls for. None of these issues has received excessive attention in the literature, but none has been completely neglected.

Human capital in utility Almost 40 years ago, in view of the then still young literature on human capital Robert Michael [1973] wrote:

> "Nearly all this research deals with the effects of the investment through the labor market. ... While the increase in market earnings is surely one aspect of the yield of investments in human capital – that is, investments in formal schooling, on-the-job training, good health, information about markets, etc.– there is no reason to suppose it is the only yield on these investments. A distinguishing characteristic of human capital is that it is embedded in an individual. It therefore accompanies him wherever he goes, not only into the labor market, but also into the theater, the voting booth, the kitchen, and so forth. So if human capital yields a flow of services through the time in the labor market, it may yield a flow of services through the time spent in other in other activities as well. Certain example seem obvious: some education yields productive services through the time spent reading books or balancing a checkbook, some improvements in health yield productive services through time spent in participating in sports, and so forth. If general forms of human capital yield such services through time spent outside the labor market, these services should also be considered 'return' on the investment in human capital."

The analysis of Michael [1973] was formalized within the Lancasterian framework, building on the concept of household production function as developed in Becker's article on the allocation of time (Becker [1965]). Apart from its standard market return, human capital also raises the productivity of household production of nonmarket goods. Michael [1973] already writes that "one could alternatively argue that education directly alters the household's utility function", as is assumed in the present paper. Becker integrates the analysis of Michael [1973] into the second edition of his book on human capital and notes in the introduction:

> "Although important studies of the effects of human capital in the market sector can be expected, I anticipate that the excitement will be generated by studies of its effects in the nonmarket sector" (Becker [1993], page 10).

As far as the economic profession is concerned, this anticipation has not yet come true, although Becker is not the only Nobel Laureate emphasizing the relevance of human capital as a source of nonmarket return. Amartya Sen [1997], summarizes why he too deems necessary a widening of the concept of human capital:

> "The use of the concept of human capital, which concentrates only on one part of the picture, is certainly an enriching move, but it needs supplementation. This is because human beings are not merely means of production, but also the end of the exercise." ..."Given her personal characteristics, social background, economic circumstances, etc., a person has the ability to do (or be) certain things that she has reason to value. The reason for valuation can be *direct* (the functioning involved may directly enrich her life, such as being well-nourished or being healthy), or *indirect* (the functioning involved may contribute to further production, or command a price in the market). ... Consider an example. If education makes a person more efficient in commodity production, then this is clearly an enhancement of human capital. This can add to the value of production in the economy and also to the income if the person who has been educated. But even with the same level of income, a person may benefit from education, in reading, communication, arguing, in being able to choose in a more informed way, in being taking more seriously by others, and so on. The benefits of education, thus, exceeds its role as human capital in commodity production."

Sen [1997, 1999] distinguishes between "human capital" and "human capabilities".[3] He uses the former term to denote the narrow conventional perspective concentrating on human qualities that can be employed as "capital" in production in the same way as physical capital. And he reserves the latter term for the broader perspective also including "the ability of human beings to lead lives they have reason to value and to enhance the substantive choices they have." In contrast I use term "human capital" in this broader sense. This is essentially a matter of terminology as also Sen remarks: "The two perspectives cannot but be related since both are concerned with the actual abilities that they achieve and acquire" ... and ..."the human capital perspective can – in principle – be defined very broadly to cover both types of valuation". Precisely because the same set of achieved and acquired

[3] While Sen's work is mainly concerned with developing countries, the present paper rather deals with inequality and missed opportunities in rich countries.

abilities can raise the level of income *and* be a direct source of happiness, the present paper summarizes these abilities in a single state variable "human capital". Sen's "means and ends" receives a simple non-self-refering interpretation: human capital simply is state variable, that is a means to raise income but also end in itself.

If Sen's concept of human capabilities has so far had little impact on the economic literature on human capital, the same can not be said for its general impact. In particular Sen's ideas have had considerable influence on the formulation of the Human Development Report, published regularly by the United Nations Development Programme.

Threshold in individual optimization To my knowledge the possibility of threshold dynamics in concave individual optimization has so far not been studied in models of human capital accumulation. A striking feature of the present model is that very little is needed to generate the history dependency of optimal individual behavior (apart from inserting human-capital into the instantaneous utility function). In particular the path-dependence also arises in a concave setting. The possibility that path-dependence can occur in concave optimization problems has been largely ignored in much of the literature which generally relates path-dependence to increasing returns to scale or to coordination failures. Nevertheless, the fact that history dependent behavior may occur in individual concave optimization problems satisfying for instance the Arrow-Kurz conditions is known at least since Kurz [1968]. Wirl and Feichtinger [2005] provide a comprehensive analysis of when multiple steady states and path-dependence arise due to control-state interactions in concave intertemporal optimization problems. The optimization problem of the present paper is very similar the one studied in Hartl et al. [2004], who study an intertemporal optimization problem with path-dependent solutions both in the concave as well as in non-concave domain of the Hamiltonian. Kurz [1968] (with the state variable physical capital or wealth in utility), Chen [2007] (with a habit stock or accumulated past consumption in utility), or the present model (with human capital in utility) are but some examples fitting in the framework of Wirl and Feichtinger [2005] or Hartl et al. [2004].

Activating welfare and negative marginal income tax Many countries have recently adopted negative marginal tax rates into their tax-transfer systems, as for instance the United States (earned income tax credit), the United Kingdom (working tax credit), Canada (working income tax benefit), Germany (combined wages), Ireland, New Zealand, Austria, Belgium, Denmark, Finland, France or the Netherlands. Some (notably Scandinavian) countries already have a tradition a activating welfare or "workfare". Surprisingly few theoretical models provide a foundation for the general sympathy for the concept or activating welfare among both applied economists, politicians and the general public. In the framework of Mirrlees' seminal article (Mirrlees [1971]) and most subsequent literature on optimal taxation, the marginal income tax rates are always non-negative. An exception is Diamond [1980], who shows that if, instead of considering the intensive margin of labor choices, one considers the extensive margin (an individual worker faces the binary choice of working or not), negative marginal income tax rates can be optimal (see also Saez [2004], Choné and Laroque [2005], Laroque [2005]). Beaudry et al. [2007] deviate from the Mirrlees framework by assuming that the government is uninformed both about households' value of time in economic and non-economic activities and by allowing the government to condition not only on total incomes but also on individual wage rates. Beaudry et al. show that an optimal tax-transfer includes a negative marginal tax rate for workers with a wage below a certain cutoff rate.

The literature on optimal taxation typically defines optimality in terms of a utilitarian social welfare function.[4] Justifying a transfer-tax system requires a redistributional argument. In contrast, in the present paper a tax-transfer system can Pareto-improve the market outcome. The aim of taxation authorities is to bring about a Pareto-improvement rather then merely redistribute. This difference is not as substantial as it might appear however. The market outcome of the present paper can be Pareto-inefficient only due to external effects. As has been explained, without such external effects, the negative marginal tax would need a redistributional argument here too: Starting from a polarized equilibrium, optimality with respect to a utilitarian social welfare requires redistribution. Simple transfers would generate higher social welfare than no transfers, but simple transfers would be Pareto-improved by contingent transfers involving a negative marginal income tax for much the same reasons as in the present setting.

The more substantial difference between the present paper and the literature on optimal taxation is that there households differ by unobservable but fixed characteristics (their types), while here household only differ in their *initial* human capital. The initial human capital is but a state variable that can be changed over time. The possibility to raise an individual household's human capital above the threshold, to activate him *once and for all* is at the very heart of the present approach and is essential aspect of the policy debate as well.

The remaining of this article is organized as follows: Section 2 introduces the individual optimization problem. Section 3 establishes the conditions under which optimal individual behavior is history-dependent. The solution does not depend on the presence of a public good. In the presence of a public good the uncoordinated market solution will of course be Pareto-inefficient. Section 4 determines the Pareto-optimal solution path in an economy assuming that all households start with the same level of human capital and constructs a first best income taxation that implements the op-

[4] An exception is Moffitt [2006,], who assumes that the government directly cares about the effort of the poor and also show that a negative marginal tax rate can be optimal.

timal solution. Section 5 gives up the assumption that all households are initially endowed with the same level of human capital and studies the effect of unconditional transfers. Section 6 shows that a contingent transfer system can Pareto-improve a fiscal system that finances the public good but does not redistribute. The Pareto-improvement requires the activation of the initially poor households and involves a negative marginal income tax. Section 7 concludes.

2 The individual household problem

Consider the intertemporal household optimization problem

$$\begin{aligned} & \max_{\{x_t\}_{t=0}} \int_0^\infty e^{-\rho t} u(c_t, x_t, h_t, G_t) dt \qquad (1) \\ \text{subject to } \dot{h}_t &= \underset{+}{g(x_t}, h_t) = (\gamma x - \delta h)h \text{ for all } t \geq 0, \\ c_t &= w h_t \text{ for all } t \geq 0 \\ & \text{given } h_0 \text{ and } \{G_t\}_{t=0}^\infty . \end{aligned}$$

where c_t (consumption) and x_t (effort) are control variables and h_t is a state variable (human capital, education, skill, fitness, capability or knowledge). The amount G_t of the public good is exogenously given in the present section and will not affect the solution. Household income wh_t depends on the household's skill and the general (net)productivity level w_t. In the absence of credit markets assumed here, the household simply consumes his current income ($c_t = wh_t$).[5]

The interaction between effort and human capital In this formulation the interaction between x and h is captured by the term $\gamma x h$ in $\dot{h}_t = g(x_t, h_t) = (\gamma x - \delta h)h$: Effort is the more effective, the more one already knows. In the present framework a necessary condition for the possibility of history dependent optimal individual behavior is that, over some domain of skill levels, the incentive to exert effort rises with rising skill. The assumption that $g_{xh} = \gamma > 0$ brings us halfway toward this requirement: The effectiveness of effort in generating future human capital rises with rising h.

Formally equivalent one could assume that effort is the less unpleasant, the more one knows: $u(c, \frac{\text{effort}}{h}, h, G)$ and $\dot{h} = \gamma \cdot \text{effort} - \delta h^2$. If we define $x := \frac{\text{effort}}{h}$ as relative effort, this leads to the maximization problem (1).

Also note that the accumulation rule $\dot{h} = (\gamma x - \delta h)h$ features stronger depreciation than the more standard form $\dot{h} = (\gamma x - \delta)h$. This is assumed to exclude the possibility of unbounded growth.[6]

[5] In a companion paper I study the model in model in the presence of a perfect credit market ..... .

[6] The threshold dynamics also occurs with the more standard accumulation rule. The difference to the present formulation would be that the strictly positive attractor of h_t in case of a household starting above the threshold would be unbounded, while it is bounded here.

The Maslow condition The assumption that $g_{xh} = \gamma > 0$ guarantees that skill accumulation becomes easier with rising skill. To complete the necessary condition for history dependence it has to be made sure, that the rising effectiveness of effort in generating skill (with rising skill) is not more than compensated by a decreasing marginal utility of skill. The total (direct and indirect) marginal utility of h should not decrease too fast with rising h. Formally, this requires that the negative of the elasticity $\frac{du}{dh}(wh, h, x, G)$ with respect to h is not larger than the elasticity of g_x with respect to h. Given $g_{xh} = \gamma > 0$, this requirement will follow from the "Maslow condition": Maslow's observation that deficiency needs must be met before the household is ready to act upon his growth needs is captured by the assumption that (1) the marginal utility of consumption tends to infinity when consumption tends to zero ($\lim_{\tau\to\infty} \frac{\partial u(c^\tau, x, h^\tau, G)}{\partial c} = \infty$ for any sequence $(c^\tau, h^\tau)_\tau$ with $c^\tau \to 0$, independently of $\lim h^\tau$) and *falls fast* with increasing consumption (the negative of the elasticity of u_c with respect to c is large: $-u_{cc}\frac{c}{u_c} \geq 1$), and (2) that the marginal (direct) utility of human capital $\frac{\partial u(c,x,h,G)}{\partial h}$ remains bounded when human capital tends to zero ($\lim_{\tau\to\infty} \frac{\partial u(c^\tau, x, h^\tau, G)}{\partial h} < \infty$ for any sequence $(c^\tau, h^\tau)_\tau$ with $h^\tau \to 0$, independently of $\lim c^\tau$) and *does not fall* when human capital rises ($\frac{\partial^2 u}{\partial h^2} = 0$). Furthermore I assume that $\lim_{\tau\to\infty} \frac{\partial u(c^\tau, x^\tau, h^\tau, G)}{\partial x} = -\infty$ for any sequence $(c^\tau, x^\tau, h^\tau)_\tau$ with $x^\tau \to \infty$ (independent of $\lim c^\tau$). Thus standard Inada-conditions hold for $c \to 0$ and $x \to \infty$ but not for $h \to 0$.

Note that the necessary effort to raise human capital by an infinitesimal unit $\dot{h} = (\gamma x - \delta h)h$ tends to infinity when h tends to zero. Thus only a very strong incentive will induce a household with low h to nevertheless raise h. This can only be the high marginal utility of c (the economic motive). The (bounded) marginal utility of h plays no role for such a household.

The public good The amount G_t cannot be influenced by the individual household and will in later sections provide a natural way to incorporate a common interest of all households in the existence of a greatest possible numbers of skilled households that are able to help finance G. To otherwise keep the influence of G as simple as possible I assume that it enters instantaneous utility in an additively separable way. In the laissez-faire economy I first consider, the presence of G then has no effect on individual behavior or equilibrium.

A simple example that captures all of the above features and that will be assumed in the remainder of the paper is

$$u(c_t, x_t, h_t, G_t) = m(c) + bh - v(x) + \xi(G),$$

with $m(c) = \kappa \ln c$,[7] $v(x) = \left(ax + \frac{\alpha}{2}x^2\right)$, $\kappa, a \geq 0$ and $\alpha > 0$. In this example, κ measures the intensity of the economic motive (for exerting effort), while b measures the intensity of the non-economic

[7] The intertemporal elasticity of subsitution is one. The consumption part of the Maslow condition thus is just met. The threshold dynamics also occurs if the elasticity of substitution is smaller than one. The lower stationary human capital level, which is zero here, would be strictly positive instead.

motive. The assumption about the hierarchy of needs is satisfied since $\lim_{c\to 0} u_c(c,x,h,G) = \infty$ and $\lim_{h\to 0} u_h(c,x,h,G) = b$.

3 History-dependent optimal individual behavior

The Hamiltonian of (1) is $H(x,h,G,\lambda) = m(wh)+bh-v(x)+\xi(G)+\lambda\cdot(\gamma x-\delta h)h$. Existence of an inner solution to $\max_x H(x,h,G,\lambda)$ requires $H_x = -v_x + \lambda\gamma h = 0$, which is sufficient since $H_{xx} = \alpha < 0$. At the maximum $v_x = \lambda\gamma h$. Taking the derivative with respect to time yields $v_{xx}\dot{x} = \dot{\lambda}\gamma h + \lambda\gamma\dot{h} = \dot{\lambda}\gamma h + v_x(\gamma x - \delta h)$. Inserting the adjoint equation $\dot{\lambda} = \rho\lambda - H_h = \rho\lambda - m_h - b - \lambda\cdot(\gamma x - 2\delta h)$ into this expression for $v_{xx}\dot{x}$ and using $v_x = \lambda\gamma h$ yields $v_{xx}\dot{x} = v_x(\rho+\delta h) - (m_h + b)\gamma h$. Together with the accumulation rule for human capital this defines the dynamic system

$$\begin{cases} \dot{x} = \frac{v_x(\rho+\delta h)-(m_h+b)\gamma h}{v_{xx}} \\ \dot{h} = (\gamma x - \delta h)h \end{cases}, \tag{2}$$

or with $m(c) = \kappa \ln c$ and $v(x) = ax + \frac{\alpha}{2}x^2$

$$\begin{cases} \dot{x} = \frac{(a+\alpha x)(\rho+\delta h)-(m_h+b)\gamma h}{\alpha} \\ \dot{h} = (\gamma x - \delta h)h \end{cases} \tag{3}$$

The two isoclines are

$$\begin{cases} \dot{x} = 0 \text{ if } x_{\dot{x}=0}(h) := \frac{(m_h+b)\gamma h}{\alpha(\rho+\delta h)} - \frac{a}{\alpha} \\ \dot{h} = 0 \text{ if } h = 0 \text{ or } x_{\dot{h}=0}(h) := \frac{\delta}{\gamma}h \end{cases} \tag{4}$$

and with $m(c) = \kappa \ln c$

$$\begin{cases} \dot{x} = 0 \text{ if } x_{\dot{x}=0}(h) := \frac{\gamma}{\alpha}\frac{\kappa+bh}{\rho+\delta h} - \frac{a}{\alpha} \\ \dot{h} = 0 \text{ if } h = 0 \text{ or } x_{\dot{h}=0}(h) := \frac{\delta}{\gamma}h \end{cases} \tag{5}$$

Stationary solutions

The two isoclines always have an intersection at $h_* = 0$ defining the trivial stationary state of (3). The other stationary states are determined by the solutions to $\frac{\gamma}{\alpha}\frac{\kappa+bh}{\rho+\delta h} - \frac{a}{\alpha} = \frac{\delta}{\gamma}h$ or, rearranging, to

$$\frac{\delta^2}{\gamma}h^2 - \left(\frac{b\gamma - a\delta}{\alpha} - \frac{\delta}{\gamma}\rho\right)h - \frac{\kappa\gamma - \rho a}{\alpha} = 0.$$

With $C := \frac{\kappa\gamma-\rho a}{\alpha}$, $B := \left(\frac{b\gamma-a\delta}{\alpha} - \frac{\delta}{\gamma}\rho\right)$, and $D = \frac{\delta^2}{\gamma} > 0$ the (non-zero) stationary states of (3) are the solutions to

$$E(h) := Dh^2 - Bh - C = 0 \tag{6}$$

Equation (6) may have no, one or two strictly positive solutions

$$h_{1,2} = \frac{B \pm \sqrt{B^2 + 4DC}}{2D}.$$

Case 1 One strictly positive solution $h^* = \frac{B+\sqrt{B^2+4DC}}{2D} > 0$ if $C > 0$.

Case 2 Two strictly positive solutions $h^{\text{th}} = \frac{B-\sqrt{B^2+4DC}}{2D} > 0$ and $h^* = \frac{B+\sqrt{B^2+4DC}}{2D} > h^{\text{th}}$ if $-\frac{B^2}{4D} < C < 0$ and $B > 0$.

Case 3 No strictly positive solution.

Case 3a No real solution if $C < -\frac{B^2}{4D} < 0$.

Case 3b Two negative solutions if $-\frac{B^2}{4D} < C < 0$ and $B < 0$.

These cases and more generally most arguments in the present paper can best be understood with the help of the (x, h)-phase diagram (see Figures 1 to 5). The $\dot{h} = 0$-isocline consists of the vertical x-axis ($h = 0$) and the increasing straight line $x_{\dot{h}=0}(h) = \frac{\delta}{\gamma}h$. A higher level of activity is needed to keep constant a higher level of human capital. This $\dot{h} = 0$-isocline is unchanged throughout the paper. The position of the $\dot{x} = 0$-isocline depends on the parameters (changing form case to case) and of the economic environment (changing from section to section). The two crucial features are the intercept with the x-axis, $x_{\dot{x}=0}(0) = \frac{\gamma}{\alpha}\frac{\kappa}{\rho} - \frac{a}{\alpha}$, and the slope at $h = 0$, $\frac{d}{dh}\left(x_{\dot{x}=0}(0)\right)$. In all cases $x_{\dot{x}=0}(h)$ has an intersection with the $\dot{h} = 0$-isocline at the trivial steady state solution $h_* = 0$. Concerning the strictly positive steady state solutions, the three above cases arise graphically:

Case 1 If $x_{\dot{x}=0}(0) = \frac{\gamma}{\alpha}\frac{\kappa}{\rho} - \frac{a}{\alpha} > x_{\dot{h}=0}(0) = 0$ as in Figures 1a and 1b then the two isoclines have exactly one intersection at a strictly positive h, say $h^* > 0$. Note that for this case, at least qualitatively, it is irrelevant whether $x_{\dot{x}=0}(h)$ is increasing (as in Figure 1a) or decreasing (as in Figure 1b).

This case occurs when $C > 0$ or equivalently $\kappa > a\frac{\rho}{\gamma}$, i.e. whenever the economic motive for effort is strong (κ large), the household is patient (ρ small), it is very easy to raise h (γ large) or the simple cost of effort is low (a small). As a short-cut, I will say in this case that the "economic motive for effort is strong".

Case 2 If $x_{\dot{x}=0}(0) < x_{\dot{h}=0}(0)$ and the $\dot{x} = 0$-isocline is rising sufficiently fast compared with the $\dot{h} = 0$-isocline as in Figure 2, then there are two strictly positive intersections, $h^{\text{th}} > 0$ and $h^* > h^{\text{th}}$. This case requires that the economic motive is not very strong ($C < 0$ or $\kappa < a\frac{\rho}{\gamma}$) and at the same time the non-economic motive b is sufficiently strong $b > \left(\frac{\alpha}{\gamma}\rho + a\right)\frac{\delta}{\gamma}$ ($B > 0$). Note that by the Maslow condition, for *small* levels of human capital, income, and consumption, the non-economic motive for effort is always dominated both by the economic motive as well as the cost of effort. Sufficiently "Strong non-economic motive" means that the constant marginal utility of human capital is high such that the non-economic motive rises sufficiently fast with rising human capital and income.

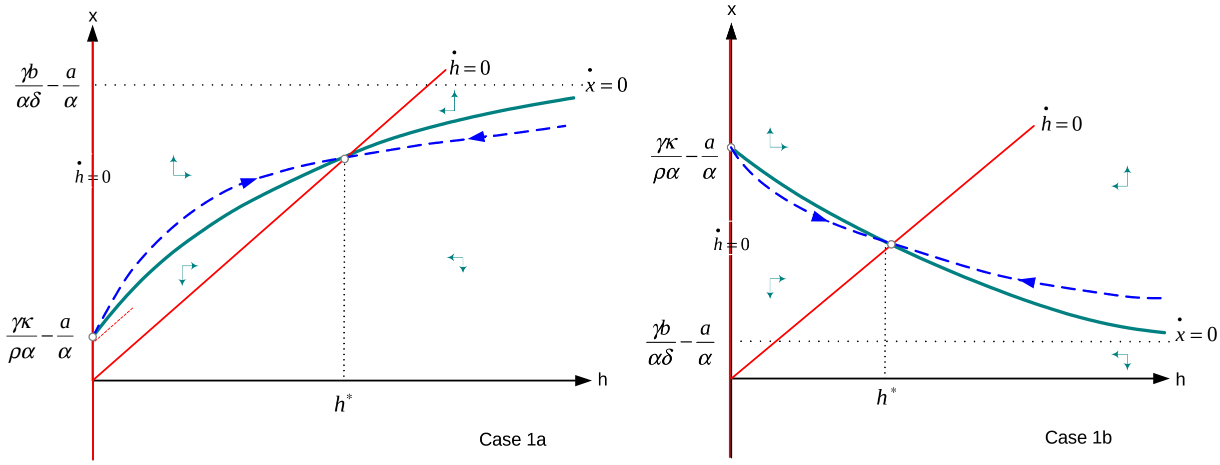


Figure 1: Case 1

Case 3 If $x_{\dot{x}=0}(0) < x_{\dot{h}=0}(0)$ and the $\dot{x} = 0$-isocline is not rising fast with h as in Figure 3, then there are no positive intersections.

If at $h = 0$ the slope of $\dot{x} = 0$-isocline is larger than the slope or the $\dot{h} = 0$-isocline (formally if $\frac{d}{dh}x_{\dot{x}=0}(0) = \frac{\gamma}{\alpha}\frac{(\rho b - \kappa\delta)}{\rho^2} > \frac{\delta}{\gamma} = \frac{d}{dh}x_{\dot{h}=0}(0)$ or if $b > \frac{\delta\rho\alpha}{\gamma^2} + \kappa\frac{\delta}{\rho}$), then a variation of the (simple) cost of effort a can generate all 3 cases. This is again best seen by considering the diagrams: In Figure 1a $\frac{d}{dh}x_{\dot{x}=0}(0) > \frac{d}{dh}x_{\dot{h}=0}(0)$. Starting from Figure 1a, raising a shifts down the function $x_{\dot{x}=0}(h)$, first generates Figure 2 and then, further raising a, Figure 3. More precisely, Case 1a occurs if $a \leq \gamma\kappa/\rho$, Case 2 occurs if $a > \gamma\kappa/\rho$, but not larger than the solution to $b = \frac{2\delta\sqrt{\alpha\gamma(a\rho - \gamma\kappa)} + \alpha\rho\delta + a\delta\gamma}{\gamma^2}$. Case 3a occurs if a is larger than this solution. Thus:

Lemma 1 *If the economic motive for effort is not dominated everywhere (for each h) by the non-economic motive, formally: if $b > \kappa\frac{\delta}{\rho} + \frac{\delta\rho\alpha}{\gamma^2}$ (at $h = 0$, $x_{\dot{x}=0}(h)$ is increasing with a slope higher than $x_{\dot{h}=0}(h)$), then a variation of the (simple) cost of effort, a, from sufficiently low to sufficiently large, generates all three cases: Case 1 for small a (see Figure 1a), Case 2 for intermediate a (see Figure 2), Case 3 for large a (see Figure 3). Case 2 never occurs if the non-economic motive for effort too small; formally: If $b < \frac{\delta}{\gamma}\left(\frac{\alpha}{\gamma}\rho + a\right)$ $(B < 0)$ then Case 2 occurs for no κ, if $b < \delta\left(\frac{\kappa}{\rho} + \frac{\rho\alpha}{\gamma^2}\right)$ then Case 2 occurs for no a.*

It is obvious from the figure that if at $h = 0$ the slope of $\dot{x} = 0$-isocline is smaller than the slope or the $\dot{h} = 0$-isocline ($\frac{d}{dh}(x_{\dot{x}=0}(0)) < \frac{d}{dh}\left(x_{\dot{h}=0}(0)\right)$ or $b < \frac{\kappa\delta}{\rho} + \frac{\delta\alpha\rho}{\gamma^2}$), then the threshold case (Case 2) occurs for no level of effort costs a.

Appendix 8.1 provides a complete characterization of all cases in the (a, b)-plane given the other parameters (See Figure 8.1). This also generates a complete proof of Lemma 1.

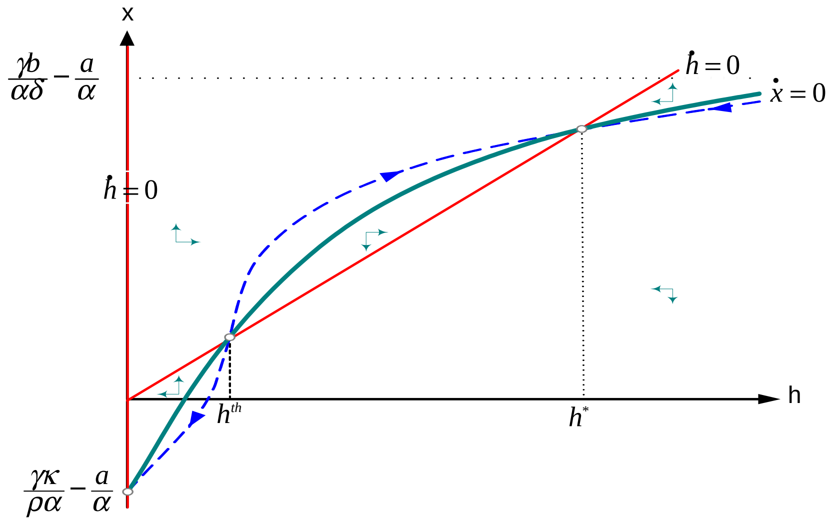


Figure 2: Case 2

Non-stationary solutions

The above considerations deal with the stationary solutions of the dynamic system (3) and to the household problem (1). Appendix 8.2 shows that if α is sufficiently large, then in all three cases and for any initial h_0 the household problem has a unique solution $\{x(h_t), h_t\}_{t\geq 0}$, where $x(h)$ is a continuous policy function. The following three propositions characterize the corresponding optimal path $\{h_t\}_{t\geq 0}$ of the state variable:

Case 1, Strong economic motive, high efficiency of learning, low discounting: Activation.

Proposition 1 *If the economic motive is sufficiently strong (Case 1), then h^* is a global attractor. Formally: If $\kappa > a\rho/\gamma$, then $\lim_{t\to\infty} h_t = h^* > 0$ for all $h_0 > 0$. See Figure 1.*

Proof. Appendix 8.2. ■

Case 2, Weak economic motive, strong non-economic motive: path-dependence Remember the Maslow condition: At small h (and hence consumption) both the marginal utility of consumption as well as the marginal disutility of effort are *always* very large compared to the marginal utility of h. Strong economic motive means high marginal utility at small c compared to the disutility of effort at small x (κ/a large), while strong non-economic motive means that the constant marginal utility of human capital is high. The economic motive for effort alone is not large compared to effort cost. The non-economic motive is not dominated everywhere by cost of effort (it is of course dominated for small h by the Maslow condition).

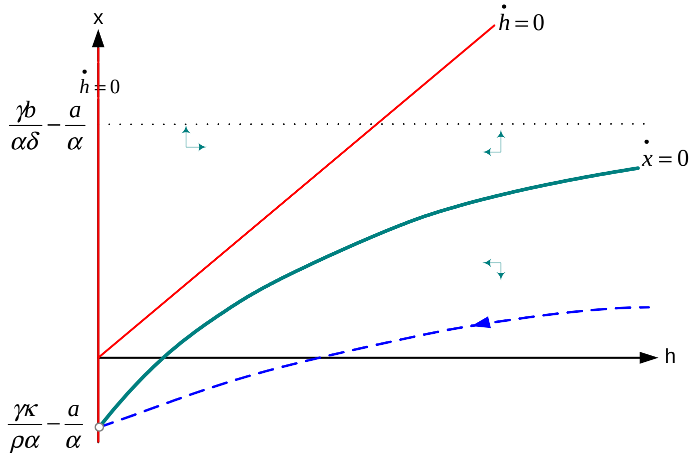


Figure 3: Case 3

Proposition 2 *If the economic motive is not dominant and the non-economic motive is sufficiently strong (Case 2), then the individually optimal policy is path-dependent. Formally: If* $\kappa < \min\{a\rho/\gamma, b\rho/\delta\}$ *and* $b > \frac{2\delta\sqrt{\alpha\gamma(a\rho-\gamma\kappa)}+\alpha\rho\delta+a\delta\gamma}{\gamma^2}$*, then* $\lim_{t\to\infty} h_t = h_* = 0$ *for all* $h_0 < h^{th}$ *and* $\lim_{t\to\infty} h_t = h^*$ *for all* $h_0 > h^{th}$*, where* $x(h)$ *is a continuous policy function. See Figure 2.*

Proof. Appendix 8.2. ■

The initially unskilled (and poor) completely loose their skill and income, while the initially sufficiently skilled converge to a high level h^* of human capital and income wh^*.

Case 3, Weak economic and non-economic motives: Deterioration to passivity The economic and the non-economic motives together are dominated by the cost of effort everywhere (even for large h).

Proposition 3 *If both the economic and the non-economic motives are weak, then no household will exert sufficient effort to raise its human capital. Formally: If* $b\rho/\delta < \kappa < a\rho/\gamma$ *or [*$\kappa < \min\{a\rho/\gamma, b\rho/\delta\}$ *and* $b < \frac{2\delta\sqrt{\alpha\gamma(a\rho-\gamma\kappa)}+\alpha\rho\delta+a\delta\gamma}{\gamma^2}$*], then* $h_* = 0$ *is the only non-negative real stationary solution and* $\lim_{t\to\infty} h_t = h_* = 0$ *for all* h_0*. See Figure 3.*

Proof. Appendix 8.2. ■

Proposition 4 *If the non-economic motive for effort is not dominated everywhere by the economic motive (*$b > \frac{\kappa\delta}{\rho} + \frac{\delta\alpha\rho}{\gamma^2}$*: at* $h = 0$*,* $x_{\dot{x}=0}(h)$ *is increasing with a slope higher than* $x_{\dot{h}=0}(h)$*), then a variation of the simple cost of effort a from a sufficiently low to a sufficiently large value generates all three cases: For small a all households converge to the same strictly positive level of human capital; for* *intermediate a the threshold case prevails; for large a all households choose a path towards increasing passivity.*

History dependent optimal individual behavior is excluded if the non-economic motive for effort is too weak. Formally: If $b < \frac{\delta}{\gamma}\left(\frac{\alpha}{\gamma}\rho + a\right)$ *(*$B < 0$*) then history dependence occurs for no* κ*, if* $b < \delta\left(\frac{\kappa}{\rho} + \frac{\rho\alpha}{\gamma^2}\right)$ *then history dependence occurs for no* a*.*

Proof. *Corollary to Lemma 1 and the three previous Propositions.* ■

The necessary conditions for history dependent optimal behavior

The second part of Proposition 4 shows that within the class of felicity and accumulation functions ($u(c,x,h) = m(c) + bh - v(x)$, $m(c) = \kappa \ln c$, and $g(x,h) = (\gamma x - \delta h)h)$), history dependence only occurs due to the presence of a sufficiently strong direct marginal utility of human capital (large b): The relevant term in Equation 3, $\gamma h\,(m_h + b) = \gamma h\left(\frac{\kappa}{h} + b\right) = \gamma\kappa + b\gamma h$, must be increasing in h. Concerning the felicity function (with $m(c) = \kappa \ln c$), this requires $b > 0$, the non economic motive has to be sufficiently strong. Concerning skill accumulation, this obviously requires that $\gamma = g_{xh} > 0$: The effectiveness of effort in raising h has to rise with the skill level.

If we consider the more general class of CIES functions $m(c) = \kappa\frac{\sigma}{\sigma-1}(c)^{\frac{\sigma-1}{\sigma}}$ for if $0 < \sigma \neq 1$ and $m(c) = \kappa \ln c$ for $\sigma = 1$, the term $\gamma h\,(m_h + b) = \gamma h\left(\kappa w^{\frac{\sigma-1}{\sigma}} h^{-\frac{1}{\sigma}} + b\right)$ is increasing in h iff $\frac{d}{dh}\gamma h\,(m_h + b) = \gamma\left[(1-\frac{1}{\sigma})\kappa w^{\frac{\sigma-1}{\sigma}} h^{-\frac{1}{\sigma}} + b\right] > 0$. Assuming $\gamma = g_{xh} > 0$, this requires $b > (\frac{1-\sigma}{\sigma})\kappa w^{\frac{\sigma-1}{\sigma}} h^{-\frac{1}{\sigma}}$ or

$$\begin{cases} \sigma < 1,\ b > 0 \text{ and } h > w^{\sigma-1}\left[(\frac{1-\sigma}{\sigma})\frac{\kappa}{b}\right]^{\sigma}, \\ \sigma = 1 \text{ and } b > 0 \\ \sigma > 1. \end{cases} \tag{7}$$

Thus if the consumption part of the Maslow condition is strengthened by raising the elasticity of m_c from $\frac{1}{\sigma} = 1$ to $\frac{1}{\sigma} > 1$, then a sufficiently large $b > 0$ remains necessary for the possibility of history dependent optimal behavior. In a companion paper, I show that for $\sigma < 1$ the threshold case occurs under similar conditions as in the present paper. The essential difference is that any household will under all circumstances avoid asymptotic starving (while for $\sigma = 1$ he does so only for large κ). This shifts the smaller saddle-point stable stationary state from $h_* = 0$ to a strictly positive value $h_* > 0$. On the other hand we see that by violating the consumption part of Maslow's condition (by assuming $\sigma > 1$), the direct felicity effect of human capital is no longer necessary for the possibility of the threshold case.

In the general case the necessary condition for history dependence the above condition $\frac{d}{dh}\gamma h\,(m_h + b) > 0$ is replaced by the condition $\frac{d}{dh}\left[g_x(x,h) \cdot \frac{du}{dh}(wh,h,x)\right] > 0$ or equivalently by

$$\varepsilon_{g_x,h} > -\varepsilon_{u_h,h}, \tag{8}$$

where $\varepsilon_{y,z}$ is the elasticity of y with respect to z. If $m(c)$ is of the CIES-class, Conditions (7) and (8) are identical.

This condition is very intuitive: $g_x(x,h) \cdot \frac{du}{dh}(wh,h,x)$ can be interpreted as the incentive to exert effort: Raising x raises $\dot{h} = g(x,h)$ which in turn raises future felicity (by both the indirect effect via income and consumption and the direct effect). Existence of a threshold requires that, over some domain of skill levels, the incentive to exert effort rises with rising skill, thus that $\frac{d}{dh}\left[g_x(x,h) \cdot \frac{du}{dh}(wh,h,x)\right] > 0$.

Also remember that the interaction between effort and skill, which in the present formulation arises through skill accumulation (effort is the more effective the more you know), can be shifted to the felicity function (effort is the less unpleasant, the more you know) without changing any of the results.

4 Symmetric optimum and first best taxation

4.1 The planner's problem in an egalitarian society

The solution of the individual household problem in the previous section does not depend on the path $\{G_t\}_{t\geq 0}$ of the public good provided. Of course this was based on the fact that the amount G_t of the public good was exogenously given for the individual household. Without any fiscal system organizing the funding of the public good, G_t will be zero for all $t \geq 0$. The market equilibrium simply consists of the collection of solutions to the individual household problems of the previous section.

Consider an economy in which all households are equally skilled at $t = 0$. If they manage to coordinate the provision of the public good (and assuming equal treatment of identical households), they agree about the optimal extent of the provision (which, depending on h_0, may be null). The corresponding cooperative household problem has to adjust the non-cooperative problem (1) by adding the additional control variable G_t (which was exogenous to the non-cooperative individual household) and by replacing the budget constraint $c_t = wh_t$ with $c_t = wh_t - G_t$:

$$\max_{\{x_t, G_t\}_{t=0}} \int_0^{\infty} e^{-\rho t}\left[m(c) + bh - v(x) + \xi(G)\right] dt \tag{9}$$

$$\begin{aligned} \text{subject to } \dot{h}_t &= (\gamma x - \delta h)h \text{ for all } t \geq 0, \\ c_t &= wh_t - G_t \text{ for all } t \geq 0, \\ &\text{given } h_0. \end{aligned}$$

Remember that the public good has been introduced to generate a common interest in the education of the greatest possible number of households. To fix attention, I assume that the public good satisfies "growth needs" (in Maslow's terminology), e.g. enters the utility function in the same way as does h:

$\xi(G) = \xi \cdot G$ (where ξ is a positive constant rather than a function).[8] Inserting the additional constraint $c_t = wh_t - G_t$, the Hamiltonian becomes $H(x, h, G, \lambda) = m(wh - G) + bh - v(x) + \xi G + \lambda \cdot (\gamma x - \delta h)h$. An inner solution to $\max_G H(x, h, G, \lambda)$ requires $-m_c + \xi = 0$, $m_c = \xi$, or, with $m(c) = \kappa \ln c$, $\frac{\kappa}{wh_t - G_t} = \xi$, $G_t = wh_t - \frac{\kappa}{\xi}$. This inner solution is only relevant for $G_t \geq 0$ or $h_t \geq \frac{\kappa}{w\xi}$. Thus

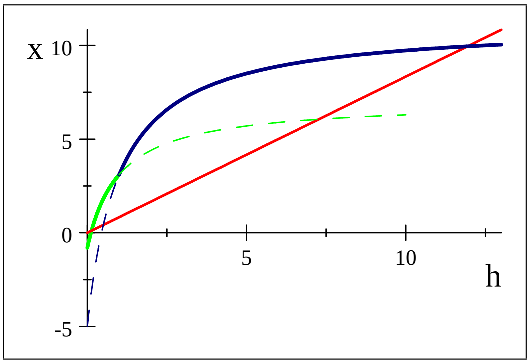
$$G_t = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} \\ wh_t - \frac{\kappa}{\xi} \text{ if } h_t \geq \frac{\kappa}{w\xi}. \end{cases} \tag{10}$$

The general form of the equations of motion (3) and of the isoclines (4) are unchanged. With (10) we now have

$$m_h = \frac{\kappa w}{wh_t - G_t} = \begin{cases} \frac{\kappa}{h} \text{ if } h_t \leq \frac{\kappa}{w\xi} \\ w\xi \text{ if } h_t \geq \frac{\kappa}{w\xi} \end{cases} \tag{11}$$

and thus

$$\begin{cases} \dot{x} = 0 \text{ if } \begin{cases} x = \frac{\gamma}{\alpha}\frac{\kappa + bh}{\rho + \delta h} - \frac{a}{\alpha} \text{ for } h_t \leq \frac{\kappa}{w\xi} \\ x = \frac{\gamma}{\alpha}\frac{(\xi w + b)h}{\rho + \delta h} - \frac{a}{\alpha} \text{ for } h_t \geq \frac{\kappa}{w\xi} \end{cases} \\ \dot{h} = 0 \text{ if } h = 0 \text{ or } x = \frac{\delta}{\gamma}h. \end{cases} \tag{12}$$



Planner's solution in Case 2

Unsurprisingly, since G satisfies "growth needs" rather that "deficiency needs", it is optimal not to provide the public good if the common initial human capital is low. More precisely, the threshold level of the cooperative solution path is the same as the solution to the non-cooperative solution ($x^{\text{th}}_{\text{opt}} = x^{\text{th}}$, $h^{\text{th}}_{\text{opt}} = h^{\text{th}}$). Only when households are sufficiently skilled to provide a minimum consumption, the public good becomes interesting and households prefer to spend a part of their income on the public good, provided they manage to coordinate its financing, as I have assumed in the present section: The positive stationary solution of the cooperative solution is larger then that of the non-cooperative solution ($x^*_{\text{opt}} > x^*, h^*_{\text{opt}} > h^*$). This also shows that if households with $h_t \leq \frac{\kappa}{w\xi}$ nevertheless accumulate h, they do it not for its own sake, but rather to be able to finance more of c.

[8] Alternatively one may consider a public good that satisfies deficiency needs (safety) as does c: $\xi(G) = \xi \cdot \ln G$.

4.2 Implementing the symmetric first best

This section shows that the symmetric first best can be implemented with a personal income tax $T(h_t^i, \overline{h}_t)$, that is, a tax which only depends on the current income wh_t^i (or equivalently on its human capital h_t^i) of household i and not on its effort x_t^i. To guarantee balanced fiscal budget, the tax must be allowed to depend on aggregate income $w\overline{h}_t = w\int_{j=0}^1 h_t^j dj$. Let $\overline{T}_t = \int_{j=0}^1 T(h_t^j, \overline{h}_t)dj = G_t$. The individual household problem

$$\max_{\{x_t\}_{t=0}} \int_0^\infty e^{-\rho t} u(c_t, x_t, h_t, G_t)dt \tag{13}$$

$$\begin{aligned} \text{subject to } \dot{h}_t &= (\gamma x_t - \delta h_t)h_t \text{ for all } t \geq 0, \\ c_t &= wh_t - T(h_t^i, \overline{h}_t) \text{ for all } t \geq 0, \\ & \text{given } h_0,\ \{\overline{h}_t\}_{t=0} \text{ and } \overline{T}_t = G_t \end{aligned}$$

Since initially $h_0^i = \overline{h}_0$ for all i and the tax only depends on h_t^i, at individual optimum $h_t^i = \overline{h}_t$ remains satisfied such that at intertemporal **equilibrium** $T(h_t^i, \overline{h}_t) = T(\overline{h}_t, \overline{h}_t) = \overline{T}_t$ must be satisfied for all t. Inserting the budget constraint, the Hamiltonian becomes $H(x, h, \lambda) = m[wh - T(h, \overline{h})] + bh - v(x) + \xi w\overline{h} + \lambda \cdot (\gamma x - \delta h)h$. The general form of the equations of motion (3) and of the isoclines (4) are again unchanged. The optimal dynamic system and its isoclines (12) are replicated iff the marginal utility of consumption $m_h = \frac{\kappa(w - T_h(h,\overline{h}))}{wh - T(h,\overline{h})}$ replicates that of the optimal solution (Equation (11), thus iff

$$m_h = \frac{\kappa(w - T_h(h,\overline{h}))}{wh - T(h,\overline{h})} = \begin{cases} \frac{\kappa}{h} \text{ if } h_t \leq \frac{\kappa}{w\xi} \\ \xi w \text{ if } h_t \geq \frac{\kappa}{w\xi} \end{cases} \tag{14}$$

This is the case for $T(h, \overline{h}) = 0$ if $h_t \leq \frac{\kappa}{w\xi}$ and $T_h = \frac{w\xi}{\kappa}[\frac{\kappa}{\xi} - wh + T(h, \overline{h})]$ if $h_t \geq \frac{\kappa}{w\xi}$. Solving this differential equation we get

$$T(h, \overline{h}) = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} \\ wh + I_t \cdot e^{\frac{w\xi}{\kappa}h} \text{ if } h_t \geq \frac{\kappa}{w\xi} \end{cases} \tag{15}$$

where I_t is a constant of integration. Note that while I_t is constant as a function of h_t^i, it will not be constant as a function of time. I_t is determined by the equilibrium condition $T(h_t^i, \overline{h}_t) = T(\overline{h}_t, \overline{h}_t)$ $= G_t = w\overline{h}_t - \frac{\kappa}{\xi}$: $w\overline{h}_t + I_t \cdot e^{\frac{w\xi}{\kappa}\overline{h}_t} = w\overline{h}_t - \frac{\kappa}{\xi}$, hence $I_t = -\frac{\kappa}{\xi}e^{-\frac{w\xi}{\kappa}\overline{h}_t}$ and

$$T(h, \overline{h}) = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} \\ wh - \frac{\kappa}{\xi}e^{\frac{w\xi}{\kappa}(h - \overline{h})} \text{ if } h_t \geq \frac{\kappa}{w\xi}. \end{cases} \tag{16}$$

For $h_t \geq \frac{\kappa}{w\xi}$ the marginal tax is

$$T_h(h, \overline{h}) = w\left[1 - e^{\frac{w\xi}{\kappa}(h - \overline{h})}\right] \begin{cases} > 0 \text{ for } \frac{\kappa}{w\xi} \leq h < \overline{h} \\ = 0 \text{ for } h = \overline{h} \\ < 0 \text{ for } h > \overline{h}. \end{cases}$$

Thus the tax reduces the incentive to exert effort for a household below average skill. This tax is not activating.

Note that the tax $T(h,\overline{h}) = wh + I_t \cdot e^{\frac{w\xi}{\kappa}h}$ induces the Pareto-optimal individual behavior for any arbitrary I_t (which is independent of the individual h): $m(c) = \kappa \ln\left(-I_t \cdot e^{\frac{w\xi}{\kappa}h}\right) = \kappa \ln(-I_t) + w\xi h$ and therefore $m_h = w\xi$, so that the equation of motion of λ is that of the symmetric first best. For the individual household, the optimal symmetric path of x, h is implemented independently from his expectations about present or future $\overline{h}_t$, since this only affects G_t which enters utility additively. Determining I_t is only relevant for guaranteeing the financing of the optimal G_t at $h_t^i = \overline{h}_t$.

5 Polarization and simple transfers

5.1 Polarization

Section 4 has assumed that all households start with the same initial human capital at $t = 0$. In contrast consider now an economy in which n_* households start with human capital below the individual threshold ($h_0^i < h^{\text{th}}$ for $i \in [0, n_*]$) and $n^* = 1 - n_*$ households start above the threshold ($h_0^i > h^{\text{th}}$ for $i \in [n_*, 1]$). After a while the n^* households initially above the threshold will cluster in the neighborhood of the stable attracting steady state h^* while the others will have lost much of their initial human capital. This trend towards increased polarization does not depend on whether the first best symmetric tax system was in effect or not, as long as the threshold case (Case 2) holds.

5.2 Simple transfers accentuate polarization

We will see in the next section, that polarization may be Pareto-inefficient. The rich may voluntarily finance a transfer that raises the disposable income of the poor to guarantee their survival and education as potential contributors to the funding of the public good. Of course this requires, that the welfare program does in fact activate the effort of the poor and raises their education. Unsurprisingly, a simple unconditioned transfer will not have this effect. Nevertheless, it seems useful to first study the effects of this simplest type of transfer. The conclusions of the present section do not depend on the motives of the funding party. As has been discussed in the introduction there may be more direct reasons for providing transfers. These may be paternalistic or altruistic preferences or simply the balance of political power that forces the rich to pay the transfer.

What happens if a household with low human capital and correspondingly low income is paid a monetary transfer Z ensuring a minimal standard of living? For simplicity of exposition assume that the group eligible to social transfer payments is exempted from any (further) tax. The individual household problem (1) is therefore unchanged except that $m(Z + wh)$ replaces $m(wh)$. The

Hamiltonian becomes $H(x,h,\lambda) = m(Z+wh) + bh - v(x) + \xi(G) + \lambda \cdot (\gamma x - \delta h)h$ and the dynamic system

$$\begin{cases} \dot{x} = \frac{v_x(\rho+\delta h)-(m_h+b)\gamma h}{v_{xx}} = \frac{(a+\alpha x)(\rho+\delta h)-(\kappa\frac{w}{Z+wh}+b)\gamma h}{\alpha} \\ \dot{h} = (\gamma x - \delta h)h \end{cases} \tag{17}$$

with isoclines

$$\begin{cases} \dot{x} = 0 \text{ if } x = x_{\dot{x}=0}(h,Z) := \frac{\gamma}{\alpha}\frac{\kappa\overbrace{\frac{wh}{wh+Z}}^{1 \text{ for } Z=0}+bh}{\rho+\delta h} - \frac{a}{\alpha} \\ \dot{h} = 0 \text{ if } h = 0 \text{ or } x = \frac{\delta}{\gamma}h \end{cases}$$

The essential difference to the system without the transfer is that $x_{\dot{x}=0}(0,Z) = -\frac{a}{\alpha} < 0 = x_{\dot{h}=0}(0)$ for any $Z > 0$ independently of how strong the economic and non-economic motives of effort, while for $Z = 0$ we had that $x_{\dot{x}=0}(0,0) = \frac{\gamma}{\alpha}\frac{\kappa}{\rho} - \frac{a}{\alpha} > 0$ for sufficiently strong economic motive or sufficiently patient households. As a consequence Case 1 is not possible if $Z > 0$. This already provides the essential intuition for the following proposition, which is proven in more detail in Appendix 8.3:

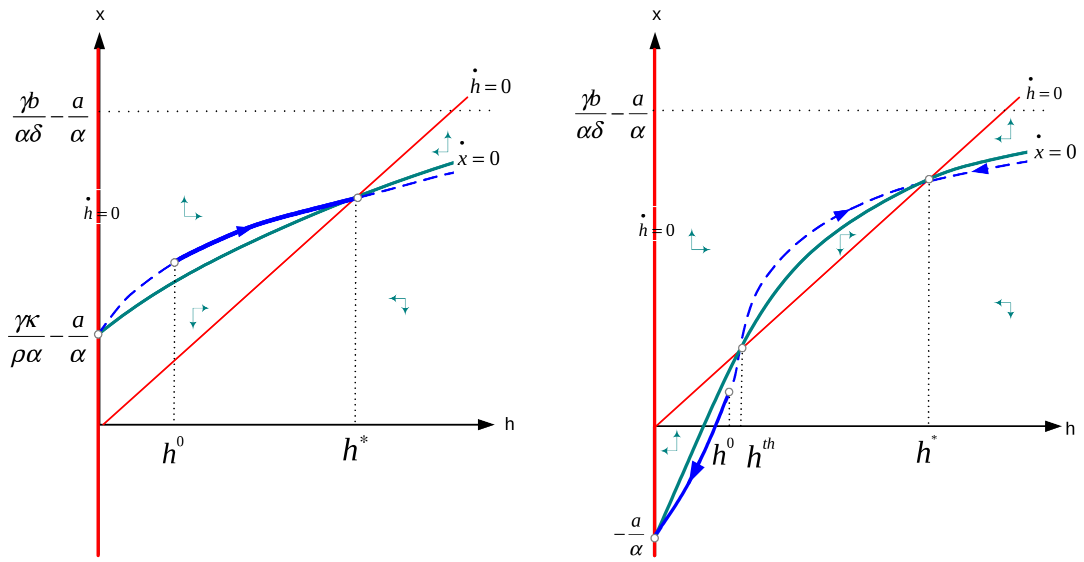


Figure 4: Left: Initially poor and motivated household Right: Same household receiving transfer.

Proposition 5 *Any transfer introduces a threshold level $h^{th}(Z) > 0$, such that a household with initial skill $h_0 < h^{th}(Z)$ gives up struggling and will eventually loose all his skill. The household will have enough to eat (his life-time utility at $t = 0$ is raised) but remain uneducated. Unskilled but patient and motivated households, initially in Case 1 (Figure 4), which would have liberated themselves from poverty and low skill, are enticed to passivity and will remain unskilled in the presence of monetary transfers. If without transfer, the household was already in Case 2, then the transfer raises the threshold ability $h^{th}(Z)$ and raises it the more, the larger the transfer ($h^{th}(Z)$ increasing).*

The transfer thus can prevent economic *and* personal emancipation towards high skill and income.

Proposition 5 does not depend on the source of the transfers[9] or on the motivation of the contributors. Suppose that, starting at a Pareto-inefficient polarized equilibrium, a hypothetical social planner or a real government wants to implement a Pareto-improving tax-transfer system. Such a system has to win the approval of the poor and at the same time motivate them to raise their effort such as to increase their human capital. If simple transfers were the only available policy, Proposition 5 can be read as a version of the "Samaritan's Dilemma". One can only reduce the likelihood that some of the uneducated will "starve" by at the same time reducing the likelihood that the poor get educated and self-sustained.

6 Activating welfare, negative marginal income tax

As we have seen, simple transfers enhance the passivity of the beneficiaries rather than raising their effort, so that the (non-altruistic) rich have no incentive to finance the transfers. To win the approval of the contributors the transfer-scheme has to activate. To activate, the transfer has to be conditioned either on sufficient effort or on the fruits of effort – human capital or income. At the same time, to win the voluntary participation by the beneficiaries, the transfer has to be sufficiently attractive to reward them for the additional effort. The present section introduces a class of activating and Pareto-improving tax-transfer systems that satisfies these conditions.

The initial situation. Consider an economy in the interesting threshold Case 2. The initial situation at time $t = 0$ is a polarized society with a fraction n_* of unskilled households with incomes below the unstable stationary $h^{\text{th}}_{\text{sym}}$ of Section 4.2 ($h^{\text{u}}_0 < h^{\text{th}}_{\text{sym}}$) and a fraction $n^* = 1 - n_*$ of skilled households with incomes close to h^*_{sym}. The n_* unskilled do not participate in the funding of the public good. The provision of the public good follows the optimal solution of the coordinated household problem of Section 4.2 restricted to the n^* initially rich and skilled households. Note that subject to the constraint that the n_* initially unskilled receive no transfers, the optimal tax of Section 4.2 remains optimal for sufficiently large t, since the maximal potential contributions of the poor tend to zero in the course of time.

[9] The model can for instance be closed by introducing a constant marginal labor income tax to balance government budget. The individual budget constraint becomes $c^i_t = T_t + wh^i_t = T_t + (1-\tau)\widetilde{w}h^i_t$, where $\widetilde{w}$ is the gross wage per labor efficiency unit, τ is the tax rate on labor income, $w = (1-\tau)\widetilde{w}$ is the after tax wage. Government budget is balanced if $\int_0^1 T^i_t di = \tau\widetilde{w}\int_0^1 h^i_t di$, such that if all other households are at steady state, then $T_t = \tau\widetilde{w}_t \int_0^1 h^i_t di$ is constant and the individual household problem is the same as before, where now $w = (1-\tau)\widetilde{w}$.

Proposition 6 *Any polarized equilibrium (with optimal funding of the public good in the absence of transfers) is Pareto-inefficient. There exists a Pareto-improving tax-transfer system which is activating in the sense that under the program all households cross the threshold level of human capital in finite time, and converge to the high stationary level h^*_{sym}. Individual taxes and transfers only depend on current individual income and involve a negative marginal tax for those households below the threshold level.*

I prove this proposition by constructing a corresponding welfare program.

General idea: *Activation:* Consider Figure 5. To activate, the $\dot{x} = 0$-isocline has to be shifted up sufficiently such that the initially unskilled choose to overcome the threshold. This will be achieved by replacing the laissez-faire income $y^{\mathrm{u}}_t = wh^{\mathrm{u}}_t$ of the initially poor for a certain time period $t \in (0, \widetilde{t})$ by the new disposable income $y^{\mathrm{dis}}(h^{\mathrm{u}}_t) = w\phi z_t\,(h^{\mathrm{u}}_t)^{\eta}$. For $\eta > 1$, y^{dis} rises faster in h than does wh, which strengthens the economic motive for effort. If η is sufficiently large, then $\widetilde{h} := h^{\mathrm{yes}}_{\widetilde{t}} > h^{\mathrm{th}}_{\mathrm{sym}}$, where h^{yes}_t describes the optimal path conditioned on participation by the poor. The larger η, the smaller the corresponding $\widetilde{t}$. The time-autonomous term $\{z_t\}_{t\geq 0}$ will be used to generate a constant transfer Z *at* equilibrium.

Voluntary participation by the contributor: The cost to each contributor is a transfer payment of $\frac{1-n^*}{n^*}Z(h^{\mathrm{u}}_t) = \frac{1-n^*}{n^*}\left[w\varphi z_t\,(h^{\mathrm{u}}_t)^{\eta} - wh^{\mathrm{u}}_t\right]$ for each $t \in (0, \widetilde{t})$. The benefit to each contributor is that there will be a higher G_t for each $t > \widetilde{t}$. The larger η, the larger the transfers but the shorter the period of transfers and the earlier the benefits. It is shown that for sufficiently small Z, the present utility value of the benefits is larger than the present utility value of the costs if η is chosen sufficiently large to raise the initially poor's rate of human capital accumulation above the discount rate ρ.

Voluntary participation by the beneficiary: The "sufficiently small Z" (for the rich to participate) should at the same time be sufficiently large for the poor to voluntarily participate. This is always the case for sufficiently small initial levels of human capital h^{u}_0. If initially, h^{u}_0 does not satisfy this requirement, the beginning of the transfer payment has to be delayed accordingly.

The n_* beneficiaries: For $t < \widetilde{t}$ each of the n_* initially unskilled households receives a transfer $Z(h^{\mathrm{u}}_t)$ that depends on its income wh^{u}_t. For $t > \widetilde{t}$ the household is no longer entitled to the transfer and participates in the optimal symmetric scheme of Section 4.2 and thus pays taxes $T(h^{\mathrm{u}}_t, \overline{h}) = wh^{\mathrm{u}}_t - \frac{\kappa}{\xi}e^{\frac{w\xi}{\kappa}(h^{\mathrm{u}}_t - \overline{h}_t)}$.

The n^* contributors: Each of the n^* contributors (with $h_0^{\mathrm{s}} \geq h_{\mathrm{sym}}^{\mathrm{th}}$) pays the n^*-optimal symmetric tax plus an equal share of social transfers:

$$T(h, \overline{h}_t^{\mathrm{s}}) = \frac{G_t + n_* Z(h_t^{\mathrm{u}})}{n^*} = \begin{cases} 0 \text{ if } wh_t^{\mathrm{s}} \leq \frac{n_*}{n^*} \frac{Z(h_t^{\mathrm{u}})}{w} \\ \frac{n_*}{n^*} Z(h_t^{\mathrm{u}}) \text{ if } \frac{n_*}{n^*} \frac{Z(h_t^{\mathrm{u}})}{w} < h_t^{\mathrm{s}} \leq \frac{\kappa}{w\xi} + \frac{n_*}{n^*} \frac{Z(h_t^{\mathrm{u}})}{w} \\ wh - \frac{\kappa}{\xi} e^{\frac{w\xi}{\kappa}(h - \overline{h}_t^{\mathrm{s}})} \text{ if } h_t^{\mathrm{s}} \geq \frac{\kappa}{w\xi} + \frac{Z(h_t^{\mathrm{u}})}{w}. \end{cases}$$

We will later make sure that $h_{\mathrm{sym}}^{\mathrm{th}} \geq \frac{\kappa}{w\xi} + \frac{n_*}{n^*} \frac{Z(h_t^{\mathrm{u}})}{w}$ such that $T(h_t^s, \overline{h}_t^{\mathrm{s}}) = wh_t^s - \frac{\kappa}{\xi} e^{\frac{w\xi}{\kappa}(h_t^s - \overline{h}_t^{\mathrm{s}})}$. It is shown in Appendix 8.4 that this tax scheme is optimal for the n^* contributor households if they are constrained to finance the transfers $n_* Z(h_t^{\mathrm{u}})$. The solution to this problem and in particular the values $h_{\mathrm{sym}}^{\mathrm{th}}$ and h_{sym}^{*} do not depend on n^* or the transfer if the n^* contributing households are sufficiently rich initially. Note that at the optimal solution of the unskilled households, the transfers will be constant $Z(h_t^{\mathrm{u}}) = Z$ for $t \leq \widetilde{t}$ and zero for $t > \widetilde{t}$.

The transfer is chosen such as to provide a disposable income that depends on h_t in a simple tractable way and allows to raise the household's reward for increased human capital:his human capital:

$$y_t^{\mathrm{dis}} = wh_t + Z(h_t) = w\phi z_t h_t^{\eta}$$

where ϕ is a constant and z_t is a time-autonomous term not depending on the households behavior. In other words, the transfer is $Z(h_t) = y_t^{\mathrm{dis}} - wh_t = w\phi z_t h_t^{\eta} - wh_t$ or equivalently, the tax 'paid' by the household is

$$T(h_t) = -Z(h_t) = wh_t - w\phi z_t h_t^{\eta} = wh_t \left[1 - \phi z_t h_t^{\eta - 1}\right].$$

The transfer is positive (tax negative) iff $w\phi z_t h_t^{\eta} > wh_t$ or $h_t > (\phi z_t)^{\frac{1}{-(\eta-1)}}$. The marginal tax for a participating beneficiary household is $T_h = w\left(1 - \eta\phi z_t h_t^{\eta-1}\right) < 0$ iff $1 < \eta\phi z_t h_t^{\eta-1}$ or iff $h^{-(\eta-1)} < \eta\phi z_t$ or (for $\eta > 1$) iff $h_t > (\eta\phi z_t)^{\frac{1}{-(\eta-1)}}$. Since $(\phi z_t)^{\frac{1}{-(\eta-1)}} > (\eta\phi z_t)^{\frac{1}{-(\eta-1)}}$ the transfer is positive *and* marginal tax negative iff $h_t > (\phi z_t)^{\frac{1}{-(\eta-1)}}$. Thus if h_0 is small ϕz_0 has to be large to induce participation by the beneficiary.

Activation The Hamiltonian of the beneficiary household is $H(x, h, c) = m(y_t^{\mathrm{dis}}) + bh - v(x) + \lambda(\gamma x - \delta h)h$, where $y_t^{\mathrm{dis}}(h_t) = w\phi z_t h_t^{\eta}$ for $t \leq \widetilde{t}$ and $y_t^{\mathrm{dis}}(h_t) = \frac{\kappa}{\xi} e^{\frac{w\xi}{\kappa}(h_t - h_{\mathrm{sym}}^{*})}$ for $t > \widetilde{t}$ (see Section 4.2). For $t \leq \widetilde{t}$, $m(y_t^{\mathrm{dis}}) = \kappa \ln(w\phi z_t h_t^{\eta})$ and thus $m_h = \frac{\kappa\eta}{h}$ does not depend on $w\phi z_t$. $m_h = \frac{\kappa\eta}{h}$ replaces $m_h = \frac{\kappa}{h}$ in (3) of the original household problem of Section 3. Correspondingly $x_{\dot{x}=0}(h) := \frac{\gamma}{\alpha} \frac{\kappa + bh}{\rho + \delta h} - \frac{a}{\alpha}$ of the original household problem is replaced by $x_{\dot{x}=0}^{\mathrm{yes}}(h) = \frac{\gamma}{\alpha} \frac{\eta\kappa + bh}{\rho + \delta h} - \frac{a}{\alpha}$ for $t \leq \widetilde{t}$. For $t > \widetilde{t}$ the household follows the equation of motion the socially optimal symmetric solution of Sections 1 and 4.2.[10]

$$\begin{cases} \dot{x}_t = 0 \text{ if } \begin{cases} x^{\text{yes}}_{\dot{x}=0}(h_t) = \frac{\gamma}{\alpha}\frac{\eta\kappa+bh_t}{\rho+\delta h_t} - \frac{a}{\alpha} \text{ for } t \leq \widetilde{t} \\ x^{\text{sym}}_{\dot{x}=0}(h_t) = \frac{\gamma}{\alpha}\frac{(\xi w+b)h_t}{\rho+\delta h_t} - \frac{a}{\alpha} \text{ for } t > \widetilde{t} \end{cases} \\ \dot{h}_t = 0 \text{ if } h_t = 0 \text{ or } x = \frac{\delta}{\gamma}h_t \end{cases} \tag{18}$$

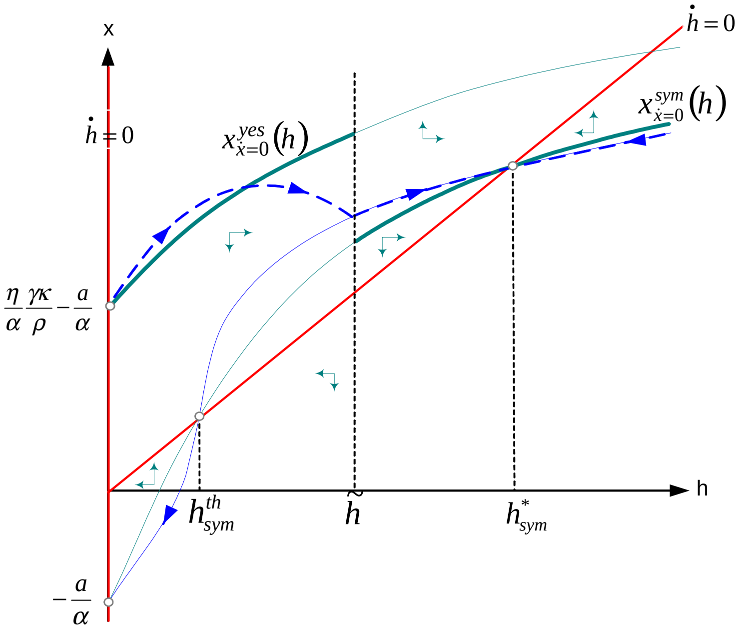


Figure 5: Activating welfare.

Each η determines a unique phase-diagram which is relevant only for $t \leq \widetilde{t}$. Because instantaneous utility is concave in x the policy function will be continuous at $\widetilde{t}$. For $t > \widetilde{t}$ a participating household follows the well defined saddle path determined by the contributor's household problem known from the previous section. Thus, given $\eta, h_0, \widetilde{t}$, the (participating) household chooses x_0 such that the corresponding solution path of the transfer regime arrives exactly in period $\widetilde{t}$ at the saddle-path of the contribution regime. Let $\widetilde{h}$ be the corresponding human capital level (see Figure 5). Raising η shifts up the $\dot{x}$-isocline $x^{\text{yes}}_{\dot{x}=0}(h)$ and therefore the policy function as well. If $\eta > \frac{\rho a}{\gamma\kappa}$, then $x^{\text{yes}}_{\dot{x}=0}(0) = \frac{\gamma}{\alpha}\frac{\eta\kappa}{\rho} - \frac{a}{\alpha} > 0$, in which case the transfer-system is activating for any household starting with initial human capital $h_0 < h^{\text{th}}_{\text{sym}}$. Thus, to make sure that the $\widetilde{h}$ reached at $\widetilde{t}$ is larger than $h^{\text{th}}_{\text{sym}}$, η has to be sufficiently large (see Figure 5).

Choose any $\widetilde{h} \in (h^{\text{th}}_{\text{sym}}, h^*_{\text{sym}})$. Consider the solution path $\left\{x^{\text{yes}}(h|\eta, \widetilde{h})\right\}_{h>0}$ of the dynamic system defined by η and crossing $x^{\text{sym}}_{\dot{x}=0}(h_t)$ at $\widetilde{h}$. Suppose given some h_0 and $\widetilde{t}$ the household has chosen

[10]Since the n_* initially unskilled households are identical no transfers have to be financed anymore. All housholds contribute to the financing of the public good now.

$x^{\text{yes}}(h|\eta,\widetilde{h})$ leading to $\widetilde{h}$. If h_0 is reduced and $\widetilde{t}$ kept constant, then the household will of course no longer choose the path $x^{\text{yes}}(h|\eta,\widetilde{h})$ leading to $\widetilde{h}$. If we want the household to stick to $x^{\text{yes}}(h|\eta,\widetilde{h})$ when reducing h_0, we have to simultaneously raise $\widetilde{t}$. Thus, keeping in mind that the household chooses $\widetilde{h}$ while $\widetilde{t}$ is given exogenously for him, we can nevertheless express the policy function $x^{\text{yes}}(h|\eta,\widetilde{h})$ given $\eta,\widetilde{h}$. Let $\Delta(\eta,\widetilde{h}) := \inf_{h\in]0,\widetilde{h}]}\widehat{h}(h)$ be the smallest growth rate of h along this path. Then $\widetilde{h}/h_0 \geq e^{\Delta(\eta,\widetilde{h})\widetilde{t}}$ and obviously $\widetilde{t}\to\infty$ when $h_0\to 0$.

Note that we are considering the threshold Case 2 in which among others $C = \frac{\kappa\gamma-\rho a}{\alpha} < 0$ or $\frac{\rho a}{\kappa\gamma} > 1$. Thus the necessary condition for activation at small h_0 ($\eta > \frac{\rho a}{\gamma\kappa}$) implies that $\eta > 1$.

Constant transfer along the solution path The solution of the individual household problem in case he participates, $\{x_t^{\text{yes}}, h_t^{\text{yes}}\}_{t\geq 0}$, does not depend on $\{w\phi z_t\}_{t\geq 0}$. The time-autonomous term $\{z_t\}_{t\geq 0}$ can thus be made a function of h_t^{yes} without involving any circularity. For instance, $\{z_t\}_{t\geq 0}$ can be determined such as to avoid increasing cost to the net contributor, when the recipients get more and more educated. A specific natural benchmark example is to determine $z_t = z(h_t^{\text{yes}})$ such as to generate a constant transfer Z. The net transfer (along the household solution!) is constant for

$$z_t = \frac{h_t^{\text{yes}} + \frac{Z}{w}}{\phi\left(h_t^{\text{yes}}\right)^\eta}.$$

Inserting this z_t into $Z(h_t) = y_t^{\text{dis}} - wh_t = w\phi z_t h_t^\eta - wh_t$ yields $Z(h_t) = \frac{wh_t^{\text{yes}}+Z}{\left(h_t^{\text{yes}}\right)^\eta}h_t^\eta - wh_t$, or for $h_t = h_t^{\text{yes}}$: $Z(h_t^{\text{yes}}) = Z$. The net income along the solution path then is $y_t^{\text{dis}} = wh_t + Z$. The contributors don't pay more and more when the poor get richer and richer and the poor keep the benefits of their additional h.

Voluntary participation by the beneficiary As has already been remarked, the fact that the program activates is not sufficient to warrant the approval of the beneficiary since the activation may be motivated by a penalty for $h_t < h_0$ if the household participates (The autonomous term $z_t = \frac{h_t^{\text{yes}}+\frac{B}{w}}{\phi\left(h_t^{\text{yes}}\right)^\eta}$ falls in the course of time).[11] Note that h_t^{yes} is the human capital along the optimal solution path *assuming* that the household participates.

Let h_t^{no} correspond to the optimal solution path assuming that the household does not participate (which is the solution to the original household problem of Section 3). The household will voluntarily participate in the program if

$$\int_0^{\widetilde{t}} e^{-\rho t}\left[\kappa\ln\left(w\phi z_t\left(h_t^{\text{yes}}\right)^\eta\right) - v(x_t^{\text{yes}})\right]dt + \int_{\widetilde{t}}^\infty e^{-\rho t}\left[\kappa\ln\left(\frac{\kappa}{\xi}\right) - v(x_t^{\text{yes}})\right]dt > \int_0^\infty e^{-\rho t}\left[\kappa\ln(wh_t^{\text{no}}) - v(x_t^{\text{no}})\right]dt.$$

[11] If, for any $h < \widetilde{h}$, the income under welfare were at least as high as without the program (even for a hypothetical $h_t < h_0$), then the welfare program would clearly be as good for the household as no welfare, since even under the program, it could choose the old effort path without loosing income (in this case, the household would even continuously vote for continuation of the program).

For this it is sufficient that $\kappa\left[\ln(w\phi z_t\left(h_t^{\text{yes}}\right)^\eta - \ln wh_t^{\text{no}}\right] \geq v(x_t^{\text{yes}}) - v(x_t^{\text{no}})$ for all $t \in (0,\widetilde{t})$ and $\kappa\left[\kappa\ln(\frac{\kappa}{\xi}) - \ln wh_t^{\text{no}}\right] \geq v(x_t^{\text{yes}}) - v(x_t^{\text{no}})$ for $t \geq \widetilde{t}$ or that $\kappa\ln\frac{\phi z_t\left(h_t^{\text{yes}}\right)^\eta}{h_t^{\text{no}}} \geq v(x_t^{\text{yes}}) - v(x_t^{\text{no}}) \geq v(x^{\max}) - 0$ for $t \in (0,\widetilde{t})$, where $x^{\max} := \sup_{t\in(0,\widetilde{t})} x_t^{\text{yes}}$ and that $\kappa\ln\frac{\kappa}{\xi h_t^{\text{no}}} \geq v(x^{\max})$ for $t \geq \widetilde{t}$. Since $h_0 > h_t^{\text{no}}$ for $t \geq \widetilde{t}$, this latter part of this condition is satisfied for $h_0^{\text{u}} \leq \frac{\kappa}{\xi}e^{-\frac{v(x^{\max})}{\kappa}}$.

With $z_t = \frac{h_t^{\text{yes}} + \frac{Z}{w}}{\phi\left(h_t^{\text{yes}}\right)^\eta}$ the former part of the condition becomes $\eta\kappa\ln\frac{\phi\frac{h_t^{\text{yes}}+\frac{Z}{w}}{\phi\left(h_t^{\text{yes}}\right)^\eta}\left(h_t^{\text{yes}}\right)^\eta}{h_t^{\text{no}}} = \eta\kappa\ln\frac{wh_t^{\text{yes}}+Z}{wh_t^{\text{no}}} \geq v(x_t^{\max})$ for all $t \in (0,\widetilde{t})$. Since h_t^{no} falls and h_t^{yes} rises for all $t \in (0,\widetilde{t})$, it is sufficient that $\eta\kappa\ln\frac{wh_0^{\text{u}}+Z}{wh_0^{\text{u}}} \geq v(x_t^{\max})$, $\frac{wh_0^{\text{u}}+Z}{wh_0^{\text{u}}} \geq e^{\frac{v(x^{\max})}{\eta\kappa}}$, or that

$$Z \geq wh_0^{\text{u}}\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right). \tag{19}$$

Voluntary participation by the contributor To guarantee that the welfare-program is Pareto-improving it remains to show that for sufficiently small h_0^{u} the rich too vote for the welfare scheme if (19) is satisfied.

In addition to the contribution to the public good, each contributing households pays the contribution $\frac{n_*}{n^*}Z$ to the social security fund during the time interval $[0,\widetilde{t}]$. First of all it has of course to be made sure that the "rich" are in fact sufficiently rich to finance the transfer and in addition coordinate on providing a positive amount of the public good, i.e. that the above assumption $h_{\text{sym}}^{\text{th}} > \frac{\kappa}{w\xi} + \frac{n_*}{n^*}\frac{Z(h_t^{\text{u}})}{w}$ is satisfied, thus

$$Z < \frac{n^*}{n_*}\left(wh_{\text{sym}}^{\text{th}} - \frac{\kappa}{\xi}\right).$$

Since $\frac{n^*}{n_*} = \frac{1-n_*}{n_*} \to \infty$ for $n_* \to \infty$, the condition can always be satisfied by assuming that their are only few beneficiaries. Instead the condition will be met below by making sure that Z is sufficiently small.

In exchange the contribution from 0 to $\widetilde{t}$, the supply of the public good will be raised for all $t > \widetilde{t}$ by $\int_0^{n_*}\left[wh_t^i - \kappa/\xi\right]di = n_*\left(wh_t^{\text{u}} - \kappa/\xi\right) > n_*\left(w\widetilde{h} - \kappa/\xi\right)$, corresponding to an individual tax-reduction of $\frac{n_*}{n^*}\left(w\widetilde{h} - \kappa/\xi\right)$ for each initial contributor.

The contributors therefore agree to the transfer scheme if $\int_0^{\widetilde{t}} e^{-\rho t}Zdt < \int_{\widetilde{t}}^\infty e^{-\rho t}\left(w\widetilde{h} - \kappa/\xi\right)dt$ or if $Z < \left(w\widetilde{h} - \kappa/\xi\right)\frac{\int_{\widetilde{t}}^\infty e^{-\rho t}dt}{\int_0^{\widetilde{t}} e^{-\rho t}dt} = \frac{w\widetilde{h}-\kappa/\xi}{e^{\rho\widetilde{t}}-1}$.

We have seen that the beneficiaries agree if $Z = wh_0\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right)$. Assuming $wh_0\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right) < \frac{n^*}{n_*}\left(wh_{\text{sym}}^{\text{th}} - \frac{\kappa}{\xi}\right)$ for a moment, the scheme is therefore Pareto-improving if $h_0\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right) < \frac{\widetilde{h}-\kappa/(w\xi)}{e^{\rho\widetilde{t}}-1}$ or if $\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right)\left(e^{\rho\widetilde{t}} - 1\right) < \frac{\widetilde{h}}{h_0}\left(1 - \frac{\kappa}{w\xi\widetilde{h}}\right)$.

With $\Delta := \inf_{t\in]0,\widetilde{t}]}\widehat{h}_t$ and thus $\frac{\widetilde{h}}{h_0} \geq e^{\Delta\widetilde{t}}$ a more demanding sufficient condition for the welfare program to be Pareto-improving becomes $\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right)\left(e^{\rho\widetilde{t}} - 1\right) < e^{\Delta\widetilde{t}}\left(1 - \frac{\kappa}{w\xi\widetilde{h}}\right)$. For $h_0 \to 0$ and

$\widetilde{h}$ constant, $\frac{\widetilde{h}}{h_0} \to \infty$ and hence $\widetilde{t} \to \infty$. Note that while $v(x^{\max})$ depends on η it is bounded for any given η. Furthermore $1-\frac{\kappa}{w\xi\widetilde{h}} > 0$, since $\widetilde{h} > \frac{\kappa}{w\xi}$. Thus for small h_0 (which will always be reached in case of non-participation), voluntary financing is guaranteed if $\Delta > \rho$. In other words, the welfare program is Pareto-improving if η is sufficiently large to ensure $\Delta = \inf_{t\in(0,\widetilde{t})} \widehat{h}_t > \rho$ (or $\widehat{h}_t > \rho$ for $h_t \in [h_0, \widetilde{h}]$), which is always possible since $h^{\text{th}}_{\text{sym}} < \widetilde{h} < h^{*}_{\text{sym}}$. Since the term $\frac{n^*}{n_*}\left(wh^{\text{th}}_{\text{sym}} - \frac{\kappa}{\xi}\right)$ does not depend on any of the policy parameters $\eta, \Delta, \widetilde{h}$, or $\widetilde{t}$ the first condition $wh_0\left(e^{\frac{v(x^{\max})}{\eta\kappa}} - 1\right) < \frac{n^*}{n_*}\left(wh^{\text{th}}_{\text{sym}} - \frac{\kappa}{\xi}\right)$ is obviously satisfied for small h_0. This proves Proposition 6.

Negative marginal tax At the individual solution the household receives a constant transfer such that its disposable income growth with gross incomes: $y^{\text{dis}}_t = wh^{\text{yes}}_t + Z$ for $t \in (0, \widetilde{t})$. Nevertheless the marginal tax $T_h = w\left(1 - \eta\phi z_t h_t^{\eta-1}\right) = w\left[1 - \eta\phi\frac{h_t + \frac{Z}{w}}{\phi\left(h^{\text{yes}}_t\right)^{\eta}} h_t^{\eta-1}\right] = w\left[1 - \eta\left(1 + \frac{Z}{wh^{\text{yes}}_t}\right)\right]$ is strictly negative if the transfer is positive since $\eta > 1$.

7 Concluding remarks and extensions

The pleasure of activity The essential deviation of the present paper from most of the literature on human capital was the introduction of the state variable human capital as an argument in the instantaneous utility function. I have argued in the Introduction that most individuals and households derive direct satisfaction from their knowledge and skill ($u_h > 0$). Now it seems equally plausible that most people enjoy a certain amount of activity and the effort that comes with it. The standard assumption (which the present paper retains) that all effort is costly ($u_x < 0$) could in fact be criticized for similar reasons than the standard assumption $u_h = 0$ (which the present paper gives up). Learning may be as much fun as knowing, jogging as much as being fit. To capture this idea, I allow for positive marginal utility of effort in a companion paper. In particular, I assume that a task is pleasant if the effort it requires from an individual is in balance with the individual's skill, while deviations from this balance reduce utility. The path-dependence of optimal individual behavior which is generated by the interaction between the state variable human capital and the control variable effort, at the centre of the present paper, also arises in this alternative framework.

Perfect capital markets A second assumption of the basic maximization problem (1) and all its variations in the present paper was that in every period, each household consumes its entire labor income. In other words, credit markets have been excluded. At first sight, one may conjecture that introducing credit markets excludes path-dependent behavior and persistent inequality. In fact, there is a large literature in which inequality with respect to education, income, and utility only persists due to capital market imperfections. In this literature, unskilled households with low overall utility only remain unskilled because they cannot finance today's education using loans they pay back tomorrow with their then high skilled labor income. In contrast, in the present framework, human capital can neither be bought by money (be this a loan, initial wealth, or a transfer) nor has its accumulation any *monetary* opportunity costs ($\dot{h}_t > 0$ does not reduce current income wh_t). The reason why some households do not raise their human capital here is that it is too hard rather than too expensive. In reality the accumulation of human capital generally has monetary costs (or opportunity cost) and at the same requires non-monetary effort. It is this non-monetary component that generates path-dependence and does so also if capital markets are perfect.

While introducing perfect capital markets does not eliminate the possibility of thresholds, it adds a further interesting dimension to the present paper: The interplay between monetary wealth (physical capital) and human capital. I tackle this issue in a related paper, where I show that, as in the case without credit market but with transfer, Case 1 is excluded for a household with initial wealth $k_0 > 0$. The possibility of saving and/or living on initial financial wealth introduces a threshold level of human capital $h^{\text{th}}(k_0)$, below which a path towards passivity is always chosen. Furthermore this threshold level $h^{\text{th}}(k_0)$ is the larger, the larger initial financial wealth.

The specific instantaneous utility function A third assumption was the specific utility function concerning consumption and human capital. While the relevant feature of this specific example was that standard Inada conditions hold for and $c \to 0$ and $x \to \infty$, the marginal (direct) utility of human is not particularly large at low levels of human capital, the details of the results of course depend on the specific example. In particular it has been assumed that the elasticity of intertemporal substitution for c was unity: $m(c) = \kappa \ln c$. This was sufficient to make sure that the economic motive could be strong enough to overcome the cost of effort at low h (this was the case for κ large enough). One may criticize this specification for also allowing the economic motive to be too weak, that is, for assuming households that – under some circumstances (small κ) – accept asymptotic starving. If one wants to avoid this possibility one has to assume that the intertemporal elasticity of substitution is smaller than one. If $m(c) = \kappa \ln c$ is replaced by $m(c) = \kappa \frac{\sigma}{\sigma-1}(c)^{\frac{\sigma-1}{\sigma}}$ for $0 < \sigma < 1$, then the household will be less reluctant to substitute present comfort (low x) with future consumption (low future wh) and struggles harder when he realizes that this is necessary to avoid starving ($c = wh \to 0$). In fact, it can be shown that there no longer is a stationary state at $h = 0$. However, optimal individual behavior remains path-dependent under similar assumptions as those of the present paper. The smaller saddle-point stable stationary state h_* is merely shifted from $h_* = 0$ to a strictly positive value and there still are three stationary states then, $0 < h_* < h^{\text{th}} < h^*$. The present paper chooses the unity elasticity of substitution because it is analytically more tractable. Even for the simplest example with $\sigma = 1/2$, the $\dot{x} = 0$-isocline $x_{\dot{x}=0}(h)$ is cubic rather than quadratic as in the present paper.

The common interest A fourth assumption of the present paper was the existence of a public good and the simple specific form it affects utility. Sections 3 and 5 were completely independent of the existence of public goods. The public good has been introduced to generate a common interest in the education of the greatest possible number of households. Proposition 6 in Section 6 can easily be adapted to an economy in which the rich have other reasons to voluntarily support the poor (a dislike for extreme poverty, for low education, for lost opportunities ...) or an economy in which the rich are forced to finance transfers to the poor. In all these cases, simple non-contingent transfers would generally be Pareto-dominated by activating transfers. I have chosen to provide the rich with an "economic motive" for helping the poor since this is probably least debatable among economist and since the economic motive will still be there if it is joined by more altruistic motives. Some less economic minded readers may prefer to put more weight on altruistic motives. In fact, while most people support modern politics for trying to promote general education, some at the same time blame politicians for doing so for purely economic reasons, hunting for the human capital's economic reward rather then valuing the human capital for its own sake, or, as Sen puts it, as an aim in itself. An altruistic rich household in the present framework would share this view since for the individual household education is in fact an aim in itself. Fortunately *any* transfer system that manages to raise the human capital of the poor beyond the threshold generates the double dividend, irrespective on how noble the motives behind the transfer.

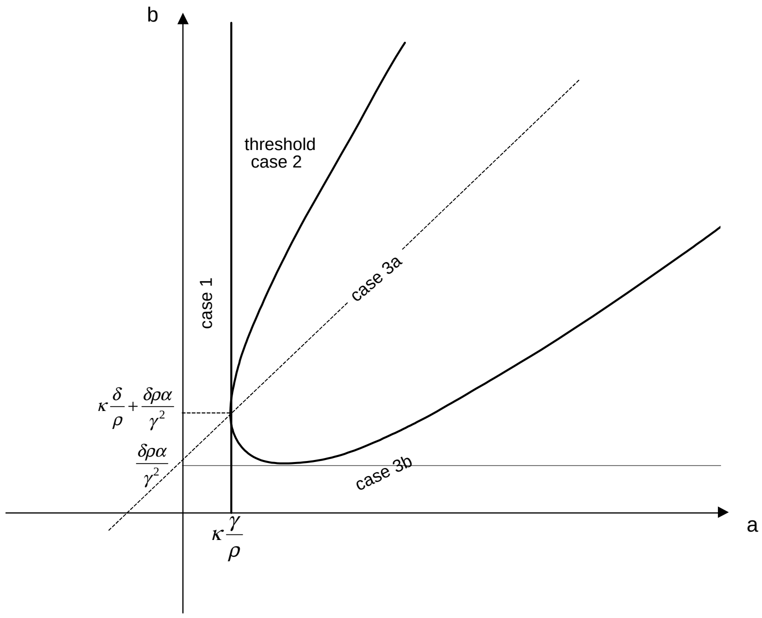
8 Appendix

8.1 Appendix: Characterization of the three cases

This appendix provides a more complete characterization of the three cases of Section 3. Consider Figure 8.1 in the (a, b)-plane that allows to classify all cases by a variation of the (direct) cost of effort a and the strength b of the non-economic motive, given all other parameters. On the vertical line defined by $a = \kappa\gamma/\rho$, the term $C = \frac{\kappa\gamma - \rho a}{\alpha}$ is zero. For any pair (a, b) left to this $C = 0$-line $a < \kappa\gamma/\rho$ or $C > 0$, such that the term $B^2 + 4DC$ is always positive and (6) has two real roots. Furthermore, since $C > 0$, $\sqrt{B^2 + 4DC} > B$ such that one of these roots is negative and one positive: $h_1 = \frac{B - \sqrt{B^2+4DC}}{2D} < 0 < h_2 := h^* = \frac{B+\sqrt{B^2+4DC}}{2D}$ (Case 1). For $C < 0$ the set of pairs (a, b) for which $B^2 + 4DC = 0$ correspond to the thick curve in Figure 8.1.

$$B^2 + 4DC > 0 \text{ if } \begin{cases} b > 2\frac{\alpha}{\gamma}\sqrt{\frac{\delta^2}{\gamma}\frac{\rho a - \kappa\gamma}{\alpha}} + \frac{\delta\alpha}{\gamma^2}\rho + a\frac{\delta}{\gamma} \text{ and } b > \frac{\alpha\delta}{\gamma^2}\rho + a\frac{\delta}{\gamma} \\ b < -2\frac{\alpha}{\gamma}\sqrt{\frac{\delta^2}{\gamma}\frac{\rho a - \kappa\gamma}{\alpha}} + \frac{\delta\alpha}{\gamma^2}\rho + a\frac{\delta}{\gamma} \text{ and } b < \frac{\alpha\delta}{\gamma^2}\rho + a\frac{\delta}{\gamma} \end{cases}$$

where $b \gtrless \frac{\alpha\delta}{\gamma^2}\rho + a\frac{\delta}{\gamma}$ corresponds to $B \gtrless 0$. The $B = 0$-curve is the increasing dashed line in Figure 8.1. For any pair (a, b) to the right of the $C = 0$-line ($C < 0$) but above the upper branch of



the $B^2+4DC=0$-curve, $B^2+4DC>0$ and $B>0$ such that (6) has two real roots, which are both strictly positive since $\sqrt{B^2+4DC}<B$: $0<h_1=\frac{B-\sqrt{B^2+4DC}}{2D}<h_2=\frac{B+\sqrt{B^2+4DC}}{2D}$ (Case 2). For any pair (a,b) to the right of the $C=0$-line ($C<0$) and in between the two branches of the $B^2+4DC=0$-curve, $B^2+4DC<0$ such that (6) has no real root (Case 3a). For any pair (a,b) to the right of the $C=0$-line ($C<0$) below the lower branch of the $B^2+4DC=0$-curve, $B^2+4DC>0$ and $B<0$ such that (6) has two real roots, which are both negative since $\sqrt{B^2+4DC}<B$: $h_1=\frac{B-\sqrt{B^2+4DC}}{2D}<h_2=\frac{B+\sqrt{B^2+4DC}}{2D}<0$ (Case 3b).

Considering the quadratic equation $E(h)=Dh^2-B(a)h-C(a)=0$, Case 2 occurs for small but strictly positive C, if for $C(a)=0$ the coefficient of the linear term $-B(a)$ is negative, i.e. if $B(a)>0$: At $C(a)=\frac{\kappa\gamma-\rho a}{\alpha}=0$, thus at $a=\frac{\kappa\gamma}{\rho}$, the equation $E(h)=Dh^2-B(a)h-C(a)=0$ always has a solution $h^{\text{th}}(a)=0$ and a second strictly positive solution $h^{*}(a)>0$. Raising a reduces $C(a)$. Since $B(a)$ remains positive for small variations of a, the solution $h^{\text{th}}(a)$ becomes positive and we move to (the interior of) Case 2. A necessary and sufficient condition for Case 2 to hold for some a therefore is that $B(a=\frac{\kappa\gamma}{\rho})=\frac{b\gamma-\frac{\kappa\gamma}{\rho}\delta}{\alpha}-\frac{\delta}{\gamma}\rho>0$ or equivalently that $b>\frac{\kappa\delta}{\rho}+\frac{\rho\alpha\delta}{\gamma^2}$. In the phase diagram this condition corresponds to the case that the slope of slope of $x_{\dot{x}=0}(h)$ at $h=0$ is larger than the slope of $x_{\dot{h}=0}(h)$: $\frac{d}{dh}x_{\dot{x}=0}(0)=\frac{\gamma}{\alpha}\frac{(\rho b-\kappa\delta)}{\rho^2}>\frac{\delta}{\gamma}$ iff $b>\frac{\delta\rho\alpha}{\gamma^2}+\kappa\frac{\delta}{\rho}$ (Cases 1a, 2, 3a). In in Figure 8.1 $b=\frac{\delta\rho\alpha}{\gamma^2}+\kappa\frac{\delta}{\rho}$ determines the upper horizontal line (dotted). If and only if b lies above this line, a variation of a generates the three cases.

Also consider the lower horizontal line (dashed) in Figure 8.1 which is defined by $b=\kappa\frac{\delta}{\rho}$. For any pair (a,b) above this line, $x_{\dot{x}=0}(h)$ is increasing and concave: $\frac{d}{dh}\left(x_{\dot{x}=0}\right)=\frac{\gamma}{\alpha}\frac{b(\rho+\delta h)-(\kappa+bh)\delta}{(\rho+\delta h)^2}=$

$\frac{\gamma}{\alpha}\frac{(\rho b-\kappa\delta)}{(\rho+\delta h)^2} > 0$ iff $b > \kappa\frac{\delta}{\rho}$. Similarly, for (a,b) below this line, $x_{\dot{x}=0}(h)$ is decreasing and convex.

- Slope $x_{\dot{x}=0}(0) >$ Slope $x_{\dot{h}=0}(h)$ iff $b > \frac{\delta\rho\alpha}{\gamma^2} + \kappa\frac{\delta}{\rho}$
- Slope $x_{\dot{x}=0}(0) \in [0,$ Slope $x_{\dot{h}=0}(h)]$ iff $b \in \left[\kappa\frac{\delta}{\rho}, \frac{\delta\rho\alpha}{\gamma^2} + \kappa\frac{\delta}{\rho}\right]$
- Slope $x_{\dot{x}=0}(0) < 0$ iff $b < \frac{\kappa\delta}{\rho}$.

Furthermore it is easy to check that the $B = 0$-curve, the upper horizontal line (dotted) and the $C = 0$-line always intersect in the same point (see Figure 8.1). This proves Lemma 1.

8.2 Appendix: The Individual Optimization Problem

Stability of the interior stationary solutions to (3) The associated Jacobian of the system (2) $J = \begin{pmatrix} \frac{\partial\dot{h}}{\partial h} & \frac{\partial\dot{h}}{\partial x} \\ \frac{\partial\dot{x}}{\partial h} & \frac{\partial\dot{x}}{\partial x} \end{pmatrix} = \begin{pmatrix} \gamma x - 2\delta h & \gamma h \\ \frac{v_x\delta-(m_h+b)\gamma-m_{hh}\gamma h}{v_{xx}} & (\rho+\delta h) - \dot{x}\frac{v_{xxx}}{v_{xx}} \end{pmatrix}$ At $\dot{h} = 0$ and $h \neq 0$: $J = \begin{pmatrix} -\delta h & \gamma h \\ \frac{(a+\alpha x)\delta-(m_h+b)\gamma-m_{hh}\gamma h}{\alpha} & (\rho+\delta h) \end{pmatrix}$ and $\det J = -\delta h(\rho+\delta h) - \frac{(a+\alpha x)\delta - b\gamma}{\alpha}\gamma h$. At $x = \frac{\delta}{\gamma}h$ $\det J = -\delta h(\rho+\delta h) - \frac{(a+\alpha\frac{\delta}{\gamma}h)\delta-b\gamma}{\alpha}\gamma h = \left[\frac{(b\gamma^2-a\delta\gamma-\alpha\delta\rho)}{\alpha} - 2\delta^2 h\right]h$. With $2\delta^2 h_{1,2} = \frac{(\gamma^2 b - a\delta\gamma - \alpha\rho\delta)}{\alpha} \pm \frac{\sqrt{(\gamma^2 b-\alpha\rho\delta-a\delta\gamma)^2-4\alpha\delta^2\gamma(a\rho-\gamma\kappa)}}{\alpha}$ $\det J(h_{1,2}) = \left\{\frac{(b\gamma^2-a\delta\gamma-\alpha\delta\rho)h}{\alpha} - \left[\frac{(\gamma^2 b-a\delta\gamma-\alpha\rho\delta)}{\alpha} \pm \frac{\sqrt{(\gamma^2 b-\alpha\rho\delta-a\delta\gamma)^2-4\alpha\delta^2\gamma(a\rho-\gamma\kappa)}}{\alpha}\right]\right\}h$ $= \det J(h_{1,2}) = -\left[\pm\frac{\sqrt{(\gamma^2 b-\alpha\rho\delta-a\delta\gamma)^2-4\alpha\delta^2\gamma(a\rho-\gamma\kappa)}}{\alpha}\right]h$. Thus $\det J(h^{\text{th}}) > 0$ and $\det J(h^*) < 0$. Therefore h^{th} is unstable (whenever it exists, thus in Case 2) and h^* is saddle-point stable (whenever it exists, thus in Cases 1 and 2).

Stability of the trivial stationary solution to (3) The associated Jacobian of the system (2) at the steady state with zero ability $h_* = 0$ is $J(h_*) = \begin{pmatrix} \gamma x - 2\delta h & \gamma h \\ \frac{(a+\alpha x)\delta-(m_h+b)\gamma-m_{hh}\gamma h}{\alpha} & (\rho+\delta h) \end{pmatrix} = \begin{pmatrix} \gamma x & 0 \\ \frac{(a+\alpha x)\delta-(m_h+b)\gamma-m_{hh}\gamma h}{\alpha} & \rho \end{pmatrix}$ with $\det J(h_*) = \gamma\rho x$. At $x_* = \frac{\gamma}{\alpha}\frac{\kappa}{\rho} - \frac{a}{\alpha}$, $\det J(h_*) = \gamma\rho\left(\frac{\gamma}{\alpha}\frac{\kappa}{\rho} - \frac{a}{\alpha}\right) > 0$ in Case 1, where (h_*, x_*) is thus unstable and $\det J(h_*) = \gamma\rho\left(\frac{\gamma}{\alpha}\frac{\kappa}{\rho} - \frac{a}{\alpha}\right) < 0$ in Case 2 and 3, where (h_*, x_*) is thus a saddle point.

$\det J = (\gamma x - 2\delta h)(\rho+\delta h) - \frac{(a+\alpha x)\delta-(\frac{\kappa}{h}+b)\gamma+\frac{\kappa}{h^2}\gamma h}{\alpha}\gamma h = (\gamma x - 2\delta h)(\rho+\delta h) - \frac{(a+\alpha x)\delta-(b)\gamma}{\alpha}\gamma h$

$\det J(h_*) = \gamma x\rho$.

$\Rightarrow$ In Case 1 $\det J(h_*) > 0$ and the trivial steady state (h_*, x_*) is unstable and in Cases 2 and 3 $\det J(h_*) < 0$ such that the trivial steady state $h_* = 0$ is a saddle point.

The transversality condition The necessary transversality condition for (1) is $\lim_{t\to\infty} e^{-\rho t}\lambda_t h_t = 0$. Inserting the FOC $\lim_{t\to\infty} e^{-\rho t}\frac{v_x}{\gamma h}h_t = \lim_{t\to\infty} e^{-\rho t}\frac{v_x}{\gamma} = \lim_{t\to\infty} e^{-\rho t}\frac{a+\alpha x}{\gamma} = \lim_{t\to\infty} e^{-\rho t}\frac{\alpha x}{\gamma} = 0$. Any path satisfying (3) and converging to one of the three possible stationary solutions (including these stationary solutions) obviously satisfy the transversality condition $\lim_{t\to\infty} e^{-\rho t}\frac{\alpha x_t}{\gamma} = 0$ since in all these cases $\{|x_t|\}_t$ is bounded.

For Arrow-Kurz sufficiency theorem (see below), one also needs that $\lim_{t\to\infty} e^{-\rho t}\widetilde{\lambda}_t h_t \geq 0$ for all admissible paths with the co-state variables $\{\widetilde{\lambda}_t\}_t$ of the proposed solution. This is satisfied because $h_t \geq 0$ and $\widetilde{\lambda}_t = \frac{a+\alpha\widetilde{x}_t}{\gamma\widetilde{h}_t} \geq 0$ ($\widetilde{x}_t \geq -a/\alpha$ and $\widetilde{h}_t \geq 0$).

Any path starting at (h_0, x_0) above the proposed policy function satisfying (3) violates the transversality condition. For such a path (h_t, x_t) tends to (∞, ∞) with $\lim_{t\to\infty}\frac{h_t}{x_t} \leq \frac{\gamma}{\delta}$. Thus $\lim_{t\to\infty}\frac{\dot{x}}{x} = \frac{(a\rho-\kappa\gamma)}{\alpha x_t} + \rho + \frac{a\delta-b\gamma}{\alpha}\lim_{t\to\infty}\frac{h_t}{x_t} + \delta\lim_{t\to\infty} h_t = \infty$ and hence $\lim_{t\to\infty} e^{-\rho t}\frac{\alpha x_t}{\gamma} = \infty$. (Note that this follows so easily only because of the quadratic depreciation, $\delta > 0$. Without this assumption we had to require that γ is not too large or that α is sufficiently large to guarantee bounded dynastic utility).

Next consider any path starting at (h_0, x_0) below the proposed policy function satisfying (3). For such a path (h_t, x_t) tends to $(0, -\infty)$. $v(x)$ takes its minimum at $x = -\frac{a}{x}$ such that any reduction of x below $-\frac{a}{x}$ not only reduces h an thus future utility but at the same time directly reduces instantaneous utility. Clearly this cannot be optimal. Thus the household will never choose $x < -\frac{a}{x}$ and any path starting at (h_0, x_0) below the proposed policy function is excluded. One can show that this argument is also reflected in a violation or the transversality condition (To induce the household to keep reducing x and h the shadow price of h has to tend to $-\infty$ sufficiently fast).

Concavity One way to proof that the proposed solution does in fact solve the maximization problem (1) is to show that it satisfies the Arrow-Kurz sufficient condition. This condition requires that $\widetilde{H}(h,\lambda) := \max_x H(x,h,\lambda)$ be concave w.r.t. h for all λ (resulting from the necessary conditions). With $\widetilde{H}(h,\lambda) = m(wh) + bh - \left[a\frac{\lambda\gamma h-a}{\alpha} + \frac{\alpha}{2}\left(\frac{\lambda\gamma h-a}{\alpha}\right)^2\right] + \xi(G) + \lambda\cdot(\gamma\frac{\lambda\gamma h-a}{\alpha} - \delta h)h$, the condition becomes $\widetilde{H}_{hh} = m_{hh} - \frac{(\lambda\gamma)^2}{\alpha} + 2\lambda\cdot(\lambda\frac{\gamma^2}{\alpha} - \delta) = -\frac{\kappa}{h^2} - 2\delta\lambda + \frac{(\lambda\gamma)^2}{\alpha} < 0$ and is satisfied iff $\frac{\kappa}{\lambda h^2} + 2\delta > \frac{\gamma^2}{\alpha}\lambda$ or, inserting the FOC, iff $\frac{\kappa}{h} + 2\delta\frac{a+\alpha x}{\gamma} > \frac{(a+\alpha x)^2}{\alpha h}$. Since $a + \alpha x \geq 0$ it is sufficient that $\alpha\kappa > (a+\alpha x)^2$ or $x < \sqrt{\frac{\kappa}{\alpha}} - \frac{a}{\alpha}$.

We know that $x < x_{\dot{x}=0}(\infty) = \frac{\gamma b}{\alpha\delta} - \frac{a}{\alpha}$ along along each path satisfying the necessary conditions. Thus the sufficient condition is satisfied if $\frac{\gamma b}{\alpha\delta} < \sqrt{\frac{\kappa}{\alpha}}$ or if α is sufficiently large: $\alpha > \left(\frac{\gamma b}{\delta}\right)^2\frac{1}{\kappa}$. The parameter a is still free to generate the three cases.

In particular Case 2 requires that $\kappa < a\rho/\gamma$ and $b > \frac{2\delta\sqrt{\alpha\gamma(a\rho-\gamma\kappa)}+\alpha\rho\delta+a\delta\gamma}{\gamma^2}$. Thus Case 2 satisfies the Arrow-Kurz conditions if

$$\sqrt{\frac{\alpha a\rho}{\gamma}} \overset{\text{Case 2}}{>} \sqrt{\alpha\kappa} \overset{\text{Concavity}}{>} \frac{\gamma b}{\delta} \overset{\text{Case 2}}{>} \frac{2\sqrt{\alpha\gamma\,(a\rho-\gamma\kappa)} + \alpha\rho + a\gamma}{\gamma}$$

To simultaneously satisfy these constraints, take any α, a, ρ, δ. Choose $\kappa < a\rho/\gamma$. Given $\alpha, a, \rho, \delta, \kappa$ choose $\overline{\gamma b} < \delta\sqrt{\alpha\kappa}$. Given $\alpha, a, \rho, \delta, \kappa$, $\overline{\gamma b}$ choose $\overline{\gamma b} > \frac{2\sqrt{\alpha\gamma(a\rho-\gamma\kappa)}+\alpha\rho+a\gamma}{\overline{\gamma b}\cdot\gamma}$ (note that either γ or b can take any value).

Wirl and Feichtinger [2005] show that if the unstable stationary solution lies in the concave domain of the Hamiltonian, then this stationary solution is a node of the dynamic system and the policy function $x(h)$ is continuous across all three stationary states as drawn in Figure 2.

More general sufficient conditions Hartl et al. [2004] show that concavity of the Hamiltonian at the unstable stationary state is not necessary for the continuity of the policy function. The policy function $x(h)$ is continuous across all three stationary states as in the concave case if the unstable stationary state is a node, which is the case if the Eigenvalues of $J(h^{\text{th}})$ are real. Even if the Arrow-Kurz sufficient condition is not satisfied, the phase diagram is qualitatively the same as in the completely concave case. (Sufficiency follows from the fact that only the proposed solution path satisfies all the necessary conditions (including the transversality conditions, which is not necessary to apply the sufficient condition above) and from the fact that the general conditions for existence of a piecewise continuous solution are satisfied).

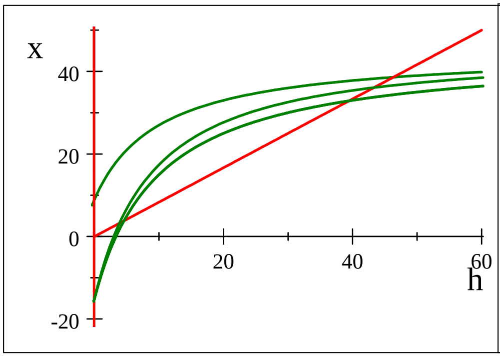
It has been shown that the stationary solutions $h_* = 0$ and $h^* > 0$ (whenever they exists) are saddle-point stable ($\det J(h^{\text{th}}) < 0$) and that h^{th} (whenever it exists, thus in Case 2) is unstable ($\det J(h^{\text{th}}) > 0$). The unstable stationary state is a node if the Eigenvalues of $J(h^{\text{th}})$ are real.

At $\dot{h} = 0$ and $h \neq 0$ and $\dot{x} = 0$: $\det \begin{pmatrix} \frac{\partial \dot{h}}{\partial h} - \xi & \frac{\partial \dot{h}}{\partial x} \\ \frac{\partial \dot{x}}{\partial h} & \frac{\partial \dot{x}}{\partial x} - \xi \end{pmatrix} = \det \begin{pmatrix} -\delta h - \xi & \gamma h \\ \frac{(a+\alpha\frac{\delta}{\gamma}h)\delta - b\gamma}{\alpha} & \rho + \delta h - \xi \end{pmatrix} =$ $-\delta h\rho - 2\delta^2 h^2 - \xi\rho + \xi^2 - \gamma h\frac{a\delta}{\alpha} + \gamma h\frac{b\gamma}{\alpha} = 0$ for $\xi_{1,2} = \frac{1}{2}\rho \pm \frac{1}{2}\sqrt{\frac{1}{\alpha}\left(8\alpha h^2\delta^2 - 4bh\gamma^2 + 4ah\gamma\delta + 4\alpha h\delta\rho + \alpha\rho^2\right)}$. The Eigenvalues $\xi_{1,2}$ are real if $8\alpha h^2\delta^2 + 4ah\gamma\delta + 4\alpha h\delta\rho + \alpha\rho^2 > 4bh\gamma^2$ or if $h > \frac{\left(\frac{b\gamma - a\delta}{\alpha} - \frac{\delta}{\gamma}\rho\right)}{2\frac{\delta^2}{\gamma}} - \frac{\rho^2}{8\delta^2 h} = \frac{B}{2D} - \frac{\rho^2}{8\delta^2 h}$. At h^{th} the condition becomes $h^{\text{th}} = \frac{B-\sqrt{B^2+4DC}}{2D} > \frac{B}{2D} - \frac{\rho^2}{8\delta^2 h^{\text{th}}}$ or $\frac{\sqrt{B^2+4DC}}{2D} < \frac{\rho^2}{8\delta^2\frac{B-\sqrt{B^2+4DC}}{2D}}$ or $B\sqrt{B^2+4DC} - \left(B^2+4DC\right) < \frac{\rho^2}{2\gamma}D$. Since $C < 0$ in Case 2, $\sqrt{B^2+4DC} < B$, such that it is sufficient that $B^2 - \left(B^2+4DC\right) < \frac{\rho^2}{2\gamma}D$ or $\frac{\rho a - \kappa\gamma}{\alpha} < \frac{\rho^2}{8\gamma}$.

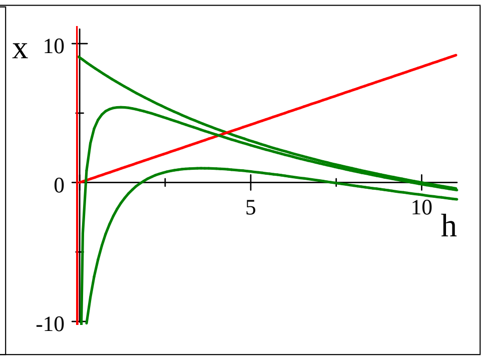
Since $\rho a - \kappa\gamma > 0$, this inequality is satisfied if α is large, $\alpha > (\rho a - \kappa\gamma)\frac{8\gamma}{\rho^2}$. (Remember that apart from $\kappa < a\rho/\gamma$ Case 2 requires that $b > \frac{2\delta\sqrt{\alpha\gamma(a\rho-\gamma\kappa)}+\alpha\rho\delta+a\delta\gamma}{\gamma^2}$, such that given α, b can always be chosen sufficiently large to satisfy Case 2).

8.3 Appendix: Simple Transfers

Apart from the stationary state at $h_* = 0$ the new dynamic system (17) has stationary states for $\frac{\gamma}{\alpha}\frac{\kappa\frac{1}{1+\frac{Z}{wh}}+bh}{\rho+\delta h} - \frac{a}{\alpha} = \frac{\delta}{\gamma}h$ or for $Q(h,Z) = \frac{\delta^2}{\gamma}h^3 - \left[B - \frac{\delta^2}{\gamma}\frac{Z}{w}\right]h^2 - \left[C + B\frac{Z}{w}\right]h + \rho\frac{a}{\alpha}\frac{Z}{w} = 0$ with as before $C = \frac{\kappa\gamma - \rho a}{\alpha}$, $B = \left(\frac{b\gamma - a\delta}{\alpha} - \frac{\delta}{\gamma}\rho\right)$ and $D = \frac{\delta^2}{\gamma} > 0$. Note that for $Z = 0$: $Q(h, 0) = E(h, 0) = \left(\frac{\delta^2}{\gamma}h^2 - Bh - C\right)h$. The polynomial function of degree 3 has up to three strictly positive real roots.



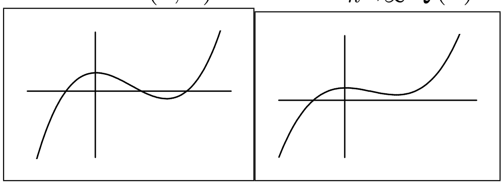
Higher T, lower $x_{\dot{x}=0}$-isocline with discontinuity at $T = 0$.(Case 1a)



Higher T, lower $x_{\dot{x}=0}$-isocline with discontinuity at $T = 0$.(Case 1b)

Lemma 2 *If $Z > 0$, then $Q(h)$ has either no or two strictly positive roots.*

Proof. Since the coefficient $D = \frac{\delta^2}{\gamma}$ of the cubic term is positive $\lim_{h\to\infty} Q(h) = \infty$ and $\lim_{h\to-\infty} Q(h) = -\infty$. Furthermore $Q(0) = \rho\frac{a}{\alpha}\frac{Z}{w} > 0$. From $\lim_{h\to-\infty} Q(h) = -\infty$ and $Q(0) > 0$ follows that one root must be negative. If $Q(\overline{h}) < 0$ for some $\overline{h} > 0$, then there must be a second root in the interval $(0, \overline{h})$. Since $\lim_{h\to\infty} Q(h) = \infty$ there must be a third root at some $h > \overline{h}$.


■

Lemma 3 *If $C > 0$ and $B < \sqrt{4\frac{\delta^2}{\gamma}\frac{a\rho}{\alpha}}$, then a small positive transfer turns Case 1 into the threshold Case 2, while a sufficiently large transfer turns Case 1 into Case 3. If $C > 0$ and $B > \sqrt{4\frac{\delta^2}{\gamma}\frac{a\rho}{\alpha}}$, then any transfer turns Case 1 into Case 2. The threshold human capital level h^{th} of Case 2 depends positively on the transfer and the strictly positive stable stationary h^* depends negatively on the transfer.*

Proof. $\lim_{h\to 0} x_{\dot{x}=0}(Z, h) = \frac{\gamma}{\alpha}\frac{\kappa\frac{1}{1+\frac{Z}{w\lim h}} + bh}{\rho + \delta h} - \frac{a}{\alpha} = -\frac{a}{\alpha}$ for all $Z > 0$. At $Z = h = 0$ the effect of raising Z is dominated by the effect of the constant term, which raises $Q(0, Z)$. Thus starting from Case 1 ($Q(h, 0) = 0$ has one strictly positive root) $Q(h, Z) = 0$ has two strictly positive solutions for small positive Z (Case 2).

$\lim_{Z\to\infty} Q(h, Z)\frac{w}{Z} = \frac{\delta^2}{\gamma}h^2 - Bh + \rho\frac{a}{\alpha} = 0$ has roots $h^{\infty}_{1,2} = \frac{B \pm \sqrt{B^2 - 4D\rho\frac{a}{\alpha}}}{2D}$. If $B > \sqrt{4D\rho\frac{a}{\alpha}}$ then both roots are positive. The household problem remains in the threshold Case 2 even for very large

Z. If $B < \sqrt{4\frac{\delta^2}{\gamma}\frac{a\rho}{\alpha}}$ then there is no real root. The household problem is in Case 3 for sufficiently large Z. ■

8.4 Appendix: Constraint first best and its implementation

There are n_* poor households and $n^* = 1 - n_*$ rich households. The rich coordinate on the optimal amount of G_t, assuming equal contributions for the rich and no contributions to the poor. In addition, the rich have to pay a transfer Z to each poor $h_t^{\text{u}} < \widetilde{h}$. Each rich thus contributes $\frac{G_t}{n^*}$ to the founding of the public good and $\frac{1-n^*}{n^*}G_t$ to the social security fund. This section determines the optimal amount of G_t for the n^* rich under the constraint that each rich household pays $\frac{1-n^*}{n^*}Z$ and also determines the tax system that implements this constraint optimum.

An inner solution to $\max_G H(x,h,G,\lambda)$ requires $-m_c + \xi = 0$, $m_c = \xi$, or, with $c = wh - \frac{1-n^*}{n^*}Z - \frac{1}{n^*}G_t$, $m(c) = \kappa \ln c$, $\frac{\kappa}{wh - \frac{1-n^*}{n^*}Z - \frac{1}{n^*}G_t} = \xi$ or $G_t = n^*\left[\left(wh - \frac{1-n^*}{n^*}Z\right) - \frac{\kappa}{\xi}\right]$. This inner solution is only relevant for $G_t \geq 0$ or for $h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$. Thus

$$G_t = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ n^*\left[\left(wh - \frac{1-n^*}{n^*}Z\right) - \frac{\kappa}{\xi}\right] = n^*\left[wh - \frac{\kappa}{\xi}\right] - (1-n^*)Z \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}. \end{cases} \tag{20}$$

$$c_t = \begin{cases} wh - \frac{1-n^*}{n^*}Z \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ wh - \frac{1-n^*}{n^*}Z - \frac{1}{n^*}n^*\left[\left(wh - \frac{1-n^*}{n^*}Z\right) - \frac{\kappa}{\xi}\right] = \frac{\kappa}{\xi} \text{ if } h_t \geq h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}. \end{cases} \tag{21}$$

The general form of the equations of motion (3) and of the isoclines (4) are unchanged. With (10) we now have

$$m_h = \frac{\kappa w}{wh - \frac{1-n^*}{n^*}Z - \frac{1}{n^*}G_t} = \begin{cases} \frac{\kappa}{h} \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ w\xi \text{ if } h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \end{cases} \tag{22}$$

and thus the dynamic system is exactly the same as in the unconstraint egalitarian optimum, except that the kink in $x^{\text{sym}}_{\dot{x}=0}(h)$ occurs at $\frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$ rather then $\frac{\kappa}{w\xi}$:

$$\begin{cases} \dot{x} = 0 \text{ if } \begin{cases} x^{\text{sym}}_{\dot{x}=0}(h) = \frac{\gamma}{\alpha}\frac{\kappa + bh}{\rho + \delta h} - \frac{a}{\alpha} \text{ for } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ x^{\text{sym}}_{\dot{x}=0}(h) = \frac{\gamma}{\alpha}\frac{(\xi w + b)h}{\rho + \delta h} - \frac{a}{\alpha} \text{ for } h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \end{cases} \\ \dot{h} = 0 \text{ if } h = 0 \text{ or } x = \frac{\delta}{\gamma}h. \end{cases} \tag{23}$$

To determine which tax-system implements the constraint optimum I solve

$$\max_{\{x_t\}_{t=0}} \int_0^\infty e^{-\rho t} u(c_t, x_t, h_t, G_t) dt \tag{24}$$

$$\begin{aligned} \text{subject to } \dot{h}_t &= (\gamma x_t - \delta h_t) h_t \text{ for all } t \geq 0, \\ c_t &= w h_t - T(h_t^i, \overline{h}_t) \text{ for all } t \geq 0, \\ &\text{given } h_0,\ \{\overline{h}_t\}_{t=0} \text{ and } \overline{T}_t = \max\left\{0, \frac{G_t}{n^*} + \frac{1-n^*}{n^*}\frac{Z}{w}\right\} \end{aligned}$$

where

$$G_t = \begin{cases} wh_t \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ n^*\left[wh - \frac{\kappa}{\xi}\right] - (1-n^*)Z \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}. \end{cases} \tag{25}$$

$$c_t = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ = wh_t - \left[\frac{n^*\left[wh-\frac{\kappa}{\xi}\right]-(1-n^*)Z}{n^*} + \frac{1-n^*}{n^*}\frac{Z}{w}\right] = \frac{\kappa}{\xi} \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}. \end{cases} \tag{26}$$

The solution is the same as in the unconstraint world. As there the Hamiltonian becomes $H(x, h, \lambda) = m[wh - T(h, \overline{h})] + bh - v(x) + \xi w\overline{h} + \lambda \cdot (\gamma x - \delta h)h$. The optimal dynamic system and its isoclines (12) are replicated iff the marginal utility of consumption $m_h = \frac{\kappa(w-T_h(h,\overline{h}))}{wh-T(h,\overline{h})}$ replicates that of the constraint optimal solution (Equation), thus iff

$$m_h = \frac{\kappa(w - T_h(h, \overline{h}))}{wh - T(h, \overline{h})} = \begin{cases} \frac{\kappa}{h} \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ \xi w \text{ if } h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \end{cases} \tag{27}$$

$T(h, \overline{h}) = 0$ if $h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$ and for $h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$: $T_h = \frac{\xi w}{\kappa}\left[\frac{\kappa}{\xi} - wh + T(h, \overline{h})\right]$. This is the same differential equation except for the threshold h. As before the solution is

$$T(h, \overline{h}) = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ wh + I_t \cdot e^{\frac{w\xi}{\kappa}h} \text{ if } h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \end{cases} \tag{28}$$

where I_t is a constant of integration. I_t is determined by the equilibrium condition which now is, for $h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$, $T(h_t, \overline{h}_t) = \frac{G_t}{n^*} + \frac{1-n^*}{n^*}\frac{Z}{w} = \frac{n^*\left[wh-\frac{\kappa}{\xi}\right]-(1-n^*)Z}{n^*} + \frac{1-n^*}{n^*}\frac{Z}{w} = wh_t - \frac{\kappa}{\xi}$.Thus $wh_t + I_t \cdot e^{\frac{w\xi}{\kappa}h} = wh_t - \frac{\kappa}{\xi}$ or $I_t = -\frac{\kappa}{\xi}e^{-\frac{w\xi}{\kappa}h}$ for $h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$ as before $w\overline{h}_t + I_t \cdot e^{\frac{w\xi}{\kappa}\overline{h}_t} = w\overline{h}_t - \frac{\kappa}{\xi}$, hence $I_t = -\frac{\kappa}{\xi}e^{-\frac{w\xi}{\kappa}\overline{h}_t}$ and

$$T(h, \overline{h}) = \begin{cases} 0 \text{ if } h_t \leq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w} \\ wh - \frac{\kappa}{\xi}e^{\frac{w\xi}{\kappa}(h-\overline{h})} \text{ if } h_t \geq \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}. \end{cases} \tag{29}$$

The tax is exactly the same as the one implementing the unconstraint optimum. However part of the tax is now spent to finance the transfer $(1 - n^*)Z$ and only the rest is used to finance G_t. The general intuition behind these results is that due to the hierarchy of needs and the linearity of u in both h and G it is always optimal to first satiate the basic needs, which is achieved at $c = \frac{\kappa}{\xi}$ and to spend all remaining disposable income on the public good. For this result the composition of the disposable income (wh or $wh - \frac{1-n^*}{n^*}Z$) it is irrelevant.

Since apart from the location of the kink ($h_t = \frac{\kappa}{w\xi} + \frac{1-n^*}{n^*}\frac{Z}{w}$) the phase diagram does not depend on Z, the levels of $h^{\text{th}}_{\text{sym}}$ and h^*_{sym} do not depend on Z.

References

[Becker, 1965] Becker, Gary S. 1965. A Theory of the Allocation of Time. *Economic Journal*, **75**, 493–517.

[Becker, 1993] Becker, Gary S. 1993. *Human Capital.* third edn. The University of Chicago Press.

[Chen, 2007] Chen, Been-Lon. 2007. Multiple BGPs in a Growth Model with Habit Persistence. *Journal of Money, Credit and Banking*, **39**(1), 25–48.

[Hartl *et al.* , 2004] Hartl, R.F., Kort, P.M., Feichtinger, G., & Wirl, F. 2004. Multiple Equilibria and Thresholds Due to Relative Investment Costs. *Journal of Optimization Theory and Applications*, **123**(1), 49–82.

[Kurz, 1968] Kurz, Mordecai. 1968. Optimal Economic Growth and Wealth Effects. *International Economic Review*, **9**(3), 348–357.

[Laroque, 2005] Laroque, G. 2005. Income Maintenance and Labor Force Participation. *Econometrica*, **73**, 341–376.

[Michael, 1973] Michael, Robert T. 1973. Education in Nonmarket Production. *Journal of Political Economy*, **81**, 306–327.

[Mirrlees, 1971] Mirrlees, J. A. 1971. An Exploration of the Theory of Optimal Taxation. *The Review of Economic Studies*, **38**, 175–208.

[Moffitt, 2006] Moffitt, R. 2006. Welfare Work Requirements with Paternalistic Government Preferences. *The Economic Journal*, **116**, F441–F458.

[Paul Beaudry & Szalay, 2007] Paul Beaudry, Charles Blackorby, & Szalay, Dezsö. 2007. *Taxes and Employment Subsidies in Optimal Redistribution Programs.* DP.

[Sen, 1997] Sen, Amartya. 1997. Editorial: Human Capital and Human Capability. *World Development*, **25**, 1959–1961.

[Wirl & Feichtinger, 2005] Wirl, Franz, & Feichtinger, Gustav. 2005. History Dependence in Concave Economies. *Journal of Economic Behavior Organization*, **57**, 390–407.