Educational Policy and the Economics of the Family

Abhijit V. Banerjee*

December 8, 2000

*I am grateful to Esther Duflo, Paras Mehta, Antonio Rangel, Debraj Ray and particularly T.N. Srinivasan for their comments.

1 Introduction

The family plays a crucial role in decisions about investment in education, at least in developing countries. The reasons are simple: education is expensive and there is only a very limited scope for borrowing in order to invest in education. Education is expensive because free and functioning public schools do not exist in many places and even when they do they are often free only in name (The Probe Report, 1999, Kochar, 2000). One cannot borrow to invest in education because human capital provides very poor collateral, and in any case credit markets function very poorly in developing countries. Most of the financial investment in education therefore has to be funded by the family.

As a result, most theoretical and empirical studies of investment in education have to take a stand on how they think the family makes its decisions. This is an issue especially with respect to decisions about education that are taken when the children are young and parents are, at least nominally, in charge of financial decisions for the family. There is, however, very little agreement about how families make these decisions. Even a cursory glance at the literature on investment in education reveals a wide variety of views ranging from the Barro-Becker model (Becker, 1981; Loury, 1981; Mookherjee and Ray, 2000) where parents and children share a single unified utility function, to the pure overlapping generation model, where the family is purely a nexus for transactions—the old "lend" to their children who repay them with old age care (Kotlikoff and Spivak, 1981; Cox, 1987; Cox, 1990; Cremer, Kessler and Pestieau, 1992; Cox and Jakubson, 1995, Barham et al, 1995). There is also a range of formulations that are inbetween these two—various forms of imperfect altruism where the parents put some weight on the income, consumption or human capital of their children[1] as well as "warm-glow" bequest preferences, where the parent values what she is giving and not what the child is getting.[2]

However, to the best of our knowledge, there have been no systematic attempts to understand whether and in what way these things matter—in other words, what are the positive and normative implications of making each of these alternative assumptions. The goal of this paper is to fill a part of this gap in the literature. We present a simple one-good, two-sector growth

[1] See Behrman, Pollak and Taubman (1982), and Ermisch and Francesconi (2000) for examples of formulations of this class.

[2] See Glomm and Ravikumar (1992), Galor and Zeira (1993), Banerjee and Newman (1994), and Galor and Moav (1999).

model where goods are produced using skilled and unskilled labor. Everyone is endowed with a unit of unskilled labor, but skill has to be accumulated by investing in human capital. We assume that there are no credit markets, and investment in human capital is either financed by the state[3] or by the parents of the child who is getting the human capital—this seems to be a reasonable approximation to the situation in most less-developed countries. In addition, we make the plausible and standard[4] assumption that the production of skills is more intensive in human capital than the production of goods. Finally, we assume that the production function for skills exhibits decreasing returns and that there are no human capital externalities. We recognize that each of these last assumptions is potentially controversial. They are made mainly for strategic reasons. The assumption of non-convexities in the educational production function has been shown by Galor and Zeira (1993) to have a number of strong implications when combined with the assumption of imperfect credit markets—in particular, there are poverty traps and public investment in education can increase steady state output. Moreover, they argue that their results are independent of the way we model parental preferences as long as the non-convexity is large enough. Since we are interested in the differences between the implications of different types of parental preferences, avoiding non-convexities clearly helps. We do, however, take comfort in the fact that the evidence for non-convexities in education is rather mixed. Psacharapoulus (1994) argues on the basis of evidence from a number of countries that the returns are highly concave. On the other hand, a couple of detailed studies from the U.S. find evidence of small non-convexities (Angrist and Krueger, 1999; Card and Krueger, 1992). Likewise, the decision not to allow for human capital externalities was driven in part by the fact that the consequences of such externalities have been widely studied and in part by the fact that there seems to be very little micro-evidence for the presence of large externalities of this type.[5]

We study this model under four different assumptions about family decision-making. At one extreme is what we call true altruism (the so-called "Barro" preferences). At the other extreme is the case we have labeled perfect functional altruism: In this model the family is just a contractual relation ("people only care about their children because children take care of them

[3] Where the state could be the local government.

[4] This assumption goes back at least to Uzawa (1965). Rebelo (1991), in his well-known study of education policy in growth models also makes this assumption.

[5] See Acemoglu and Angrist (1999) and Duflo (2000). Bils and Klenow (1998) make a case against educational externalities using cross-country data.

in old age"), and we assume that there are no restrictions on the possibilities for contracting between the generations. Of the two intermediate cases, one is a model of limited altruism: parents get "warm-glow" utility from the amount they invest in their children. The other is a model of limited functional altruism: the family relation is purely contractual but the set of enforceable contracts is quite limited.

Because we assume that there are no credit markets, all these models look rather similar in their short predictions. In all of them, a dynasty that has more education today will have more education tomorrow. This is because there is an income effect on the demand for education—more educated parents are richer and therefore buy more education for their children. Because of these income effects, even lump-sum taxes and subsidies will have an effect on educational investment in the short run.

While the data supports these predictions (see Jacoby, 1994; Glewwe and Jacoby, 2000; and Carvalho, 2000, for example) this type of evidence obviously cannot help us discriminate between the four alternative models. We therefore focus on the long run predictions of the models: under each of the assumptions about family decision making, we study the long-run properties of the resulting economy (is there long-run inequality? is the steady state unique? etc.) as well as the effects of a number of alternative policies including lump-sum and proportional taxes and subsidies and changes in the returns to human capital.

The results reveal an intriguing symmetry. The two polar models, which represent true altruism (the so-called "Barro" preferences) and purely functional altruism ("people only care about their children because children take care of them in old age"), turn out to have very similar long-run properties—there is no inequality in the long-run and the steady state is unique. Lump-sum taxes and subsidies have no long-run effects, while a proportional tax on human capital reduces human capital investment even if it is then redistributed as a lump-sum educational subsidy. On the other hand, policies that increase the rate of return on human capital raise investment. Finally, a proportional subsidy to education spending raises investment in both these models. However, one key difference between the models is that in the true altruism case this increase in investment reduces the welfare of the representative agent, while in the functional altruism case, since there is not a representative agent, it is not possible to make a clear welfare statement.

The two intermediate models also turn out to deliver relatively similar results, though these results are very different from what we learned from the polar models. One of these models is a model of limited altruism, where there is a warm-glow bequest motive. The other is a model of limited functional altruism which differs from the earlier model because of incomplete contracting between parents and children. Both of these models generate the possibility of long-run inequality and multiplicity of steady states, for related but distinct reasons. Both models also raise the possibility that lump-sum taxes and subsidies would have long-run effect. In the warm-glow model, and perhaps in the other model as well, proportional taxes on human capital may sometimes raise average investment and increases in the returns to human capital may reduce investment by making teachers expensive.

The similarity between the results from the polar models is not accidental. In the model of perfect functional altruism, the fact that there are no restrictions on the contracting possibilities within the family, makes sure that in equilibrium the interests of all generations are aligned. In the true altruism case there are no conflicts of interest to begin with. By contrast, two intermediate models are chosen to capture two important reasons why there may be a conflict of interest even in equilibrium. In the warm-glow model the parent is completely unresponsive to what the child actually wants—to take a not unfamiliar example, the parent is willing to pay to send the child to an expensive school but not to give the child the money instead. Basically since the parent is only interested in giving along this one dimension and expects nothing in return, the child has no way of persuading the parent to do what she wants.[6] In the limited functional altruism model, by contrast, parents and children do want to cut a deal, but the problem is enforceability.[7]

The question of whether we should expect the interests of parents and children to be aligned, goes back at least to Becker's famous Rotten Kid Theorem (Becker, 1981). Our main contri-

[6] Things would be different, for example, if the parent was planning to leave the child some cash in addition to the expenditure on schooling. Then the child could bargain with the parent along the lines—if you just gave me a little bit more cash, then you could save yourself the entire expense of sending me to that school where I know I will miserable. But for the average parent in an LDC, it may be reasonable to assume that all that parents ever plan to give their child is an education so that this kind of bargaining is ruled out.

[7] We use the two intermediate models to illustrate two reasons why the interests of generations may not be aligned. The interpretation we give them is meant only to be suggestive. For example, one could imagine warm glow preferences which take exactly the form of our limited functional altruism case and conversely, one form of limited functional altruism may be that children are obliged to take care of their parents in proportion to how much their parents have spent on their education (it may be this is all that is observable to society at large).

bution is to connect these issues to questions about the long-run effects of education policy. The gap between the answers we get about education policy from the different models is rather substantial. The fact that there is no long-run inequality in the two polar models amounts to ruling out poverty traps; the uniqueness of the steady state in those models says that even if there is a shock which wipes out a huge part of a country's human capital (as AIDS seems to be in the process of doing in sub-Saharan Africa), the country will recover.[8] The fact that a lump-sum educational subsidy financed by a proportional tax on human capital always reduces investment is a strong argument for *laissez-faire,* while the opposite conclusion lends support to a more interventionist stance. That higher returns on human capital promote more investment in the polar models corresponds to the intuition people have from models with perfect credit markets. It therefore lends support to the so-called demand-side view of education—the view that the reason why people in developing countries have stayed away from schools is because these countries have not invested enough in raising the returns to education.[9] The intermediate models suggest a more sceptical view: not that one should be opposed to new technologies and other policies that make education more rewarding, but that such policies may not be adequate in themselves and may need to be complemented by other supply-side policies.

The paper begins by laying out the basic model and examining the properties of the model under the assumption of perfect credit markets. The fourth section is the heart of the paper—the alternative models of the family decision-making are laid out and analyzed in a setting where there are no credit markets. We conclude with a discussion of the alternative models and what separates them.

2 The Basic Model

Consider a world where there is only one final good but two types of human inputs—skill and unskilled labor. The final good is produced by combining the two types of inputs using a technology:

$$y = f(H, L, \alpha)$$

[8] This assumes that the unique steady state is an attractor, which is of course not always true.

[9] See for example Foster and Rosenzweig (2000).

where H and L are, respectively, the amount of skill and the amount of unskilled labor used in production, and α is a parameter describing the nature of the technology. We assume that the production function exhibits constant returns with respect to the two inputs together but diminishing returns with respect to each individual input. This formulation also imposes the restriction that all levels of human capital are perfect substitutes—one of the insights of the paper by Mookherjee and Ray, cited above, is that relaxing this assumption introduces an additional source of persistent inequality.

Each person in the economy is assumed to own one unit of unskilled labor. In addition they own a certain number of units of skills, which correspond to the amounts they invested in human capital.

Agents in this economy live for three periods. In the first period of their lives there is no consumption or work: all they do is acquire skills. The second period is when people work and earn an income. This income can be spent on current consumption or saved and invested for consumption in the next period, when they no longer have wage income. In the second period of their lives they also give birth to exactly one child. The population therefore stays unchanged over time and each cohort is assumed to be size one.

Human capital is produced using a combination of human capital and unskilled labor. The cost of acquiring h units of human capital is $\gamma s(h, h^-)$ units of human capital and $(1-\gamma)s(h, h^-)$ units of unskilled labor, where $0 \leq \gamma \leq 1$ and h^- is the level of human capital of the parent of the person who is acquiring the human capital. We assume that $s(0, h^-) = 0$, $\frac{\partial s}{\partial h} > 0$, $\frac{\partial^2 s}{\partial h^2} > 0$, $\frac{\partial s}{\partial h^-} < 0, \frac{\partial^2 s}{\partial h^{-2}} > 0$, and $\frac{dS(h)}{dh} > 0$, where $S(h) = s(h, h)$. The second and third assumptions tell us that the cost function is increasing and convex, therefore ruling out non-convexities. The next two assumptions capture the idea that the family atmosphere matters—children of skilled parents find it easier to acquire skills—though at a diminishing rate. The last assumption imposes the condition that the increased cost from increasing h is not outweighed by the reduction in cost coming from the fact that the next generation will now have more education and therefore find it easier to educate their children.

Throughout this paper we will make the assumption of perfectly competitive markets for labor and skills, with the price of labor at time t being w_t^L and that of skill w_t^H. Except in the next section we will make the extreme assumption that there are no capital markets and no assets other than human capital.

Finally in terms of policy instruments, we assume that there is an educational subsidy of e_0+e_1E, where $E=(\gamma w_t^H+(1-\gamma)w_t^L)s(h_{t+1},h_t)$ is the amount the family spends on education. This is financed partly or entirely by a tax on the earning members of society that is partly lump-sum and partly a function of the taxpayer's human capital: $T=\tau_0+\tau_1 h_t w_t^H$: Lump-sum taxes are of course standard while the tax on human capital earnings will turn out to be a simple way to introduce redistributive taxation. In order to get sharper results and limit the number of cases, we will, for the most part, focus on the case where we start from a situation where the family was already spending some amount on education and then look at the effects very small changes in the taxes and subsidies. This allows us to avoid the issue of corner solutions.[10]

3 The Benchmark: Altruistic Families and Perfect Credit Markets

In this section I establish a benchmark for the results in the next section by examining what happens under the standard assumptions of perfect altruism and perfect credit markets. I adopt the standard formulation where perfect altruism is represented by a single utility function for the entire dynasty:

$$\sum_{t=0}^{\infty}\delta^t u(c_t).$$

The instantaneous utility function $u(c_t)$ is assumed to be increasing and concave and c_t is interpreted as the total consumption of all the generations present at time t.

Perfect credit markets amount to assuming that anyone can borrow and lend as much as they want at the market interest rate r. Under this assumption, each dynasty faces the intertemporal budget constraint

$$\omega_{t+1}=r_t(\omega_t-c_t-(\gamma w_t^H+(1-\gamma)w_t^L)s(h_{t+1},h_t)(1-e_1)+w_t^L+h_t w_t^H(1-\tau_1)+e_0-\tau_0), \tag{1}$$

[10] This formulation does however come with the cost that there is no way to distinguish between educational subsidies and general income subsidies since even if the subsidy were earmarked for education, the family could always cut back on what it was already spending on education.

where ω_t is the starting wealth of the tth generation. The term $(\gamma w_t^H + (1-\gamma) w_t^L) s(h_{t+1}, h_t)(1-e_1)$ represents the investment in the next generation's human capital measured in units of the good. Maximizing the above utility function under this set of constraints gives us the first-order conditions:

$$u'(c_t) = \delta r_t u'(c_{t+1}) \tag{2}$$

$$\begin{aligned} & r_t(\gamma w_t^H + (1-\gamma) w_t^L) s_1(h_{t+1}, h_t)(1-e_1) \\ & w_{t+1}^H(1-\tau_1) - (\gamma w_{t+1}^H + (1-\gamma) w_{t+1}^L) s_2(h_{t+2}, h_{t+1})(1-e_1). \end{aligned} \tag{3}$$

It is apparent from these conditions that, not surprisingly, there are no income effects in this economy—the amount invested in education does not depend on parental wealth. However, there can be a parental human capital effect which in the data may look like a wealth effect: for example, if $s_{12} < 0$, an increase in h_t, keeping h_{t+2} fixed, lowers the marginal cost of investing in education (without affecting the benefits) and therefore raises h_{t+1}.[11]

As I say in the introduction, the focus here is on what happens in the steady state. In steady state $r_t = r_{t+1} = r$, $c_t = c_{t+1}$, $w_t^H = w_{t+1}^H = w^H$,$w_t^L = w_{t+1}^L = w^L$ and $h_t = h_{t+1} = h_{t+2} = h$ so that the above conditions reduce to one key condition:

$$s_1(h,h) + \delta s_2(h,h) = \delta \frac{(1-\tau_1)/(1-e_1)}{\gamma + (1-\gamma) w^L/w^H}. \tag{4}$$

We are now in a position to ask a number of questions about the nature of the steady state.

Inequality: Is the steady state perfectly egalitarian or it possible inequality persists in the sense that different dynasties within the same economy converge to different steady state levels of human capital? Or in other words, are there multiple steady states for a fixed value of w^L/w^H?[12]

The argument, made above, suggesting that the children of educated parents get more education, tells us that such multiplicity may be possible and indeed this is the case if δ is small enough and the ratio s_{12}/s_{11} is positive and large enough. However i*n the rest of this paper we*

[11] Actually, the amount of human capital in all generations will go up. The increase in h_{t+1} causes h_{t+2} to go up, which causes $s_2(h_{t+2}, h_{t+1})$ to go down, which encourages further increases in h_{t+1}, etc.

[12] We also have to check that the value of w^L/w^H so generated is consistent with the amount of human capital that we would have in the steady state, but we can always choose the production function to make that fit.

focus on the case where such multiplicity cannot arise, namely the case where $s_{12} = 0$. There are two reasons why we make this choice: first Galor and Tsiddon (1997) have already developed the possibility of multiple steady states when $s_{12} < 0$ and we have nothing to add to their results. Second, it can be shown that the multiplicity only exists when the dynasties are sufficiently impatient—for δ close to 1, we have the equivalent of a Turnpike Theorem, which tells us that the steady state is independent of initial conditions.

Under the assumption that $s_{12} = 0$, all dynasties converge to the same level of h. To see this, observe first that under this condition, neither h_t nor h_{t+2} enters Equations 2 and 3, which tells us that every dynasty chooses the same level of h_{t+1}. Therefore, everyone must have the same level of human capital in every period other than the very first.

Uniqueness: Do otherwise identical economies that start with very different average levels of human capital eventually end up with same average level? To answer this question we need to show that the steady state equation:

$$s_1(h,h) + \delta s_2(h,h) - \delta \frac{(1-\tau_1)/(1-e_1)}{\gamma + (1-\gamma)g(h,\alpha)} = 0 \tag{5}$$

has a unique solution where the function $g(h,\alpha) \equiv \frac{w^L}{w^H}$ denotes the steady state relative wages. To derive the function $g(h,\alpha)$ note that in a steady state where everyone has h units of human capital, the net supply of unskilled labor and human capital to the production sector are given by $1 - S(h)(1-\gamma)$ and $h - S(h)\gamma$, respectively. Therefore

$$w_H = f_H(h - S(h)\gamma, 1 - S(h)(1-\gamma), \alpha) \text{ and } w_L = f_L(h - S(h)\gamma, 1 - S(h)(1-\gamma), \alpha).$$

What happens to $g(h,\alpha)$ when h goes up? Since $f(\cdot)$ is homogeneous of degree one, $g(\cdot)$ depends only on the ratio $\frac{h-S(h)\gamma}{1-S(h)(1-\gamma)}$. Since, by assumption $S(h)$ is increasing in h, the denominator of this expression clearly goes down when h goes up. The effect on the numerator is, however, potentially ambiguous. However, if h is in the neighborhood of a steady state,

$$\frac{d(h - S(h)\gamma)}{dh} = 1 - s_1\gamma - s_2\gamma > 1 - s_1\gamma - \delta s_2\gamma = 1 - \frac{\delta(1-\tau_1)/(1-e_1)}{1 + \frac{(1-\gamma)}{\gamma}g(h,\alpha)}.$$

Therefore, in the neighborhood of a steady state, $g(h)$ is increasing in h as long as the expression on the right is positive. This is true as long as $(1-\tau_1)/\,(1-e_1)$ is not too much larger than 1; in other words the subsidy cannot be too large. Basically, if the subsidy is too large there may be over-investment in education to the point where an extra unit of human capital costs more than one unit of human capital to produce. For the rest of the paper we will assume that subsidy is never so large that this we get into this case and therefore we will assume *that $g(h,\alpha)$ is always increasing as a function of h, the per capita endowment of human capital in the economy..*

To complete the argument we differentiate the steady state map (Equation 5) with respect to h and use the fact that $s_{12} = 0$, to get the expression

$$s_{11} + \delta s_{22} + \delta \frac{1-\gamma}{\gamma} \frac{(1-\tau_1)/(1-e_1)}{(\gamma + (1-\gamma) g(h,\alpha))^2} g_1(h,a).$$

In the neighborhood of a steady state this expression is always positive since s_{11}, s_{22} and $g_1(h)$ are all positive. Therefore, the steady state has to be unique.

The Effects of Taxes and Subsidies

We are now in a position to look at some policy experiments. Consider first a small increase in the lump-sum educational subsidy e_0 or a fall in the lump-sum tax τ_0. Since neither enters Equation 5, they clearly have no effect on the long-run level of investment in education, as long as the non-negativity constraint on investment in education is not binding.[13]

Changes in e_1 and τ_1 clearly do have an impact on educational investment. It is easily checked that the first-best level of investment will be achieved when $\tau_1 = e_1$. Since tax collection si costly, the optimum is presumably $\tau_1 = e_1 = 0$. Consequently, the only reason to subsidize education is to counteract the effects of pre-existing taxes on human capital. Moreover, an increase in e_0 financed by an increase in τ_1 clearly reduces educational investment. We interpret this as saying that if education subsidies are harder to target than taxes on human capital, then the net effect of the subsidy may be to discourage education.[14]

Raising Returns to Education

Finally, we can look at what happens when the rate of return to education goes up. We capture this by assuming that raising α lowers $g(h,\alpha)$ for all values of h: this amounts to an increase in the relative price of skills, for each level of h, and can be interpreted as the result, for example, of a government policy promoting the adoption of a new skill-intensive technology.[15]

[13] In the short run, however, changes in $e_0 - \tau_0$ keeping all other policies fixed does affect the interest rate and through it investment in education.

[14] Once again, this is only true if the non-negativity constraint does not bind.

[15] Foster and Rosenzweig (2000) suggest that the green revolution in India was an example of such a policy.

It is clear from Equation 5 that if $g(\cdot)$ goes down, h will have to up. Therefore, an increase in the rate of return on education does increase investment in education in this model.

4 Models with No Credit Markets

In this world there are no credit markets, or at least the existing credits markets are too inefficient to be of relevance to most people. People invest in their children's schooling using their family resources. Moreover, investing in education is the only investment opportunity available to the family. There are, however, also taxes and subsidies, just as in the previous section. Therefore, its budget constraint in any period t can be written as:

$$w_t^L + h_t w_t^H(1-\tau_1) - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)(1-e_1) + e_0 - \tau_0 = c_t.$$

This constraint holds for every t. Moreover, consumption is assumed to be always non-negative in order to make the credit constraint meaningful.

In the rest of the paper we will examine the implications of specific assumptions about altruism within the framework defined above. In this discussion income effects will play an important role, though, as we will see, they have rather different implications depending on what we assume about altruism.

4.1 True Altruism

In this case, each generation maximizes the present value of the consumption utilities of all the following generations. As is well known, this leads to a time-consistent decision rule, with each generation choosing what the previous generations would have wanted it to choose. In other words, the family acts as single decision maker. More specifically, let the consumption utility of the t^{th} generation be $u(c_t)$,$u' > 0$, $u'' < 0$. The t^{th} generation then maximizes $\sum_{s=t}^{\infty} \delta^{s-t} u(c_s)$, given their endowment of human capital, h_t, and a budget constraint

$$w_t^L + h_t(1-\tau_1)w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)(1-e_1) + e_0 - \tau_0 = c_t, \forall t \qquad (6)$$

If $v_t(h_t)$ is the indirect utility function so generated, it is well known that this is equivalent to assuming that each generation maximizes $u(c_t) + v_{t+1}(h_{t+1})$. Using the budget constraint, this can be rewritten as $u(h_t, h_{t+1}) + v(h_{t+1})$ which has the form $U^t(h_t, h_{t+1})$. Moreover, it is easy to see that $U_1^t = [w_t^H(1-\tau_1)-(\gamma w_t^H+(1-\gamma)w_t^L)s_2(h_{t+1},h_t)(1-e_1)]u'(c_t) > 0$ and $U_{12}^t = -(\gamma w_t^H + (1-\gamma)w_t^L)[w_t^H(1-\tau_1) - (\gamma w_t^H + (1-\gamma)w_t^L)s_2(h_{t+1},h_t)(1-e_1)]s_1(h_{t+1},h_t)(1-e_1)u''(c_t) > 0$ by the concavity of $u(c_t)$. $U^t(h_t, h_{t+1})$ is therefore supermodular. For such cases we have the following (trivial) result:

Claim 1: *Suppose each generation maximizes a function of the form $U^t(h_t, h_{t+1})$. Then dynasties with more human capital at time t also have more human capital at time $t+1$ as long as U^t is supermodular.*

Proof. At time t each dynasty maximizes U^t. Let h_t' and h_t'' be two values of h_t and h_{t+1}' and h_{t+1}'' be the corresponding optimal choices for h_{t+1}. Let $h_t' < h_t''$ but $h_{t+1}^{\text{`}\prime\prime} \leq h_t'$. By revealed preference $U^t(h_t', h_{t+1}') \geq U^t(h_t', h_{t+1}^{''})$ and $U^t(h_t'', h_{t+1}'') \geq U^t(h_t'', h_{t+1}')$. These two inequalities together imply $U^t(h_t', h_{t+1}') - U^t(h_t', h_{t+1}'') \geq 0 \geq U^t(h_t'', h_{t+1}') - U^t(h_t'', h_{t+1}'')$. This contradicts the fact that U^t is supermodular. Therefore, $h_t' < h_t''$ always implies $h_{t+1}'' > h_t'$. ■

It follows from this result that in the model in this sub-section the children of the educated will tend to be more educated as well, $\frac{dh_{t+1}}{dh_t} > 0$, despite the fact that we have now assumed $s_{12} = 0$ to rule out a direct effect of the education of the parents on the education of their children.

Next, using the budget constraint we can rewrite $\sum_{t=0}^{\infty} \delta^t u(c_t)$ as $\sum_{t=0}^{\infty} \delta^t u(w_t^L + h_t(1-\tau_1)w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1},h_t)(1-e_1) + e_0 - \tau_0)$. Maximizing this with respect to the sequence $\{h_t\}$ yields the first order conditions:

$$
\begin{aligned}
& (\gamma w_{t-1}^H + (1-\gamma)w_{t-1}^L)s_1(h_t, h_{t-1})(1-e_1) \\
= \; & \delta \frac{u'(c_t)}{u'(c_{t-1})}[w_t^H(1-\tau_1) - (\gamma w_t^H + (1-\gamma)w_t^L)s_2(h_{t+1},h_t)(1-e_1)].
\end{aligned}
$$

Using the fact that in the steady state $c_t = c_{t+1}$, $w_t^H = w^H$ and $w_t^L = w^L$ $\forall t$,this reduces to the condition

$$
s_1(h,h) + \delta s_2(h,h) - \delta \frac{(1-\tau_1)/(1-e_1)}{\gamma + (1-\gamma)g(h,\alpha)} = 0
$$

which will be recognized as Equation 5, the steady state condition under altruism and perfect credit markets. It follows that under the assumption that $s_{12} = 0$, there is complete equalization in the long-run, even though the rich invest more than the poor in the human capital of their children. Moreover, the steady state is unique and the level of investment and consumption in the limit are what they would be in a first-best world.[16]

The long-run effects of policy in this model, not surprisingly, are the same as in the benchmark case—lump-sum taxes and subsidies have no effect, and while proportional taxes and subsidies do have an effect, no taxes and no subsidies remains optimal. The general policy bias that comes from this model is that, at least in the long-run, what matters is removing barriers that prevent the returns to human capital from being as high as possible.[17]

4.2 Perfect Functional Altruism

In this subsection we consider the exact opposite model—a model where people are completely selfish in the sense that parents do not care at all about what happens to their children. The reason why there appears to be altruism, in the sense of parents investing in the education of their children, is that families act as imperfect substitutes for the missing credit market. In other words, parents invest in the education of their children in return for old age care. There are several plausible reasons for why such a transaction may be possible in an economy where there are no other credit transactions: first, parents typically know a lot more about their children, which may limit the amount of adverse selection; second, there may be social norms that punish children who do not take care of their parents, but do not apply to those who fail to honor their debts; finally, there may be emotional ties between parents and children which enable parents to "punish" children who do not take care of them.

To do justice to this case we need to start with a model where there are three generations at any point in time—the young, whose only role is to get an education; the middle-aged who work, pay for their children's education and support their parents; and the old who simply consume what their children give them. Suppose that people only get utility from how much they consume when they are middle-aged and when they are old. Their preferences are therefore captured by the utility function $u_m(c_t) + u_o(p_{t+1})$.

To complete the model we also need to say something about exactly how the intergenerational

[16] This is essentially a deterministic version of the result in Loury (1981) showing that in the long run all dynasties have the same distribution of consumption even in the presence of credit constraints.

[17] The short-run effects of policy are, however, potentially different from the first-best case. In particular because human capital is not instantly equalized, a subsidy financed by a tax on human capital does redistribute and the resulting increase in the income of the poor will typically increase their investment in human capital and thus hasten convergence.

"implicit" contract gets chosen. Here we make the extreme assumption that there is one grand contract encompassing all future generations, which gets chosen at time 0. The choice of the contract is based on maximizing a weighted average of the utilities of all the generations: $\Sigma_{t=0}^{\infty}\lambda^t(u_m(c_t) + u_o(p_{t+1}))$, $0 < \lambda < 1$.[18] Presumably the idea is not that there is actually contractual negotiation between all the generations, but rather that there is a social norm which mimics the optimal grand contract.

There are several motives behind the choice of this particular formulation: First, it provides us with the closest parallel to the case of true altruism. Second, it is significantly simpler than many of the alternatives as there is no role for strategic interactions between generations.[19] Moreover, the steady state of this model coincides with the steady state of a model where the contract is only between adjacent generations, but each generation assumes that the contractual choice of all future generations is unaffected by their own choice.[20]

Given these assumptions, we can now write down the exact maximization problem facing a particular dynasty at time 0, given that it started with h_0 amount of human capital and an obligation of p_0 towards its parent. It has to choose $\{h_t, p_t\}_{t=1}^{t=\infty}$ to maximize $\Sigma_{t=0}^{\infty}\lambda^t(u_m(c_t) + u_o(p_{t+1}))$ given the initial values $h_0$and p_0 and the budget constraint

$$w_t^L + h_t(1-\tau_1)w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)(1-e_1) + e_0 - \tau_0 - p_t = c_t, \forall t. \quad (7)$$

This maximization yields a value function $v(h_0, p_0)$. At any point in time, the dynasty can be thought of as maximizing $u_m(c_t) + u_o(p_{t+1}) + \lambda v(h_{t+1}, p_{t+1})$, subject to a budget constraint which takes the form of Equation 7. While this is not exactly in the form $U^t(h_t, h_{t+1})$, the argument made in Claim 1 is easily extended to this case. Therefore we have:

Claim 2: *Those whose parents have more human capital also have more human capital of their own.*

In other words, there are income effects in this model and inequality in education will tend

[18] Since the natural interpretation of λ is that it is the bargaining power of the next generation relative to that of the current generation, the restriction that $\lambda < 1$ is not entirely reasonable, but the gain in analytical tractability from making this assumption is significant.

[19] For example, an alternative would have been to allow only adjacent generations to contract with each other. In this case, we might have to keep track of how changes in the contract between this generation and the next affect the contract between the next generation and the one after.

[20] This makes it a Nash equilibrium but does not require that it be sub-game perfect.

to be relatively persistent. To see what happens in the long-run, observe that the first order conditions for the above maximization problem (assuming an interior solution exists) are:

$$\lambda u_m'(c_t) = u_o'(p_t)$$

$$\begin{aligned} & (\gamma w_t^H + (1-\gamma) w_t^L) s_1(h_{t+1}, h_t)(1-e_1) u_m'(c_t) \\ = \; & \lambda[(1-\tau_1) w_{t+1}^H - (\gamma w_{t+1}^H + (1-\gamma) w_{t+1}^L) s_2(h_{t+2}, h_{t+1})(1-e_1)] u_m'(c_{t+1}). \end{aligned}$$

In a steady state the second equation reduces to the much more compact condition:

$$s_1(h,h) + \lambda s_2(h,h) = \lambda \frac{(1-\tau_1)/(1-e_1)}{\gamma + (1-\gamma) g(h,\alpha)}.$$

The expression defining the steady state level of human capital is exactly the same condition as in the case of true altruism, with λ replacing δ. An fall in λ, much like a fall in δ, discourages investment in human capital—greater power in the hands of the older generation leads to less education.

The other properties of the steady state are relatively unsurprising: Under the maintained assumption that $s_{12} = 0$, there is no inequality in the long-run, and the steady state is unique. Moreover, the changes in policy parameters have effects in the same direction as they would have in the case of true altruism.

However, a number of new policy issues arise: first, if the government initiates a pay-as-you-go social security system, investment in education is likely to collapse since parents will no longer feel the need for transfers from their children. Second, it raises the issue of whether the government should try to alter the balance of power within the family: there is no *a priori* reason why λ, which measures the bargaining power of the young, should also be the weight children get in the social welfare function. For example, the government may want to maximize steady state welfare rather than the maximand that actually gets maximized. In this case, unlike in the case of true altruism, a proportional subsidy to education or a mandatory schooling law may be optimal. In other words, the government may push people to invest more in education in order to safe-guard the interests of the young.

4.3 Limited Altruism: Warm-Glow Bequest Preferences

Parents may actually care about how they have treated their children even when they are indifferent to the child's welfare. The standard form of such preferences is given by the so-called warm-glow model (Andreoni, 1989). In this model, parents care about their own consumption and the level of bequest they leave their children (rather than their consumption or utility), which, in our model, is simply the amount invested in the child's human capital. There is, however, an issue about how we measure the amount invested by parents, as a part of the expenditure is financed by subsidies and subsidies in turn are financed by taxes. Rather than joining this issue now, we assume that there are no taxes and subsidies and postpone the discussion until the end of this sub-section. For the time being, we assume that each parent maximizes a utility function of the form $u(c_t) + v((\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t))$ subject to the budget constraint (Equation 6) (with all taxes and subsidies set to zero), where both $u(\cdot)$ and $v(\cdot)$ are increasing, concave functions. The choice variable is still h_{t+1}.

Using the budget constraint we can rewrite the maximand in the form:

$$u(w_t^L + h_t w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)) + v((\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)).$$

Since and w_t^H and w_t^L are parameters from the point of view of an individual, this has the form $U^t(h_t, h_{t+1})$. Moreover:

$$\begin{aligned} U_1^t(h_t, h_{t+1}) &= u'(\cdot)[(w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s_2(h_{t+1}, h_t)] \\ &\quad + v'(\cdot)(\gamma w_t^H + (1-\gamma)w_t^L)s_2(h_{t+1}, h_t) \end{aligned}$$

and

$$\begin{aligned} U_{12}^t(h_t, h_{t+1}) &= -u''(\cdot)[w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s_2(h_{t+1}, h_t)](\gamma w_t^H + (1-\gamma)w_t^L)s_1(h_{t+1}, h_t) \\ &\quad - u'(\cdot)[(\gamma w_t^H + (1-\gamma)w_t^L)s_{12}(h_{t+1}, h_t)] + v'(\cdot)(\gamma w_t^H + (1-\gamma)w_t^L)s_{12}(h_{t+1}, h_t) \\ &\quad + v''(\cdot)(\gamma w_t^H + (1-\gamma)w_t^L)s_2(h_{t+1}, h_t)(\gamma w_t^H + (1-\gamma)w_t^L s_1(h_{t+1}, h_t). \end{aligned}$$

Because $s_{12} = 0$, this expression is clearly positive and therefore Claim 1 applies to this case—the children of the educated will tend to be more educated. Maximizing these preferences subject to the budget constraint (Equation 6) tells us that at an interior maximum it must be the case that:

$$\begin{aligned} & u'(w_t^L + h_t w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t))(\gamma w_t^H + (1-\gamma)w_t^L)s_1(h_{t+1}, h_t) \\ = \; & v'((\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t))(\gamma w_t^H + (1-\gamma)w_t^L)s_1(h_{t+1}, h_t). \end{aligned}$$

In steady state this reduces to:

$$\frac{u'(w^L + hw^H - (\gamma w^H + (1-\gamma)w^L)S(h))}{v'((\gamma w^H + (1-\gamma)w^L)S(h))} = 1. \tag{8}$$

Inequality: Both the numerator and the denominator of the expression on the left are typically going to be decreasing as a function of h.[21] Whether the curve given by this expression slopes up or down depends on which of the two decreases faster. In particular, it is not hard to find cases where it goes up and down several times, generating several solutions in the process, which generate the possibility of there being inequality in the steady state.

To see more precisely what is going on here, consider first the case where there is constant relative risk aversion, i.e., $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, and moreover $v(\cdot) = \delta u(\cdot)$. In this case, the condition 8 can be rewritten in the form:

$$\frac{w^L + hw^H}{(\gamma w^H + (1-\gamma)w^L)S(h)} = 1 + (\frac{1}{\delta})^{\frac{1}{\sigma}}. \tag{9}$$

It is easily checked, using the fact that the function $S(h) \equiv s(h, h)$ increasing and convex, and the fact that $S(0) = 0$,that the left-hand side is decreasing as a function of h, and therefore the value of h that solves this equation is unique.

Consider next the other standard formulation for preferences, namely the constant absolute risk aversion family: i.e., $u(c) = 1 - e^{-\sigma c}$ and $v(\cdot) = \delta u(\cdot)$. In this case, Equation 8 can be written in the form:

$$-(w^L + hw^H) + 2((\gamma w^H + (1-\gamma)w^L)S(h)) = \frac{1}{\sigma}\log_e \delta. \tag{10}$$

The left-hand side of this equation defines a convex function of h. If $S'(0) = 0$, it always starts by being downward-sloping but eventually slopes up (because $S(h)$ grows faster than h). Assuming $-w^L > \frac{1}{\sigma}\log_e \delta$ and $S(h)/h \to \infty$ as $h \to \infty$, we have the case that is in Figure 1. There are

[21] This is not always true because for high enough values of h, it is conceivable that the numerator increases with h. This requires, however, that the family is heavily over-investing in education. In other words, it must be the case that an increase in h actually *reduces* output net of educational expenditure.

three potential stationary values of h in this case—$h = 0$, $h = h_1$ and $h = h_2$. Equation 8 does not actually hold at $h = 0$: what happens is that the marginal utility of bequests is strictly less than the marginal utility of consumption, but since bequests cannot be negative, $h = 0$ is the optimal choice. Of the other two steady states, the one at h_2 is stable and therefore perhaps more worthy of interest.

In both the CARA and CRRA case the basic forces are the same—there is an income effect of being more educated, which makes educated parents more willing to invest. This, however, is counteracted by the fact that an educated parent has to invest more just to make sure that his child is not less educated than he is, and, moreover, the marginal cost of investment is rising. Which of these effects dominates and therefore whether there is one or more stationary level of h, depends on the exact curvature of the utility functions.[22]

Uniqueness: The argument in the previous paragraphs shows that there can be inequality in the long-run for CARA preferences. Dynasties which start with more human capital will also tend to end up with more human capital. Given that there is inequality, it is likely that convergence will also fail: an economy that starts with a large number of high human capital dynasties will tend to have a high long-run level of human capital.

There is actually a second, independent reason why convergence might fail: to highlight this reason we choose CRRA preferences so that in the long-run there is perfect equality. Moreover assume (purely for convenience) that $\gamma = 1$. Under these conditions we can rewrite Equation 9 in the form:

$$\frac{g(h,a) + h}{S(h)} = 1 + (\frac{1}{\delta})^{\frac{1}{\alpha}}. \tag{11}$$

Both the denominator and the numerator of the ratio $\frac{g(h,a)+h}{S(h)}$ are increasing in h. We know that the ratio $h/S(h)$ declines as h goes up since S is convex. However, it is entirely possible that $\frac{g(h,a)}{S(h)}$ increases in h at least over a range—this will be the case when the degree of substitutability between skilled and unskilled labor is relatively small. It is therefore quite straightforward to construct examples where the function $\frac{g(h,a)+h}{S(h)}$ has the form given in Figure 2. There are clearly three steady states in this case, of which the two extreme ones are stable. The multiplicity of steady states reflects the following simple intuition: teachers are cheap in

[22] This source of long run inequality is closely related to that in Galor and Moav (1999), though their model is based on Stone-Geary type preferences for bequest and is therefore analytically more demanding.

an economy where there is a lot of human capital and therefore the same level of bequests buys more human capital, and as a result the initial high level of human capital is reproduced.[23]

Taxes and Subsidies.

We now need to address the question of how taxes and subsidies enter the preferences. One possibility is that $v = v((\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)(1-e_1) - e_0)$. This assumption says that people count as their bequest the entire amount they have spent on education, but not the taxes they have paid which are then spent on subsidizing education. An alternative possibility would be that people count the taxes they have paid as part of their bequest. In this case $v = v((\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)(1-e_1) - e_0 + \tau_0 + \tau_1 h_t w_t^H)$. While the second assumption is more conventional, we prefer the first of these assumptions on the grounds that it is very difficult to imagine that people can identify the part of their tax payments that is going to pay for the educational subsidies, but we allow for both.

Under the first assumption, the steady state condition turns out to be:

$$\frac{u'(w^L + h(1-\tau_1)w^H - (\gamma w^H + (1-\gamma)w^L)S(h)(1-e_1) + e_0 - \tau_0)}{v'((\gamma w^H + (1-\gamma)w^L)S(h)(1-e_1) - e_0)} = 1.$$

It is clear from this expression that changes in lump-sum taxes and subsidies have long-run effects in this model: an increase in e_0 and a cut in τ_0 promote investment in education. Moreover, an increase in e_0 and τ_0 that leaves the government deficit unchanged ($\Delta e_0 = \Delta\tau_0$) increases investment in human capital. Indeed, any increase in subsidies and taxes that keep everyone's net income unchanged (at the original steady state value of h) will increase investment. This is because people in this model do not make the connection between the increases in subsidy and the increase in taxes.

Under the second assumption, the steady state condition is going to be:

$$\frac{u'(w^L + h(1-\tau_1)w^H - (\gamma w^H + (1-\gamma)w^L)S(h)(1-e_1) + e_0 - \tau_0)}{v'((\gamma w^H + (1-\gamma)w^L)S(h)(1-e_1) - e_0 + \tau_0 + \tau_1 h_t w_t^H)} = 1.$$

In this case, changes in taxes and subsidies that do not change the public deficit have no effect on investment as long as there is no inequality in the steady state. On the other hand, an increase in e_0 without any other changes in taxes or subsidies does increase investment in

[23] This type of multiplicity of steady states, sustained by price effects, is similar to the multiplicity in Banerjee and Newman, 1993.

human capital (at the expense of increasing the public deficit). If, however, there is inequality in the steady state, changes in proportional taxes and subsidies have distributional effects—and as a result there may be an effect on average investment even if there is no change in the public deficit. Consider the case in Figure 1, where we have CARA preferences and multiple steady states. To simplify matters, assume that subsidies are lump-sum but taxes are proportional, i.e., $e_1 = 0$ and $\tau_0 = 0$, and that there is a balanced budget, i.e., $e_0 = \tau_1 \overline{h} w^H$, where $\overline{h}$ is the mean level of human capital. With these assumptions we can rewrite Equation 10 as

$$-w^L - hw^H + 2(\gamma w^H + (1-\gamma)w^L)S(h) + 2(h - \overline{h})\tau_1 w^H = \frac{1}{\sigma}\log_e \delta.$$

First consider what happens for fixed values of w^L and w^H. The effect of an increase in τ_1 is shown in Figure 1 An increase in τ_1 shifts the left-hand side of the above equation down for low values of h and up for high values of h. For small changes in τ_1 starting at $\tau_1 = 0$, the stable steady state at $h = 0$ remains and the stable "good" state as h_2 moves left. If we think of the population as being initially distributed across these two steady states, it is clear that the effect of increasing τ_1 is to reduce the mean level of human capital. However, for larger increases in τ_1, the steady states at $h = 0$ and h_1 vanish, and the entire population ends up at h_2. This tells us that higher taxes can bring about more investment. However the argument is incomplete—so far we have kept the values of w^L and w^H fixed. However note that we can minimize the changes in w^L and w^H by making the elasticity of substitution in production high enough. Therefore the same result, namely that a lump-sum subsidy financed by a proportional tax on human capital, can increase investment in human capital holds even when. w^L and w^H are endogenous. Essentially, redistribution here helps start a virtuous cycle and this leads to a large increase in human capital.

The fact that the steady state is not unique even when there is no inequality in the steady state suggests a different possibility for a virtuous cycle. Suppose starting at a "bad" steady state, the government borrows money abroad and invests it in a proportional subsidy to human capital. If the subsidy is large enough, this can set off a virtuous cycle where there is more and more human capital and therefore investment in human capital becomes cheaper and cheaper. As the economy gets richer, the government can raise taxes and repay the initial debt.

Raising returns to education.

The effect of changes in returns to education are best seen by looking at the steady state condition with CRRA preferences (with $\gamma = 1$):

$$\frac{g(h,a)+h}{S(h)} = 1 + (\frac{1}{\beta})^{\frac{1}{\alpha}}.$$

It is clear that a shift in α that shifts $g(h,\alpha)$ down (a skill-biased change), has to reduce investment in education, since at any stable steady state the function $\frac{g(h,a)+h}{S(h)}$ has to be a declining function of h. An increase in the rewards for human capital always reduces investment in human capital in this case, because it makes teachers more expensive and education is produced using teachers. Clearly the strength of this result derives from the fact that given the specification of preferences, the size of the rewards from education do not matter to the parents. While this is clearly an extreme assumption, it is not implausible that parents will care less about the returns to education than about the costs, given that they pay the costs and their children get the returns.

4.4 Limited Functional Altruism: The Effects of Incomplete Inter-generational Contracts

The model of functional altruism based on the idea of a grand family contract clearly stretches the limits of enforceability. Perhaps a more plausible assumption is that the inter-generational contract is incomplete: it cannot specify the exact obligations of parents and children, all it can do is require that children take care of their parents at a level commensurate with their own consumption levels.[24] This gives parents the incentive to invest in their children's education as a way of increasing the children's consumption and therefore their own consumption.

We capture this type of bequest preference by assuming that each generation cares abut its share of total family consumption when it is middle-aged and when it is old: $u(\mu c_t) + \delta u((1-\mu)c_{t+1})$ where $u(\cdot)$ is an increasing, concave and twice differentiable function and μ is the share of family consumption that goes to the middle-aged. The budget constraint is still given by Equation 6.

What makes this case somewhat less tractable than previous ones is that there is a conflict of interest between the generations—the current generation cares only about the next, while

[24] In many countries parents live with their grown up children and share the standard of living of their children.

the next generation cares about the one after.[25] This raises the issue of strategic interactions between the generations—*ex ante* each generation would like to promise that it will consume every dollar that it gets from its parent's investment (this maximizes the parent's willingness to invest). *Ex post*, it will want to share some of that money with their own children. Assuming that binding contracts are not possible, each generation's maximization problem should take account of the next generation's *equilibrium* decision rule for allocating their earnings between consumption and human capital investment. Unfortunately, this gives rise to the possibility of multiple equilibria, which differ in the decision rule adopted by each generation. To limit the number of equilibria, we confine ourselves to looking only at the set of Markovian equilibria. These are equilibria where each generation uses a decision rule of the form $h_{t+1}(h_t)$. In other words, the human capital investment by each generation depends only on the amount of human capital that it got from its parent and the index of time.[26] It is easily checked that such an equilibrium exists.

However, this by itself does not tell us very much about the steady states of this economy since everything depends on the function $h_{t+1}(h_t)$, which itself is determined in equilibrium. In order to understand the possibilities that arise in this case, we need to look at a parametric example: We show that for a specific choice of parameters, $h_{t+1}(h_t)$ is linear and this allows us to characterize the solution in more detail.

4.4.1 A linear-quadratic example

Assume that $u(x) = x - \frac{1}{2}\phi x^2$. Assume also that $s(h_{t+1}, h_t) = s \cdot h_{t+1}$ and that $\gamma = 1$. Using the budget constraint we rewrite the family's maximand as:

$$\begin{aligned}
& u(\mu(w_t^L + h_t(1-\tau_1)w_t^H - (\gamma w_t^H + (1-\gamma)w_t^L)s(h_{t+1}, h_t)(1-e_1) + e_0 - \tau_0)) \\
& +\delta u((1-\mu)(w_{t+1}^L + h_{t+1}(1-\tau_1)w_{t+1}^H - (\gamma w_{t+1}^H + (1-\gamma)w_{t+1}^L)s(h_{t+2}, h_{t+1})(1-e_1) + e_0 - \tau_0)).
\end{aligned}$$

The first order condition for utility maximization with respect to h_{t+1} turns out to be:

[25] See Bernheim and Ray (1987) for a different model of this class, which, once again illustrates the difficulties of working with such models.

[26] The index of time is there to allow for the possibility that w^H and w^L change along the equilibrium path.

$$
\begin{aligned}
& u'(\mu c_t)\mu(\gamma w_t^H + (1-\gamma)w_t^L)s(1-e_1) \qquad (12) \\
= \;& \delta u'((1-\mu)c_{t+1})(1-\mu)[w_{t+1}^H(1-\tau_1) - s(\gamma w_{t+1}^H + (1-\gamma)w_{t+1}^L)(1-e_1)\frac{dh_{t+2}}{dh_{t+1}}].
\end{aligned}
$$

We now claim, that given our assumptions about preferences and the production technology for human capital, there is a linear solution to this maximization problem. In other words, there is a solution that takes the form $\frac{dh_{t+1}}{dh_t} = \eta_t, \forall t$. To show this, assume that $\frac{dh_{t+2}}{dh_{t+1}} = \eta_{t+1}$ and differentiate Equation 12 with respect to h_t to get (using also the fact that $u''(c_t) = u''(c_{t+1}) = \phi$):

$$
\begin{aligned}
& \mu^2[w_t^H(\gamma w_t^H + (1-\gamma)w_t^L)s(1-e_1)(1-\tau_1) - (\gamma w_t^H + (1-\gamma)w_t^L)^2 s^2(1-e_1)^2\frac{dh_{t+1}}{dh_t}] \qquad (13) \\
= \;& (1-\mu)^2\delta\frac{dh_{t+1}}{dh_t}[w_{t+1}^H(1-\tau_1) - s(1-e_1)(\gamma w_t^H + (1-\gamma)w_t^L)\eta_{t+1}]^2.
\end{aligned}
$$

This equation defines $\frac{dh_{t+1}}{dh_t}$ in terms of a set of constants and η_{t+1}. As long as η_{t+1} is a constant, $\frac{dh_{t+1}}{dh_t}$ will also be a constant, which confirms that there is a solution with the property that $\frac{dh_{t+1}}{dh_t} = \eta_t$, $\forall t$.

Using the steady state conditions (including the fact that $\eta_t = \eta_{t+1} = \eta$), Equation 13 reduces to:

$$
s\frac{1-e_1}{1-\tau_1}[\frac{1}{(\gamma + (1-\gamma)g(h))} - s(\frac{1-e_1}{1-\tau_1})\eta] = \delta\eta(\frac{1-\mu}{\mu})^2[\frac{1}{(\gamma + (1-\gamma)g(h))} - s(\frac{1-e_1}{1-\tau_1})\eta]^2
$$

or

$$
s = \frac{\delta\eta}{(\frac{\mu}{1-\mu})^2 + \delta\eta^2}\frac{1-\tau_1}{1-e_1}[\frac{1}{(\gamma + (1-\gamma)g(h))}] \qquad (14)
$$

where h is the average amount of human capital in the economy in the steady state.

The two sides of this equation are represented in Figure 3. The right-hand side, represented as $H(\eta)$, defines a non-monotonic function of η for any fixed value of h. It is close to zero and increasing for η in the neighborhood of zero but decreases towards zero for high values of η. Therefore, if there is at least one value of η that solves this equation there must be two. In Figure 3, these two steady states are identified as $\underline{\eta}$ and $\overline{\eta}$. Note that both $\underline{\eta}$ and $\overline{\eta}$ are positive: as in the previous models, parents with more education will have children with more education.

The multiplicity of possible values for η has a simple explanation—a high value of η means that each generation responds to an increase in its own human capital by substantially increasing investment in its children's human capital. This amounts to saying that the current generation of middle-aged get a low rate of return on their investment in children—as is well-known, people may invest more when the rate of return is low, if the income effect is strong enough.[27]

We do not take a stand on which of $\underline{\eta}$ and $\overline{\eta}$ is the right solution. The usual method of using the stability properties to eliminate solutions is problematic here given the complex, forward-looking dynamics associated with Equation 13.

However we can say something more about the properties of η. Rewrite steady state condition 12 in the form

$$s = \frac{\delta}{\frac{\mu u'((\mu c)}{(1-\mu)u'((1-\mu)c)} + \delta\eta} \frac{1-\tau_1}{1-e_1} [\frac{1}{(\gamma + (1-\gamma)g(h))}]. \tag{15}$$

and combine it with equation 14 to get.

$$\eta \frac{u'(\mu c)}{u'((1-\mu)c)} = \frac{\mu}{1-\mu}. \tag{16}$$

This immediately implies that $\eta \geq 1$ if and only if $\mu \geq 1/2$. In the case where $\eta > 1$ dynastic trajectories will tend to diverge—a positive shock to h_t, starting at the steady state, will generate an even larger increase in h_{t+1} and so on. The case where $\mu < 1/2$ which corresponds to the case, not implausible in the context of developing countries, where the balance favors the elderly, is better behaved since it implies $\eta < 1$. In this case dynasties will return to their steady state values following a shock. Henceforth we will focus on the case where $\mu < 1/2$ and $\eta < 1$.

Finally observe that the steady state condition 15 is not unlike the corresponding condition for the first-best case, the one difference being that δ has been replaced by $\frac{\delta}{\frac{\mu u'((\mu c)}{(1-\mu)u'((1-\mu)c)}+\delta\eta}$. It is not clear how these numbers compare since there is no reason why the $\delta's$ should be the same, but if they were the same, then, at least for μ not too different from 1/2, it can be shown that $\frac{\delta}{\frac{\mu u'((\mu c)}{(1-\mu)u'((1-\mu)c)}+\delta\eta} > \delta$. In other words, a result of the strategic interaction between the generations is that there is underinvestment in education.

Inequality: These different values of η represent different dynastic strategies. They will typically lead to long-run inequality in the sense that those dynasties will have different steady state levels of consumption and human capital. To see this, observe that as long as $\mu < 1/2$,

[27] The right-hand side of Equation 12 must be non-negative, which implies that η must be such that $\eta^{-1} \geq s(\gamma+(1-\gamma)g(\overline{h})(1-e_1))/(1-\tau_1)$. However, this is always the case, since from Equation 15 $s(\gamma+(1-\gamma)g(\overline{h})(1-e_1))/(1-\tau_1) = \frac{\delta\eta}{(\frac{\mu}{1-\mu})^2+\delta\eta^2} < \eta^{-1}$.

$\frac{u'(\mu c)}{u'((1-\mu)c)} = \frac{1-\phi\mu c}{1-\phi(1-\mu)c}$ is increasing in c. It follows from Equation 16 that in this case, high values of η will be associated with low values of steady state consumption. [28]

Uniqueness: As suggested in the previous sub-section, once there is inequality in the steady state, the presumption is that there will also be multiple steady states: intuitively, an economy which starts with only high η dynasties will be different from an economy which starts with only low η dynasties, even in the long-run. However, as before, it is also worth checking whether there can be multiple steady states without any inequality. In equation 15, there is yet another possible source of multiplicity. However actually demonstrating multiplicity will require a more involved discussion than we feel is appropriate here. We therefore confine ourselves to explaining the logic of the construction: Suppose h were to go down for some reason. This pushes the curve up in Figure 3 because $g(h)$ goes down. As we know, the effect this has on η depends on the initial value of η. Assuming all dynasties started with $\eta = \underline{\eta}$, η goes down, which means that in the steady state, h has to go up (assuming $\mu < 1/2$), contradicting our premise that h had gone up. This tells us that there cannot be multiple steady states in this case. Now consider the case where initially everyone was at $\overline{\eta}$ The reasoning is exactly the same but now η goes up, pushing h down. There is "positive feed-back", creating the possibility of multiplicity.

The basic intuition for the multiplicity is closely to related to the explanation for why there are two equilibrium values of η. An increase in h lowers the rate of return to human capital, which has both an income effect and a substitution effect. Investment goes up when the income effect dominates.

Taxes, subsidies and returns to human capital: Lump-sum taxes and subsidies do not enter Equations 14 and 16 and therefore have no direct effect on the steady state values of η and c. However since in steady state $c = w^L + hw^H(1-\tau_1) - sh(\gamma w^H + (1-\gamma)w^L)(1-e_1) + e_0 - \tau_0$, an increase in the net (lump-sum) subsidy to the household sector will lead to a decrease in h for any fixed c,[29] which raises the curve in Figure 3. This will lead to a fall in $\underline{\eta}$ and an increase in $\overline{\eta}$. For those who were initially at $\underline{\eta}$, the effect will be to increase c and h (again assuming $\mu < 1/2$). For those who were initially at $\overline{\eta}$, there will be a fall in both c and h.

[28] Since c is typically increasing as a function of h, this implies that high values of the marginal propensity to invest in education will go with less overall investment. Moreover, since $h_{t+1}(h_t) = h_0 + \eta h_t$, this implies that high values of η will go with low values of h_0.

[29] This follows from our maintained assumption that

An increase in τ_1 has a direct negative effect on h, which can be seen from equation 15, under the assumption that c and η are fixed. This is the standard disincentive effect. However it also pushes the curve down in Figure 3, which raises $\underline{\eta}$ and reduces $\overline{\eta}$. For those who started at $\underline{\eta}$ this will lead to a further fall in h (assuming $\mu < 1/2$). For those who started at $\overline{\eta}$, this would counteract the initial fall in h though we do not yet know whether the net effect can be positive. The same logic also applies to the case of an increase in the return to human capital. Starting at $\underline{\eta}$ the effect is always positive, but we cannot rule the possibility that it is negative when $\eta = \overline{\eta}$.

5 Conclusion

The main point of this paper was to underscore the important differences between the implications of alternative ways of modeling the family in a world where educational investment is largely financed by the family. The two polar models turn out to share many of the features of the model with credit markets. This contrasts with what we find in the two intermediate models—the incomplete contract model and the warm-glow model—though it should be emphasized that the results from the incomplete model are quite tentative.

Emphasizing these differences is important as many questions of fundamental importance in making educational policy turn on them. However, it should be noted that the main differences are in the long-run. In the short run, all the models have in common the fact that there are income effects—the more educated will invest more in their children's education. These arise from the absence of credit markets for financing human capital investment that all the models share, a relatively uncontroversial assumption at least in the context of LDCs. But these income effects seem to be less important for the long-run outcome in the two models than they are in the other two models. In other words, while the presence of income effects does justify a degree of interventionism even in the polar models, such interventionism comes with a long-run cost, whereas in the two intermediate models both long-run and short-run considerations might justify a more interventionist stance.

Of course, the observation of the mere fact of difference is ultimately less satisfying than understanding exactly why these differences arise. As we said in the introduction, one key difference between these models comes from the extent of alignment of interests between the generations.

The formal analysis above partially reinforces this claim. The key source of similarity between the two polar models comes from the fact that, in the end, the investment decision effectively comes down to a trade-off between current and future consumption for the same generation. Since consumption in every period is constant in the steady state, the current marginal utility of consumption is equal to that in the future and therefore the marginal utility terms cancel out and do not enter the steady state condition. By contrast, in the two intermediate models, because of the inability to align incentives, the marginal utility terms do not cancel out and are in fact the source of all the interesting results. In particular, the one case where there is always a unique steady state value of η in the incomplete contracts model is when $\mu = 1/2$ (from Equation 16 it is immediate that in this case $\eta = 1$). But $\mu = 1/2$ also makes the current and future consumption of every generation equal, in effect making the steady state value of the marginal rate of substitution between current consumption and investment in children a constant.

If this intuition is correct, it suggests that the two polar models are actually quite special—if for example there is exogenous growth and steady state consumption grows, the steady state value of the marginal rate of substitution between current consumption and investment in children will no longer be a constant, even in these models where there seems to be alignment of interests, and this will weaken the rather stark properties of these classes of models. Moreover, any model where the exact shape of people's own utility function and the utility they get, directly or indirectly, from their bequests to their children are different, has the potential to be very different from the two polar models.

In part this suggests that we need more theoretical research aimed towards putting some structure on the nature of family decision-making. Rangel (1999) is one example of this type of work, arguing that the requirement that the implicit contract within the family be enforceable imposes important restrictions on the family's choices. In part it suggests an empirical agenda aimed towards identifying and testing the peculiar implications of these and other models of the family. As pointed out already, because all of these models display income effects, their short run implications tend to be quite similar. However, as pointed out by Ermisch (1996) some of the implications of functional altruism, especially about the reaction to an increase in the child's income, are quite different from the implications of true altruism.[30] Both this kind of empirical work and related experimental work, have an enormous contribution to make to the way we

[30] His conclusion on the basis of his empirical analysis favors altruism.

think about education policy.

References

- - - - - Public report on basic education in India. Delhi; Oxford: Oxford University Press, 2000.

Acemoglu, D. and J. Angrist (1999), "How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws", mimeo, MIT.

Andreoni, J. (1989), "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence", *Journal of Political Economy*, Vol. 97, pp. 1447-1458.

Angrist, J. and A. Krueger (1999), "Empirical Strategies in Labor Economics", in *Handbook of Labor Economics, Volume 3,* Ashenfelter, O. and D. Card (eds.), Amsterdam; New York: Elsevier Scientific, pp1277-1366.

Banerjee, A. and A. Newman (1993), "Occupational Choice and the Process of Development", *Journal of Political Economy*, Vol. 101 (2), pp. 274-98.

Banerjee, A. and A. Newman (1994), "Poverty and Well-Being in Developing Countries: Poverty, Incentives and Development", *American Economic Review Papers and Proceedings*, May, pp. 211-215.

Barham, V. R. Broadway, M. Marchand and P. Pestieau (1995), "Education and the Poverty Trap", *European Economic Review*, Vol. 39, pp. 1257-1275.

Becker, Gary, A Treatise on the Family. Cambridge, MA: Harvard University Press, 1981.

Behrman, J., R. Pollak and P. Taubman (1982), "Parental Preferences and Provision for Progeny", *Journal of Political Economy*, Vol. 90(1), pp. 52-73.

Bernheim, B. and D. Ray (1987), "Economic Growth and Intergenerational Altruism", *Review of Economic Studies*, Vol. 54, pp. 227-243.

Bils, M. and P. Klenow (1997), "Does Schooling Cause Growth or the Other Way Around?", mimeo, University of Rochester.

Card, D. and A. Krueger (1992), "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, Vol. 100 (1), pp. 1-40.

Carvalho, I. (2000), "Household Income as a Determinant of Child Labor and School Enrollment in Brazil: Evidence from a Social Security Reform", mimeo.

Cox, D. (1987), "Motives for Private Income Transfers", *Journal of Political Economy*, Vol. 95, pp. 508-546.

Cox, D. (1990), "Intergenerational Transfers and Liquidity Constraints", *Quarterly Journal of Economics*, Vol. 105, pp. 187-217.

Cox, D. and G. Jakubson (1995), "The Connection Between Public Transfers and Private Interfamily Transfers", *Journal of Public Economics*, Vol. 57, pp. 129-167.

Cremer H., D. Kessler and P. Pestieau (1992), "Intergenerational Transfers within the Family", *European Economic Review*, Vol. 36, pp. 1-16.

Duflo, E. (2000), "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment", NBER Working Paper #7860.

Ermisch, John (1996), "Parental Support for Human Capital Investment by Young Adults", Centre for Economic Policy Research Discussion Paper: 1536.

Ermisch, J. and M. Francesconi (2000), "Educational Choice, Families and Young People's Earnings", *The Journal of Human Resources*, Vol. 35(1), pp. 143-176.

Foster, A. and M. Rosenzweig (2000), "Does Economic Growth Increase the Demand for Schools? Evidence from Rural India, 1960-1999", mimeo, Center for Economic Development and Policy Reform and Univ. of Pennsylvania.

Galor, O. and O. Moav (1999), "From Physical to Human Capital Accumulation: Inequality in the Process of Development", mimeo, Brown University.

Galor, O. and D. Tsiddon (1997), "The Distribution of Human Capital and Economic Growth", *Journal of Economic Growth*, Vol. 2 (1), pp. 93-124.

Galor, O. and J. Zeira (1993), "Income Distribution and Macroeconomics", *Review of Economic Studies*, Vol. 60 (1), pp 35-52.

Glewwe, P. and H. Jacoby (2000), "Economic Growth and the Demand for Education: Is there a Wealth Effect", mimeo, University of Minnesota.

Glomm, G. and B. Ravikumar (1992), "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality", *Journal of Political Economy*, Vol. 100 (4), pp 818-834

Jacoby, H. (1994), " Borrowing Constraints and Progress through School: Evidence from Peru", *Review of Economics and Statistics*, Vol. 76(1), pp. 151-60.

Kochar, Anjini (2000), "Emerging Challenges for Indian Education Policy", mimeo, Stanford.

Kotlikoff, L. and A. Spivak (1981), "The Family as an Incomplete Annuities Market", *Journal of Political Economy*, Vol. 89 (2), pp 372-391.

Loury, G. (1981), "Intergenerational Transfers and the Distribution of Earnings", *Econometrica,* Vol. 49, pp 843-867.

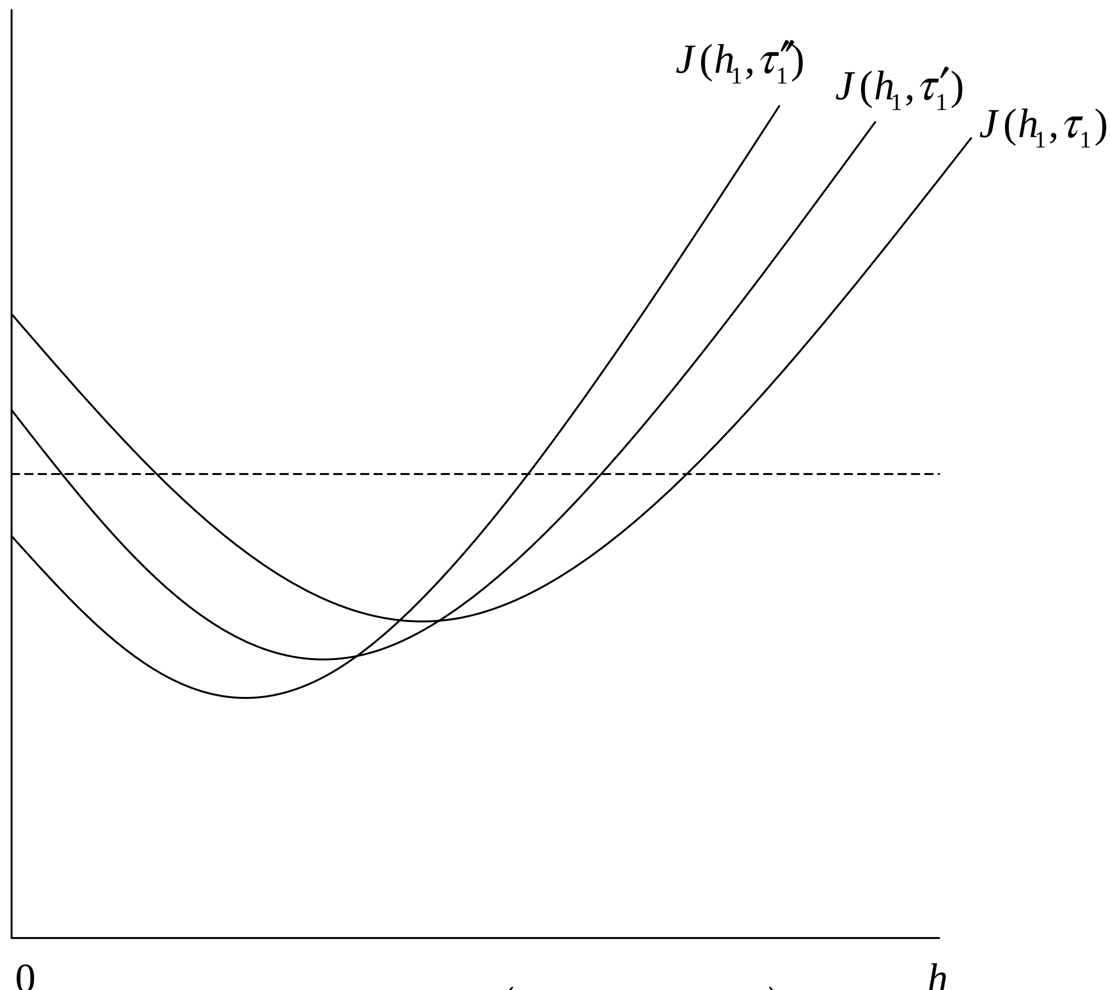
Mookherjee, D. and D. Ray (2000), "Persistent Inequality", mimeo, Boston University.

Psacharopoulos, G. (1994), "Returns to Investments in Education: A Global Update", *World Development*, Vol. 22(9), pp. 1325-1343.

Rangel, Antonio (1999), "Forward and Backward Intergenerational Goods: A Theory of Intergenerational Exchange", mimeo, Stanford University.

Rebelo, Sergio (1991), "Long-Run Policy Analysis and Long-Run Growth", *Journal of Political Economy*, Vol. 99 (3), pp 500-521.

Uzawa, H. (1965), "Optimal Technical Change in an Aggregative Model of Economic Growth", *International Economic Review*, Vol. 6, pp. 18-31.



$$J(n,\tau_1) = -w^L - hw^H + 2\left(\gamma w^H + (1-\gamma)w^L\right)S(h)$$
$$+ 2(h-\bar{h})\tau_1 w^H$$

$$\tau < \tau' < \tau''$$

Figure 1

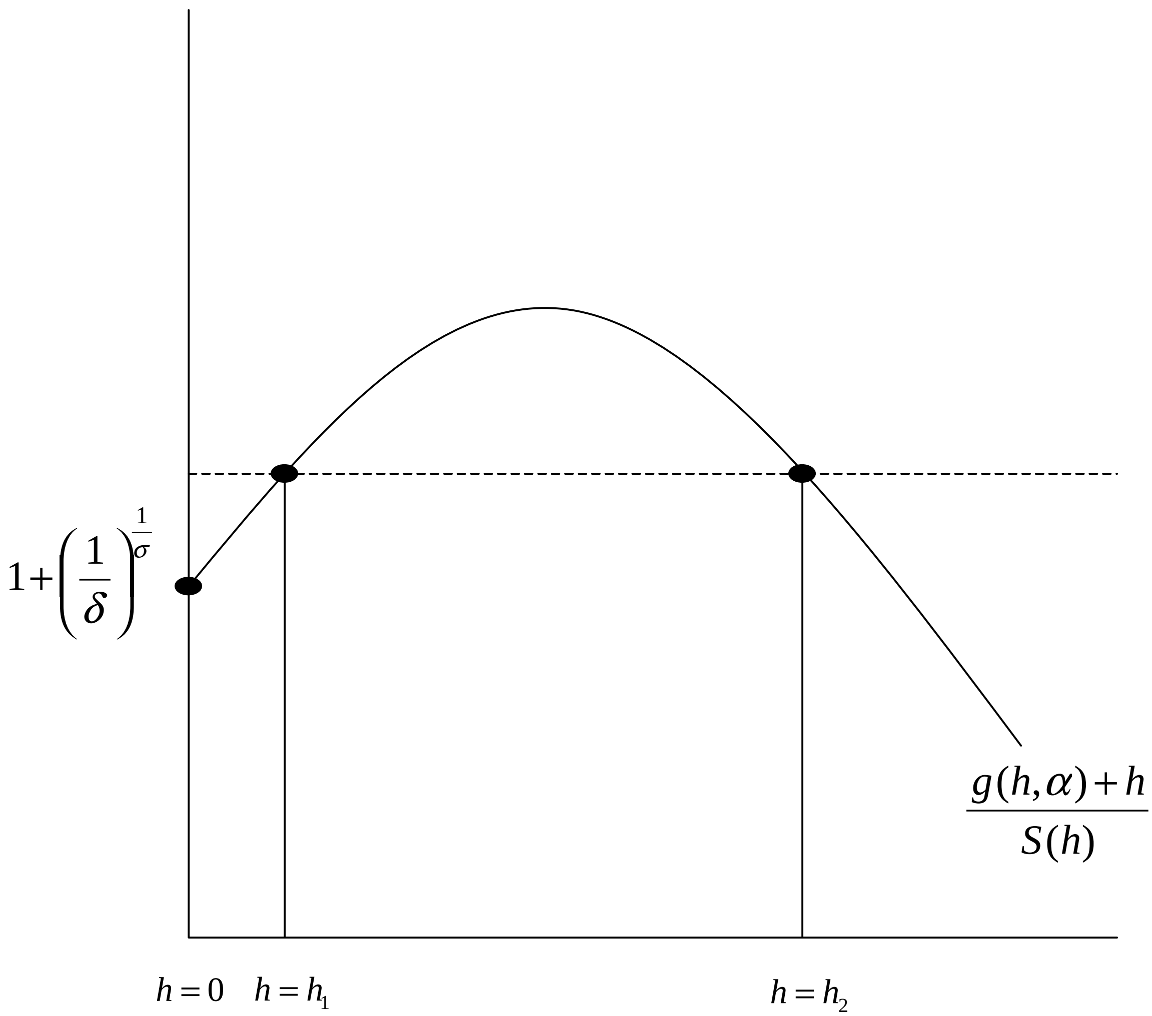

$1+\left(\frac{1}{\delta}\right)^{\frac{1}{\sigma}}$
$\frac{g(h,\alpha)+h}{S(h)}$
$h=0$
$h=h_1$
$h=h_2$

Figure 2

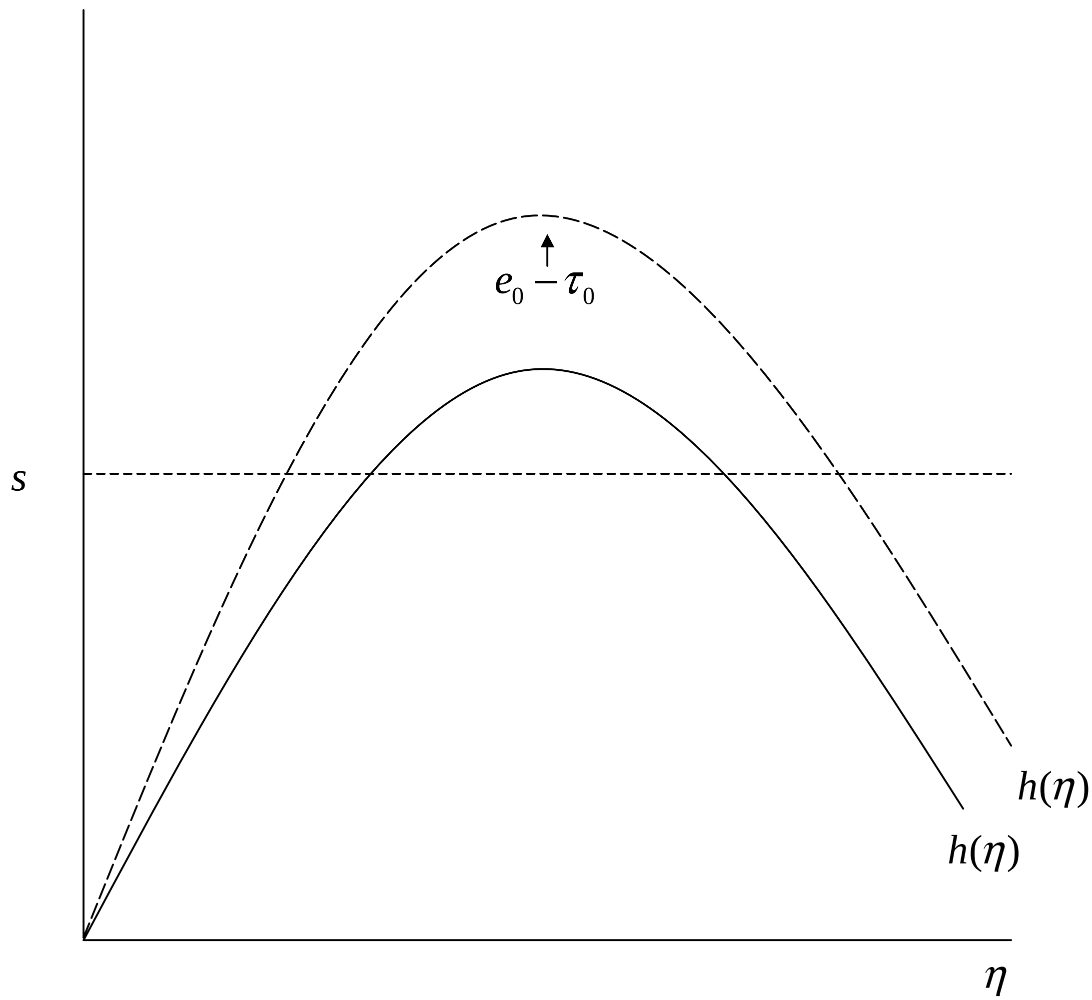


Figure 3