Education, Income Distribution and Public Policy

Robert A.J. 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: 15 February 2002.

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 con…ned to primary education; secondary and higher eduaction are also heavily subsidized. Moreover, governments do not only support schooling opportunities of the disadvantaged. In practice, government programs that encourage education also favour the rich. This chapter is concerned with the question of whether governments should subsidize education so heavily and comprehensively?

The economics literature o¤ers 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 e¤ects in production (see, e.g., 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, recent empirical studies cast doubt on the importance of positive externalities (Acemoglu and Angrist, 1999, Bils and Klenow, 2000, and Krueger and Lindahl, 1999). Second, capital market imperfections may hinder poor individuals to …nance educational expenditures and cost of living while at school (see, e.g., Saint-Paul and Verdier, 1993, Perotti, 1993, and Benabou, 1999). 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 2000, Keane and Wolpin, 1999, Shea, 2000, and Cameron and Taber, 2000).

In this chapter, we discuss a new rationale for education subsidies. 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 and pre-tax wage inequality 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 strong support to this hypothesis. The level of this correlation and the average level of education subsidies, 5 % 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 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 (e.g. 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 within each of the types are ignored.

Our analysis calls for subsidies to all levels of education because an increase in the mean level of education gives rise to general equilibrium effects that reduce wage inequality. This redistribution policy, that relies on indirect, general equilibrium effects, 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, that relies on direct, partial equilibrium effects, might run into trouble due to adverse general equilibrium effects. The empirical evidence supports these ideas. There is a strong negative relation between the first moment of the human capital distribution and the second moment of the wage distribution, but there seems to be no relation between the second moments of both distributions.

However, there is 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 (see for example 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 income and substitution effects in redistribution is reversed. Usually, redistribution is brought about by the income effects of a policy (e.g. progressive income taxation), while the substitution effects (less productive effort) reduce their effectiveness. For education subsidies, it is the other way around. Substitution effects contribute to redistribution, while income effects work in the opposite direction. We derive the precise condition for education subsidies to be redistributive. Furthermore, we discuss policies which may reduce the adverse income effects, while maintaining their e¤ect on the wage distribution.

The potential role of education subsidies in redistribution policy has become even more important in recent decades because of the widening of the pre-tax income distribution in many countries. Increases in pre-tax income inequality puts pressure on politicians to make the tax system more progressive and, hence, more distortionary (see Milanovic, 2000, for recent empirical evidence). We argue that promoting education may help to reduce the political demand for additional redistributive measures. Many studies relate the increase in income inequality to rapid skilled biased technological progress. The increase in average years of schooling has not been su¢cient to prevent income dispersion from rising. In Tinbergen's (1975) terminology, education has lagged behind in the race between technology and education.

The structure of this paper is as follows. Section 2 reviews the empirical evidence on two crucial factors in our model, the degree of substitutability between workers with di¤erent educational attainment and the sensitivity of this educational attainment to …nancial incentives. Recently, Teulings and Van Rens (2002) have analyzed simultaneously the evolution of log GDP per capita, income dispersion and the supply of human capital in a panel of 100 countries over the period 1960-1990. We discuss their results in more detail. Section 3 deals with our theoretical analysis, based on previous work, see Dur and Teulings (2001). First we discuss the case where the complementarity of schooling and innate ability is ignored. Here, we also discuss some empirical evidence regarding the relation between education subsidies and the progressivity of the tax system. This evidence suggest that actual economies are indeed close to Becker's constrained e¢cient redistribution policy. Next, we extend the model to allow for the complementarity between ability and education. We discuss a simple rule of thumb which determines whether or not subsidies to education are part of an optimal redistribution policy. The empirical evidence suggests that this condition is critical. It requires only slight variations in the parameter values for the condition to change sign. Section 4 discusses some further policy implications of our analysis. Four issues are at stake here. First, we discuss the proper level of centralization of education policy. The e¤ectiveness of education as a redistribution policy rests on an externality in schooling decisions. Hence, education policy should be set at a su¢ciently centralized level for all externalities to be internalized. The minimal level of centralization goes up when labour becomes more mobile. Second, we analyze the consequences of a lack of commitment on the side of the government regarding income policy. We show that credibility problems bring politically feasible education subsidies half way between a full internalization of all externalities and a complete ignorance of externalities. Next, we address the adverse effects of education subsidies on the income distribution as a result of the complementarity between education and innate ability. A clever policy design might help to overcome these adverse effects, while at the same time preserving the favourable effect of the average level of education on the income distribution. Some of these designs are observed in practice. Finally, we discuss the role of non-linearities in the income policy. Our theoretical analysis is based on the idea that income policy shifts the mean of the skill distribution to the right, while it leaves higher moments unaffected. This policy works by its indirect, general equilibrium effects. However, actual policies often aim at changing higher moments, in particular by raising the educational attainment of the least skilled. Typical examples are programs like the EITC and the New Deal. These policies are intended to compress the distribution of gross income by compressing the distribution of human capital. However, as we will argue, they are likely to become the victim of their own success. Since these policies raise the supply of human capital in the lower tail of the distribution, the general equilibrium effect of this policy will reduce wages for these workers. Section 5 concludes.

2 Some empirical evidence

2.1 Education and income inequality

A crucial mechanism in our model is that an increase in the mean level of human capital causes its return and, hence, wage inequality to fall. Raising the average years of education in the economy makes low skilled workers more scarce, raising their wages, while high skilled workers become more abundant, reducing their wages. A necessary condition for this substitution effect is that workers with different levels of skills are imperfect substitutes in production. With perfect substitution between skill types, relative wages would be independent of the supply of human capital in the economy. Katz and Murphy (1992) provide some direct evidence for this effect. 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.

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, 1999). 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.

Figure 1, taken from Teulings and Van Rens (2001) 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 strong positive relationship. Taken together, the results suggest that by increasing average years of schooling, income inequality may be reduced. Some simple regressions 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 1

Teulings and Van Rens (2001) analyze simultaneously the evolution of log GDP per capita, the Gini coefficient, and the stock of human capital in a panel of 100 countries over the period 1960-1990. A theory of imperfect substitution between skill types has simultaneous implications for the effect of human capital on income dispersion and on GDP. An increase in the economy's stock of human capital reduces the return to human capital and hence income inequality. With respect to GDP, taking the Mincer equation as a starting point for the contribution of education to GDP, the marginal effect of education should be equal to the return to human capital. Since this return is negatively related to the mean level of education, so should be the marginal effect on GDP. 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. Note that 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.

Table 1 provides evidence from regressions on panel data for log GDP per capita, the Gini coe¢ cient, and the stock of human capital using data from some 100 countries over the period 1960-1995. The model for GDP allows for both imperfect substitution between types of education (by adding a second order e¤ect in education) and skill biased technological progress (by adding cross e¤ects of time and education). Combined with the former feature, the gradual increase in the level of education in most countries leads to a fall in the return to human capital. The latter feature yields a rise in the return to human capital. The average e¤ect of a year of education on GDP according to these regressions is rather high, about 20 %. However, it is highly sensitive to the mean level of education as it falls by even 5 % point for each additional year of education. This evidence is largely consistent with the Mincer equation, in line with the conclusions by Krueger and Lindahl (2001), and in contracst to those of Barro and Sala-i-Martin (1999). The regressions on the variance of log income (which is derived from the Gini's) provide strong evidence for a negative e¤ect of schooling on income inequality: a one year increase in average years of education reduces the variance of log income by 0.05. The evidence strongly supports the notion that the rate of return to education declines in the mean level of education of the workforce. On average over evidence, a one year increase in the mean years of education reduces the return to education by 2%.

The estimations provide information on the pace of technological progress. Keeping constant the average education level, the return to education has gone up strongly over the period. The pace of technological progress increased dramatically during the eighties, raising the return by as much as 6.7% in one decade. To o¤set this increase in the return, the average years of schooling should have increased by as much as $\frac{6:7\%}{5\% \text{ per year}} = 1:3$ year. In fact, the increase in the level of education was more moderate, resulting in a dramatic increase in income dispersion. In terms of Tinbergen's (1975) famous race between technology and education, technology has been in the winning mood during the last decades. One can therefore conclude, with some exaggeration, that Tinbergen's race (1975) between education and technology and Mincer's earnings function rule the world.

One can use these numbers to calculate a simple summary statistic, to which we refer as the compression elasticity γ: the percentage decrease in the return to human capital per percentage increase in the value of the stock of human capital. Then, starting from a return to human capital of 10 % and using the 2 % decline per year of additional human capital, the compression elasticity reads:

$$\gamma \equiv \frac{\text{\% fall return}}{\text{\% increase stock}} = \frac{\frac{\text{2\% point}}{\text{10\%}} \text{ per year}}{\text{10\% per year}} = 2$$

So, when the mean years of schooling increases from 10 to 11 years, and the initial return to education is 10%, the return drops to 8%. If $\gamma = 0$, then the return to human capital is independent of its supply, so that we are back in world of perfect substitution between types of labor. Teulings (2001) derives a relation between this compression elasticity and Katz and Murphy's elasticity of substitution between high and low skilled workers:

$$\gamma = \frac{1}{Var[w] \times \sigma_{\text{low-high}}} \approx \frac{1}{0.60^2 \times 1.4} \approx 2$$

where Var[w] is the variance of wages and $\sigma_{\text{low-high}}$ is the substitution elasticity between low and high skilled labour, and we take the standard deviation of wages to be 0.60. Hence, Teulings and Van Rens's estimate of the size of the compression effect is broadly consistent with Katz and Murphy's (1992) estimate of the elasticity of substitution between low and high skilled labour. However, our interpretation of the evidence is different from the usual interpretation in terms of a CES production function with two types of labor. In this two type frame work, only the relative wage of high and low skilled workers can change. 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 (2001). 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. The basic mechanism at work is that a worker with a particular skill level will end up in a less complex job when highly-skilled workers are abundant than when they are scarce, since the distribution of job complexity is fixed (a Leontief technology). 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, 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

We discuss the implications of this model for relative wages in greater detail. The higher the human capital of a worker, s, the higher her wage rate, w. Hence, there are relative wages for infinite number of worker types (not just two), which are all affected by the mean level of human capital. A change in the mean level of human capital μ reduces the return to that capital:

$$w(s;\mu) = w_0(\mu) + w_s(\mu)\,s \qquad (1)$$
$$w_s(\mu) = e^{-\gamma\mu}$$

where $w(s;\mu)$ denotes the log wage of a worker with human capital s when the mean level of human capital in the economy is μ. Hence, $w_s(\mu)$ denotes the return to human capital. The second equation shows how this returns declines when μ goes up. As is easily checked, the compression elasticity γ is equal to the complexity dispersion parameter for $\mu = 0$.

The situation is illustrated in Figure 2. When an individual worker raises her human capital from s to $s + h$ while all other workers keep their human capital constant, her wage goes up by $w_s(\mu)\,h$, which is a shift along the curve. However, when all workers do the same, there is an additional, indirect effect, shown as a shift of the curve as it twists to a flatter position. The mean μ goes up by h. 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 on its return play a crucial role in our analysis. They 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. It is this externality that provides a rationale for subsidies to education.

FIGURE 2

The "twisting" of the wage function, equation (1), implies that the greater the "difference" in human capital between two worker types, $|s_1 - s_2|$, the larger is the change in relative wages due to an increase in the mean level of human capital $d\mu$: $\left[e^{-\gamma\mu} - e^{-\gamma(\mu+d\mu)}\right]|s_1 - s_2|$. This feature is due to the DIDES structure. Apart from this "twisting" of the wage function in response to a general increase in the level of human capital, there is another implication of the DIDES model that can be tested empirically. 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 3, 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 (1998, 2000). Teulings (1998,2000) shows that a decrease of the minimum wage by 10 % causes the wages of workers earning slightly more than the minimum to go down by 8 %. Apart from the fact that this evidence offers strong support for the DIDES type of production structure that we use, this mechanism will also play an important role in Section 4, where we discuss the implications for policies like the EITC and the New Deal.

FIGURE 3

2.2 The sensitivity of education to financial incentives

Another important parameter in our analysis is the elasticity of demand for schooling with respect to its cost. This elasticity determines the effectiveness of subsidies in increasing the stock of human capital. The sensitivity of education to financial incentives is particularly important when innate ability and education are complementary. Then, education subsidies favour the rich because they take up most education, which reduces the overall effect of stimulating education on income inequality. The larger is the elasticity of schooling, the lower are the subsidies that are required for a given increase in the mean level of human capital, and hence, to compress the wage distribution.

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 entittled 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). The 40 % that used the subsidies have increased their educational attainment on average by 1/3 year. Were the subsidy effective for all education levels, the effect would have been $\frac{1}{3 \times 40\%} \cong 0.85$ year. For WW II veterans, the data allow a more refined desaggregation by SES. There, the upper quintile of the SES distribution achieves a gain in educational attainment of even 2 years. Hence, a reasonable estimate for the elasticity of demand for schooling with respect to its cost is 0.25. One might suppose that for this 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. Hence, we view these estimates as a lower bound.

3 Education and efficient redistribution policy

3.1 The basic model

The empirical evidence discussed in the previous section suggests that by promoting education, income inequality may be reduced. Education subsidies may therefore serve as a redistributive instrument, complementary to progressive income taxes. To explore this idea, we discuss the results of a theoretical analysis, based on previous work, see Dur and Teulings (2001).

Consider a society consisting of individuals who are born with different levels of innate ability. They spend the first years of their life at school. These two factors, innate ability and schooling, jointly determine the skill level with which the worker enters the labor market. For the sake simplicity, we ignore the direct cost of education and focus completely on the cost of foregone labor income. This fits the observation that the direct cost of education are, relatively, of minor importance. Individuals choose their years of schooling as to maximize their life-time utility. The optimal years of schooling may vary between ability types. When education and innate ability are complementary, people with higher abilities go to school longer because they benefit more from education. The large literature on the ability bias in the return to education shows that education and innate ability are indeed complementary (see for example Angrist and Krueger, 1991). This is important for the redistributive effect of education subsidies: the more ability and education are complementary, the more the high ability types benefit from education subsidies.

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. Hence, individuals invest in human capital up to the point where the marginal cost equal the market rate of return to human capital. We also abstract from production externalities in schooling decisions, like knowledge spillovers.

After the investment in human capital, individuals start their working career. For the transparency of the analysis, we abstract from imperfections on the labor market. Workers therefore earn their marginal product of labor and there is no unemployment. The individual's log wage rate $w(s;\mu)$ per unit of effort depends on the individual's human capital s and the mean level of human capital in the economy μ, as discussed in Section 2. Gross income is the product of this wage rate and the effort the individual chooses to provide.

A worker's log disposable income d depends on her wage rate, her choice of effort, and the government's income policy. 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 indivual, h, and her gross log income y. The income policy can therefore be contingent on these two factors only. We simplify our analysis at this point by considering log linear income policies only:

$$d = d_0 + d_y y + d_h h$$

In a non-interventionist, redistribution free equilibrium, we have: $d_0 = 0; d_y = 1; d_h = 0$, so that $d = y$. d_y is Musgrave and Musgrave's coefficient of residual income progression and measures the progressivity of the income tax. 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, 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 that we actually tax education. Let λ be the cost of a year of education relative to the net discounted value of disposable income. It has two components, the real interest rate and the rate of depreciation (individuals die sooner or later); λ will typically be of the order of magnitude of 10 %. Hence, d_h/λ measures the subsidies to education as a share of total cost, or alternatively, the marginal subsidy rate.

Individuals set effort as to maximize their utility. Hence, as in Mirrlees (1971), redistributive income taxes distort productive 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 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. We refer to these combinations as constrained Pareto efficient. The adjective "constrained" refers to information constraints on effort, ability, and skill, which limit the policy options that are available to the government.

3.2 Efficient redistribution policy without complementarity of ability and education

We first discuss the results for the case where innate ability and education are not complementary. In that case, 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 the efficient level of education subsidies for the case where ability and education are not complementary:

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

where η is the wage elasticity of effort supply and where γ is the compression elasticity. Equation (2) describes for any level of income tax progression the constrained Pareto efficient level of education subsidies as a percentage of the cost of education.

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: investment in human capital in a market economy 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 (2), 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 ´. The more elastic the supply of e¤ort ´, 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 a¤ect the ratio between income taxes and subsidies to education.

The subsidy to education can be interpreted as a Pigouvian subsidy to o¤set 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 di¤erentials. This generates both positive and negative income e¤ects for other workers. The value weighted sum of these e¤ects is exactly zero (as applies always for substitution e¤ects in constant returns to scale economy). However, this compression e¤ect 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 e¢ciency cost when using other instruments. We refer to this e¤ect as a distributional externality.

3.3 Some empirical evidence on e¢cient 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 4, taken from Van Ewijk and Tang (2000), provides some evidence. There is a clear negative relation between the progressivity of the income tax and the level of education subsidies.

FIGURE 4

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 4. The model allows a crude calculation of the optimal level of subsidies to education as a share of GDP. The e¢cient level of education subsidies for redistributive purposes depends on the values of ´ and ° (see equation (2)). As discussed in the previous section, an empirically plausible value for ° is 2. Similar to Diamond (1998), we assume the supply elasticity of e¤ort ´ to be equal to a half. The coe¢cient or residual income progression (d_y) is on average 0:85 in OECD countries (see also Figure 4). Hence, for the average OECD country, imperfect substitution justi…es a subsidy to education of approximately 7 % of the cost of foregone labor income:

$$\frac{d_h}{\lambda} \approx 0.44\,(1 - 0.85) \approx 7\%$$

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 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.

3.4 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 disproportionally favor the people with high ability. Hence, the complementarity of education and ability may cause education subsidies to increase income dispersion.

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

$$\frac{d_h}{\lambda} = \left[1 + \frac{\eta}{(1+\eta)} \frac{(1-\xi)\gamma - \psi\xi}{1-\xi}\right] (1 - d_y) \tag{3}$$

where ψ is the inverse of the elasticity of educational attainment with respect to the cost of education and where the parameter ξ measures the degree of complementarity of innate ability and years of schooling in the production of human capital. The higher ξ, the greater the difference in the take up of education between high and low ability types. Hence, the higher this parameter, the greater the direct, regressive effect of subsidies to education. There is an alternative interpretation of ξ: ξ^2 is the share of education, measured at its "true" (non ability biased) return, in the total variance of wages. Teulings and Hartog (1998) report values of about 0.30 for the United States. Angrist and Krueger (1991) find both the 'true return' λ and the ability bias to be 10 %. Our model implies that educational attainment and ability are perfectly correlated for $\xi > 0$, and that there are no other sources of variation in wages. Then, Angrist and Krueger's results suggest $\xi \approx 0.50$. However, ability is not perfectly correlated with educational attainment. There are other sources of wage variation, apart from the 'true return' and ability bias. Hence, a value of a half is an upper bound for ξ. We use: $\xi \approx 0.33$.

The first term between square brackets in equation (3) 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 implies that if $\psi\xi < (1 - \xi)\gamma$, then education subsidies are an efficient redistributive instrument next to progressive income taxation. 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 is the inverse of the elasticity of educational attainment with respect to the cost 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 aquisition 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 is the compression elasticity, measuring the relative decrease in the return to these ability differentials per value unit increase in human capital.

Whether the condition $\psi\xi < (1 - \xi)\gamma$ is satisfied is extremely 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. Hence, education subsidies make most sense if one can find policies that limit the direct income effect while at the same time maximizing the effect on the mean level of human capital. Such more sohpisticated policies are observed in practice. For instance, in the present model, the only cost of education are 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 primary education. Raising the quality has no adverse income effects and is likely to raise the average skill level in the economy. A further option is to include intergenerational information in the subsidization scheme. The social economic status of the previous generation can be used as an indicator for the expected educational attainment of the next generation. These policies shift the balance between adverse income effects and the intended substitution effects in favor of the latter. This boils down to subsidies that are conditional on parental income, an institution that is widely applied in practice. Before discussing these issues in greater detail, see section 4.3, we shall consider what aggregation level is appropriate for the internalization of the distributional externalities of schooling decisions.

4 Further implications

4.1 The adequate level of centralization

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 education 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 inef-

...ciently 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. (Coen: desalniettemin, education subsidies in de VS zijn hoog. Hoe gaan we hiermee om?) The main di¤erence 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 decentral 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.

4.2 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 di¤er along two dimensions. First, they are either at school or working. Second, they di¤er 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 di¤erences between the interest of those at school and those working. The main interest is within generations: the low income people have an interest in past accumulation of human capital (because of general equilibrium e¤ects on relative wages) and today expropriation of the fruits of human capital (for redistribution).

Suppose that an income policy at a particular point in time is bound to be uniform across generations. Then, the temptation to expropriate the fruits of past human capital formation con‡icts with the desire to stimulate current human capital formation by young generations. In particular, consider the median voter at a particular point in time. He is tempted to ignore the e¤ect 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 policy making without commitment brings an economy exactly half way between complete internalization of redistributional externalities of schooling decisions and complete decentralization, where externalities are fully ignored. See Dur and Teulings (2001).

Policies to counteract: do not allow years of education to be a variable in the tax system, and frame subsidies to the education system of the form irreversible grants during the years at school.

4.3 Policies to reduce adverse income effects

Coen zie ook einde Section 3.3

Even more important, the result that the direct and the indirect effect more or less cancel does not imply that we should forget about the argument. For reasons of tractability of the model, we have restricted income policy to be a simple log-linear function. In practice, we observe much more sophisticated schemes that increase the substitution effects of education subsidies, while at the same time reduce the adverse income effects. For instance, in the present model, the only cost of education are 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 primary education. Raising the quality has no adverse income effects and is likely to raise the average skill level in the economy. A further option is to include intergenerational information in the subsidization scheme. The social economic status of the previous generation can be used as an indicator for the expected educational attainment of the next generation. This might shift the balance between adverse income effects of education subsidies and the desired substitution effects in favor of the latter. In the practice of income policy, this boils down to subsidies that are conditional on parental income, an institution that is widely applied in practice. These issues are further discussed in section 4.3.

Investment in quality of education (for everybody, less sensitive to shirking at the university, presenting evidence by Leuven Oosterbeek en Ophem)

Grants related to parental income distribution (maxizing the marginal effect at the lowest incentive effect, the reverse of the standard optimal taxation problem)

4.4 Direct compression of the human capital distribution?

Many policies geared towards direct compression of the human capital distribution.

Luxembourg and Lisbon summit: strong focus on raising the level of education of least skilled

Direct evidence: variance of education does not affect dispersion (Teulings and Van Rens)

Theoretical analysis, related to minimum wage argument (Teulings 2001)

Counterproductive

EITC and New Deal

Heckman Lochner and Taber

5 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 born 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 favour high ability types. Our overview of some empirical studies suggest 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. However, 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 reducing the adverse income effects. We offer a number of suggestions.

First, only foregone earnings are analyzed in this paper, 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 primary education: raising the quality affects everybody, but does not have the adverse income effects. This policy has no adverse income effects, but is likely to raises the average skill level in the economy.

A further option is to include intergenerational information in the subsidization scheme. The social economic status of the previous generation can be used as an indicator for the expected educational attainment of the next generation. This might shift the balance between adverse income effects of education subsidies and the desired substitution effects in favor of the latter. In the practice of income policy, this boils down to subsidies that are conditional on parental income, an institution that is widely applied in practice.

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 rate 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 income effects 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 income effects 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 income effects 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 income effects (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 bene…t from the general equilibrium e¤ects of this policy, see Teulings (2001).

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 will 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 e¤ect of the subsidy. Stated more crudely: there is limit to the demand for hamburger ‡ippers. If we use tax policy to increase their supply, sooner or later their gross wages will fall. 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 aggrevates the problem. This points to the need of a more formal analysis of the functional form of the optimal policy.

6 Literature

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

Aghion, Philippe, and Peter Howitt (1998), Endogenous Growth Theory, Cambridge, MA: MIT Press.

Alesina, Alberto, and Dani Rodrik (1994), Distributive Politics and Economic Growth, Quarterly Journal of Economics, 109(2), pp. 465-490.

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

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

Atkinson A.B. and A. Brandolini (1999) Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries. mimeo, available at: http://www.nu¤.ox.ac.uk/economics/people/atkinson.htm

Autor, D.H., L.F. Katz, and A.B. Krueger (1998), "Computing inequality: have computers changed the labor market?", Quarterly Journal of Economics, p. 1169-1213.

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

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.

Bénabou, Roland (1996). Inequality and Growth. In: NBER Macroeconomics Annual 1996, Ben S. Bernanke and Julio Rotemberg (eds.). Cambridge: MIT Press. Also available as NBER Working Paper No.5658 July 1996.

Benabou, Roland (1999), Tax and Education Policy in a Heterogeneous Agent Economy: What Levels of Redistribution Maximize Growth and E¢ciency?, Econometrica, forthcoming.

MOETEN WE NIET NAAR ZIJN AER PAPER VERWIJZEN?

Benhabib, Jess and M. Spiegel (1994). The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. Journal of Monetary Economics, vol.34 no.2, pp.143-174.

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

Blundell, Richard and Stephen Bond (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics, vol. 87, pp.115-143.

Boadway, Robin, Nicolas Marceau, and Maurice Marchand (1996), Investment in Education and the Time Inconsistency of Redistributive Tax Policy, Economica, 63, pp. 171-189.

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 E¤ect of Education on Earnings. Chapter 30 in the Handbook of Labor Economics, Volume 3, O. Ashenfelter and D. Card (eds.). Amsterdam: Elsevier.Becker, Gary S. (1983), A Theory of Competition among Pressure Groups for Political In‡uence, Quarterly Journal of Economics, 98(3), pp. 371-400.

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.

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.), Di¤erences 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.

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.

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.

Heckman, James J. and Peter J. Klenow (1997) Human Capital Policy. mimeo, University of Chicago. Available at: http://www.klenow.com/.

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 E¤ect 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.), Di¤erences and Changes in Wage Structures, pp. 227-264, Chicago: University of Chicago Press.

Konrad, Kai A. (2001), Privacy and Time-Consistent Optimal Labor Income Taxation, Journal of Public Economics, 79(3), pp. 503-519.

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

Lee, D.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 oosterbeek ophem

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

Meltzer, Allan H., and Scott F. Richard (1981), A Rational Theory of the Size of Government, Journal of Political Economy, 89(5), pp. 914-927.

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.

O'Neill, Donal (1995), Education and Income Growth: Implications for Cross-country Inequality, 103(6), pp. 1289-1301.

OECD (1987), Employment Outlook, Paris.

Pencavel, John (1986), Labor Supply of Men: a Survey, in: O. Ashenfelter and R. Layard, Handbook of Labor Economics, Amsterdam, North Holland, 1986, pp. 3-102.

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

Persson, Torsten, and Guido Tabellini (1994), Is Inequality Harmful for Growth?, American Economic Review, 84(3), pp. 600-621.

Psacharopoulos, G. (1994). Returns to Investment in Education: A Global Update. World Development, vol.22 no.9, pp.1325-1343.

Romer, Thomas (1975), Individual Welfare, Majority Voting, and the Properties of a Linear Income Tax, Journal of Public Economics, 4(2), pp. 163-185.

Roberts, Kevin W.S. (1977), Voting over Income Tax Schedules, Journal of Public Economics, 8(3), pp. 329-340.

Sadka, Efraim (1976), On Income Distribution, Incentive E¤ects 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.

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 E¢cient 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, C.N. (2001), Comparative Advantage, Relative Wages, and the Accumulation of Human Capital, discussion paper TI 99-049/3, Tinbergen Institute.

Teulings, Coen N., and Joop Hartog (1998), Corporatism or Competition? Labour Contracts, Institutions and Wage Structures in International Comparison, Cambridge: Cambridge University Press.

Teulings, C.N. and T. van Rens (2001), Education, growth and income inequality, Tinbergen Institute/Princeton University.

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.

Trostel, Philip A. (1993), The E¤ect of Taxation on Human Capital, Journal of Political Economy 101(2), 327-350.

Trostel, Philip A. (1996), Should Education Be Subsidized?, Public Finance Quarterly, 24(1), pp. 3-24.

Ulph, D. (1982) untitled paper presented at the Econometric Society meeting 1982.

Van Ewijk, C. and P.J.G. Tang (2000), E¢cient progressive taxes and education subsidies, CPB Research Memorandum 170, The Hague.

FIGURE 1 Human capital and its return, and income inequality
(Teulings and Van Rens)

FIGURE 2 Spill over e¤ects of minimum wages
(Teulings 2000)

FIGURE 3 Human capital acquistion and the "twisting" of the wage function
(our …gure)

FIGURE 4 Progressivity of taxes and subsidies to education
(Van Ewijk and Tang)

TABLE 1Years of education and the return to human capital in various countries
(Teulings and Van Rens, Table 2)

TABLE 2 Human capital, GDP, and inequality
(key regression results from Teulings and Van Rens, Table 1,3,5,6)

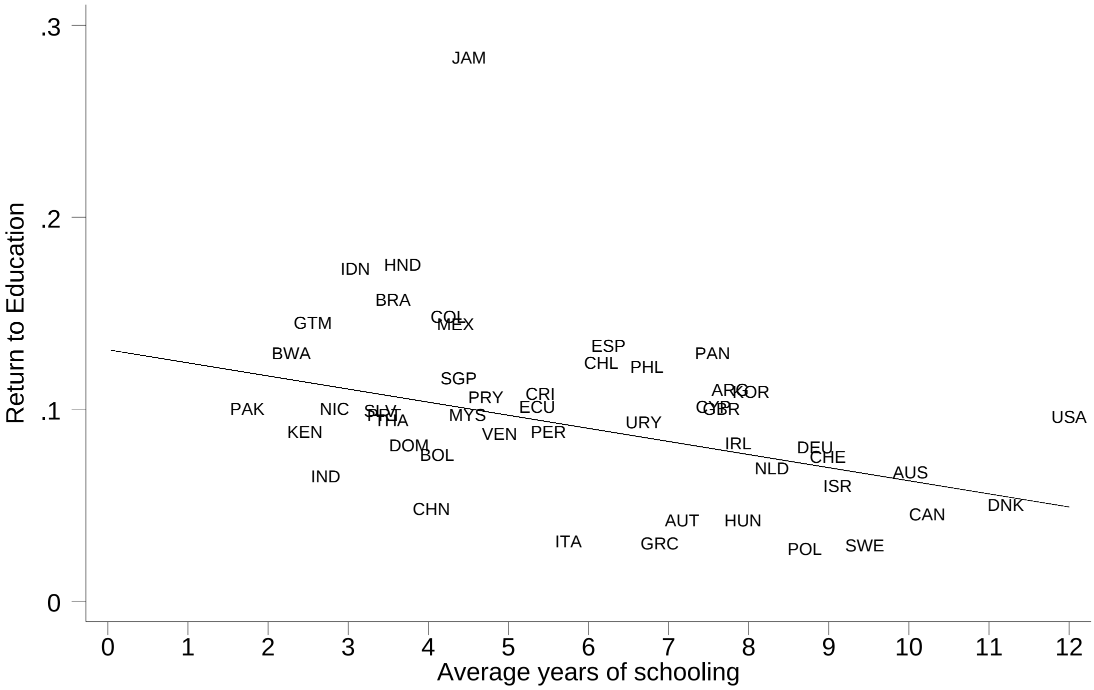
TABLE 3 Inequality, progressivity of taxes, and education subsidies for various countries
(Van Ewijk and Tang, inequality from Teulings and Van Rens)

TABLE 4 Quality weighted years of education and inequality
(Leuven, Oosterbeek and Van Ophem)

Can we obtain information on the progressivity of educational grants?

Figure 1. Return to education, education and inequality

A. Diminishing returns to education



B. Returns to education and inequality

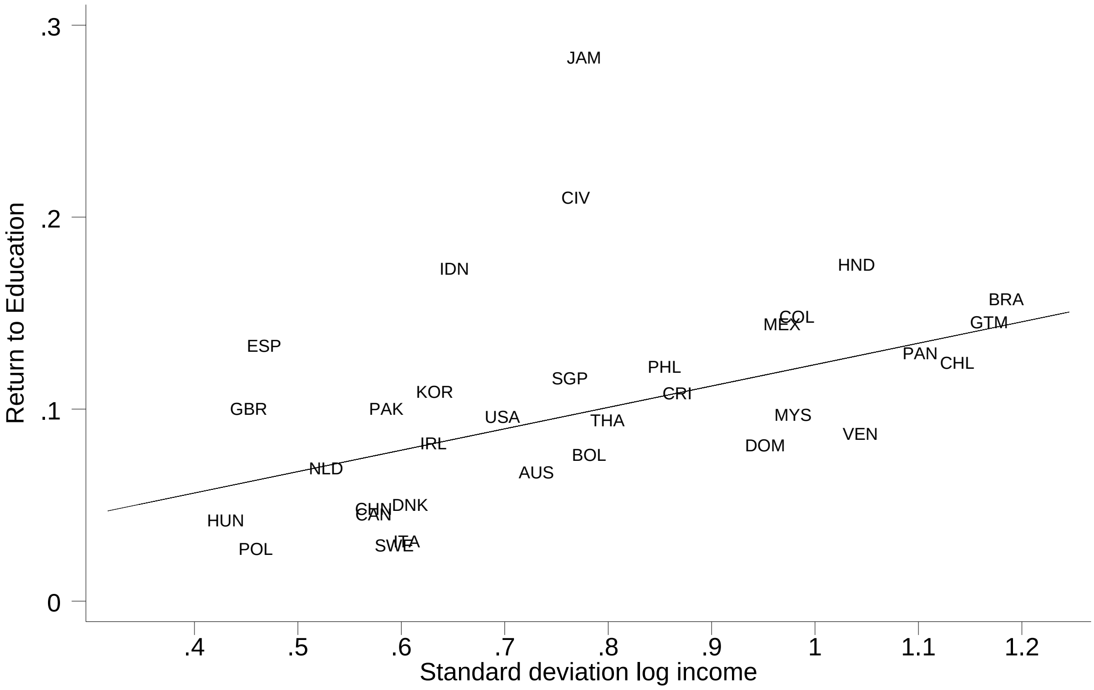


Figure 2. Twisting of the wage function

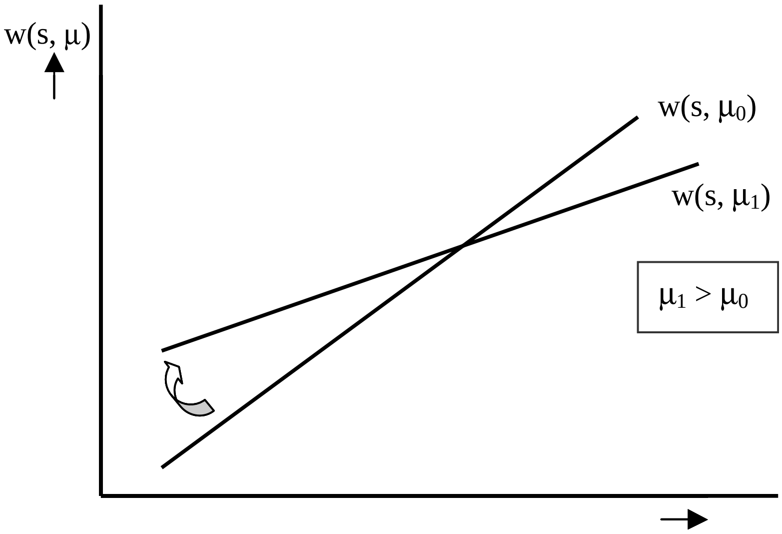


Figure 3: The wage function before and after a minimum wage increase

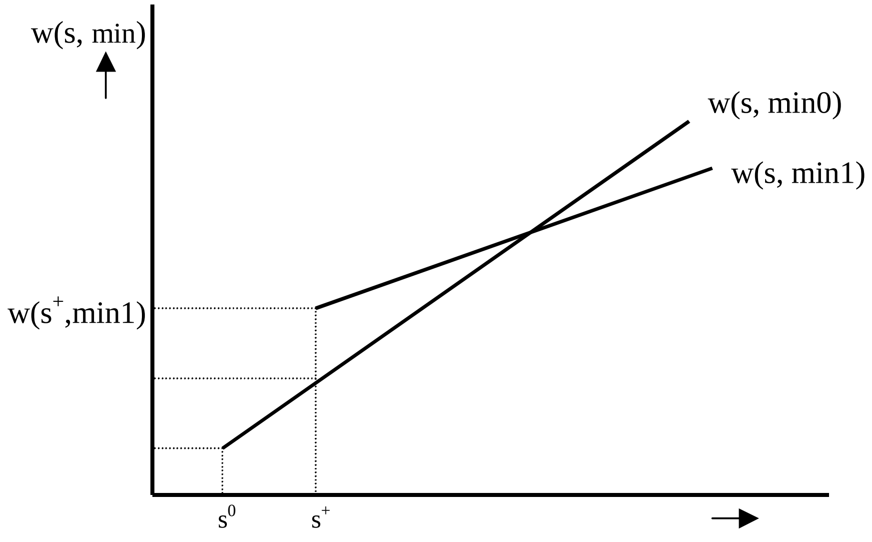
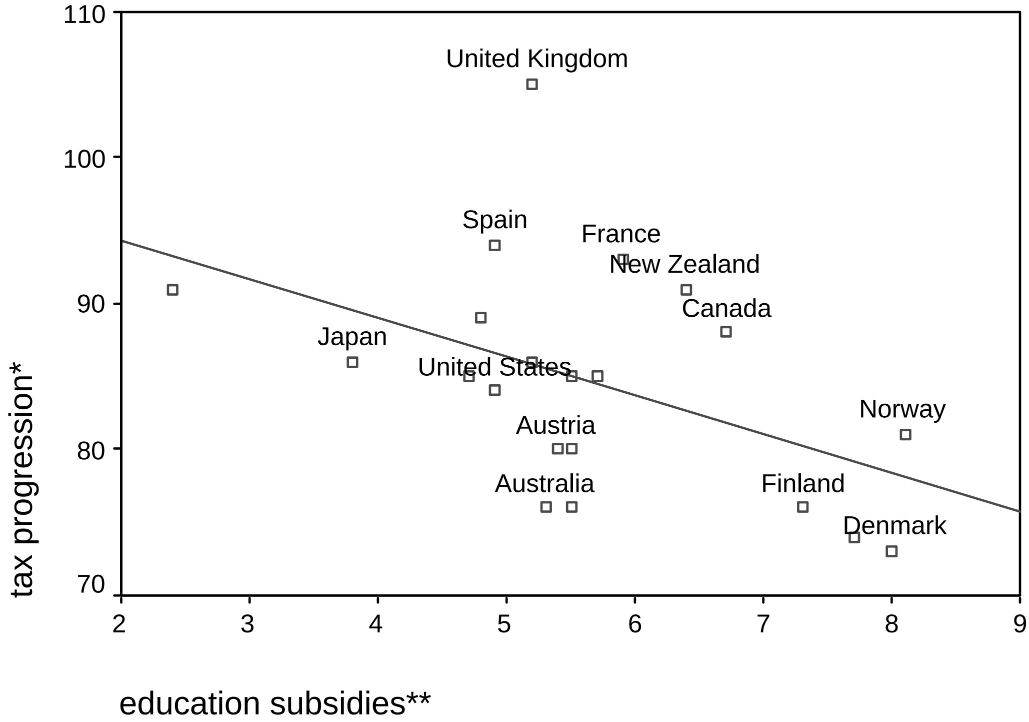


Figure 4: Education subsidies and progressivity of the income tax in OECD countries



* 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 over from Van Ewijk & Tang (2000), right panel of figure 1

Table 1

Description of the main variables in the dataset

Variable	Obs	Mean	Std.Dev.	Min	Max	Description and source
y_t	1060	8.611	1.037	6.122	11.172	Log real GDP per worker, 1985 intl. prices, Chain index (PWT 5.6a).
Δy_t	429	0.021	0.027	-0.066	0.101	10 year changes in real GDP per worker. (annualized)
D_t	370	0.560	0.319	0.100	1.552	Variance of log income. Calculated from Gini coefficient income distribution (Deininger and Squire).
ΔD_t	92	0.000	0.017	-0.052	0.051	10 year changes in variance of income. (annualized)
S_t	775	4.240	2.848	0.040	12.000	Average years of education attained by the population over 25 years of age (Barro and Lee).
ΔS_t	328	0.066	0.066	-0.225	0.387	10 year changes in average years of education. (annualized)
V_t	662	12.657	5.834	1.043	35.823	Variance of the education distribution (rough estimate constructed on the basis of Barro and Lee data).
ΔV_t	273	0.249	0.297	-0.888	1.361	10 year changes in variance of education. (annualized)

Table 1 (continued)

GDP growth equation

	(1) OLS	(2) OLS	(3) OLS (baseline model)	(4) OLS with V[educ]	(5) WLS (GDP/w)	(6) WLS (population)
ΔS_t	0.08546	0.17025	0.24335	0.24508	0.24717	0.24814
	(4.11)	(3.25)	(3.84)	(3.09)	(3.94)	(3.99)
$\Delta (S_t^2)$		-0.00780	-0.00848	-0.00881	-0.00840	-0.00898
		(2.07)	(2.16)	(1.75)	(2.18)	(2.32)
ΔS_t (year=70)			-0.09705	-0.07495	-0.09901	-0.09956
			(1.87)	(1.34)	(1.95)	(1.89)
ΔS_t (year=80)			-0.06732	-0.07423	-0.07728	-0.06933
			(1.35)	(1.42)	(1.60)	(1.39)
ΔV_t				-0.00461		
				(0.73)		
S_{t-1}	0.00297	0.00857	0.01217	0.00902	0.01231	0.01218
	(4.31)	(4.45)	(5.42)	(3.21)	(5.51)	(5.57)
S_{t-1}^2		-0.00045	-0.00058	-0.00034	-0.00058	-0.00059
		(2.67)	(3.29)	(1.60)	(3.36)	(3.46)
S_{t-1}(year=70)			-0.00349	-0.00325	-0.00386	-0.00323
			(2.80)	(2.39)	(3.14)	(2.65)
S_{t-1}(year=80)			-0.00300	-0.00391	-0.00339	-0.00276
			(2.63)	(3.14)	(3.02)	(2.48)
V_{t-1}				0.00037		
				(1.03)		
y_{t-1}	-0.00616	-0.00787	-0.00839	-0.00723	-0.00848	-0.00812
	(2.99)	(3.74)	(4.04)	(2.81)	(4.07)	(4.08)
(year=70)	0.03449	0.03506	0.05590	0.05516	0.05769	0.05427
	(10.21)	(10.34)	(8.17)	(7.25)	(8.21)	(8.05)
(year=80)	0.02120	0.02179	0.04017	0.04659	0.04269	0.03832
	(6.54)	(6.82)	(5.77)	(6.12)	(5.99)	(5.61)
Constant	0.03816	0.04033	0.02715	0.02040	0.02735	0.02601
	(2.34)	(2.51)	(1.67)	(1.04)	(1.66)	(1.66)
Observations	292	292	292	250	292	292
R-squared	0.32	0.34	0.37	0.38	0.37	0.37
F-statistic[1]	11.29	9.56	11.08	4.65	11.27	11.29
p-value	0.0009	0.0022	0.0010	0.0321	0.0009	0.0009

Absolute value of t statistics in parentheses.

[1] H_0: Long-run effect (coefficient S_{t-1} divided by minus coefficient y_{t-1}) equals short-run effect (coefficient ΔS_t). The F-tests reject the null when the p-value is smaller than 0.05.

Table 1 (continued)

Income inequality

	(1) OLS in levels	(2) OLS in levels	(3) FE in levels		(4) OLS in first difs	(5) OLS in first difs	(6) WLS (GDP/w)	(7) WLS (popul)	(8) OLS with dums for def. ch.
S_t	-0.07192	-0.08573	-0.05534	ΔS_t	-0.09820	-0.05611	-0.05718	-0.05394	-0.01934
	(2.47)	(3.05)	(1.62)		(1.40)	(1.96)	(2.01)	(1.94)	(0.77)
S_t^2	0.00085	0.00170	0.00365	$\Delta (S_t^2)$	0.00320				
	(0.38)	(0.78)	(1.56)		(0.66)				
S_t^{60}	-0.03155								
	(0.95)								
S_t^{70}	-0.02715								
	(1.51)								
S_t^{80}	0.00623								
	(0.41)								
V_t	0.00065	0.00105	-0.00070	ΔV_t	0.00094	-0.00176	-0.00169	-0.00269	0.00065
	(0.20)	(0.33)	(0.20)		(0.13)	(0.29)	(0.29)	(0.47)	(0.13)
(yr=60)	0.11027	-0.05257	-0.00012						
	(0.60)	(0.76)	(0.00)						
(yr=70)	0.13346	0.00491	-0.01062	(yr=70)	-0.00801	-0.00846	-0.00835	-0.00779	-0.00474
	(1.37)	(0.10)	(0.44)		(1.49)	(1.59)	(1.57)	(1.53)	(1.03)
(yr=80)	-0.06782	-0.02754	-0.03376	(yr=80)	-0.00554	-0.00562	-0.00560	-0.00459	-0.00351
	(0.67)	(0.68)	(1.79)		(1.28)	(1.30)	(1.30)	(1.11)	(0.88)
1{inc}	0.09302	0.09840	0.25144	Δ1{inc}	0.04095	0.04155	0.04169	0.04030	
	(1.70)	(1.81)	(4.08)		(3.89)	(3.98)	(3.89)	(3.93)	
1{hh}	-0.04313	-0.03647	-0.00107	Δ1{hh}	-0.00059	0.00007	-0.00008	-0.00065	
	(1.20)	(1.03)	(0.04)		(0.11)	(0.01)	(0.01)	(0.12)	
1{gr}	0.26680	0.26782	0.00693	Δ1{gr}	-0.00487	-0.00527	-0.00550	-0.00458	
	(6.98)	(7.00)	(0.13)		(0.42)	(0.46)	(0.48)	(0.42)	
				dumms					yes
Const.	0.73888	0.76879	0.55529	Const.	0.01011	0.01056	0.01039	0.01008	0.00571
	(10.48)	(11.35)	(5.48)		(2.72)	(2.90)	(2.90)	(2.85)	(1.71)
Obs.	262	262	262	Obs.	77	77	77	77	77
R-sq	0.47	0.46	0.21	R-sq	0.29	0.29	0.28	0.27	0.63
Nr of countries			71	Nr of countries					
				F-stat[1]					4.34
				p-value					0.0000

Absolute value of t statistics in parentheses.

[1] H_0: Dummies for definitional changes jointly insignificant. The F-tests reject the null when the p-value is smaller than 0.05.