TI 2003-024/3
Tinbergen Institute Discussion Paper

Are Education Subsidies an Efficient Redistributive Device?

Robert A.J. Dur
Coen N. Teulings

Faculty of Economics, Erasmus University Rotterdam, and Tinbergen Institute.

Are education subsidies an efficient redistributive device?

Robert Dur and Coen N. Teulings

Tinbergen Institute, Erasmus University Rotterdam

Address: P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

E-mail: dur@few.eur.nl; teulings@few.eur.nl.
This version: August 2003.

Abstract

We argue that promoting education may be a means to reduce income inequality. When workers of different skill levels are imperfect substitutes in production, an increase in the level of human capital in the economy reduces the return to education. Hence, a given compression of after-tax incomes can be achieved at lower marginal tax rates. Optimal redistribution policy faces a trade-off between the distortions of taxes on effort and the distortions of education subsidies on the investment in human capital. We discuss empirical evidence on three crucial elasticities. Our argument explains the actual pattern of education subsidies in OECD countries quite well.

However, there is an offsetting effect. When education and ability are complements, high ability types take up more education. A subsidy to education will then favor these types. We discuss the condition for the net effect of education subsidies to be progressive. The empirical evidence suggest that this condition is critical for a simple education subsidy. We consider some more elaborate schemes for education subsidies.

JEL codes: H21, H52, J24, J31.

Keywords: income inequality, optimal taxation, education.

1 Introduction

Throughout the Western world, education is heavily subsidized. Public expenditures on education amount to some 6% of GDP on average in OECD countries and make up a considerable share of total public expenditures. Public policy regarding education typically has a broad character. Subsidies are not confined to primary education; secondary and higher education are also heavily subsidized. Moreover, governments do not only support schooling opportunities of the disadvantaged. In practice, government programs that encourage education also favor the rich. This chapter is concerned with the question of whether governments should subsidize education so heavily and comprehensively?.

The economics literature offers two main arguments for subsidies to education. Neither of them can fully account for the wide prevalence of education subsidies. First, the endogenous growth literature has emphasized that investment in human capital may have positive spillover effects in production (Lucas, 1988, and Tamura, 1991). As these externalities are not taken into account in individual schooling decisions, education subsidies are needed to prevent underinvestment in education and to promote economic growth. The externality argument calls for subsidies directed to all educational levels and all individuals in as far as externalities are present. However, the evidence for positive externalities is mixed (Acemoglu and Angrist, 1999, Bils and Klenow, 2000, Krueger and Lindahl, 2000, and Teulings and Van Rens, 2002). Second, capital market imperfections may hinder poor individuals to finance educational expenditures and cost of living while at school (Saint-Paul and Verdier, 1993, Perotti, 1993, and Benabou, 2000 and 2002). This argument is hard to reconcile with the comprehensiveness of government subsidies to education. If education subsidies only serve to attain equality of opportunity, subsidies targeted at the disadvantaged would be sufficient. Moreover, the empirical evidence for borrowing constraints for educational choices is limited (Cameron and Heckman, 1998 and 1999, Keane and Wolpin, 2001, Shea, 2000, and Cameron and Taber, 2000).

In this chapter, we discuss a new rationale for education subsidies. In the spirit of Tinbergen (1975), we argue that education subsidies may be a part of an optimal redistribution policy. Our argument hinges on general equilibrium effects of an increase in human capital formation. When workers of different skill levels are imperfect substitutes in production, an increase in the mean level of human capital in the economy reduces the return to human capital. The supply of high-skilled workers goes up, reducing their relative wages, while the supply of low-skilled goes down, increasing their relative wages. Hence, the return to human capital (the difference between both wage levels) and pre-tax wage inequal-

ity go down. The reduction in pre-tax income inequality implies that a given after-tax income distribution can be reached with less progressive income taxes. Hence, by promoting education, the distortionary cost of progressive taxation may be reduced. Optimal redistribution policy faces a trade-off between the distortions arising from education subsidies and the distortionary effect of income taxation.

Following Becker's (1983) efficient redistribution hypothesis, our analysis contributes to the understanding of observed institutions. Insofar as the political system has an incentive to consume Pareto improving policy reforms, our model provides a positive theory of the tax structure: observed institutions should be constrained Pareto efficient. We present some empirical evidence that observed institutions fit our model reasonably well. Our theory predicts a correlation between the progressivity of the income tax and the level of education subsidies. We present data which give some support to this hypothesis. The level of this correlation and the average level of education subsidies, 6 % of GDP, correspond surprisingly well with the predictions of the model for reasonable parameter values. Also, our model explains why cross country differences in the dispersion of disposable income are primarily due to differences in the dispersion of gross income, not to differences in the progressivity of the tax system.

Our theoretical analysis stands in the tradition of Mirrlees' (1971) Noble prize winning paper on optimal income taxation. Mirrlees considers the case where worker types are perfect substitutes, so that relative wages for various ability types are independent of supply and demand. Imperfect substitution between worker types is crucial for our analysis. Previously, Feldstein (1973) has analyzed this problem, and a whole 1982 issue of the *Journal of Public Economics* is devoted to the issue (Allen, 1982, Stern, 1982, and Stiglitz, 1982). The conclusion of these early contributions is that imperfect substitution between types of labor does not make a great deal of difference for realistic values of the elasticity of substitution. Our claim is that this conclusion is largely due to an unresolved technical problem. Where Mirrlees applied a continuous type distribution for the perfect substitution case, a continuous type production function with imperfect substitution was not available. Hence, a production function with a discrete number of types (in practice: two types, see Johnson, 1984, for a model with three types) was applied. Teulings (2000) shows that using a production function with only two instead of a continuum of types yields a seriously downwardly biassed estimate of the spill-over effects of minimum wages. Our claim is that the same problem applies for general equilibrium effects of an increase in the mean level of human capital, since large shifts in relative wages

3

within each of the types are ignored.

Our analysis calls for subsidies to all levels of education. This redistribution policy contrasts sharply with the usual idea of compressing the wage distribution via compression of the distribution of human capital, that is by putting special policy effort in raising the education of the least skilled. This latter policy relies on direct, partial equilibrium effects of investment in human capital. However, it is unclear why a subsidy to the human capital acquisition of the poor would dominate generic income support. Standard economic theory suggests it does not. Moreover, subsidies to the human capital acquisition of the poor might run into trouble due to adverse general equilibrium effects. By a standard supply and demand argument, the additional human capital in the lower tail of the skill distribution is likely to reduce the wages for these skill types. Empirical evidence on the relation between the income distribution and the distribution of human capital in a panel of countries supports these ideas. There is a strong negative relation between the mean level of education and income inequality, supporting the general equilibrium effect, but there seems to be no relation between the variance of education and income inequality, not supporting the direct effect. This arguments suggests that the promotion of education at all levels contributes more to a progressive income policy than a policy focussed at the low levels only.

However, there is a counterforce that limits the redistributive virtues of subsidies to education. The large literature on the ability bias in the return to education shows that education and innate ability are complementary (Angrist and Krueger, 1991). Subsidies to all levels of education favor therefore predominantly the high ability types, leading to a widening instead of a compression of the income distribution. We face the remarkable situation that the role of the direct effect on income (or: utility) and the substitution effect in redistribution are reversed. Usually, redistribution is brought about by the direct effect of a policy on income (e.g. in the case of progressive income taxation), while the substitution effect (less productive effort) reduces its effectiveness. For education subsidies, it is the other way around. Substitution effects contribute to redistribution, while the direct effect on income works in the opposite direction. We derive the condition for education subsidies to be redistributive. Furthermore, we discuss some more elaborate schemes for education subsidies which reduce the adverse direct effect on income, while maintaining their favorable substitution effect.

The structure of this chapter is as follows. Section 2 presents the main elements of our theoretical model, based on previous work, see Dur and Teulings (2001). Next, we review in section 3 the empirical

evidence on three crucial elasticities in our model: elasticities of supply and demand for human capital and the degree of complementarity between education and innate ability. Section 4 analyzes the optimal redistribution policy. First, we discuss efficient redistribution policy in the absence of complementarity between ability and education. We present some crude empirical evidence regarding the relation between education subsidies and the progressivity of the tax system. Next, we examine how optimal policy is affected by the complementarity between ability and education. Section 5 discusses some further implications of our analysis. Four issues are at stake here. First, we discuss the proper level of centralization of education policy. As the effectiveness of education as a redistribution policy rests on an externality in schooling decisions, education policy should be sufficiently centralized for all externalities to be internalized. Second, we analyze the consequences of a lack of commitment on the side of the government regarding its income policy, which limits its capability to credibly internalize the externalities. Third, we discuss more elaborate schemes for education subsidies that help to reduce the adverse direct effect on the income distribution due to the complementarity between ability and education. Finally, we discuss the role of non-linearities in the income policy. Section 6 concludes.

2 The model of the economy

Our theoretical analysis follows previous work, see Dur and Teulings (2001). The reader who wants to understand the formal details of the analysis should consult that reference. Here, we use a more intuitive presentation. In order to provide a clear cut separation between our model and models based on capital market imperfections, we assume perfect capital and insurance markets. Individuals can borrow sufficient funds to finance their consumption during their years of education at the going interest rate. Also, they can insure perfectly the risk on their investment in human capital due to the uncertainty about their life expectancy. For the sake of simplicity, we ignore the direct cost of education and focus on the cost of foregone labor income. This fits the empirical observation that the direct cost of education are, relatively, of minor importance.[1] Individuals choose their years of schooling as to maximize their lifetime utility. Hence, individuals invest in human capital up to the point where the marginal cost equals the market rate of return to human capital. For

---

[1]A referee suggested that this might seriously bias our conclusions. If students' hours and other inputs are fully complementary in the production of human capital, the direct cost of education can be modelled as a simple surcharge on the Mincerian rate of return. If not, new and interesting questions arise, which are discussed tentatively in section 5.

the transparency of the analysis, we abstract from imperfections on the labor market. We also abstract from production externalities in schooling decisions, like knowledge spillovers. Workers therefore earn their marginal product of labor and there is no unemployment. Furthermore, these assumptions imply that the private return to education (i.e.: the effect on wages) is equal to the social return (i.e.: the effect on GDP). Finally, types of labor are the only factors of production.

Our economy consists of individuals who are born with different levels of innate ability, which we denote by a. They spend the first years h of their life at school. These two factors, innate ability and years of schooling, jointly determine the human capital or skill level s with which the worker enters the labor market. After the investment in human capital, individuals start their working career. The individual's log wage rate $w(s, \mu)$ per unit of effort depends on the individual's human capital s and on μ, the mean level of human capital among the workers in the economy. The partial derivative of log wages $\partial w / \partial s$ is the Mincerian rate of return to human capital. Gross income is the product of this wage rate and the effort the individual chooses to provide. Individuals set effort as to maximize their utility. We follow Diamond (1998) in simplifying the analysis by applying a utility function that is additive in the cost of effort and consumption. This additive specification rules out income effects on the supply of effort. The supply of effort of a worker is therefore increasing in her (net) wage rate.

Subsequently, we discuss the two most important building blocks of our model, the production function of human capital and relative wages, more extensively. For this discussion, it is useful to define a benchmark equilibrium of this economy. For this, we take the non-interventionist, redistribution free equilibrium, where the government does not implement any income policy, so that the income distribution is fully determined by market forces. Without loss of generality, we normalize our measure of human capital s such that in this redistribution free equilibrium, its Mincerian rate of return to human capital is unity, $\partial w / \partial s \equiv w_s = 1$, and its mean is equal to zero, $\mu = 0$. We shall apply one further normalization in this redistribution free equilibrium below.

2.1 Human capital production

The human capital of a worker is a function of two inputs, her innate ability a and the years of education h:

$$s(a, h) = a + (\beta h - \xi a) - \frac{1}{2} \psi (\beta h - \xi a)^2 \tag{1}$$

where β, ξ, and ψ are (weakly) positive parameters. Without loss of generality, we normalize the ability measure a such that in this redistribution free equilibrium, $\beta h - \xi a = 0$. The marginal return of an additional year of schooling is given by the partial derivative of $s(a, h)$ with respect to h:

$$s_h = \beta \left[1 - \psi \left(\beta h - \xi a\right)\right] \tag{2}$$

Since $\beta h - \xi a = 0$ in the redistribution free equilibrium, β measures the marginal return to a year of education, $s_h = \beta$. The term $-\psi \left(\beta h - \xi a\right)$ in equation (2) captures two essential features of the model, the decreasing returns to education, $s_{hh} \leq 0$, and complementarity of ability and education, $s_{ha} \geq 0$. Both features are discussed in some detail.

The decreasing returns to education provide an interpretation of the parameter ψ in the redistribution free equilibrium:

$$-\frac{d \ln s_h}{w_s s_h dh} = -\frac{s_{hh}}{w_s s_h^2} = \psi \tag{3}$$

since $w_s = 1$ and $s_h = \beta$ in the redistribution free equilibrium. The numerator of the left hand side is the relative change in the return to education, its denominator is the relative change in the value of the human capital of the worker, evaluated at market prices (since $w_s s_h dh$ measures the effect a change dh in the years of education on the log wage w). In the redistribution free equilibrium, workers set Mincerian rate of return to a year of education, $w_s s_h$, equal to its price. Hence, ψ^{-1} is the *price elasticity of the supply of human capital*: the percentage change in the value of human capital per percent change in its price.[2]

The parameter ξ captures the complementarity between ability and education. When $\xi = 0$, the marginal return to a year of education (keeping h fixed) does not depend on innate ability. Hence, optimal years of schooling is independent of innate ability. When $\xi > 0$, people with higher innate ability benefit more from schooling and, hence, they take up more years of schooling. We offer a simple interpretation of ξ. Consider equation (2). Workers choose their optimal years of education as to equate the marginal cost and revenue of a year of education. In

---

[2]Usually, the price elasticity of the supply of human capital θ is defined not with respect to the value of human capital, but with respect to the years of education:

$$\theta \equiv -\frac{d \ln h}{d \ln s_h} = -\frac{s_h}{s_{hh} h} = \frac{1}{w_s s_h h} \psi^{-1}$$

However, since the average years of education $\mathrm{E}[h] \cong 10$ and since the Mincerian rate of return $w_s s_h \cong 10\%$ in developed economies, both concepts happen to be numerically the same.

the redistribution free equilibrium:

$$\frac{w_s s_h}{\lambda} = 1 \qquad (4)$$

where λ is the sum of the interest rate and the rate of depreciation of the human. Most studie suggest λ to be around 10%. Let W be the wage level. The wage gain of an additional year of education is $W w_s s_h$. The lifetime revenue of an additional year of education is the net discounted value of this wage increase, using λ as the discount rate. The cost of an additional year is the foregone labor income W. Dividing both sides by W yields equation (4).[3] Workers set h such that $s_h = \lambda / w_s$, which is independent of h in the redistribution free equilibrium, since w_s is independent of h by assumption. Since s_h is independent of h, $\beta h - \xi a$ must a constant, see equation (2), or equivalently, years of education is a function of ability $h(a)$ with $dh(a)/da = \xi/\beta > 0$. By equation (1):

$$s(a, h) = a + \text{ constant}$$
$$\frac{ds[a, h(a)]}{da} = s_a + s_h \frac{dh}{da} = s_a + \frac{\xi}{\beta} s_h = 1$$

In this world, we would never be able to estimate the "true" return to education s_h, even if we had perfect information on the worker's ability. This is the case of a complete *ability bias in the return to education.* Ability and schooling are perfectly collinear in market equilibirum, since $h(a)$ is a function of a. Hence, we have no independent variation of h. Any variation in $s(a, h)$ can be equally well attributed to variation in year of education h as to variation in ability a.

Let us suppose that, from the point of view of the researcher, this is too gloomy a picture of the real world. Workers do not set their years of education exactly as predicted by the model, so that there is some random variation in h independent of a. Some workers have some kind of special preference for education, and choose therefore to take more education. Others simply make small optimization mistakes. This type of randomness in h would resolve the multicollinearity problem and allow us to identify s_a and s_h separately, with $s_a = 1 - \xi$.[4] This suggests

---

[3]Equation (4) ignores some terms that appear to be irrelevant in the full model, like difference in the level of effort when working and when at school. We omitted these here for the sake transparency. See Dur and Teulings (2001) for these issues.

[4]We apply a first order Taylor expansion of h around its equilibrium value $h(a)$. Since $\beta h(a) - \xi a$ is a constant, the second order term $-\frac{1}{2}\psi[\beta h(a) - \xi a]^2$ in equation (1) can be ignored. Rewriting the remaining terms yields:

$$s(a, h) = (1 - \xi) a + \beta h + \text{ constant}$$

a simple trick to estimate ξ. We can approximate s by log wages w. Consider the coefficients on ability in two log wage regressions, one with only ability as regressor, and another with both ability and years of education. The ratio of both regression coefficients should be equal to:

$$\frac{s_a}{\frac{ds[a,h(a)]}{da}} = 1 - \xi$$

We apply this methodology to obtain a rough estimate of ξ in Section 3.3. However, this equation also provides an interpretation of ξ: it is the share of the return to ability that can only be realized by exploiting the complementarity of ability and education. Without education, a fraction ξ of the potential return to ability would be lost.

2.2 Production and relative wages

The wage rates for each skill type s depend on their supply and demand. Labor supply is determined by the distribution of innate ability and the production of human capital. Labor demand is determined by the aggregate production function, such that the wages for each skill type are equal to their marginal productivity in market equilibrium. We simplify the analysis by constructing the production function such that log wages $w(s, \mu)$ are linear in s:

$$w(s, \mu) = w_0(\mu) + \exp(-\gamma\mu)s \qquad (5)$$

where γ is a (weakly) positive parameter. This relation is derived from assignment or matching models in the tradition of Rosen (1974), Sattinger (1975), and Teulings (1995). In the appendix, we offer a formal derivation. The Mincerian rate of return to human capital is equal to the partial derivative of log wages:

$$w_s = \exp(-\gamma\mu)$$

The rate of return is a decreasing function of the mean level of human capital in the economy, μ. This captures the imperfect substitutability between workers with various degrees of human capital. Were workers perfect substitutes, γ would be equal to zero and relative wages would be independent of the supply of human capital in the economy.

This representation of imperfect substitution deviates from the commonly used CES production function with two broad types of labor, high and low skilled (see for example the much cited empirical study by Katz and Murphy (1992)). In such a two type framework, the full burden of imperfect substitution is put a the borderline between both broad types, since within each broad type workers are perfect substitutes. Hence, relative wages can only change between broad types, not

9

within. Our interpretation allows for a continuum of worker types, each endowed with its own level of human capital s and with its own wage rate. In our economy these workers have to be assigned to jobs, which differ by their complexity, see Teulings (2002). The driving force of this model is the Ricardian concept of comparative advantage: highly-skilled workers have a comparative advantage in complex jobs since skills have a greater effect on worker productivity in more complex jobs. In the Walrasian equilibrium, highly-skilled workers will therefore be assigned to more complex jobs, where their skills yield the highest return. A general increase in the level of human capital reduces the return to human capital, since high skilled workers have to do less complex jobs, where their human capital has a lower return. Alternatively, an increase in the mean of the skill distribution μ raises the supply of high skilled workers, thereby reducing wages for the high skilled, and reduces the supply of low skilled workers, thereby raising wages for the low skilled. Hence, the return to human capital, and thus wage dispersion, decrease when the mean level of human capital goes up. The size of this effect depends on the degree of substitutability between skill types. The smaller the degree of substitutability, as measured by γ, the more the return to human capital decreases for a given increase in the stock of human capital.

This model exhibits the Distance Dependent Elasticity of Substitution (DIDES) structure: the larger the "distance" of two types in terms of their level of human capital, the lower the substitutability between worker types. The situation is illustrated in Figure 1. When an individual worker raises her human capital from s to $s+d\mu$ while all other workers keep their human capital constant, her wage goes up by $w_s(\mu) d\mu$, which is a shift along the curve. However, when all workers do the same, there is an additional, general equilibrium effect, shown as a twist of the curve to a flatter position. The mean μ goes up by $d\mu$. Hence, the wage function twists and the return to human capital falls. This "twisting" is due to substitution processes. Since substitution effects sum to zero (for a constant return to scale economy), the workers with an above average level of human capital will lose, while the workers in the lower tier of the labor market gain. Somewhere in the middle, there is a break even point. These indirect effects of human capital acquisition can be interpreted as a distributive externality of schooling decisions. An individual's decision to invest in human capital increases the stock of human capital in the economy, and reduces, therefore, the return on this capital.

FIGURE 1

We can offer an interpretation of parameter γ in the redistribution

free equilibrium, similar to that of ψ:

$$-\frac{d \ln w_s}{w_s d\mu} = -\frac{w_{s\mu}}{w_s^2} = \gamma \qquad (6)$$

since $w_s = 1$ in the redistribution free equilibrium. The parameter γ can therefore be interpreted as the *compression elasticity*: the percentage reduction in the return to human capital, $d \ln w_s$, per percent increase in the value of the stock of human capital, $w_s d\mu$, evaluated at market prices. Alternatively, its inverse γ^{-1} is the *price elasticity of demand for human capital*, the counterpart of ψ^{-1} for the demand side of the market.

3 Some empirical evidence

3.1 The demand-elasticity for human capital: γ^{-1}

The crucial mechanism in our model is that an increase in the mean level of human capital causes its return to fall, due to imperfect substitution between types of labor. A lower return to education leads to a compression of the wage distribution. With perfect substitution between skill types, relative wages, and hence wage dispersion would be independent of the supply of human capital in the economy. Since this mechanism is crucial for our analysis, we review a number of studies documenting this relation.

There is substantial direct evidence for a negative relation between the stock of human capital in the economy and income dispersion. Tilak (1989) provides some early cross country evidence. In addition, there are a number of case studies for various countries. Goldin and Margo (1992), Goldin and Katz (1999), and Goldin (1999) examine the returns to schooling and the dispersion of the wage structure in the US between World War I and II. Educational returns clearly decreased during this period and the wage structure narrowed. Goldin and coauthors relate these developments to the enormous expansion of secondary schooling beginning in the 1910s. Only after 1980, following a period of low inflow into the university system, the return to education and the dispersion of wages started to increase again (see Card and Lemieux, 2000). In most other countries, the education revolution started later. Consequently, the fall in the return to education and the narrowing of the wage structure also lagged behind (see Hartog, Oosterbeek, and Teulings (1993) for the Netherlands 1960-1985; Edin and Holmlund (1995) for Sweden; Kim and Topel (1995) for South Korea in sixties and seventies). All these studies find that income dispersion is negatively related to the supply of human capital.

Katz and Murphy (1992) provide evidence for imperfect substitutability for the post-war period in the United States, using a two types CES function. They estimate the elasticity of substitution between high and low skilled workers from time series data for the US. They estimate the elasticity to be 1.4, supporting the idea of imperfect substitution between worker types. This elasticity drives the negative relation between the return to human capital and its supply in the post war economic history of the United States. This elasticity of substitution between high and low skilled workers translates into a compression elasticity of about 2.[5] Hence, an increase of the average education level by one year at an initial return to education of 10% leads to an increase in the value of the stock of human capital by 10%. That increase causes the return to human capital to fall by $2 \times 10\% = 20\%$, that is, from 10% to 8%.

Figure 2, taken from Teulings and Van Rens (2002), provides some direct evidence on the relation between the return to human capital for some 50 countries as measured directly from individual data on the one hand, and on the mean level of education and on income inequality on the other hand. There is a clear negative relation between the return to education and average years of schooling, suggesting that skill types are indeed imperfect substitutes in production, see Panel A. Panel B relates the return to education to income inequality. There is a strong positive relationship. Taken together, the results suggest that by increasing average years of schooling, income inequality may be reduced. Some simple regressions based on the data presented in Figure 2 reveal that the return to education is about 16% for countries with no education at all, and decreases by about 0.7% for every increase of one year in the average year of schooling.

FIGURE 2

Teulings and Van Rens (2002) analyze the evolution of log GDP per capita and the variance of log wages due to shifts in the stock of human capital in a panel of 100 countries over the period 1960-1990. In a perfect Walrasian world, the effect of the average years of education S on GDP should be equal to the Mincerian rate of return to human capital, since the Mincerian rate of return measures the effect of an additional year of education on wages, and since wages are equal to

---

[5]Teulings (2002) derives a relation between the compression elasticity and the elasticity of substitution between high and low skilled workers:

$$\gamma = \frac{1}{\mathrm{Var}\,[w] \times \eta_{\text{low-high}}} \cong \frac{1}{0.60^2 \times 1.4} \cong 2$$

where $\mathrm{Var}[w]$ is the variance of log wages and $\eta_{\text{low-high}}$ is the substitution elasticity between low and high skilled labor.

(marginal) productivity. If this return is negatively related to the mean level of education in the economy due to imperfect substitution, then the same should for the marginal effect of education on GDP:

$$\log GDP = \beta_0 + \beta_1 S - \frac{1}{2}\beta_2 S^2 \tag{7}$$

$$\frac{d\log GDP}{dS} = \frac{dw}{dh} = \beta_1 - \beta_2 S$$

where S is the average number of years of education h in the economy. Hence, in a growth regression, we expect a positive effect of increases in the mean level of education, and a negative second order effect of increases in the mean level of education.[6] For the variance of log wages, the following relation applies:

$$\mathrm{Var}\,[w] = \left(\frac{dw}{dh}\right)^2 \mathrm{Var}\,[h] = \left(\beta_1 - 2\beta_1\beta_2 S + \beta_2 S^2\right) \mathrm{Var}\,[h] \tag{8}$$

The variance is a negative function of the average years of education in the economy.

Table 1 provides evidence from regressions on panel data for log GDP per capita and for the variance of log wages based on equation (7) and (8). Panel A presents a regression for log GDP on average years of schooling and years of schooling squared. The regression is run in first differences, using a ten year time frame and controlling for the level of log GDP in the previous period. The effect of a year of education on GDP according to these regressions is much higher than the Mincerian rate of human capital as usually measured, about 24%. However, due to the presence of the second order effect, this return is measured at an average level of human capital of zero. The coefficient of the second order term suggests $\beta_2 = 2 \times 0.8\% = 1.6\%$ per year increase in the stock of human capital. The regressions presented in panel B on the variance of log income provide evidence for a negative effect of schooling on income inequality. We present three regression, one in levels, one in first differences (to eliminate fixed country effects), and one where we allow only for a first order effect in S, since the first and the second order effect are highly collinear. Using the latter regression, a one year increase

---

[6]This interpretation of the role of human capital in the evolution of GDP contrasts sharply with the endogenous growth literature, where the relation between schooling and growth is driven by externalities. For instance, in Barro and Sala-i-Martin (1999), a higher level of education makes the labor force more able to deal with technological innovations. This yields a relation between the level of education and *growth,* not the *level* of GDP, see Krueger and Lindahl (2000) for a discussion.

in average years of education reduces the variance of log income by 0.05. The effect is somewhat larger in the equation measured in levels.[7]

TABLE 1

One can use these numbers to calculate the value of the compression elasticity γ implied by the estimates. Starting from a return to human capital of 10 % and using the 1.5 % decline per year of additional human capital, the compression elasticity reads:[8]

$$\gamma = \frac{0.015}{0.10^2} = 1.5$$

Hence, Teulings and Van Rens's estimate of the size of the compression elasticity is broadly consistent with Katz and Murphy's (1992) estimate of the elasticity of substitution between low and high skilled labor. We apply a value of γ of 2 in our subsequent calculations.

3.2 The supply-elasticity of human capital: ψ^{-1}

The second important parameter in our analysis is the price elasticity of demand. This elasticity determines the effectiveness of subsidies in increasing the stock of human capital. The larger is the elasticity of schooling, the larger the effect of subsidies to education on the average level of human capital. Stanley (1999) analyzes the effects of the GI Bill education subsidies for veterans from WW II and the Korean war on their educational attainment. The GI Bill reduced the cost of education by 50 % for Korean veterans and even by 60 % for WW II veterans. However, not all veterans of the Korean war were entitled to this subsidy, depending on a completely arbitrary rule regarding the date of enlistment. This "random" selection provides a natural experiment for testing the effect of financial incentives on educational attainment. The effects of these subsidies are necessarily limited to veterans with higher educational attainment, since they left the military at the age of 23. At that age, a substantial fraction has completed its investment in human capital. This fits the observation: only 40 % of the Korean veterans who were eligible took up any grants at all. Consistent with this observation, the veterans that took up grants descended predominantly from parents with a higher social economic status (SES). Among those eligible, educational attainment has increased on average by 1/3 year. Were the subsidy effective for all education levels, the effect would have

---

[7]We can backout an estimate for β_2 from these regressions. Since $\text{Var}[h] \cong 12.6$ and using $\beta_1 = 0.24$, the coefficient of the first order term suggests $\beta_2 = \frac{0.09}{2 \times 12.6 \times 0.24} = 1.5\%$ per year.

[8]Recall, by (6): $\gamma = -\frac{d \ln w_s}{w_s d\mu}$. Hence, $\gamma = \frac{d\left(\log \frac{dw}{dh}\right)/dS}{\frac{d \log GDP}{dS}} = \frac{\beta_2}{(\beta_1 - \beta_2 S)^2} = \frac{0.015}{0.10^2} = 1.5$

been $\frac{1}{3 \times 40\%} \cong 0.85$ year. Hence, the price-elasticity of the supply education equals: $\frac{0.85}{10 \times 50\%} = 0.17$, or $\psi = 6$. For WW II veterans, the data allow a more refined disaggregation by SES. There, the upper quintile of the SES distribution achieves a gain in educational attainment of even 2 years. Hence, the semi-elasticity is: $\frac{2}{10 \times 60\%} = 0.33$, or $\psi = 3$. One might suppose that for the group of veterans the cost of education has been higher than it would have been in case their educational career had not been interrupted by the war. We use a value for ψ of 4 in the subsequent calculations.

3.3 Complementarity of education and ability: ξ

In section 2.1, we discussed a simple trick to establish ξ: run two log wage regressions, one with only ability as explanatory variable, and another with both ability and years of education. The ratio of the regression coefficients on ability should be equal to $1 - \xi$. Table 2 presents the results of this type of regressions. Ability is measured by test scores. We use only males to avoid all the selectivity issues that arise due to women's labor supply decisions. We have data for three countries, for each at two points in time: the UK (NCD dataset and BCS dataset), the Netherlands (1983 and 1993), and the US (1974 and 1992).[9] For each dataset, we present two regressions: one with ability as the only explanatory variable, the other with both ability and years of schooling as explanatory variables. For the UK datasets, we have two variables for ability, one based on a math test the other on a reading test. Both are included in the regressions. This evidence suggest a value for ξ of about 0.3 to 0.6.

TABLE 2

4 Efficient redistribution policy

4.1 Government's policy instruments

We can now analyze what would be the optimal policy rule for a government that wants to redistribute income, and wants to do so in the most efficient way. We take for granted that this government wants to redistribute progressively, from the rich to the poor, but our analysis applies likewise for a degressive redistribution, from the poor to the rich. Let d be the worker's log disposable income made available by the government. The government is assumed not to provide grants to students still at school. Their net income is zero and they must finance their consumption by borrowing. At first sight, this seems to be an impor-

---

[9]We thank Peter Dolton for the regressions on UK data, Hessel Oosterbeek for those on Dutch data, and Erik Plug for those on US data.

tant limitation to our analysis. However, it is not. Due to the perfect nature of capital markets, the introduction of a grant financed from a reduction of d_h would be offset by a reduction of the take up of credit by individuals during their years at school, leaving their lifetime consumption path, their years of education h, and their level of productive effort unaffected. Hence, the effect of grants for students is equivalent to subsidies for working individuals based on their educational attainment. We incorporate education subsidies in the income tax system only for analytical convenience.

As in Mirrlees' (1971) seminal paper on optimal income taxation, the government can observe neither effort, nor innate ability, nor the skill level that is obtained by taking up education. It can only observe the years of schooling taken by an individual, h, and her gross log labor income y. The latter is equal to the wage rate per unit of effort times effort, in logs $y = w + e$, where e is log effort. The income policy can therefore be contingent on h and y only. We simplify our analysis at this point by considering log linear income policies only:

$$d = d_0 + d_y y + d_h h = d_0 + d_y (w + e) + d_h h$$

The parameter d_0, d_y, and d_h are the policy instruments of the government. In a non-interventionist, redistribution free equilibrium, we have: $d_0 = 0, d_y = 1, d_h = 0$, so that $d = y$. The parameter d_y is Musgrave and Musgrave's coefficient of residual income progression and measures the progressivity of the income tax. The special case $d_y = 1$ yields a proportional tax system. Progressive income taxation implies $d_y < 1$. The log linear specification implies a constant elasticity of net with respect to gross income. This constant elasticity specification implies that the marginal tax rate is increasing for $d_y < 1$, a feature which turns out to be important for the subsequent discussion. The parameter d_h measures the subsidy for taking up an additional year of education relative to the net discounted value of disposable income; $d_h < 0$ implies a tax on education. Analogous to equation (4) for the redistribution free equilibrium, the first order condition for the optimal years of education in this new, redistributive equilibrium reads:[10]

$$\frac{d_y w_s s_h}{\lambda} = 1 - \frac{d_h}{\lambda}$$

Setting the policy parameters equal to the redistribution free case, $d_y = 1, d_h = 0$, yields condition (4). The left hand side measure the return to

---

[10]Like in equation (4), some slight complications that arise in the full model are ignored here.

an additional year of education. A progressive tax, $d_y < 1$, reduces the private return to education. Hence, workers' privately optimal choice is to take up less education. The right hand side measures the cost of education. The education subsidy raises yearly disposable income by a fraction d_h. This benefit is discounted at a value λ. Where the first term on the right hand side measure total social cost of education (the foregone labor income), d_h/λ measures the subsidies to education as a share of total cost, or alternatively, the marginal subsidy rate. A subsidy to education d_h raises the level of educational attainment.

As in Mirrlees (1971), redistributive income taxes distort the choice of effort, as marginal revenue of effort from the point of view of the individual is below that for the society as a whole. Similarly, a subsidy or a tax to education distorts the take up of education. Policy makers face therefore the trade off between efficiency and redistribution. The question of interest is what combination of education subsidies (or: taxes) and income taxation yields the lowest distortion for a given amount of redistribution. The government's choice of feasible income policies, characterized by d_0, d_y, and d_h, is obviously constrained by a budget constraint, which tells that the sum of labor income for all ability types, $\exp(w+e)$, must be equal to the sum of the disposable income for all types, $\exp(y)$. This budget constraint is used to eliminate d_0 from the set of available policy instruments. Intuitively, the policy maker first chooses d_y and d_h to realize his goals with respect to the income distribution, and he then adjusts d_0 to balance the budget. It turns out that, given the convenient, log linear structure of all our equations, the life time utility of worker of type a is also linear in a:

$$u(a; d_y, d_h) = u_0(d_y, d_h) + u_a(d_y, d_h)a \tag{9}$$

where $u(a; d_y, d_h)$ is the lifetime utility of worker type a conditional on the policy variables d_y and d_h. The function $u_0(\cdot)$ is the utility level of a worker with innate ability zero. The function $u_a(\cdot)$ describes how a worker's utility depends on his innate ability. It should be stressed that (9) is not a social welfare function. However, we can think of the parameters $u_0(\cdot)$ and $u_a(\cdot)$ as the goals of the policy maker, where $u_a(\cdot)$ measures income (or better: utility) inequality and $u_0(\cdot)$ measures the efficiency of the economy. A more egalitarian income policy reduces $u_a(\cdot)$ at the cost of a lower $u_0(\cdot)$. An efficient redistribution policy sets d_y and d_h to maximize $u_0(\cdot)$, taking as given the value of $u_a(\cdot)$ (i.e., the distributional preference of the policy maker). In the subsequent sections, we characterize for any feasible level of inequality $u_a(\cdot)$, the mix of policy instruments d_y and d_h that minimizes efficiency losses, i.e. maximizes $u_0(\cdot)$. We refer to these combinations d_y and d_h as constrained Pareto ef-

ficient. The adjective "constrained" refers to information constraints on effort, ability, and skill, which limit the policy options that are available to the government. The subsequent sections will analyze these efficient redistribution policies, first, ignoring the complementarity between ability and education, and, second, allowing for this complementarity.

4.2 Efficient redistribution without complementarity: $\xi = 0$

Without complementarity of ability and education, all individuals choose the same years of schooling and education subsidies have no direct effect on the income distribution. The only way in which education subsidies affect the income distribution is through general equilibrium effects on the labor market. If the compression elasticity is positive (i.e., skill types are imperfect substitutes), education subsidies contribute to redistribution. The optimal level of education subsidies depends on the political demand for redistribution as well as the distortionary effects of the policy instruments.

The following equation describes for any level of distributional preference $u_a(\cdot)$, the constrained Pareto efficient ratio of tax progressivity $1 - d_y$ and education subsidies d_h/λ:

$$\frac{d_h}{\lambda} = \left[1 + \frac{\eta}{(1+\eta)^2}\gamma\right](1 - d_y) \tag{10}$$

where η is the wage elasticity of effort supply. In the absence of a strive for redistribution, $d_y = 1$, optimal education subsidies are equal to zero. Hence, the redistribution free equilibrium, $d_y = 1, d_h = 0$, is constrained Pareto efficient. This mirrors the first theorem of Welfare economics: with perfect markets, investment in human capital is Pareto efficient. If there is no demand for redistribution, the best a policy maker can do is not to intervene in the market mechanism.

When the government wants to redistribute income from rich to poor, both progressive taxes and education subsidies should be used, since in (10), if $d_y < 1$ then $d_h > 0$. Education subsidies are optimal in our model for two reasons, corresponding to the two terms within square brackets. The first term captures the effect that education subsidies offset the disincentive effects of increasing marginal tax rates on schooling. Progressive income taxation implies that the benefits of education (higher future earnings) are taxed at a higher rate than foregone earnings. Therefore, individuals underinvest in human capital, which should be corrected by providing education subsidies, see Bovenberg and Jacobs (2001). The relevance of this effect depends on the functional form of the tax scheme. Our log linear system does indeed imply increasing marginal

rates. However, a linear scheme would not yield this effect, since then marginal rates were constant. Hence, we do not want to stress this effect here. It just shows up due to the convenient log linear specification of income policy.

The second term refers to the general equilibrium effects of education, which are relevant when types of labor are less than perfect substitutes, $\gamma > 0$. Then, a constrained Pareto efficient income policy requires a subsidy to education *above* the subsidy required to offset the distortions of the income tax. By encouraging schooling, wages are compressed, implying smaller pre-tax income inequality. Hence, a given after-tax income distribution can be reached with less progressive income taxes, and hence less distortionary cost of progressive taxation. Just like progressive income taxes, education subsidies entail distortions. The optimal subsidy to education induces individuals to *overinvest* in education. The distortion in the schooling decision due to the education subsidy is traded off against the distortion in the effort decision due to marginal tax rates. The optimal redistribution policy mixes both distortions, in line with the principles of tax smoothing. The higher the compression elasticity, γ, the stronger the compression of relative wages by additional investment in human capital, and hence the higher is the optimal value of education subsidies. For the relevant range of $\eta < 1$, the optimal subsidy is increasing in η. The more elastic the supply of effort η, the higher the distortion caused by marginal tax rates, and hence the higher is the optimal subsidy to education. Note that the price elasticity of the demand for schooling does not show up in this equation. Since the schooling decision is distorted by both progressive taxation and subsidies to education, the elasticity (measuring the size of the welfare loss) does not affect the ratio between income taxes and subsidies to education.

The subsidy to education can be interpreted as a Pigouvian subsidy to offset an externality in individual schooling decisions. When deciding to take up an additional year of education, the individual raises the mean level of human capital in the economy and therefore compresses wage differentials. This generates both positive and negative income effects for other workers. The value weighted sum of these effects is exactly zero (as applies always for substitution effects in a constant returns to scale economy). However, this compression effect is a positive externality from the point of view of the policy maker, who wants to redistribute income from the rich to the poor and who can do so only at an efficiency cost when using other instruments. We refer to this effect as a distributional externality.

19

4.3 Some empirical evidence on efficient redistribution

When we follow Becker (1983) and interpret our model as a positive theory of the policy mix used for redistribution, the model predicts that countries with a stronger preference for redistribution and hence, a stronger progressivity of the tax system, have higher public spending on education. Figure 3, taken from Van Ewijk and Tang (2000), provides some crude evidence on these issues, just to give an idea of the orders of magnitude implied by the model. The horizontal axis plots education subsidies as a percentage of GDP while the vertical axis plots d_y, the percentage change in after-tax income when the before-tax wage rises with 1%.[11] There is a clear relation between the progressivity of the income tax and the level of education subsidies: countries with a more progressive income tax (a lower value of d_y) tend to have higher education subsidies.

FIGURE 3

Remarkably, the actual level of subsidies to education and its relation to the progressivity of income taxes is close to what our model predicts. Clearly, when taxes are proportional $(d_y = 1)$, subsidies to education should be zero. This is consistent with the data in Figure 3. The model allows a crude calculation of the optimal level of subsidies to education as a share of GDP, ignoring the effect of the increasing marginal tax rate (the first term between square brackets in equation (10)). The efficient level of education subsidies for redistributive purposes depends on the values of η and γ. As discussed in the previous section, an empirically plausible value for γ is 2. Similar to Diamond (1998), we assume the supply elasticity of effort η to be equal to a half. The coefficient of residual income progression (d_y) is on average 0.85 in OECD countries (see also Figure 3). Hence, for the average OECD country, imperfect substitution justifies a subsidy to education of approximately 7% of the cost of foregone labor income:

$$\frac{d_h}{\lambda} = \frac{\eta\gamma}{(1+\eta)^2}\left(1 - d_y\right) \cong 0.44 \times (1 - 0.85) \cong 7\%$$

Suppose that the average worker takes up 10 years of education, which is a reasonable value for OECD countries, and suppose that labor accounts for $\frac{2}{3}$ of GDP. Then, subsidies to education as a percentage of GDP

---

[11] Clearly, this is only a crude measure of tax progressivity. For instance, Figure 3 does not take into account differences in tax progressivity between countries as a result of differences in VAT.

should be:
$$\frac{2}{3} \times 10 \times 0.44 \times 0.10 \times (1 - 0.85) = 4.4\%$$

For the mean level of progressivity of income taxes in OECD countries, subsidies to education for the purpose of redistribution should account for approximately 4.4 % of GDP. This is close to the actual value of 5.5%. Our argument for education subsidies thus goes a long way towards explaining the actual pattern and level of education subsidies in OECD countries.

4.4 Allowing for complementarity of ability and education

When we allow for complementarity between education and innate ability, it is no longer clear whether education subsidies contribute to redistribution of income from rich to poor. On the one hand, by stimulating human capital formation, education subsidies reduce wage dispersion because skill types are imperfect substitutes in production. On the other hand, the complementarity between education and ability implies that individuals with high ability go to school longer. Since the amount of education subsidies is increasing in the years of education an individual takes up, education subsidies disproportionately favor the people with high ability. Hence, the complementarity of education and ability may cause education subsidies to increase income dispersion.

The constrained Pareto efficient level of education subsidies allowing for complementarity between ability and education is described by:

$$\frac{d_h}{\lambda} = \left[1 + \frac{\eta}{(1+\eta)^2} \frac{(1-\xi)\gamma - \psi\xi}{1-\xi} \right] (1 - d_y) \tag{11}$$

The first term between square brackets in equation (11) is again the subsidy to education needed to correct for the distortionary effect of progressive taxation on the schooling decisions. Like in the previous section, we ignore this effect in the subsequent discussion. The second term consist of two opposing force, the progressive general equilibrium effect of education subsidies and the degressive direct effect on income due to the higher take up of education by high ability / high income types. The condition for the former effect to dominate the latter reads:

$$\psi\xi < (1-\xi)\gamma$$

The condition has a simple economic interpretation. The parameter ξ is the share of wage dispersion that is attributable to the cost of human capital acquisition, while ψ^{-1} is the price elasticity of the supply

of education. Hence, the left hand side is the adverse direct effect of the subsidy: the increase in inequality due to a subsidization of the cost of human capital acquisition per value unit increase in the average human capital. The right hand side measures the reduction in inequality: $1 - \xi$ is the share of wage dispersion that is directly attributable to ability differentials, while γ^{-1} is the price elasticity of the demand for education. Whether this condition is satisfied is sensitive to the exact empirical values of the relevant parameters. For the values discussed before, both sides of the inequality are just equal, which implies that the direct income effect of education subsidies is as large as the indirect substitution effect. Hence, education subsidies do not contribute to redistribution. Much depends, however, on what one believes about the price elasticity of the demand for education. The higher the elasticity, the more education should be subsidized. The intuition is straightforward: the higher the elasticity, the lower education subsidies need to be for a given compression of wages, the smaller is the direct income effect. Moreover, a clever policy design may mitigate the direct income effects while maintaining the indirect substitution effect on income inequality. Examples of this will be discussed in subsections 5.3 and 5.4.

5 Further implications

5.1 The adequate level of centralization

As discussed before, our argument for subsidizing education rests on an externality in individual schooling decisions. Individuals do not take into account the effect of their schooling on pre-tax wage inequality and, thus, on the distortions arising from progressive income taxation. Decision making must be sufficiently centralized to internalize externalities.

Consider the case of a small district in a large country. Either labor is mobile or there is free trade of products between districts, or both. Hence, by the Heckscher-Ohlin factor price equalization theorem, relative wages are then determined by the nation wide skill distribution, not that in the own district. Evaluated at the decentralized level, education subsidies increase the dispersion of utility when ability and education are complementary. Since the district is too small to have an effect on relative wages in the economy, the only distributive effect stems from the complementarity between ability and education in skill formation. Without complementarity, education subsidies are only used to offset the distortionary effect of increasing marginal tax rates on schooling decisions. With complementarity, progressive taxation is combined with a subsidy to education which is *lower* than the subsidy needed to offset tax distortions. When there is strong complementarity, even a tax on ed-

22

ucation may become constrained Pareto efficient at the decentral level. Clearly, taxing education contributes to redistribution as high-ability types take up more education than low-ability types. The (local) distortionary effect on schooling decisions is traded off against the disincentive effect of the other redistributive instrument, progressive taxation. Since the general equilibrium effect of education subsidies on relative wages is not taken into account at the decentralized level, subsidies are inefficiently low. Hence, decentralization yields underinvestment in human capital.

The case discussed above matches closely the US institutional structure, where decisions on education are made at the level of school districts. The main difference is that the tax policy is decided predominantly at a federal level. This feature of the US system may strengthen our result that decentralized bodies provide too low subsidies to education. The reason is that central decision making on taxes introduces an additional externality in decentralized decision making, discouraging investment in human capital. While in the analysis above the direct consequences of underinvestment in human capital for the government budget are fully taken into account, this is no longer the case if local income is subject to federal taxes. Studying these issue more fully would require the introduction of separate budget constraints for the school district and the federal government.

5.2 Time consistency of the policy

So far, we have studied optimal income policy from the perspective of an individual at the beginning of his life. Moreover, we have assumed that the optimal income policy is set once and for all. In this section, we relax both assumptions to gain insight into the political viability of education subsidies in a world where the decisive voter has already started his working career and cannot commit to future policies.

Consider a dynamic economy where old generations die and new generations enter the labor force. Inhabitants differ along two dimensions. First, they are either at school or working. Second, they differ according to their ability level. For simplicity, we assume that while at school, inhabitants vote as if they are working. In this way we ignore slight differences between the interest of those at school and those working. The main interest is within generations: the low ability people have an interest in past accumulation of human capital (because of general equilibrium effects on relative wages) and today expropriation of the fruits of human capital (for redistribution).

The temptation to expropriate the fruits of past human capital formation conflicts with the desire to stimulate current human capital for-

mation by young generations. In particular, consider the median voter at a particular point in time. He is tempted to ignore the effect of income policy on schooling decisions. Since years of schooling are assumed to be observable, this implies that he can fully expropriate the high ability types who have taken up more education (since innate ability and education are complements). However, in that case, future generations of new entrants will no longer invest in education. This will gradually depress the mean education level among the workforce, thereby raising gross wage differentials, at the expense of the median voter. Since the median voter expects to live beyond today, he is also negatively affected by this long run negative effect on his gross wage rate.

Interestingly, one can prove that when voters cannot commit on their future voting behavior, the political process brings the economy exactly half way between complete internalization of redistributive externalities of schooling decisions and complete decentralization, where externalities are fully ignored, see Dur and Teulings (2001). At that point, the temptation to expropriate past investments in human capital is exactly offset by the fear of adverse general equilibrium effects by lower future investments. The lack of ability to commit to future voting behavior works to the detriment of the lower half of the income distribution, which gets less redistribution than with commitment. It is therefore in their interest to seek ways to commit not to tax investments in human capital in the future. For that reason, it may be important not to allow years of education to be a variable in the tax system. As soon as that variable enters the system, it opens the door for debate on heavier taxation of human capital in the future. The debate alone is enough to undermine the credibility of the incentives for investment in human capital, and thereby their effectiveness. Therefore, it may be much better to frame subsidies to the education system in the form of irreversible grants during the years at school or of direct subsidization of the schools themselves, since this type of subsidies are much more credible.

5.3 Subsidies to schools versus grants for students

The conclusion of section 4.3 that the direct and the indirect general equilibrium effects tend to cancel in the simple log linear set up, does not imply that we should forget about raising the level of education as an efficient redistribution instrument. Education subsidies make most sense if one can find policies that limit the direct effect on income while at the same time maximize the substitution effect on the mean level of human capital. Such more sophisticated policies are observed in practice. For instance, in the present model, the only cost of education is foregone earnings, keeping the quality of the education system fixed. One could

extend the analysis to the trade off between the quality and the direct cost of the education system. Then, a typical policy parameter might be the quality of education in general, and of primary education in particular. Leuven, Oosterbeek, and Van Ophem (2002) show that there are considerable differences in the quality of education across countries. An eyeball test suggests that these differences are related to the amount of government subsidy to the education system. The big advantage of raising the quality of primary education is that it has no adverse direct effect on income and is likely to raise the average skill level in the economy. However, depending on the exact specification of the education production function $s(a, h)$, the greater quality of primary education might be just offset as people reduce their years of schooling such that marginal cost and revenues remain equal. A general increase in the quality of education might be a more attractive alternative, as it opens quality of education as a second dimension for substitution next to years of education. The greater the total elasticity of educational attainment to incentives, the cheaper it is for the government to increase the average level of education by subsidies, and in particular, the less subsidies need to be paid to high ability types who take up a lot of education.

5.4 Subsidies based on parental income

Another option for improving the effect of a given amount of subsidy on the average education level in the economy is to include intergenerational information in the subsidization scheme. The social economic status of the previous generation is a good indicator of the expected educational attainment of the next generation, partly by nature effects, partly by nurture, see Plug and Vijverberg (2002). On average, kids of low education families drop out the education system at a younger age and with a lower skill level. An optimal subsidy to education operates at the margin, to invoke people to stay at school longer. The problem is that the margin is located at a different point for each skill group, so that high skilled workers benefit along the whole range. By using the educational attainment of the previous generation, subsidies can be tailored more precisely to the margin, improving the ratio of beneficial incentive effects versus adverse direct effect on income. In practice, this boils down to subsidies that are conditional on parental income, an institution that is widely applied.

5.5 Direct compression of the human capital distribution?

Many policies are geared towards direct compression of the human capital distribution. The recent Luxembourg and Lisbon summit of the

EU have again focussed the efforts for investment in human capital on raising the level of education of least skilled. From a distribution point of view, this seems to be an obvious idea as it raises the human capital of the most disadvantaged group. However, a second thought reveals a number of complications. First, relative to the free market outcome, it is much simpler to raise everybody's skill level by a bit than to raise a particular group's skill level by a lot, since the dead weight loss increases quadratically with the deviation from the market outcome. Second, the general equilibrium effect of stimulating human capital accumulation in the lowest strata of the distribution work perverse: they raise supply in the lower part of the distribution, thereby reducing relative wages of the least skilled, see Teulings (2002) for a detailed analysis. Table 1 provides some, though far from conclusive empirical evidence on this issue. The regressions in Panel B include the variance of education as explanatory variable (V_t). The effect of this variable on the variance of log wages is close to zero. Hence, the direct effect of a compression of the variance of education on the variance of earnings is fully offset by adverse general equilibrium effects.

The argument is very much comparable to the discussion on minimum wages. An increase in minimum wages reduces labor supply at the bottom of the labor market, thereby increasing the relative wages of their best substitutes, slightly better skilled workers. The argument is illustrated in Figure 4. Suppose that we introduce a minimum wage, that eliminates the left tail of the human capital distribution, reducing the effective supply of low skilled workers. Firms will shift their demand for these low skilled workers to the closest available substitute, slightly better skilled workers, type s^+. Hence, the introduction of the minimum wage will increase the wages of type s^+ workers substantially. Firms that used these type s^+ workers before the introduction of the minimum wage will find their cost having been increased. They will substitute to the closest substitute, type s^{++} workers, s^{++} being slightly higher than s^+. Hence, their wages go up, but by slightly less than the wages of type s^+. This yields the type of pattern shown in Figure 4, with large spillover effects of an increase in the minimum wage to workers earning wages just above the minimum. This type of pattern has been documented for the United States by Lee (1999) and Teulings (2000, 2003). A decrease of the minimum wage by 10 % causes the wages of workers earning slightly more than the minimum to go down by 8 % (Teulings, 2003). Basically, all of the increase in inequality in the lower half of the labor market in the United States during the eighties can be explained by the fall in minimum wages.

FIGURE 4

Similarly, programs like the EITC and New Deal can be victim of their own success. The large subsidies to the employment of the least skilled raise their supply, and thereby invoke adverse general equilibrium effects, which might undo the gains of the initial subsidies. The only net effect would be an increase in the marginal tax rates for the better skilled workers to finance the subsidies to the least skilled.

6   Conclusion

The general equilibrium effect of investment in human capital provide a forceful argument for the subsidization of education for a government that wants to redistribute income. Previous studies on optimal taxation have always downplayed the importance of general equilibrium effects. The reason that these effects show up much more prominently in this study is that we use a more realistic production technology, based on comparative advantage of high skilled workers in complex job types. Contrary to for example a two type CES technology, this production technology implies that the whole wage schedule becomes flatter as a result of an increase in the average stock of human capital. An efficient redistribution policy should therefore combine progressive income taxation and subsidies to the formation of human capital. Crude calculations suggest that this model provides a rationale for subsidies to the education system of about the level that we observe empirically. Moreover, the model suggests positive cross country relation between the progressivity of income taxes and the rate of subsidization of the education system: the more redistributive a country's income policy, the higher will be both the progressivity of the tax system and the subsidy to education system. This relation is also borne out by the data, with a slope that fits the theoretical predictions closely.

However, there is an effect working in the opposite direction. Since the take up of schooling is complementary to innate ability, the direct effect of a subsidy to education tends to favor high ability types. Our overview of some empirical studies suggests that both effects cancel. Much depends on what one believes about the price elasticity of the demand for education. The higher the elasticity, the more education should be subsidized. Moreover, the result that the direct and the indirect more or less cancel does not imply that we should forget about the argument. The simple log-linear income policy analyzed in this paper is applied merely for reasons of tractability. One can think of more elaborate schemes that increase the substitution effects of education subsidies, while at the same time reduce the adverse direct effect on income, in particular policies aimed at raising the quality of education and grants for students which depend on parental income.

The log linearity of the income policy imposes another strong restriction. It implies increasing marginal tax rates (for $d_y < 1$), offending the logic of the Sadka (1976) argument for low marginal rates at both ends of the income distribution. Interestingly, this argument can be extended towards education subsidies, but then reversed. Where in the case of income taxation, the direct effect on income are desired for the purpose of redistribution while the substitution effects only cause efficiency losses, here the substitution effects contribute to the redistribution while the direct effect on income work in the opposite direction. Hence, the marginal rate of education subsidies should be high at the bottom and at the top, where they do not cause substantial direct effects on income since there are no people earning less than the lowest or more than the highest income. The previous argument regarding the quality of primary education exploits this idea at the lower end of the distribution. Where this idea fits the layman's intuition, its counterpart is more surprising. A subsidy for top education programs has little adverse direct effect on income (since there are not many people taking up more years of education), while it raises the average level of education. The production function applied in this paper implies that all lower ability types will benefit from the general equilibrium effects of this policy, see Teulings (2002).

The analysis of the optimal functional form of taxes and education subsidies has strong policy implication for programs like the EITC and the New Deal, along the lines suggested by Heckman, Lochner, and Taber (1999). These programs aim at a reduction of marginal tax rates for the lowest ability types in order to combat low-skilled unemployment. The government budget constraint then dictates that marginal rates should be increased for higher ability types. The logic of the argument in this paper suggests that this policy can be victim of its own success. To the extent that the subsidies induce low ability types to go to work, the relative increase in low skilled labor supply will reduce their wages, thereby partially undoing the initial effect of the subsidy. Stated more crudely: there is limit to the demand for hamburger flippers. At the same time, the increase in marginal rates for somewhat higher skill types, which is necessary to satisfy the government budget constraint, reduces the incentive for investment in human capital, which further aggravates the problem. This points to the need of a more formal analysis of the functional form of the optimal policy.

7 Appendix: Production technology and wages

Consider an economy where the production of one unit of output requires the input of an infinite number of tasks, indexed by their level of

complexity, c. The price of a unit of output is taken as the numeraire and is therefore normalized to unity. Both the skill level s and the complexity level c vary continuously and can take any real number. The transformation of tasks into output takes place by a Leontieff technology: tasks are required in fixed proportions.[12] The skill distribution is normal, $s \sim N(\mu, \sigma^2)$. The size of the workforce is normalized to unity. The input requirements of c-type tasks per unit of output are described by a normal distribution, with the same variance as the skill distribution, $c \sim N(0, \sigma^2)$. A c-type task can be produced by any s-type worker. However, the relative productivities of various worker types differ according to the complexity of the task:

$$g(s, c) = -\frac{1}{\gamma} e^{\gamma(c-s)}$$

where $g(s, c)$ is the log productivity of skill type s in a task of complexity c. This specification implies comparative advantage of high skilled workers in complex jobs, since $g_{sc} > 0$: the productivity ratio of type s_1 and $s_2, s_2 < s_1$, is increasing in c. Teulings (1995) shows that this set up implies that every task type c is uniquely assigned to a single worker type $s(c)$, and vice versa. Furthermore, better skilled workers are assigned to more complex jobs, $s'(c) > 0$, due to comparative advantage. The equilibrium of supply and demand for each task type c requires, in logs:

$$-\left(\frac{s(c) - \mu}{\sigma}\right)^2 + g[s(c), c] + \ln s'(c) = -\left(\frac{c}{\sigma}\right)^2 + Y \qquad (12)$$

where Y is log output per unit of the workforce. The left hand side is the log supply of labor of type $s(c)$ (the log of the normal density function) plus its log productivity in task type c plus the log Jacobian $\frac{ds}{dc} = s'(c)$. The two terms on the right hand side measure the log demand for task type c: the Leontieff coefficient (again the log of the normal density) plus log output. Equation (12) is a differential equation in $s(c)$. The special case, where the variances of the skill distribution and the complexity demand distribution are equal, is the only for which

---

[12]Contrary to suggestion by some commentators, the assumption of a Leontieff technology is not crucial for our results. Teulings (2000) shows that replacing the Leontieff technology by a Cobb Douglas technology is (almost) observationally equivalent to halving the value of γ. As long as there is a slight imperfection in the substitutability between c types in the production of output, the distribution of c does not fully adjust to shifts in the distribution of s, and hence relative wages have to change. The advantage of using the Leontieff technology is that the differential equation (12) can be solved analytically.

this differential equation has an analytical solution:

$$s\left(c\right) = c + \mu \tag{13}$$

$$Y = L - \frac{1}{\gamma}e^{-\gamma\mu}$$

For the derivation of the (unique) wage equation that is consistent with this assignment, we define $x(s,c)$ to be the log production cost of c-type tasks by a s-type worker:

$$x(s,c) = w(s,\mu) - g(s,c) \tag{14}$$

that is, log production cost in log wages per worker minus log productivity per worker. In a market equilibrium, workers are assigned to tasks such that production cost for task c is minimized. The first-order condition reads:

$$x_s\left[s(c),c\right] = w_s\left[s\left(c\right),\mu\right] - g_s\left[s\left(c\right),c\right] = 0 \implies \tag{15}$$

$$w_s\left(\mu\right) = g_s\left[c - \mu, c\right] = e^{-\gamma\mu}$$

The intercept $w_0\left(\mu\right)$ remains to be determined. This is derived from the numeraire. Production cost per task weighted by their Leontieff coefficient add up to unity:

$$1 = \int_{-\infty}^{\infty}\frac{1}{\sigma}\phi\left(\frac{c}{\sigma}\right)e^{x[s(c),c]}dc$$

$$= \int_{-\infty}^{\infty}\frac{1}{\sigma}\phi\left(\frac{c - \sigma^2 e^{-\gamma\mu}}{\sigma}\right)e^{w_0 + \left(\mu+\gamma^{-1}\right)e^{-\gamma\mu}+\frac{1}{2}\sigma^2 e^{-2\gamma\mu}}dc$$

where the second line follows from the substitution of equation (13), (14), and (15), and some rearrangement. Hence:

$$w_0\left(\mu\right) = -\left(\mu + \gamma^{-1}\right)e^{-\gamma\mu} - \mu e^{-\gamma\mu} - \frac{1}{2}e^{-2\gamma\mu}\sigma^2 \tag{16}$$

8 Literature

Acemoglu, Daron, and Joshua Angrist (1999), *How large are Social Returns to Education? Evidence from Compulsory Schooling Laws*, NBER Working Paper MP/ W7444.

Allen, Franklin (1982), Optimal Linear Income Taxation with General Equilibrium Effects on Wages, *Journal of Public Economics*, 17(2), pp. 135-143.

Angrist, Joshua D., and Alan B. Krueger (1991), Does Compulsory School Attendance Affect Schooling and Earnings?, *Quarterly Journal of Economics*, 106(4), pp. 979-1014.

Barro, Robert J., and Jong Wha Lee (1993), International Comparisons of Educational Attainment, *Journal of Monetary Economics*, vol.32 no.3, pp. 363-394.

Barro, Robert J., and Jong Wha Lee (1996), International Measures of Schooling Years and Schooling Quality, *American Economic Review*, vol.86 issue 2 Papers and Proceedings May 1996, pp. 218-223.

Barro, Robert J. and Xavier Sala-i-Martin (1999). *Economic Growth.* Cambridge, MA: The MIT Press.

Becker, Gary S. (1983), A Theory of Competition among Pressure Groups for Political Influence, *Quarterly Journal of Economics*, 98(3), pp. 371-400.

Benabou, Roland (2000), Unequal Societies: Income Distribution and the Social Contract, *American Economic Review*, 90, March 2000, pp. 96-129.

Benabou, Roland (2002), Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?, *Econometrica*, 70(2), March 2002, pp. 481-517.

Bils, Mark, and Peter J. Klenow (2000), Does Schooling Cause Growth?, *American Economic Review*, 90(5), pp. 1160-1183.

Bovenberg, A. Lans, and Bas Jacobs (2001), *Redistribution and Education Subsidies*, Mimeo, University of Amsterdam.

Cameron, Stephen V., and James J. Heckman (1998), Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males, *Journal of Political Economy*, 106(2), pp. 262-333.

Cameron, Stephen V., and James J. Heckman (1999), The Dynamics of Educational Attainment for Blacks, Hispanics, and Whites, NBER Working Paper No. W7249.

Cameron, Stephen V., and Christopher Taber (2000), Borrowing Constraints and the Returns to Schooling, NBER Working Paper No. W7761.

Card, David (1999), The Causal Effect of Education on Earnings. Chapter 30 in the *Handbook of Labor Economics,* Volume 3, O. Ashenfelter and D. Card (eds.). Amsterdam: Elsevier.

Card, David, and Thomas Lemieux (2000), Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis, NBER Working Paper No. W7655.

Deininger, Klaus, and Lyn Squire (1996), A New Data Set Measuring Income Inequality, *World Bank Economic Review*, vol.10 no.3, pp.565-591.

Diamond, Peter A. (1998), Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates, *American*

Economic Review, 88(1), pp. 83-95.

Dur, Robert A.J., and Coen N. Teulings (2001), *Education and Efficient Redistribution*, Tinbergen Institute discussion paper 01-090/3, Amsterdam/Rotterdam

Edin, Per-Anders, and Bertil Holmlund (1995), The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy?, In: Freeman, Richard B., and Lawrence F. Katz (Eds.), *Differences and Changes in Wage Structures*, pp. 307-343, Chicago: University of Chicago Press.

Feldstein, Martin S. (1973), On the Optimal Progressivity of the Income Tax, *Journal of Public Economics*, 2(4), pp. 357-376.

Fernandez, Raquel, and Richard Rogerson (1996), Income Distribution, Communities, and the Quality of Public Education, *Quarterly Journal of Economics*, 111(1), pp. 135-164.

Goldin, Claudia (1999), Egalitarianism and the Returns to Education during the Great Transformation of American Education, *Journal of Political Economy*, 107(6), pp. S65-94.

Goldin, Claudia, and Lawrence F. Katz (1999), The Returns to Skill in the United States across the Twentieth Century, NBER Working Paper No. 7126.

Goldin, Claudia, and Robert A. Margo (1992), The Great Compression: The Wage Structure in the United States at Mid-century, *Quarterly Journal of Economics*, 107(1), pp. 1-34.

Hartog, Joop, Hessel Oosterbeek, and Coen Teulings (1993), Age, Wages and Education in the Netherlands, In: Johnson, Paul, Klaus F. Zimmermann (Eds.), *Labour Markets in an Ageing Europe*, pp.182-211, Cambridge: Cambridge University Press.

Heckman, J.J., L. Lochner, en C. Taber (1999), *Explaining rising wage inequality: explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents*, NBER working paper 6384.

Johnson, George E. (1984), Subsidies for Higher Education, *Journal of Labor Economics*, 2(3), pp. 303-318.

Katz, Lawrence F., and Kevin M. Murphy (1992), Changes in Relative Wages, 1963-1987: Supply and Demand Factors, *Quarterly Journal of Economics*, 107(1), pp. 35-78.

Keane, Michael P., and Kenneth I. Wolpin (2001), The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment, Penn Institute for Economic Research Working Paper 01-018.

Kim, Dae-Il, and Robert H. Topel (1995), Labor Markets and Economic Growth: Lessons from Korea's Industrialization, 1970-1990, In: Freeman, Richard B., and Lawrence F. Katz (Eds.), *Differences and Changes in Wage Structures*, pp. 227-264, Chicago: University of Chicago Press.

Krueger, Alan B., and Mikael Lindahl (2000), Education for Growth: Why and For Whom?, NBER Working Paper No. W7591.

Lee, David S. (1999), "Wage Inequality in the U.S. during the 1980s: Rising Dispersion or Falling Minimum Wage?", Quarterly Journal of Economics, No. 3, Vol. 114: 941-1023.

Leuven, Edwin, Hessel Oosterbeek, and Hans van Ophem (2002), *Explaining international differences in male wage inequality by differences in demand and supply of skill*, Mimeo, Tinbergen Institute**,** Amsterdam/Rotterdam

Lucas, Robert E., Jr. (1988), On the Mechanics of Economic Development, *Journal of Monetary Economics*, 22(1), pp. 3-42.

Milanovic, Branko (2000), The Median-Voter Hypothesis, Income Inequality, and Income Redistribution: An Empirical Test with the Required Data, *European Journal of Political Economy*, 16(3), pp. 367-410.

Mirrlees, James A. (1971), An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies*, 38(114), pp. 175-208.

Perotti, Roberto (1993), Political Equilibrium, Income Distribution, and Growth, *Review of Economic Studies*, 60(4), pp. 755-776.

Plug, Erik, and Wim Vijverberg (2002) Schooling, Family Background, and Adoption: Is it Nature or is it Nurture?, *Journal of Political Economy*, forthcoming.

Rosen, S. (1974) Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, *Journal of Political Economy*, 82(1), pp. 34-55.

Sadka, Efraim (1976), On Income Distribution, Incentive Effects and Optimal Income Taxation, *Review of Economic Studies*, 43(2), pp. 261-267.

Saint-Paul, Gilles, and Thierry Verdier (1993), Education, Democracy and Growth, *Journal of Development Economics*, 42(2), pp. 399-407.

Sattinger, M. (1975) Comparative Advantage and the Distribution of Earnings and Abilities, *Econometrica*, 43(3), pp. 455-468.

Shea, John (2000), Does Parents' Money Matter? *Journal of Public Economics*, 77(2), pp. 155-184.

Stanley, Marcus (1999), College Education and the Mid-century GI Bills, Mimeo, Harvard University.

Stern, Nicolas (1982), Optimum Income Taxation with Errors in Administration, *Journal of Public Economics*, 17(2), pp. 181-211.

Stiglitz, Joseph E. (1982), Self-Selection and Pareto Efficient Taxation, *Journal of Public Economics*, 17(2), pp. 213-240.

Tamura, Robert (1991), Income Convergence in an Endogenous Growth Model, *Journal of Political Economy*, 99(3), pp. 522-540.

Teulings, Coen N. (1995), The Wage Distribution in a Model of the Assignment of Skills to Jobs, *Journal of Political Economy*, 103(2), pp. 280-315.

Teulings, Coen N. (2000), Aggregation Bias in Elasticities of Substitution and the Minimum Wage Paradox, *International Economic Review*, 41(2), pp. 359-398.

Teulings, Coen N. (2003), The Contribution of Minimum Wages to Increasing Wage Inequality, *Economic Journal* (forthcoming).

Teulings, Coen N. (2002), *Comparative Advantage, Relative Wages, and the Accumulation of Human Capital*, Tinbergen Institute discussion paper TI 02-081/3, Amsterdam/Rotterdam.

Teulings, C.N. and T. van Rens (2002), *Education, growth and income inequality*, Tinbergen Institute/Princeton University, Tinbergen Institute discussion paper 02-001/3, Amsterdam/Rotterdam.

Tilak, Jandhyala B.G. (1989), *Education and its Relation to Economic Growth, Poverty and Income Distribution*, World Bank, Washington.

Tinbergen, Jan (1975). *Income Distribution: Analysis and Policies.* Amsterdam: North Holland.

Van Ewijk, C. and P.J.G. Tang (2000), *Efficient progressive taxes and education subsidies*, CPB Research Memorandum 170, The Hague.
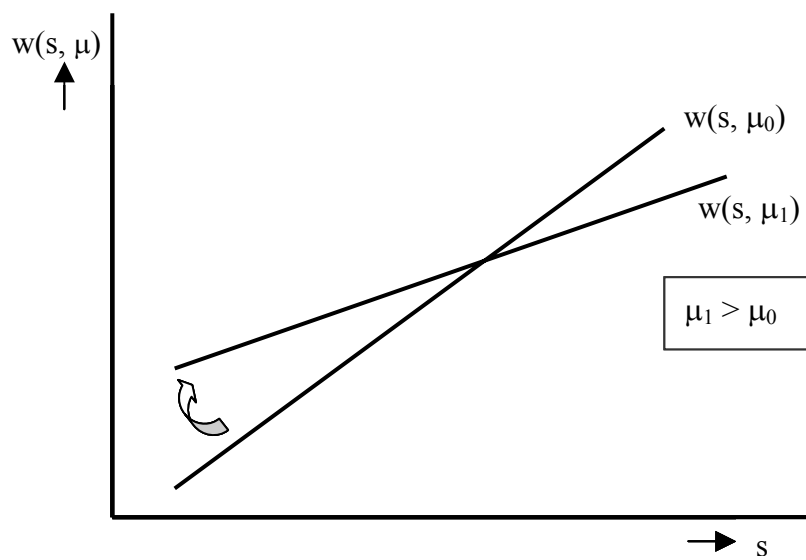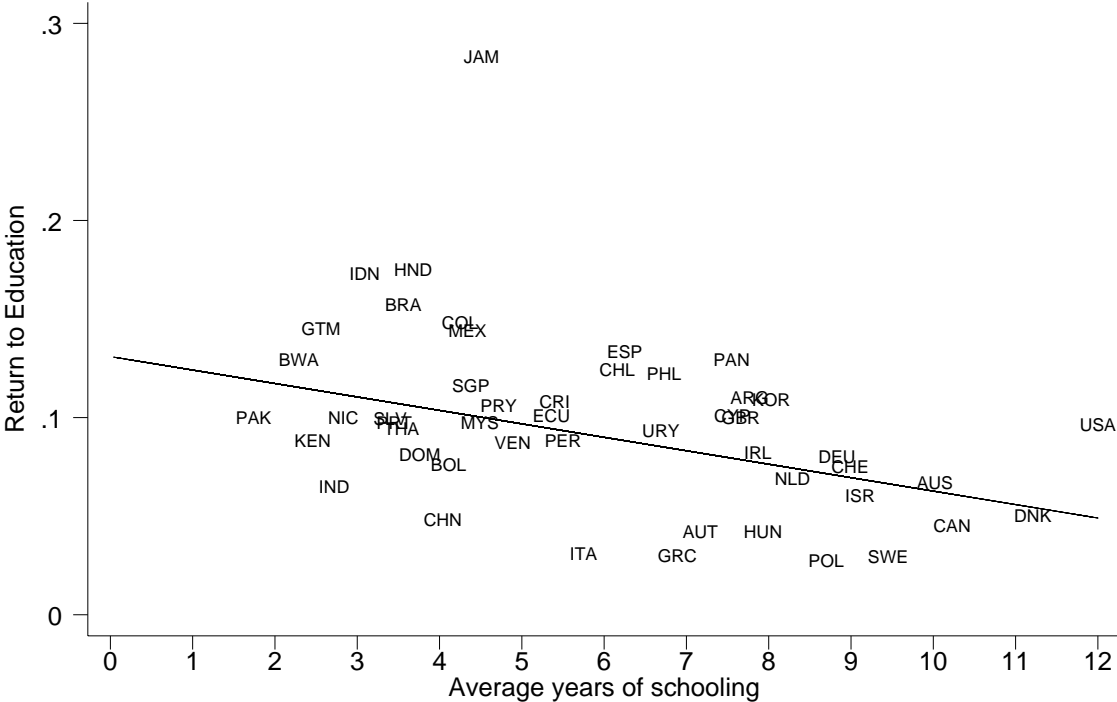
Figure 1. Twisting of the wage function

Figure 2. Return to education, education and inequality

A. Diminishing returns to education
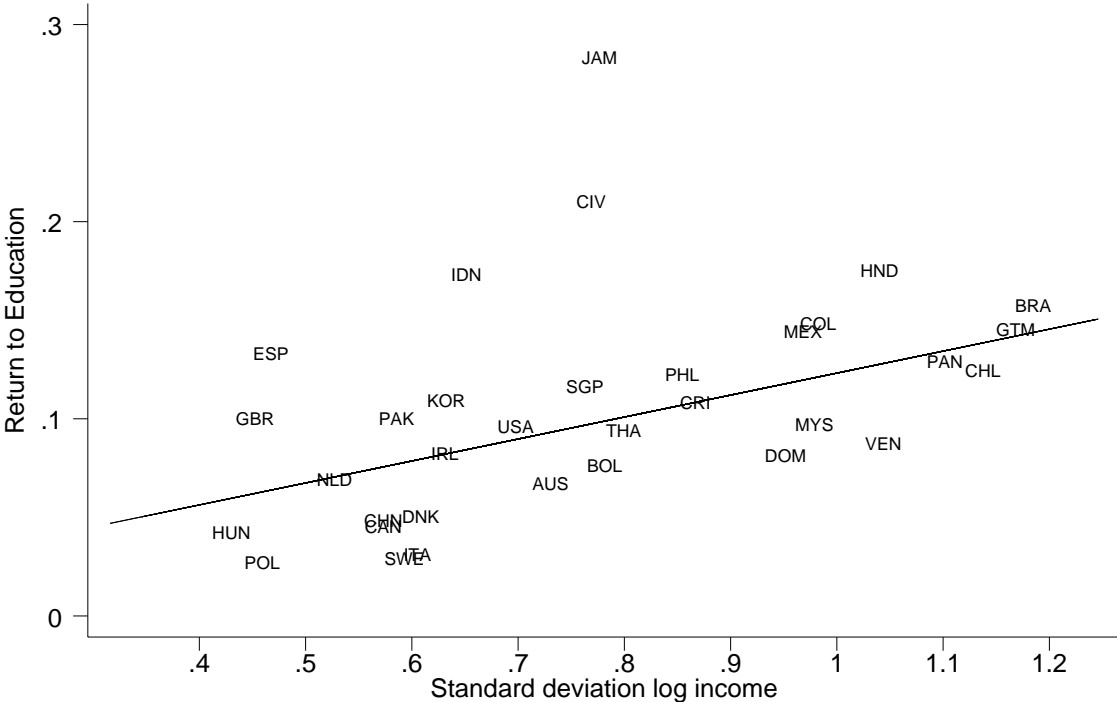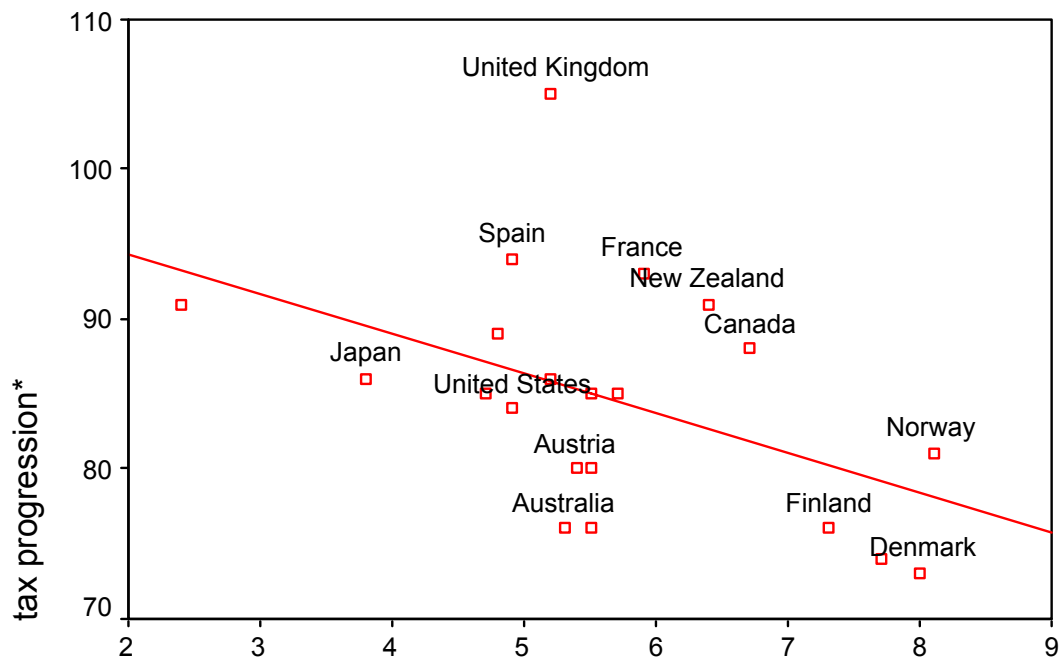

B. Returns to education and inequality

Figure 3: Education subsidies and progressivity of the income tax in OECD countries

tax progression*

education subsidies**

*   change in the after-tax wage, % of change in the before-tax wage.

** public expenditure, % of gross domestic product, in 1994.

source: OECD (1996), *Life long learning for all,* Paris; Table 1.12
             OECD (1997), *Education at a Glance – OECD indicators,*
             Paris
             OECD (1997), *Implementing the jobs study: Member
             countries experience,* Paris (for Belgium on page 91, Table
             28)

This figure has been taken from Van Ewijk & Tang (2000), page 8, right
panel of figure 1.

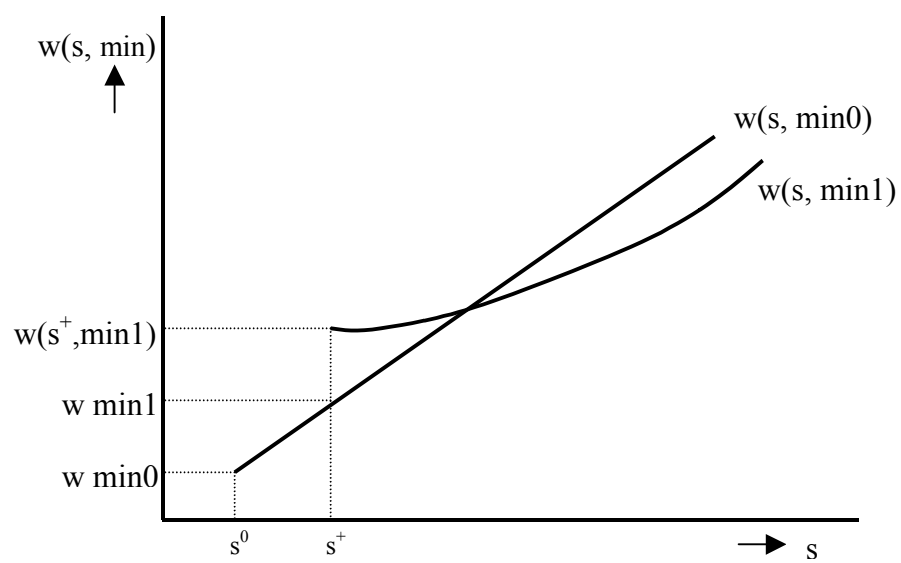**Figure 4: The wage function before and after a minimum wage increase**

Table 1

A. GDP[1]

Variable	equation in first differences
$S(\beta_1)$	0.24335
	(3.84)
$S^2(\frac{1}{2}\beta_2)$	-0.00848
	(2.16)

B. Income inequality[2]

variable	equation in levels	equation in first difs (1)	(2)
$S(\beta_1\beta_2 Var[S])$	-0.08573	-0.09820	-0.05611
	(3.05)	(1.40)	(1.96)
$S^2(\beta_2^2 Var[S])$	0.00170	0.00320	
	(0.78)	(0.66)	
V_t	0.00105	0.00094	-0.00176
	(0.33)	(0.13)	(0.29)

[1] Controls for year effects, year x S, and all variables in levels one year lagged.
[2] Controls for year effects, year x S, and type of income data.

Robust t-statistics in parentheses.
Data sources: GDP: Penn World Table 5.6a; Income inequality: Deininger and Squire (1996);
Education data: Barro and Lee (1993, 1996). See Teulings and Van Rens (2002) for more details.

Table 2: Regressions of wages on individual's ability and years of schooling for males

Regressions	with schooling			without schooling		ξ	
	Ability (math)	Ability (reading)	Schooling	Ability (math)	Ability (reading)	math	reading
UK NCDS	0.0088 (8.76)	0.0079 (4.89)	0.0453 (11.60)	0.0121 (12.47)	0.0110 (6.00)	0.28	0.29
UK BCS	0.0082 (7.90)	0.0038 (3.77)	0.0410 (11.09)	0.0105 (10.15)	0.0052 (5.17)	0.22	0.28
Netherlands 1983	0.0041 (2.92)	-	0.0399 (8.67)	0.0092 (6.96)	-	0.56	-
Netherlands 1993	0.0090 (5.11)	-	0.0444 (8.15)	0.0150 (8.84)	-	0.40	-
US 1974	0.0366 (5.90)	-	0.0496 (13.58)	0.0757 (13.32)	-	0.52	-
US 1992	0.0612 (7.39)	-	0.0920 (18.87)	0.1337 (17.13)	-	054	-

UK: NCDS Data: Ability measured at the age of 11; endogenous variable: log hourly wages at age 33; N=3202; BCS Data: Ability measured at the age of 10; endogenous variable: log hourly wages at age 30; N=2661.

Netherlands: Brabant Data: Ability measured at the age of 12 in 1952; endogenous variable: log wages; N=837 for 1983, N=505 for 1993

US: Wisconsin Longitudinal Survey: Ability measured at the age of 16/7 in 1956; endogenous variable: log family income (selected data include only complete families); N=2742 for both 1974 and 1992

T-statistics in parentheses.