Investments in the Human Capital of the Socially Disadvantaged Children – Effects on Redistribution

Tim Lohse, Peter F. Lutz and Christian Thomann*
Leibniz Universität Hannover Discussion Paper No. 484
ISSN: 0949-9962

November 24, 2011

Abstract

Recently, early investments in the human capital of children from socially disadvantaged environments have attracted a great deal of attention. In a discrete version of the Mirrlees model with a parents' and a children's generation we show the intra-generational and the inter-generational redistributional consequences of such intervention programs. It turns out that the parents' generation always loses when such intervention programs are implemented. Among the children's generation it is the rich who always benefit. Despite the expectation that early intervention puts the poor descendants in a better position, our analysis reveals that the poor among the children's generation may even be worse off if the effect of early intervention on their productivity is not large enough.

Keywords: Early Intervention, Welfare, Redistribution, Taxation
JEL classification: I38, J13, H21, I14

*Berlin School of Economics and Law, Alt-Friedrichsfelde 60, D-10315 Berlin, and Social Science Research Center Berlin (WZB), Tim.Lohse@hwr-berlin.de (corresponding author); Leibniz University Hannover; Swedish Ministry of Finance and Leibniz University Hannover.

1 Introduction

Health economics considers health as a component of human capital that is improved by investments and depreciated over time (Grossman, 2000). The economic theory of human development describes skill formation as a dynamic process where early investments largely determine the productivity of inputs later in the life cycle, integrating knowledge of other sciences like developmental psychology (Cunha and Heckman, 2008). Recently, the human capital model of health economics and the theory of skill formation were assembled to a lifecycle investment framework (Heckman, 2007). This unified framework points to the importance of early investments in the human capital of children growing up in socially disadvantaged environments. However, threefold market failure keeps these children from getting an efficient amount of investments (Cunha and Heckman, 2007). Therefore early childhood intervention programs were developed which are publicly financed inputs into the skill formation processes of very young children. These interventions are carried out, for example, by specialized family nurses whose work starts with monitoring the mother's health conditions even before she gives birth.[1] Evaluations in randomized trials indicate that some of these programs give children growing up in socially disadvantaged environments a better life, and, at the same time, provide a pay-off for the taxpayer (Karoly et al., 1998). At the moment only a small fraction of those populations that could benefit from these programs is receiving them. However, these programs are spreading with considerable speed in the U.S. (Olds, 2002).[2] Attracted from the proven outcomes in the U.S. several European countries started testing early intervention programs in randomized controlled trials (RCTs).[3]

[1] The positive effects pf prenatal home visitation by nurses on pregnancy outcomes have been documented e.g. by Kitzman et al. (1997).

[2] Nurse Family Partnership, for instance, started its replication process in 1996 and is now operating in 324 counties in 25 U.S. States (Nurse Family Partnership Foundation, 2008).

[3] In Germany "Pro Kind" is in operation since 2006, carrying out a RCT with 740 participants (of whom 370 are controls), in France "CAPEDP" is undertaking a RCT with 440 participants, in Italy "Scommettiamo sui giovani" is a RCT with 100 participants and Ireland's "preparing for life" has 200 participants. In the United Kingdom, "Family Nurse Partnership", a RCT with no less than 2400 participants was announced by the Department of Health in 2008.

If the outcomes of the RCTs in Europe are as positive as in the U.S., an implementation across a broad front can be expected. In this case, these implementations will affect the redistributive equilibrium of society which is characterized by optimally balancing equity gains from redistribution against efficiency losses from taxation where taxes are used to finance redistribution. The aim of this paper is to analyze in a Mirrleesian framework (Mirrlees, 1971) the redistributional consequences of broadly applied early childhood intervention. We use a discrete variant of the Mirrlees model (Stiglitz, 1982; Homburg, 2001) to build an economy with two generations, namely a parents' generation (first generation) and a children's generation (second generation) whose welfare is jointly maximized. In this model we can determine the welfare maximizing amount of early intervention. Redistributional consequences occur not only within the same generation (intra-generational), but also between the two generations (inter-generational). In both generations there are poor and rich individuals. The parents' generation finances early childhood intervention and thus increases the skills and therefore the productivity of their eligible children later in life. The benefits from intervention are twofold: First, there is a positive output effect on those who receive intervention programs because their productivity increases. Second, there is a positive incentive effect, i.e. early intervention relaxes the incentive compatibility constraint between rich and poor individuals in the children's generation. To specify the changes in the utility positions of the two generations we compare two optimal tax-transfer schemes, one without, and one with early childhood intervention. Our main results can be summarized as follows: The parents' generation always loses in welfare terms when such intervention programs are optimally implemented. However, the children's generation always benefits in total. But, despite the expectation. that early intervention improves the position of the children sponsored, i.e. the poor children, our analysis shows that it is the children who are not sponsored, i.e. the rich ones, benefit most. In contrast, those who are sponsored are faced with a transfer reduction and must participate fully in the labor market which decreases their leisure. This could even result in their being worse off.

The remainder of the paper is organized as follows. Section 2 refers to the related literature and section 3 presents the model. Some basic results are provided in section 4 and our main theoretical findings are given in section 5. Some illustrative examples are provided in section 6, and section 7 summarizes the findings and concludes.

2 Human capital formation and the Mirrleesian economy

Early parental investment in their children's skills and health seem to be very important in promoting noncognitive skills (Cunha and Heckman, 2008). Noncognitive skills in turn promote the acquisition of cognitive skills e.g. by making children more curious and more open to learning. Hence, noncognitive skills and cognitive skills are equally important with respect to a child's future social and economic success (Heckman et al., 2006). Moreover, there is a longterm influence fo initial conditions in childhood on health for the rest of the life (Tubeuf and Jusot, 2011).

But, even if childhood is an important and sensitive period for investment in skills, this does not justify publicly financed intervention in the skill formation process. For this purpose, some kind of market failure must be taken into account. There are, as Cunha and Heckman (2007) point out, three distinct constraints that operate on the family and its children. (1) The inability of a child to choose its parents. This is the fundamental constraint imposed by the accident of birth. (2) The inability of parents to borrow against their children's future income in order to finance investment in them. (3) The inability of parents to borrow against their own income to finance investment in their children.

While the first constraint justifies child welfare activities in general, the other two also justify special intervention in families where the credit constraint is already binding. This conclusion corresponds with those of several studies in the U.S. that show that early intervention programs only generate a high fiscal surplus if they are restricted to low income families. Olds et al. (1993), for example, who tested the Nurse Family Partnership Program in several RCTs, found that, in that case, the costs of the program were outweighed by additional tax revenues and public savings four years after the children were born. If the program was not restricted to low income families, only about half of the costs would be regained after four years. Karoly et al. (1998) studied the effects of this program fifteen years after the birth of children, where the benefits exceed a multiple of the costs in the low income group, while the benefits in the low risk group are still less than the costs. While these studies confirm the theoretical underinvestment in socially disadvantaged children, they are not indicative of what will happen if the programs are broadly applied and productivity of the disadvantaged is raised at an aggregate level.[4]

Problems of this kind can be analyzed with help of the Mirrlees model, and, in fact, there are several extensions of this model in which human capital formation is considered. In an early contribution by Ulph (1977), income redistribution and educational policies are optimized simultaneously. Ulph's model assumes that the government is informed about individual abilities when allocating educational resources, while our model assumes that such information is not known for either the purposes of taxation or for early intervention. There appears to be no reason for assuming that government can observe abilities accurately when it comes to early intervention, but, at the same time, is unable to observe abilities for the purpose of taxation. As in practice, we assume early intervention programs applying screening instruments or inclusion criteria to identify promising intervention candidates. These screening instruments contribute to the overall efficiency of an early intervention technology. The efficiency serves as a parameter in our model. Again as in practice, the instruments are far from being perfect and not usable for the purpose of lump-sum taxation later in life.[5]

Tuomala (1986) considers three ways of incorporating human capital formation into the design of optimal income tax policy. First he takes the case of uniform compulsory provision of education by the government. This does not apply here,

[4] O'Neill et al. (2011) report about favourable long-run economic returns of early intervention programs which reduce childhood health inequalities. However, it is not clear that these results carry over to the broader context of socially disadvanteged children.

[5] An extension of Ulph's model by Hare and Ulph (1979) introduces a private education market that provides perfect substitutes for publicly provided education. This facet of the problem is not relevant in our case, since the two credit constraints mentioned above prevent recipients of early intervention from buying skill investment in the private market.

since early intervention is restricted to children at risk. In his second case, each individual makes his own educational choice without knowing his own ability, while our families do not make educational choices and know their abilities. In a third case, individuals make distinct labor supply and education decisions on the basis of known abilities. In our model, the grown-ups decide about their labor supply while the government implements early intervention before, namely when they are still children.

In a one-period model, Brett and Weymark (2003) consider government expenditures on education as an instrument for redistribution in addition to the income tax. A compulsory amount of education is exogenously determined, and individuals may choose additional training. Using a type aggregator, individuals differ not only with respect to their innate ability, but also with respect to their aptitude to acquire skills. Brett and Weymark focus on the sign of marginal tax rates and show that these can even be negative. However, they do not explain specific utility changes that occur due to the introduction of education. Blumkin and Sadka (2008) extend Brett and Weymark's (2003) model and show how the discretionary individual decisions to acquire education can serve as another signal, in addition to income, for the unobserved productivity. They show, in a linear tax system, that taxing (rather than subsidizing) education can therefore be optimal.

Boadway and Marchand (1995) analyze, in a Stiglitz (1982) type model, the use of public expenditures such as education for redistributive purposes. They find that public education can improve welfare but do not specify how the utility positions of the types affected change. The level of education, moreover, is not endogenous.

A more recent contribution is by Jacobs (2005), who incorporates human capital formation into the theory of optimal linear income taxation by introducing a trade-off between time spent on labor and the time and goods spent on private education. He shows that, in that case, for the class of iso-elastic utility functions the costs of distortionary taxation increase substantially. But this scenario does not apply to early intervention, as this takes place in the early childhood where no trade-off with labor supply is possible. A new variant of the Mirrlees model is therefore required to analyze our problem.

3 The model

A finite variant of the standard optimal non-linear taxation model (Stiglitz, 1982; Homburg, 2001) is extended by a second period. The variety of skills in society consists of two types, rich and poor. In the first period, poor and rich ($h = p, r$) individuals of the parents' generation are working whereas their children are not. In the second period the parents' generation is no longer alive and the children are grown up and working. Poor and rich types among the children's generation are denoted by $h = P, R$. To keep things as general as possible, we make no assumption about the transmission of types across generations. As a result, the model is open to the coherency described in the skill formation literature (e.g. Cunha et al. 2006), in which rich parents are more likely to raise highly skilled children because parental investment is larger. On the other side, the model also allows for exceptions to that rule, since these might also occur. Because of this, it is not possible for the government to derive the descendant's type from that of the parents. The shares of the grown up poor and rich remain unchanged over the periods and denoted by $\underline{f}$ and $\overline{f}$, respectively, with $\underline{f}+\overline{f} = 1$.

Poor and rich types differ in their productivities. The productivity of the rich is exogenous and constant over time, $w^r = w^R$. Conversely, the productivity of the second-generation poor can benefit from an early intervention program $g \geq 0$ which takes place in the first period, when they are children. This program has to be financed by taxes. In the second period, when the poor among the second generation are able to work, their productivity is improved by an intervention technology $i(g)$ and is given by $w^P(g) = w^p + i(g)$ with $w^{P\prime}(g) > 0$, but $w^{P\prime}(g) = 0$ for $g \to \infty$, $w^{P\prime\prime}(g) < 0$. As in reality, early intervention may mitigate, but not eliminate, social inequality, such that $w^P(g) < w^r$ for all g. Early intervention for the rich does not pay, i.e. $w^R(g) = w^R$. The resulting productivities correspond to wage rates in the competitive labor market. An individual consuming c^h and earning a gross labor income y^h enjoys utility $u(c^h, y^h/w^h)$. The argument y^h/w^h is the time type h spends in the labor market to generate the income y^h. It must hold that y^h/w^h is smaller than an exogenous upper time limit. As the government is uninformed about the produc-

tivities and cannot observe working time, the problem is denoted in the variables that the government can observe, that is, by using c and y as choice variables. Consumption corresponds to net income since the only kind of taxation is the taxation of labor income. Income taxes follow as $T^h = y^h - c^h$, where a negative difference indicates a transfer. For convenience, partial derivatives of the utility function u are denoted with subscripts that refer to the respective arguments, i.e. u_c and u_y. The utility function satisfies the usual properties (strictly monotonically increasing in c, strictly monotonically decreasing in y; and $Hess$ u is negative definite). For simplicity we assume vanishing cross-derivatives $u_{cy} = u_{yc} = 0$. Consequently, leisure is non-inferior, and this ensures that redistribution is from the rich to the poor. The intra-periodical marginal rate of substitution of an individual with productivity w^h is defined in an income-consumption space as $mrs^h(c^h, y^h/w^h) := -u_y(c^h, y^h/w^h)/(u_c(c^h, y^h/w^h)w^h)$. We assume $mrs(c, 0) < w^r$ for c small.
The government imposes an optimal g for the first period and chooses an optimal tax-transfer scheme for both periods simultaneously. We assume that the government can perfectly commit.[6]
The government's welfare maximization problem reads:

$$\begin{aligned}
\max_{g,(c^h,y^h)_{h=p,r,P,R}}! \quad W &= u(c^p, y^p/w^p)\underline{f} + u(c^r, y^r/w^r)\overline{f} && (1)\\
&\quad + u(c^P, y^P/w^P(g))\underline{f} + u(c^R, y^R/w^R)\overline{f} \\
\text{s.t. } g &\leq (y^p - c^p)\underline{f} + (y^r - c^r)\overline{f}, && (2)\\
0 &\leq (y^P - c^P)\underline{f} + (y^R - c^R)\overline{f}, && (3)\\
u(c^r, y^r/w^r) &\geq u\left(c^p, \frac{y^p}{w^r}\right) \text{ and } u(c^p, y^p/w^p) \geq u\left(c^r, \frac{y^r}{w^p}\right), && (4)\\
u(c^R, y^R/w^R) &\geq u\left(c^P, \frac{y^P}{w^r}\right) \text{ and } u(c^P, y^P/w^P(g) \geq u\left(c^R, \frac{y^R}{w^P(g)}\right) && (5)
\end{aligned}$$

This problem can be interpreted as follows: the government maximizes the expected utility of an individual from behind a veil of ignorance, i.e. of someone who knows the entire model but neither knows the period in which he is going to live nor does he know to which of the two types he will belong. Someone

[6] Commitment problems in the context of educational investments are analysed e.g. by Boadway, Marceau and Marchand (1996).

deciding under a veil of ignorance will consider each feasible allocation to be a lottery in which he will receive a bundle (c^h, y^h) with a certain probability after the veil is lifted. The resulting scheme, and the optimal level of early intervention, are determined by risk aversion and the deadweight losses from taxation.

For the sake of clarity, any other inter-generational redistribution, that possibly blurs the redistributional effects of early intervention, is excluded. That is, we assume a situation of Ricardian equivalence, where inter-generational redistribution is perfectly impeded via borrowing or saving. For simplicity, we implement this assumption by refraining from of any saving or borrowing. Therefore, maximization has to occur subject to the two budget constraints (2) and (3), one for each period. The parents' generation has to finance the provision of early childhood intervention at the level of g with its taxes.

The second-best approach to optimal income taxation holds that the government is unable to observe productivities and working hours directly.[7] So, stipulating a tax schedule means confronting the individuals with legal choices of (c, y). For example, the rich in the second period only accept the bundle (c^R, y^R) intended for them if they are not better off by pretending to be the poor, i.e. by choosing consumption c^P and a labor time of y^P/w^R. The opposite applies to the poor. As these have participated in the early intervention program during childhood, i.e. in the first period, they mimic the rich by choosing consumption c^R and working time $y^R/w^P(g)$. Therefore, every feasible allocation must also satisfy the incentive compatibility constraints (4) and (5). Of course, an inter-generational imitation, i.e. a mimicking of members of the first generation by those of the second and vice versa, is not possible.

As usual in discrete tax models, we have two different marginal tax rates. The

[7] One could argue, that, because the type of the parents is revealed in the first period, a clever government would try to infer the type of the children from the type of the parents and, hence, realize first best taxation in the second period. But if we take a closer look on the requirements of taxation, the type would have to be not only observable but also verifiable. It is in fact a classical case of an observable but not verifiable event as described in Hart and Moore (1990). Taxation affects the basic right of property and therefore has to be in accordance with the rules of law. Since the government could only guess but never prove that a descendant of a rich parent is of the type r, it is still confined to the second best world.

discrete marginal tax rates, which in our model only apply to the rich, are defined for any $y^h \neq y^k$ as

$$m^r = \frac{T^r - T^p}{y^r - y^p} \quad \text{and} \quad m^R = \frac{T^R - T^P}{y^R - y^P}. \tag{6}$$

The discrete marginal tax rate is the relevant marginal rate for the purpose of redistribution. For conventional reasons, we will denote the local (or implicit) marginal tax rates by T'^h. These are given by

$$T'^h = 1 - mrs^h. \tag{7}$$

They are the crucial marginal tax rates for incentive considerations and specify by how much the marginal rate of substitution falls short of the wage rate.

4 Basic features of optimal allocations

Proposition 1 answers the fundamental question whether the optimization problem stated is solvable.

Proposition 1 *There exists a non-trivial solution to the maximization problem specified in (1) to (5).*

Proof. The set of feasible bundles $(g, c^h, y^h)_{h=p,r,P,R}$, which is determined by the resource constraints, is compact. The self-selection constraints define a closed subset of this compact set, which in turn is itself compact. The set is not empty: in the absence of taxes, at least r and R are at work, since $mrs(c, 0) < w^r$ by assumption. Introducing a tax with small revenue does not change the decision of the rich and, by continuity, does not violate any self-selection constraint. The existence of a non-trivial optimum follows from the Weierstraß theorem. ■

For the optimal level of early childhood intervention g^*, the following proposition holds.

Proposition 2 *The welfare maximizing level of early childhood intervention, g^*, is characterized by*

$$\underline{f}\frac{w^{P'}(g^*)}{w^P(g^*)}\frac{y^P}{w^P(g^*)}\left(-u_y(c^P, \frac{y^P}{w^P(g^*)})\right) = \frac{1}{\underline{f}/u_c(c^p, y^p) + \overline{f}/u_c(c^r, y^r)}. \tag{8}$$

Proof. Since we assume that inter-generational redistribution occurs only via early intervention g, some results from the standard optimal tax analysis apply. There is no output surplus in any optimum , i.e. the governmental budget constraints (2) and (3) are binding. Moreover, due to the redistribution from the rich to the poor, the downward incentive compatibility constraints in (4) and (5) are binding, whereas there is some slack in the upward constraints. As the upward constraints are satisfied automatically (so called chain property), they can be ignored. As the gradients of the constraints are linearly independent, a Lagrangian approach can be formulated. Let λ_1 and λ_2 be the non-negative multipliers for the two binding budget constraints (2) and (3), and μ_1 and μ_2 multipliers of the downward binding incentive compatibility constraints of (4) and (5), respectively (Bertsekas, 1999; Homburg, 2003). Let

$$\begin{aligned}
\mathcal{L} &= u(c^p, y^p/w^p)\underline{f} + u(c^r, y^r/w^r)\overline{f} + u(c^P, y^P/w^P(g))\underline{f} + u(c^R, y^R/w^R)\overline{f} \\
&\quad +\lambda_1 \left((y^p - c^p)\underline{f} + (y^r - c^r)\overline{f} - g\right) + \lambda_2 \left((y^P - c^P)\underline{f} + (y^R - c^R)\overline{f}\right) \quad (9) \\
&\quad +\mu_1 \left(u(c^r, y^r/w^r) - u(c^p, \frac{y^p}{w^r})\right) + \mu_2 \left(u(c^R, y^R/w^R) - u(c^P, \frac{y^P}{w^R})\right)
\end{aligned}$$

be the associated Lagrangian. The partial derivatives of (9) with respect to $g > 0$, c^p and c^r read, respectively,

$$\frac{\partial \mathcal{L}}{\partial g} = -\underline{f} u_y(c^P, \frac{y^P}{w^P(g)}) \frac{y^P}{w^P(g)^2} w^{P'}(g) - \lambda_1 = 0, \qquad (10)$$

$$\frac{\partial \mathcal{L}}{\partial c^p} = u_c(c^p, y^p) \left(\underline{f} - \mu_1\right) = \lambda_1 \underline{f}, \qquad (11)$$

$$\frac{\partial \mathcal{L}}{\partial c^r} = u_c(c^r, y^r) \left(\overline{f} + \mu_1\right) = \lambda_1 \overline{f}. \qquad (12)$$

Dividing (11) and (12) by the respective partial derivative of the utility function, adding up the equations, and rearranging terms yield

$$\lambda_1 = \frac{1}{\underline{f}/u_c(c^p, y^p) + \overline{f}/u_c(c^r, y^r)}. \qquad (13)$$

Substituting for λ_1 according to equation (13) into (10) and rearranging terms leads to (8). ■

The optimality condition (8) specifies a cost-benefit equalization. The left-hand side of (8) describes a positive leisure effect for the second-generation poor.

The first fraction is the percentage increase in the productivity that is due to early intervention. This increase is scaled by its working time measured in its negative disutility of labor. The entire effect is weighted by the size of the poor in the population. The right-hand side shows the marginal budgetary costs of early intervention which only accrue to the parents' generation. The marginal costs are given as an average of the marginal utility of consumption of both parental types which has to be given up to finance the intervention. This average is precisely the harmonic mean of the marginal utility of consumption in the parents' generation.[8]

The crucial requirement for early intervention to be welfare increasing is its productivity. Clearly, if the first euro invested does not yield anything, i.e. $\underline{f}w^{P\prime}(0) = 0$, then the introduction of such a program can never be optimal. But condition (8) elucidates that it is on the other hand not necessary for the program itself lead to a return in output higher than the amount invested, i.e. $\underline{f}w^{P\prime}(g) > 1$ for $g > 0$. In fact, it is the reduction in the poor's disutility of labor that benefits the society. Therefore, a sufficiently productive skill formation technology leads to $g^* > 0$. Interestingly, early intervention at a first step does not entail an incentive cost nor an incentive benefit as the first-order conditions reveal.

Optimal intervention has another interesting effect on the second-generation poor in addition to an increase in their productivity.

Remark *In a welfare optimum with early intervention $g^* > 0$, the poor in the children's generation are never unemployed.*

Proof. If the poor in the children's generation were unemployed in the optimum, $y^P = 0$, the condition for optimal early intervention (8) would be violated since the left-hand side would be zero and the right-hand side would be strictly positive. ■

The increase in the productivity of the poor increases the opportunity costs of being jobless. If a poor individual were not employed regularly afterwards, the intervention program would have created only costs but no benefits and

[8] An intuition for the harmonic rather than the arithmetic or geometric mean is provided by Kocherlakota (2005, p. 1593 and 1594).

therefore a welfare loss.

5 Welfare effects

If in the optimum $g^* > 0$, then we can be sure that welfare has increased compared to an optimum without early intervention. This is due to a core insight from optimization theory: the above optimization problem is technically identical to the standard problem of optimal income taxation without intervention plus the constraint $g = 0$. Removing this binding constrain leads to an increase in the target value.

The decisive questions are which generation benefits or loses and, more specifically, who among each generation is better or worse off due to early childhood intervention. The common belief is that these programs are introduced in the interest of the disadvantaged children.

Knowing that the rich are always subject to a zero local marginal tax rate,[9] we can state the following proposition with respect to the parents' generation:

Proposition 3 *Both rich and poor parents are worse off if there is optimal early childhood intervention $g^* > 0$ for the childrend generation.*

Proof. Consider the sub-maximization problem for the first period. The maximum value M of the parents' generation's expected utility, $u(c^p, y^p/w^p)\underline{f} + u(c^r, y^r/w^r)\overline{f}$, which is maximized subject to (2) and the left-hand inequality of (4), decreases because for it the only effect of g is the negative budgetary effect, i.e. $dM/dg = -\lambda_1 < 0$.

There are 3^2 possibilities with respect to the individual utility changes. The cases where u^r remains constant and u^p increases, u^r and u^p remain constant, u^r and u^p increase, and u^r increases whereas u^p remains constant can be excluded right from the start since the maximum value M would not decrease.

[9] Local marginal tax rates are used to distort the choice of consumption-income-bundles in order to prevent more productive types from mimicking the adjacent, less productive type. Since the rich do not have a more productive type who might imitate them, there is no need to tax the rich in a distortionary way. This is one of the very few robust insights of the standard Mirrlees model (Hellwig, 2007).

A scenario with u^r remaining constant and u^p decreasing is not possible either because r always faces $T'^r = 0$. On each indifference curve there is only one income-consumption combination which corresponds to this undistorted taxation because indifference curves are strictly monotonically increasing in income-consumption-space (Homburg, 2001). Therefore, if u^p decreases but r still enjoys its undistorted bundle, either the left-hand incentive compatibility constraint in (4) becomes slack if the $p's$ bundle is below $r's$ indifference curve or the budget constraint (2) becomes slack if $p's$ bundle is on $r's$ indifference curve. Neither of the situations can be an optimum. A bundle for p above $r's$ indifference curve would violate the left-hand constraint (4) and is therefore not feasible.

An increase in u^r along a decrease in u^p would imply that r enjoys more leisure or more consumption, both of which lead to a lower amount of T^r despite the increase in g. If this were optimal, such an utility increase for r would have also been feasible in a setting with $g = 0$. Since it did not turn out to be optimal previously, such a utility change is by no means optimal now. A similar argument holds for a decrease in u^r, but an increase in u^p which would imply an increase in the poor's leisure or its consumption, both of which would require a higher tax load for the rich in addition to g which also needs to be financed. A decrease in u^r while u^p remains constant is not optimal either: suppose it were optimal, then a decrease in g would leave the poor's utility unaffected whereas the rich would be better off. However, this would make it impossible for the transfer system to redistribute from top to bottom, which is ensured by the assumption that leisure is non-inferior.

Hence, the only possible case left is that both u^r and u^p decrease which proves the claim. ■

By the strict concavity of the parents' utility as a sum of strictly concave functions, we have the following effects: due to the increased revenue requirement for both types, leisure becomes more expensive and this drives up both types' workload. Moreover, as g increases the rich lose because now, besides redistribution, they also have to finance the provision of intervention. The poor are also worse off since the extent of redistribution decreases due to provision of g. There is no positive effect for both parental types that might countervail their utility losses.[10]

Considering the children's generation, we can state the following:

Proposition 4 *The children's generation in total always benefits from optimal early intervention $g^* > 0$. More specificly, the rich among the children's generation gain from optimal childhood intervention in either case, whereas the effect on the poor among the children's generation indetermined.*

Proof. Consider the sub-maximization problem for the second period. The maximum value N of the second generation is given by

$$\begin{aligned} N &= u(c^P, y^P/w^P(g))\underline{f} + u(c^R, y^R/w^R)\overline{f} + \lambda_2\left((y^P - c^P)\underline{f} + (y^R - c^R)\overline{f}\right) \\ &\quad + \mu_2\left(u(c^R, y^R/w^R) - u(c^P, \frac{y^P}{w^R})\right). \end{aligned}$$

The maximum value N increases in w^P as g drives up the productivity of the poor

$$\frac{dN}{dg} = -\underline{f}u_y(c^P, y^P/w^P(g))\frac{y^P}{(w^P)^2}w^{P'}(g) > 0, \tag{14}$$

provided $y^P > 0$. Hence, the second generation benefits from $g > 0$.

To infer the group specific utility effects consider the following two extreme cases. First, suppose that $y^P = 0$ for a certain productivity $\widetilde{w}^P(g) < w^R$. The incentive compatibility constraint would require that the utility levels u^P and u^R are, in fact, the same. Second, suppose that $\widetilde{\widetilde{w}}^P(g) = w^R = w$ where again $u^P = u^R$. In such a situation it must hold that $y^P > 0$ since we assumed $mrs(c, 0) < w^R$ implying $y^R > 0$.

Starting from the first situation, an increase in $w^P(g)$ above a threshold $\overline{w}^P$ such that $y^P > 0$ has one of the following utility effects:[11]

a) u^P increases and u^R increases.

b) u^P decreases and u^R increases.

c) u^P remains constant and u^R increases.

Out of the 3^2 potential situations, the cases with u^P decreasing and u^R remaining constant, u^P and u^R decreasing, u^P remaining and u^R decreasing, u^P and

[10] We abstain from modelling altruism since our main interest is the children's generation, not the parents' generation.

[11] The threshold $\overline{w}^P$ exists since the second extreme case occurs at the very end.

u^R remaining constant can be excluded right away since N would not increase. The cases where u^P increases and u^R remains constant and u^P increases whereas u^R decreases can also be excluded since utility would no longer be monotone in type as is required by the downward binding incentive compatibility constraint (5). ■

The intuition for the ambiguous effect on the poor among the children's generation is as follows. Due to their higher productivity, leisure is more expensive now than it was for the poor parents. Since the distribution of productivities has become more equal, the pressure for redistribution has also decreased. Therefore, the second-generation poor face a double utility loss compared to a situation without early intervention: firstly, they receive fewer transfers and secondly they have to increase their labor supply which creates a higher disutility from work. Second-generation poor may only gain from early intervention if the increase in their productivity is sufficiently high to enjoy consumption at a level which compensates them for these utility losses. That is, early intervention must be sufficiently effective to ensure that second-generation poor's productivity exceeds an implicit threshold. If we think about the extreme case of equal productivities of second-generation poor and rich, it is clear that such a threshold does exist. However, it is the dilemma of the efficient provision of the early intervention program that exceeding this threshold is not always welfare maximizing (refer to the illustrative simulations below). In fact, optimizing the extent of redistribution and the size of early intervention programs simultaneously means that, in sum, both generations are better off. But this does not necessarily imply that the second-generation poor are better off.
However, the second-generation rich always benefit from a productivity increase of their poor counterparts. The intuition is straightforward: since their tax burden decreases as the poor become more productive, they can enjoy more consumption and need to work less. This effect is monotone in the poor's productivity.
The overall welfare increase is - as usual in these extended optimal tax models - twofold.[12] Firstly, there are gains from the skill formation technology by

[12] See Lohse (2008) for a twofold gain from introducing workfare in an optimal tax-transfer-

which an increase in the economy's output can be reached. Secondly, there are efficiency gains. Since the skill distribution becomes more equal, the level of distortionary taxation can be decreased.

6 Numerical illustration

To illustrate the findings, some simulations of optimal tax-transfer schemes without and with early intervention are provided. They clarify the welfare effects of such programs and show how assumptions about the programs productivity affect second-best allocations.
Assume a utility function $u(c^h, l^h) = 1,000[\ln(c^h) + \ln(500 - l^h) - 11]$ with 500 as the maximum time per month. In a scenario without early intervention (which can be achieved by imposing $g = 0$ as an additional constraint to the above maximization problem) the tax system is purely redistributive. Table 1 depicts the standard optimum without such a program.

Table 1: Standard optimum

Period 1

h	w	f	c	y	T	m	T′	u
p	1	30%	315	11	−304	–	36%	945
r	3	70%	685	815	130	54%	0%	960

W$_1$= **955**

Period 2

h	w	f	c	y	T	m	T′	u
P	1	30%	315	11	−304	–	36%	945
R	3	70%	685	815	130	54%	0%	960

W$_2$= **955**

W = **W**$_1$ + **W**$_2$= **1911**

The table shows in the first three columns the type, the type's productivity, and the population size of the poor and the rich, respectively. The next two

scheme.

columns show consumption and income which are followed by the description of the tax schedule, i.e. the tax or transfer payments, the discrete and the local marginal tax rate. The last column shows the utility of the different types. In the last line, the overall welfare as the sum of the welfare in each period is given. Obviously, the allocation is the same in both periods.

Allowing for an early intervention program with $i(g) = \frac{\sqrt{g}}{7}$, the second generation poor are able to increase their productivity thanks to the program. However, g is costly and has to be financed by taxes in the first period. Table 2 gives the resulting optimum with an early intervention.

Table 2: Optimal early intervention

Period 1								
h	**w**	**f**	**c**	**y**	**T**	**m**	**T′**	**u**
p	1	30%	310	16	−295	–	36%	920
r	3	70%	679	821	142	54%	0%	942
$\mathbf{g = 11}$								$\mathbf{W_1 = 936}$
Period 2								
h	**w**	**f**	**c**	**y**	**T**	**m**	**T′**	**u**
P	1,48	30%	394	215	−179	–	25%	848
R	3	70%	712	788	77	45%	0%	1036
								$\mathbf{W_2 = 980}$
								$\mathbf{W = W_1 + W_2 = 1915}$

The second column denotes the poor's productivity which has increased by 48% due to the implementation of an early intervention program with size $g = 11$. The second-generation rich are now better off, since their tax burden has substantially decreased. By contrast, the second-generation poor are worse off. They fully participate in the labor market and their transfer income is shortened by half. Although the parents' generation has also lost, overall welfare has increased.

7 Conclusion

This paper supports the positive view of investments in the human capital of the socially disadvantaged children which is derived from a lifecycle investment framework as a unified approach of health economics and skill formation theory. Given that the return from such early intervention programs is sufficiently large, their implementation is welfare increasing. We provide an optimality condition that characterizes the interaction of output, utility, and incentive effects that go along with such programs. Our results indicate that, while the parents' generation always loses from early intervention, the rich among the children's generation always benefit. Contrary to the widely hold expectation, the effect on the poor among the children's generation - the target group of such programs - is ambiguous. They may have to work more in return for their beneficial support. Consequently, they receive fewer transfers and face an increased disutility of work which may lead to a decrease in their well-being if the impact of early intervention on their productivity is not large enough to compensate for this.

In evaluating the model, three caveats come to mind. Firstly, early childhood intervention programs can easily cover more than two generations, that is, the second generation would have to finance early intervention for a third generation and so on. But, in that case, our model can be seen as a simplification of a multi-generational intervention process showing the redistributional effects at the beginning and at the end of that process. Secondly, the present model takes for granted a positive outcome for early intervention. However, there is uncertainty about the parents' cooperation and the individual development. Hence, the present model may be extended by some stochastic component of intervention. And thirdly, the Mirrleesian framework does not capture explicitly all externalities of early intervention (as shown e.g. by Gomby, 2007). This refers to parential altruism or to beneficial effects with regard to an improved mental health or a reducation of criminality. To draw a complete picture, these effects should be added in future interdisciplinary research.

References

[1] Bertsekas, D. P. (1999), *Nonlinear Programming*, 2nd ed., Athena Scientific.

[2] Blumkin, T. & Sadka, E. (2008), 'A Case for Taxing Education', *International Tax and Public Finance* **15**, 145–163.

[3] Boadway, R. & Marchand, M. (1995), 'The use of public expenditures for redistributive purposes', *Oxford Economic Papers* **47**, 45–59.

[4] Boadway, R. & Marceau, N. & Marchand, M. (1996), 'Investment in Education and the Time Inconsistency of Redistributive Tax Policy', *Economica* **63**, 171–189.

[5] Brett, C. & Weymark, J. (2003), 'Financing education using optimal redistributive taxation', *Journal of Public Economics* **87**, 2549–2569.

[6] Cunha, F. & Heckman, J. (2007), 'The technology of skill formation', *American Economic Review* **97**, 31–47.

[7] Cunha, F. & Heckman, J. (2008), 'Formulating, Identifying, and Estimating the Technology of Cognitive and Noncognitive Skill Formation', *Journal of Human Resources* **43**, 738–782.

[8] Cunha, F. & Heckman, J. & Lochner, L. J. & Masterov, D. V. (2006), 'Interpreting the evidence on life cycle skill formation'. In: Hanushek, E. A. & Welch, F. (eds.), *Handbook of the Economics of Education*, North-Holland, forthcoming.

[9] Gomby, D. S. (2007), 'The promise and limitations of home visiting: Implementing effective programs', *Child Abuse & Neglect* **31**, 791–799.

[10] Grossman, M. (2000), *The human capital model* in: A. J. Culyer & J. P. Newhouse (eds.), Handbook of Health Economics, vol.1, ch.7, 347–408.

[11] Hare, P. & Ulph, D. (1979), 'On education and distribution', *Journal of Political Economy* **87**, 193–212.

[12] Hart, O. & Moore, J. (1990), 'Property Rights and the Nature of the Firm', *Journal of Political Economy* **98**, 1119–1159.

[13] Heckman, J. J. (1976), ‘A life-cycle model of earnings, learning and consumption’, *Journal of Political Economy* **4**, 11–44.

[14] Heckman, J. J. (2007), ‘The Economics, Technology, and Neuroscience of Human Capability Formation’, *Proceedings of the National Academy of Sciences of the United States of America* **104**, 13250–13255.

[15] Heckman, J. J. & Stixrud, J. & Urzua, S. (2006), ‘The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior’, *Journal of Labor Economics* **24**, 411–482.

[16] Hellwig, M. F. (2007), ‘A contribution to the theory of optimal utilitarian income taxation’, *Journal of Public Economics* **91**, 1449–1477.

[17] Homburg, S. (2001), ‘The optimal income tax: restatement and extensions’, *Finanzarchiv* **58**, 363–395.

[18] Homburg, S. (2003), ‘An Axiomatic Proof of Mirrlees’ Formula’, *Public Finance/Finances Publiques* **53**, 285–295.

[19] Jacobs, B. (2005), ‘Optimal income taxation with endogenous human capital’, *Journal of Public Economic Theory* **7**, 295–315.

[20] Karoly, L. A. & Greenwood, P. W. & Everingham, S. S. & Hoube, J. & Kilburn, M. R. & Rydell, C. P. & Sanders, M. & Chiesa, J. (1998), *Investing in our children: What we know and don’t know about the costs and benefits of early childhood interventions*, RAND Corporation.

[21] Kitzman, H. et al. (1997), ‘Effect of Prenatal and Infancy Home Visitaion by Nurses on Pregnancy Outcomes, Childhood Injuries, and Repeated Childbearing’, *Journal of the American Medical Association* **278**, 644–652.

[22] Kocherlakota, N. R. (2005), ‘Zero expected wealth taxes: a Mirrlees approach to dynamic optimal taxation’, *Econometrica* **73**, 1587–1621.

[23] Lohse, T. (2008), *Arbeitsverpflichtungen und ihre steuertheoretische Beurteilung*, Gabler.

[24] Mirrlees, J. (1971), 'An exploration in the theory of optimum income taxation', *Review of Economic Studies* **38**, 175–208.

[25] Nurse Familiy Partnership Foundation (2008), *Current Number of States, Counties, Clients, June 2008*, Retrieved July 7th, 2008 from http://www.nursefamilypartnership.org/resources/files/PDF/Fact_Sheets /NFP_Snapshot.pdf..

[26] O'Neill, D. & McGilloway, S. & Donnelly, M. & Bywater, T. & Kelly, P. (2011), 'A cost-effectiveness analysis of the Incredible Years parenting programme in reducing childhood health inequalities', *European Journal of Health Economics*, forthcoming.

[27] Olds, D. L. (2002), 'Prenatal and Infancy Home Visiting by Nurses: From Randomized Trials to Community Replication', *Prevention Science* **3**, 153–172.

[28] Olds, D. L. & Henderson, C. R. Jr & Phelps, C. & Kitzman, H. & Hanks, C. (1993), 'Effect of prenatal and infancy nurse home visitation on government spending', *Med Care* **31**, 155–174..

[29] Stiglitz, J. E. (1982), 'Self-Selection and Pareto Efficient Taxation', *Journal of Public Economics* **17**, 213–240.

[30] Tubeuf, S. & Jusot, F. (2011), 'Social health inequalities among older Europeans: the contribution of social and family background', *European Journal of Health Economics* **12**, 61–77.

[31] Tuomala, M. (1986), 'On the optimal income taxation and educational decisions', *Journal of Public Economics* **30**, 183–198.

[32] Ulph, D. (1977), 'On the optimal distribution of income and educational expenditure', *Journal of Public Economics* **8**, 341–356.