Does a Generous Welfare State Crowd out Student Achievement? Panel Data Evidence from International Student Tests

TORBERG FALCH
JUSTINA AV FISCHER

CESIFO WORKING PAPER NO. 2383
CATEGORY 4: LABOUR MARKETS
SEPTEMBER 2008

PRESENTED AT CESIFO AREA CONFERENCE ON EMPLOYMENT AND SOCIAL PROTECTION, MAY 2008

CESifo *Working Paper No. 2383*

Does a Generous Welfare State Crowd out Student Achievement? Panel Data Evidence from International Student Tests

Abstract

Student achievement has been identified as important contributor to economic growth. This paper investigates the relationship between redistributive government activities and investment in human capital measured by student performance in international comparative tests in Mathematics and Science during the period 1980 to 2003. In fixed effects panel models, government consumption, government social expenditures, and the progressivity of the income tax system have negative effects on student achievement. We report results from a variety of model specifications and social expenditure components, and our best estimate indicates that increased government size by 10 percent reduces student achievement by 0.1 standard deviations.

JEL Code: H2, I2, C33.

Keywords: student achievement, welfare state, government size, tax system, panel data, international tests.

Torberg Falch
Department of Economics
Norwegian University of Science and Technology
Dragvoll
7491 Trondheim
Norway
Torberg.Falch@svt.ntnu.no

Justina AV Fischer
Thurgau Institute for Economics
University of Konstanz
Hauptstrasse 90
8280 Kreuzlingen
Switzerland
fischer@twi-kreuzlingen.ch or
Justina.Fischer@unisg.ch

Version: 14 July 2008

1. Introduction

The equity-efficiency quandary of the welfare state is usually attributed to perverse incentive effects in the labor market. The welfare state includes 'unproductive' government spending which reduces the return to work and is financed by distortionary taxes. Apparently, little evidence exists on the effect of welfare state arrangements on investment in human capital.

The welfare state can be seen as a social insurance mechanism, see for example Sinn (1995). When the insurance terms for the insured improve, her incentives to invest to avoid capture are weakened. In a macroeconomic context, this moral hazard problem may have detrimental effects on investment in human capital, saving, and, ultimately, economic growth.[1] Ehrlich and Zhong (1998) and Ehrlich and Kim (2007) find a negative effect of old-age pension benefits on secondary school enrolment rates, particularly for developed countries.[2] In addition, Bjørnskov, Dreher and Fischer (2007) find that higher government consumption spending is related to less subjective well-being, perhaps through misallocation of resources, the inefficiencies generated through modern taxation schemes or other transmission channels described above.

To our best knowledge, we are among the first to empirically investigate to what extent government involvement in the economy through public goods creation and income redistribution affects individual investment in human capital. In this study, we approximate the former by both government consumption, social expenditures, and progressivity of the tax system, but the latter by international student test scores, which we adjust to ensure comparability across countries and time. Most of the existing empirical analyses on economic growth include as a proxy of human capital some measure of *quantity* of education in the population. This is obviously a crude measure, and we follow Wössmann (2003a) who argues that the number of *quality*-education-years varies across countries stronger than the mere duration of education, with which it might even be uncorrelated. Indeed, Hanushek and Kimko (2000) find that average student achievement in compulsory schooling is a much more sizable determinant of economic growth than years of education in the population. The strong effect of student achievement is confirmed by Hanushek and Wössmann (2007) and Jamison, Jamison and Hanushek (2007).

[1] Welfare state arrangements may also be seen as interventions in imperfect markets, working in the opposite direction. Indeed, the empirical cross-country literature indicates that there is no relationship between government expenditures and growth, although the results vary somewhat across studies, see for example Fölster and Henrekson (2001) and Agell, Ohlsson and Thoursie (2006). Kneller, Bleaney and Gemmell (1999) distinguish between different types of taxes and spending categories, and find that overall government expenditures induce growth, but with welfare expenditures having significantly lower effects compared to what they call "productive expenditures".

[2] Zhang and Zhang (2004) find the opposite relationship. Ehrlich and Kim (2007) report that - not unsurprisingly in a growth context - the estimates are sensitive to whether the models condition on initial GDP or not.

We construct an unbalanced panel that utilizes data on comparative international tests in Mathematics and Science for the age group 13-15 years and includes eight tests in the period 1980-2003 for a maximum of 79 countries. This panel allows the application of panel data estimation methods, considerably improving the methodology of previous analyses in education economics. Existing studies utilizing cross-country variation in student achievement are either almost exclusively based on a cross-section of individual test performances in the same test of a single year or the country average in performance across many years as in Hanushek and Kimko (2000), Hanushek and Wössmann (2007), and Jamison et al. (2007).[3] We follow to some extent the latter methodological approach, but exploit the panel structure in the data and estimate country fixed effects models that account for unobserved heterogeneity across countries.

The empirical analysis is preceded by a simple theoretical model that relates the size and the scope of the welfare state to human capital investment in terms of student effort. The insurance aspect of the welfare state manifests in a system that both reduces the risk of future income and that redistributes from high income individuals to low income individuals. In traditional human capital models (Becker, 1964), where educational outcomes are determined by rational individuals weighting costs and benefits, increased redistribution of income is predicted to weaken the incentive to invest in education. The prediction of the effect of reduced idiosyncratic risk in future income is more complicated and ambiguous.

The paper proceeds as follows. The next section presents the theoretical considerations. Section 3 describes the international student tests data and our measure of adjusted average student performance, section 4 presents the empirical model, and section 5 discusses the empirical results. Section 6 offers some concluding comments.

2. Theoretical considerations

We present a simple partial two-period model that illustrates how redistribution of income and income uncertainty may affect students' incentives in schools. The model builds on Glomm and Ravikumar (1992), but differs by including income redistribution and uncertainty.[4] At the outset we consider a model with one decision-maker, the student, but below we discuss its real-world applicability where other agents such as parents and teachers have an additional influence on student effort.

[3] The only exception seems to be Barro and Lee (2001a), who employ a panel of countries participating in international tests up to 1990 to estimate the effect of school inputs on student achievement.
[4] Another related theoretical paper is Poutvaara (2007) who relates the old-age pension system and migration prospects to investment in human capital.

Consider a life-time utility maximizing individual living in two periods. In period one, she invests in her human capital. The investment is modeled as the time devoted to education, Z, instead of leisure L, where Z + L is normalized to unity. Effort at school has opportunity costs in terms of foregone leisure, but is an investment in future income. In period two, she consumes her return to education. Assuming separability over time, the life-time utility in expectational terms is

$$\mathrm{E}\{\mathrm{V}\} = \mathrm{u}(1-\mathrm{Z}) + \mathrm{rE}\{\mathrm{U}(\mathrm{C})\}$$

where r is the discount rate, C is consumption, and utility functions u and U are concave.

The welfare state is an institution that, in period two, transfers a fraction τ of the income from high income earners to individuals with low income, and thus reduces the consumption inequality in society. The uncertainty of the representative agent's future income is captured by the stochastic return of her education investment (ε), where $\mathrm{E}(\varepsilon) = 0$ and $\mathrm{Var}(\varepsilon) = \sigma^2$. We write consumption in period two as

$$\mathrm{C} = (1-\tau)\,\mathrm{y}(\mathrm{Z})(1+\varepsilon) + \tau\bar{\mathrm{y}}$$

where the right hand side is the representative individual's expected income after redistribution. The deterministic part of the income is the productivity that depends on effort in school, y(Z), with diminishing returns ($\mathrm{y}'(\mathrm{Z}) > 0,\ \mathrm{y}''(\mathrm{Z}) < 0$). Because of income redistribution only a fraction $(1-\tau)$ of the consumption is related to own productivity, y, while the fraction τ is related to the average productivity in society, $\bar{\mathrm{y}}$. This formulation implies that transfer-related income component can be written as $\tau(\bar{\mathrm{y}} - \mathrm{y}(\mathrm{Z})(1+\varepsilon))$, and that individuals with the stochastic productivity $\mathrm{y}(1+\varepsilon)$ below (above) the societal mean productivity $\bar{\mathrm{y}}$ will have a positive (negative) transfer. Thus, the redistribution factor τ is an indicator of the generosity of the welfare state. For simplicity, the implicit taxation and transfer rules are not written down in the model.

The individual maximizes equation (1) with respect to effort Z subject to constraint (2). The first order condition is

$$\mathrm{u}'(1-\mathrm{Z}) = (1-\tau)\,\mathrm{y}'(\mathrm{Z})\,\mathrm{rE}\{\mathrm{U}'(\mathrm{C})(1+\varepsilon)\}. \qquad (3)$$

In optimum, the marginal cost of effort is equal to the expected marginal return to effort. To keep the analysis simple, we continue with the quadratic utility function $\mathrm{U}_2(\mathrm{C}) = \alpha\mathrm{C} - \beta\mathrm{C}^2/2$. Then the first order condition (3) can be written as

$u'(1-Z) = (1-\tau)\,y'(Z)\,r\left[\alpha - \beta\left((1-\tau)\left(1+\sigma^2\right)y(Z) + \tau\bar{y}\right)\right]$. Calculation of the partial derivates is straightforward. Regarding the redistribution parameter τ, we have

$$\frac{dZ}{d\tau} = -\frac{1}{D}\left[\frac{u'(1-Z)}{1-\tau} + \beta(1-\tau)\,y'(Z)\,r\left(\bar{y} - \left(1+\sigma^2\right)y(Z)\right)\right] \quad (4)$$

where $D = \beta(1-\tau)^2\,r\left(1+\sigma^2\right)\left(y'(Z)\right)^2 - u''(1-Z) - y''(Z)\,u'(1-Z)/y'(Z) > 0$. The first term in the square bracket of equation (4) reflects that the incentive to invest in education is reduced when τ rises and the return to education declines. The sign of the second term depends on the relative income position. For individuals with income below the mean $\bar{y}$, more redistribution increases income and decreases the marginal utility of consumption, which partially lowers investment in education. For rich people, however, income decreases and thus the marginal utility of consumption increases, partially working in the direction of higher investment. In principle, this indirect effect of income redistribution may be so strong that its total effect on education investment is positive. However, for a representative individual with income close to mean income,[5] the effect of redistribution on her optimal effort level is negative. Similarly, in a two-country model with two differing social security systems and migration of labor, Poutvaara (2007) finds a less redistributive, earnings-based pension system of the target country to increase investment in human capital of prospective migrants compared to flat-rate pension benefits.

Regarding uncertainty, it follows that

$$\frac{dZ}{d\sigma^2} = -\frac{\beta(1-\tau)^2\,r y'(Z)\,y(Z)}{D} < 0 \quad (5)$$

Increased uncertainty in the return to education, σ^2 , decreases investment in education. The model of, for example, Charles and Luoh (2003) predicts a similar relationship, by showing that individuals prefer less risky investments, all else equal.

The result that volatility in the return to education reduces education investment is, however, not universally true, but depends on theoretical assumptions. While our model considers the investment in effort at school as an asset, education investment may also have similarities with real options. In a model where education investment is the time devoted to non-compulsory education, and assuming that, after having left education for the labor market, the individual cannot return to education, Hogan and Walker (2007) show that investment is

[5] Since mean income seems to be higher than median income in all income distributions, it is more reasonable to assume that the representative individual has income below the mean than above the mean.

positively related to the uncertainty in the labor market: If a bad state of the labor market occurs, the individual can continue at school, but if a good state occurs, she can always switch to the labor market. As increased variability of the state of the labor market increases the upside payoff more than the downside payoff, the expected return to education investment rises. Jacobs (2007) reaches the same conclusion in a somewhat different model, in which he assumes that one can always leave the labor market again for education. Then increased variability in the return to education increases the probability that education investment is profitable at one point in time.

In this paper, termination of education is not regarded as a choice variable, as we analyze determinants of education outcome at the compulsory level of schooling. For this reason, the mechanisms described by Hogan and Walker (2007) and Jacobs (2007), which rely on duration of education as an individual choice variable, are not directly applicable. However, the probability to enroll in post-compulsory education and the range of higher education institutions an individual student can choose among is typically related to preceding performance at the compulsory level. Thus, low effort in compulsory schooling reduces the probability to attain the real option inherent in higher education.

Parents and teachers may influence their childrens' and students' effort levels through instructing and bargaining. This is an important aspect because parents and teachers are concerned about student effort at school, which may come about by altruistic, dynastic preferences of parents, and a mission or wage maximization of teachers. For reasons of simplicity, our model abstracts from these additional influences. In principle, however, the optimal effort level from the viewpoint of parents and teachers may be derived in the same way as the optimal student effort above, albeit some parameters of the model may be different. For example, children are likely to be more short-sighted than adults. In our model, their discount rate r may be lower, perhaps because of more pronounced hyperbolic discounting (see for example Laibson, 1997, or O'Donoghue and Rabin, 1999). Then it follows from equation (3) that the optimal effort level is higher for parents and teachers than for students, which results in attempts of the first to positively influence effort of the latter. In addition, it follows from equation (4) that the response to the welfare state is positively related to r, making the optimal effort level from the parental and teacher point of view more responsive to income redistribution than the optimal effort level from the student's point of view. Finally, the general comprehension in society of the importance of skills and knowledge, and the societal degree of risk aversion, also affect the parameters of the objective function. Thus, the actually observed student effort is likely to be a result of not only students' behavior, but also parental and teacher behavior, and country-specific factors such as culture and population risk aversion. In our econometric analysis, we do not aim at determining specific parameters of the model above. Instead, we are the first make an attempt to

empirically identify mechanisms through which governments' welfare and redistributive policies may affect aggregated student achievement.

From our model, the implicit function of education investment is

$$Z = f\left(\tau, \sigma_l^2, \bar{y}\right) = Z\left(g, \bar{y}\right)$$

where $g = g\left(\tau, \sigma_l^2\right)$ captures the common effects of the welfare state. The investment Z decreases in $\bar{y}$ because the utility function is concave, while the effect of the welfare state g is in general ambiguous. However, redistribution of income, τ, is regarded as the dominating aspect of the welfare state and has a negative effect on Z. Our testable hypothesis is that a more generous welfare state affects individuals' educational investments negatively.

3. International measures of student achievement

We rely on comparative international tests of student achievement conducted by different international organizations. The International Association for the Evaluation of Educational Achievement (IEA) has been responsible for the largest number of such tests, among them the TIMSS tests, but also the OECD has developed a Programme for International Student Assessment (PISA).

We construct a synthetic panel data set of the national averages of international tests on student performance covering a period of almost 25 years (1980 – 2003). These tests cover the core subjects Reading, Mathematics and Natural Science separately, but we restrict our attention to student assessments in Mathematics and Science for several reasons. First, these two subjects have more similarities with each other than with Reading and are thus more suitable for constructing a synthetic panel. Second, reading skills are tested less regularly internationally, and even within the same test and year, potentially differ considerably by language. Third, performance in Mathematics and Natural Science are more likely to determine a country's innovativeness in an economic growth context, as empirically tested in Hanushek and Wössmann (2007). Comparability of test results is also given in the age dimension, as all tests included are conducted on middle-aged students (13-15 years). Choosing this age group has also the advantage that compulsory schooling still applies, mitigating selection out of education issues. The tests we utilize are described in Table 1.

Even though all tests are in the fields 'Mathematics' and 'Science', they do not necessarily test the same cognitive skill: The IEA tests are related to common elements of school curricula across countries while IAEP is geared towards the curriculum in USA building on the national testing procedures developed by the National Assessment of Education Progress NAEP. The OECD PISA test has a more real-world approach and claims to assess the skills that are considered to be essential for full participation in the society. These differences do not, however, seem to be very important with respect to measured student performance. For example, the correlation coefficient between the test results for the 18 countries participating both in TIMSS 2003 and PISA 2003 is 0.94.[6]

Table 1. Data sources description

Year	Test organization	Acronym	Test subjects	Test age or grade	Countries	Data source
1980-81	IEA	SIMS	Mathematics	13 years	3 in 1980 14 in 1981	Lee and Barro (1997) Travers and Westbury (1989)
1983-85	IEA	SISS	Science	14 years	11 in 1983 11 in 1984 1 in 1985	Postlethwaite and Wiley (1992)
1990-91	IAEP	IAEP	Mathematics and Science	13 years	2 in 1990 17 in 1991	Lee and Barro (1997)
1994-95	IEA	TIMSS	Mathematics and Science	Grade 8	4 in 1994 36 in 1995	timss.bc.edu/
1998-99	IEA	TIMSS-repeat	Mathematics and Science	Grade 8	6 in 1998 31 in 1999	timss.bc.edu/
2000-02	OECD	PISA 2000	Mathematics and Science	15 years	32 in 2000 9 in 2002	www.pisa.oecd.org
2002-03	IEA	TIMSS 2003	Mathematics and Science	Grade 8	7 in 2002 38 in 2003	timss.bc.edu/
2003	OECD	PISA 2003	Mathematics and Science	15 years	40 in 2003	www.pisa.oecd.org

Note. For some countries separate scores are reported for different parts of the country. We have calculated mean country averages by using population as weight. IEA (except the 1983/84 test) and IAEP tests are conducted in the fall in the southern hemisphere and in the spring in the northern hemisphere. PISA 2000 originally only included five non-OECD countries, but nine additionally non-OECD countries conducted the same test in 2002.

Recently, it has become common to report national averages based on Items Response Theory which weights the different questions by their difficulty ("Warm estimates", Warm, 1989), and standardizes the scores such that the average across all students across countries participating is 500 with standard deviation of 100. Particularly the PISA studies employ this

[6] The correlation coefficient between the average Science and Mathematics score in TIMSS-repeat 1999 and PISA 2000 is 0.87 and for IAEP 1991 and TIMSS 1995 the correlation coefficient is 0.80. The correlation coefficients are calculated using the adjusted test score described below. Interestingly, as can be seen from Figure 2 below, USA has its poorest performance in the IAEP test that was based on the US curriculum.

student performance measure. With this approach, the average score of a particular country will depend on the achievement of the students in the other participating countries. Thus, the test scores for a particular country are not necessarily comparable over time, in particular since the country sample changes. More importantly, for the tests prior to 1991, "Warm estimates" were not calculated, so that we have to rely on the share of correct answers for these tests.[7]

To make the scores on the different international tests comparable on a common metric, we have re-scaled the average scores for each international test by the following procedure. First we calculate the average of the Mathematics and Science tests when both subjects are tested. Second, we standardize the average score for each test to have mean zero and standard deviation equal to unity for a "core" group of 15 countries. The "core" is defined as the countries that have participated in at least six out of the eight international tests reported in Table 1, namely Australia, Canada, Hong Kong, Hungary, Israel, Italy, Japan, Korea, Netherlands, New Zealand, Russia, Sweden, Thailand, UK, and USA.[8] Third, we re-scale the scores for each of the other countries using the same parameters as for the "core" countries. Finally, since there are two tests for many countries in 2003 (TIMSS and PISA), we calculate the average of those tests in 2003.

Making the results from different tests comparable across time has been a challenge also for previous empirical studies. For example, Hanushek and Kimko (2000) calculate a measure of labor-force quality based on the percent of correct answers in international student achievement tests for the period 1965-1991. They adjust the mean for each test, but not the variance (except the linear scaling that follows from the adjustment of the mean). Adjusting the means across tests is crucial in their analysis because they subsequently calculate an aggregated 30-year average quality measure for each country. More recently, Hanushek and Wössmann (2007) utilize for their cross-section of national student performance tests from TIMSS, PISA and the IEA up to 2003 and, in addition to adjusting the means, they correct the dispersion of each single test in a similar way as ours.[9]

Figure 1a shows that the density of our measure of student achievement across the 15 "core" countries observations is close to the normal distribution. The density for all observations

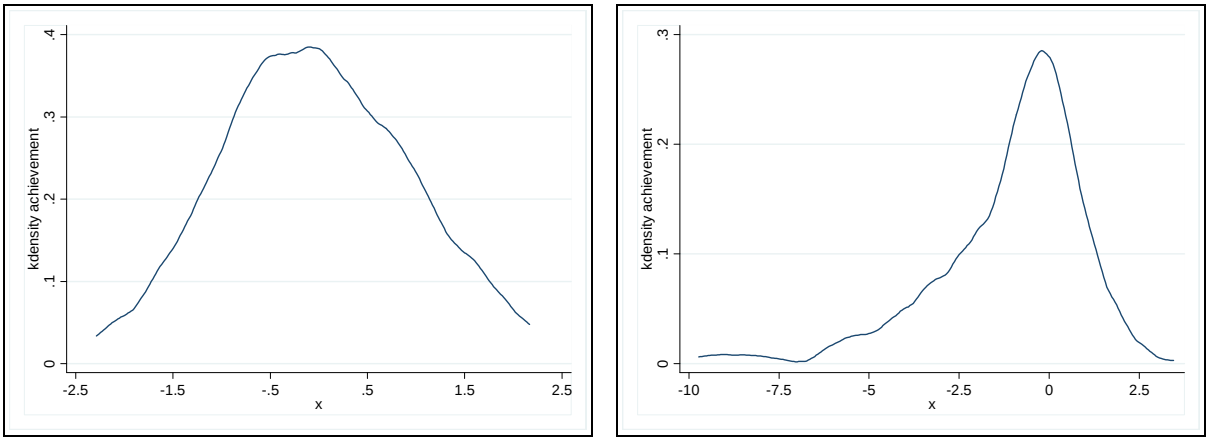
[7] We have compared the Warm estimates and percent correct answers for the IEA tests in 1994-95 and 1998-99 for which both measures are available. The correlation coefficients for Mathematics are 0.997 and 0.982, respectively, and for Science 0.994 and 0.977, respectively. Thus, the differences across countries do not seem to be influenced in any important way by the choice of scale.

[8] More precisely, we standardize the score for those of the "core" countries that participated in the particular test. Out of the 15 "core" countries used to standardize the test scores, the data sources reports results for 11 countries in 1980-81, 12 in 1983-84, 8 in 1990-91, 15 in 1994-1995, 14 in 1998-99, 15 in OECD 2000, 13 in TIMSS 2003 and 13 in OECD 2003. Only USA has test scores for all tests.

presented in Figure 1b has a long left tail, illustrating that some countries, mostly developing countries that participate less frequently in international tests, have low student achievement.

Figure 1. Kernel density of student achievement

a) "Core" country observations b) All observations



In a fixed effect model, identification is only based on within-country variation. Figure 2 shows the development over time for the "core" countries. The figure indicates that there are some systematic changes. For example, the relative achievement in the more neo-liberal Western economies USA, Canada and UK increased during the 1990s, while the achievement declined in Israel and in the transition countries Russia and Hungary. Some countries perform consistently better than others. For example, Italy performs below average and Netherlands performs above average in each test. However, Figure 3 shows that there is quite some variation in the *change* in student achievement, although the variation is lower than that for the distribution in *levels* of achievement.[10]

4. Model specification

We estimate the following model for student achievement Z of country i in year t

$$Z_{it} = \beta_1 g_{it} + \beta_2 \log(GDP/POP)_{it} + \beta_3 \log(POP)_{it} + \beta_4 EDU_{it} + \phi_i + \varphi_t + \varepsilon_{it} \qquad (7)$$

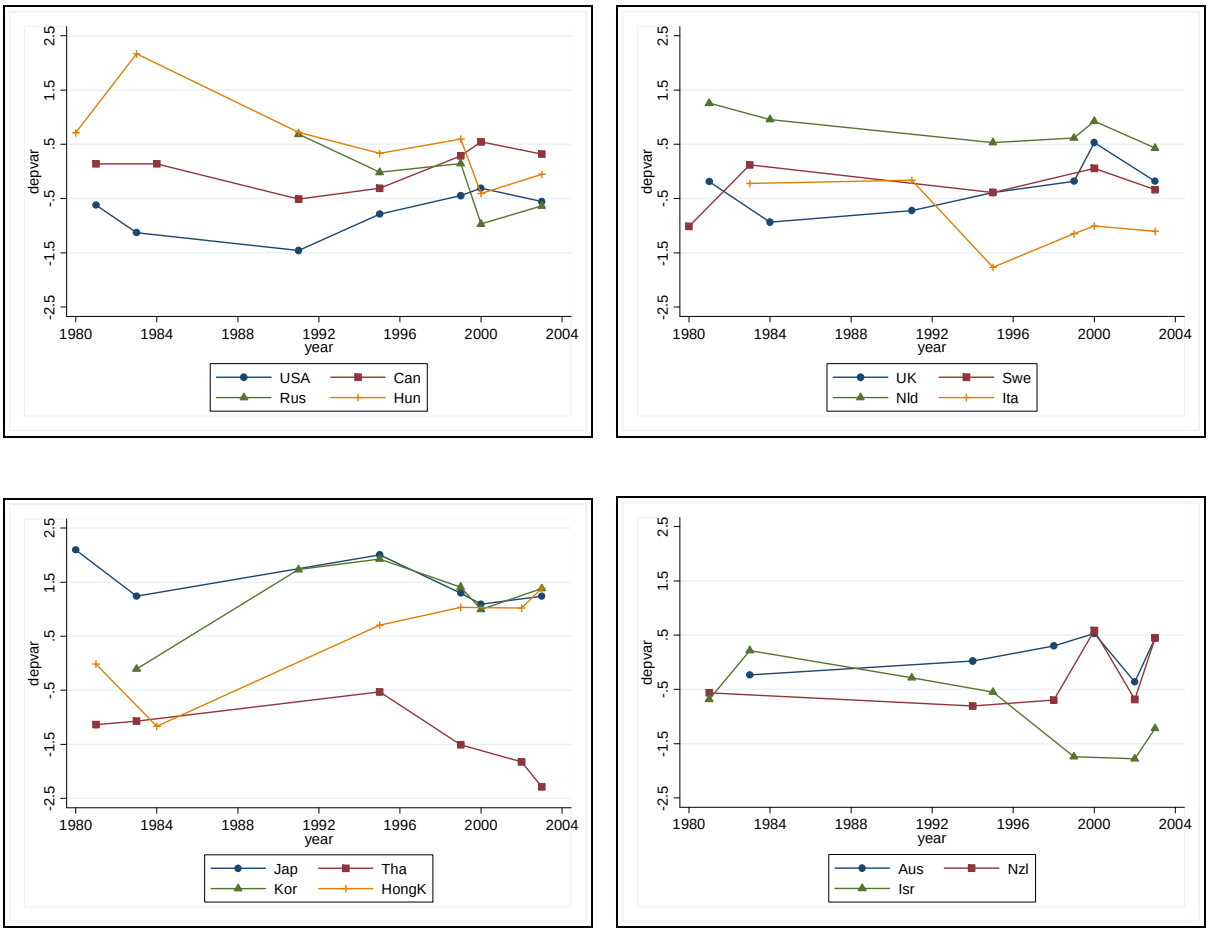
where g_{it} is a measure of the welfare state, discussed in detail below, while GDP per capita (GDP/POP) is the proxy for $\bar{y}$ in (6). Family characteristics as parental income and education have strong effects in micro studies of student achievement, which is why we also include the

[9] Hanushek and Wössmann (2007) use as their "core" countries 13 OECD countries with "stable education systems", but they do not report which countries.

[10] In Figure 3, only observations with at most eight years interval are utilized.

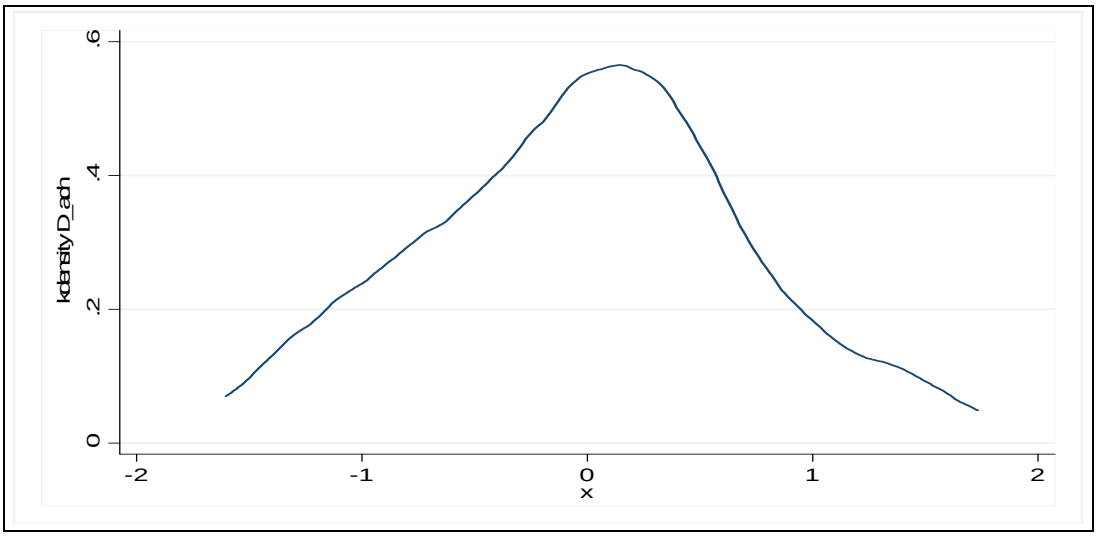
share of adult population with at least some secondary education (EDU). Since we measure GDP in per capita terms, we also employ population size (POP). Country-specific fixed effects ϕ_i account for time-invariant differences between countries such as culture for learning, risk aversion, but also school system characteristics, while time fixed effects φ_t account for the fact that the independent variables have positive trends.[11]

Figure 2. Country specific development in relative student achievement.



[11] See Wössmann, 2003b, for an analysis of educational institutions and student performance. Notice that including time-specific effects may influence the interpretation of the results, as we discuss later. The scaling of the test scores makes the scores comparable over time so that, by including time fixed effects, the model in essence draws inference on which other countries that participated on the different test and year. Thus, we also report results where the time-specific effects are replaced by a simple trend.

Figure 3. Kernel density of change in student achievement



Correct specification of the time structure of the model will prove essential. In our theoretical model, it is assumed that it is the welfare state arrangements in the future that affect students' investment decisions made today. We argue that the contemporaneous societal situation may be the best proxy for students' expectations. On the other hand, educational production is cumulative such that expectations of students and parents at earlier grades in the past may be important for observed achievement today at the age of 13-15 years. To some extent this is taken into account by the fixed effects approach because identification is based on within-country changes, but we will also investigate this issue by using five-years moving averages in the independent variables in some model specifications.

Our focal determinant in this analysis, the welfare state g_{it}, is made operational in three ways: Firstly, we employ general government consumption spending (in percentage of GDP), obtained from the WDI (2007) database of the World Bank, a widely used measure of government involvement in the economy that has been employed in various cross-country growth studies (Fölster and Henrekson, 2001, Agell et al., 2006) and happiness studies (Bjørnskov, Dreher, and Fischer, 2007 and 2008). Government consumption excludes financial transfer arrangements of the welfare state, but includes the production of services and public goods that is in the responsibility of the government and mostly financed by taxes. Given that most public goods are financed through progressive tax systems, they entail a strong consumption redistribution aspect, equalizing feasible consumption levels for heterogeneous incomes. Following the traditional public finance literature, we will refer to this measure as 'government size'. However, conditioning the model on GDP and population size, measuring the number of users and beneficiaries of public goods' creation, implies that the effect of per capita government consumption spending (government size) can be interpreted as the effect of welfare state *generosity* towards the general population.

Secondly, we use public sector social expenditures (in percentage of GDP) that are obtained from OECD *Social Expenditure database* (SOCX) and include aggregated public expenditures of all government tiers.[12] This measure captures transfers from government institutions, namely "benefits to, and financial contributions targeted at, households and individuals in order to provide support during circumstances which adversely affect their welfare, provided that the provision of the benefits and financial contributions constitutes neither a direct payment for a particular good or service nor an individual contract or transfer" (OECD, 2007, p. 7). OECD defines expenditures as 'social' if they satisfy two criteria: first, they have to intend a social purpose, and, second, these programs must be based on either inter-personal redistribution or compulsory participation (OECD, 2007, p. 8). They take the form of "cash benefits (e.g. pensions, income support during maternity leave, and social assistance payments), social services (e.g. childcare, care for the elderly and disabled) and tax breaks with a social purpose (e.g. tax expenditures towards families with children, or favorable tax treatment of contributions to private health plans)" (ibidem, p. 7), excluding the administrative costs of executing them.

Figure 4 presents within country variation in social expenditures as a share of GDP for the "core" countries with data available. There is a tendency of increased social expenditures during the period of investigation. The average share of social expenditures for the core countries in the Figure 4 increased from 0.17 in 1980 to 0.19 in 2003.[13] The Netherlands is the only country with reduced social expenditures, while Japan has the largest growth. Notice that social expenditures as a share of GDP serve as automatic stabilizers and, thus, typically shrink in a boom and expand in a recession. Thus, it is important to include GDP in the empirical model in order to avoid identification on variation in national income.

We also employ separate components of public social expenditure that differentiate government transfers by social policy area such as, e.g., health, old-age, unemployment, active labor market policies, housing, family, and 'other'. Table 2 provides an overview of spending programs that are attributed to each policy area. The major population is, in principle, entitled to all those spending categories so that each may exert an independent effect of its own.[14] The correlation coefficient between government consumption spending (from the WDI) and total social spending (from the OECD) is equal to 0.67 in our sample.

[12] The OECD defines expenditures as 'public' (as opposed to being 'private') when institutions of the 'General Government' control the relevant financial flows. The 'General Government' in this context includes different levels of government and social security funds. This definition of 'public' includes, often by tradition, transfers by compulsory social insurances and social assistance schemes (see also OECD 2007, p.8-10).

[13] For all 29 OECD countries included in the empirical analyses, social expenditures increase from 17 percent of GDP in 1980 to 21 percent in 2003.

[14] The category 'other social expenses' includes 'WWII survivor' and 'incapacity' benefits as well as 'other social expenses'. 'Other social spending' is a kitchen-sink variable, comprising benefits that could not be

Figure 4. Country specific development in social expenditures as share of GDP

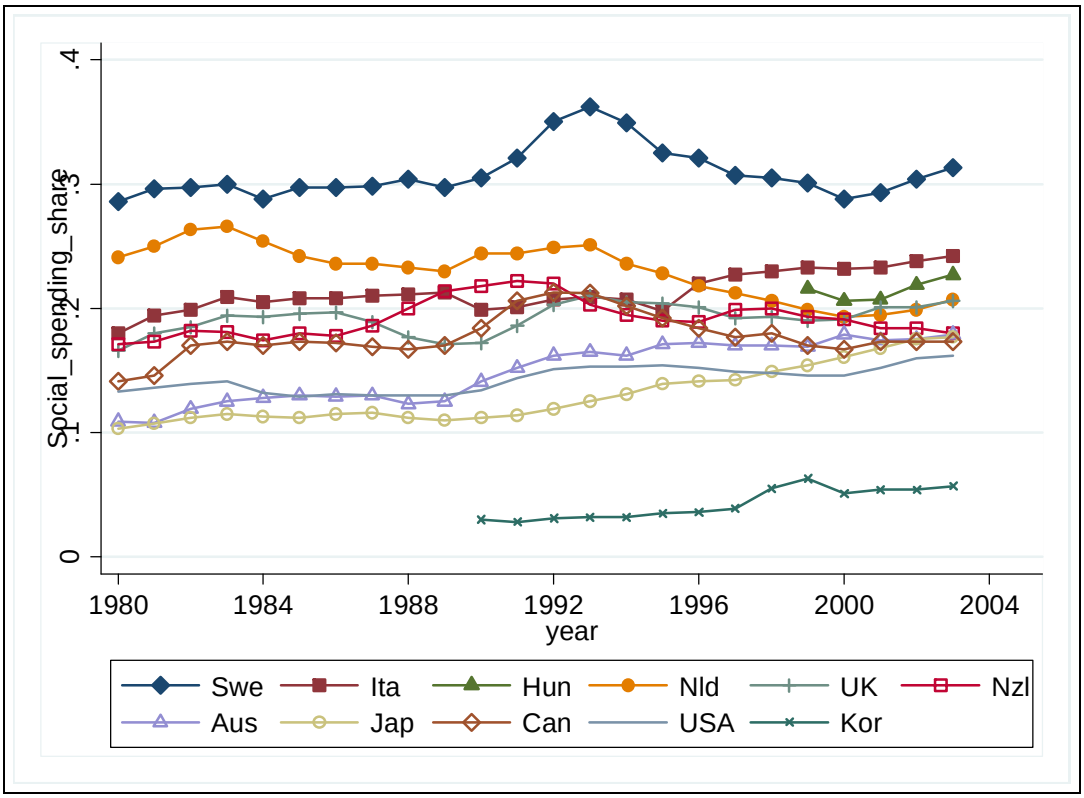


Table 2. Types of social expenditures in OECD countries

Policy area	Programs
Old-age	Pensions, early retirement pensions, home-help, residential services for the elderly.
Survivors	Pensions and funeral payments.
Incapacity-related	Care services, disability benefits, benefits accruing from occupational injury and accident legislation, employee sickness payments.
Health	Spending on in- and out-patient care, medical goods, prevention.
Family	Child allowances and credits, childcare support, income support during leave, sole parent payments.
Active labour market policies	Employment services, training youth measures subsidised employment, employment measures for the disabled.
Unemployment	Unemployment compensation, severance pay, early retirement for labour market reasons.
Housing	Housing allowances and rent subsidies.
Other social policy areas	Non-categorical cash benefits to low-income households, other social services; i.e. support programmes such as, food subsidies, which are prevalent in some non-OECD countries.

Note. Source is Social Expenditure 1980-2003, OECD 2007, p.8.

attributed to any of the major categories. If there was an effect in the mode, its interpretation would be highly questionable.

Table 3 presents some descriptive statistics on government consumption in the world sample, and social expenditures for OECD countries, including its single components. The variance in social expenditures is slightly higher than that for government consumption, both overall and within countries. The within country variation, for which we identify the effects on student achievement, constitutes 7-8 percent of the overall variance. Pension spending is the largest component of social expenditures, followed by public health spending.

Table 3. Descriptive statistics of government consumption and social expenditures

	Obser-vations	Mean	Standard deviation overall	Standard deviation within countries	Minimum value	Maximum value
General government consumption spending, percent of GDP	232 (OECD)	17.65	5.39	1.46	5.69	41.47
General government consumption spending, percent of GDP	124 (OECD)	18.90	4.22	1.05	10.08	29.62
Public sector social expenditures, percent of GDP	124 (OECD)	19.62	5.61	1.62	2.8	32.5
Active labor market policy spending, share of GDP	120 (OECD)	0.61	0.44	0.20	0	2.2
Public health spending, share of GDP	124 (OECD)	5.56	1.28	0.55	1.4	8.3
Family allowance spending, share of GDP	124 (OECD)	1.90	1.08	0.31	0	4.1
Unemployment benefit spending, share of GDP	120 (OECD)	1.17	0.90	0.44	0	4.4
Pension spending, share of GDP	124 (OECD)	6.38	2.76	0.76	0.6	12.8
Housing spending, share of GDP	102 (OECD)	0.42	0.39	0.17	0	1.8
Other social spending, share of GDP	99 (OECD)	4.00	1.38	0.64	1.50	8.90

The third measure of the generosity of the welfare state we employ is an index of income tax rate progressivity developed by the *Fraser Institute* (Gwartney and Lawson, 2002). The index constitutes an income-bracket adjusted marginal tax rate levied in the highest income bracket in one country, adjusted for the lowest income threshold for this income bracket. The redistributive impact of a given tax rate depends on the financial threshold from which on the rate applies. The top tax rate index of the Fraser Institute is therefore adjusted for threshold effects to facilitate comparability of the marginal top income tax rate across countries and

time. The index ranges from 0 to 10, with higher values representing a higher (income bracket adjusted) top tax rate, and, thus, more redistribution of income. Between 1970 and 2000, data have been collected every five years, and from 2000 on, annually. We have linearly interpolated missing values to maximize the number of observations in our regressions.

GDP and population are taken from the WDI 2007 database, while adult education attainment is taken from Barro and Lee (2001b). We use the percentage of the population over age 25 with secondary school attained. Up to 2000, the data are available every five year, and we have linearly interpolated missing values. Since the last year with information is 2000, we lag the variable by three years in order to include all the international tests in the analysis. In some robustness analyses, we also use current data on primary school educational expenditures (% GDP) and pupil-teacher ratios from the World Bank. Again, we linearly interpolate the variables when there are missing values for at most five years.[15]

5. Empirical Results

5.1. Government consumption

Table 4 presents results for government consumption spending including up to 72 countries and 232 observations. The first column simply presents the correlation between student achievement and government size, measured by government consumption spending as share of GDP. There is no unconditional correlation. Column (2) includes the control variables GDP per capita, adult educational attainment w.r.t. secondary education, population size, and time-specific fixed effects. The number of observations drops due to missing observations for adult education for some countries. As expected, we find strong positive effects of GDP per capita and adult education. The positive income effect mirrors Hanushek and Kimko (2000) who report a positive effect of student achievement on economic growth, and is in accordance with micro evidence on the effect of parental income on student performance. However, the result for GDP is not in accordance with our theoretical model where the utility function is concave. The positive effect of adult education mirrors results from previous micro-econometric studies.

The conditional effect of government consumption appears large and is significant at the 1 percent level. Conditional on income, a small public sector is favorable, and conditional on public sector size, students in rich and well-educated countries perform better than those living in poor and low-educated countries. When the share of government consumption

[15] From 1990 on, these data have been roughly collected on an annual basis.

increases by 0.1 log-points (approximately 10 percent) student achievement declines by 0.15 "core" country standard deviations.

Table 4. The effect of government consumption on student achievement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gov. consumption spending, percent of GDP (log)	0.107	-1.462**	-0.650	-1.021*	-1.596**	-1.100**	-1.060	-1.086
	(0.414)	(0.378)	(0.509)	(0.486)	(0.558)	(0.377)	(0.945)	(1.551)
GDP per capita (log)	-	1.502**	1.234*	1.123*	2.056**	0.483	1.451	3.735*
		(0.213)	(0.485)	(0.475)	(0.748)	(0.292)	(0.698)	(1.473)
Percentage secondary school attained among adults (log)	-	0.763**	0.510	0.736*	2.238**	1.268**	0.424	1.454
		(0.292)	(0.409)	(0.371)	(0.836)	(0.271)	(0.639)	(1.052)
Population size (log)	-	0.040	2.279*	1.515	-8.492	-0.048	5.221*	-11.04
		(0.77)	(1.046)	(0.978)	(6.758)	(0.62)	(2.241)	(8.728)
Trend	-	-	-	-0.041**	-	-	-	-
				(0.014)				
Country fixed effects	No	No	Yes	Yes	Yes	No	Yes	Yes
Time fixed effects	No	Yes	Yes	No	Yes	Yes	Yes	Yes
County specific trends	No	No	No	No	Yes	No	No	Yes
Observations	232	208	197	197	187	128	128	124
No of countries	72	59	48	48	43	28	28	26
Sample	All	All	All	All	All	OECD	OECD	OECD
R^2	0.0003	0.455	0.943	0.937	0.982	0.301	0.851	0.932
R^2 (within)	-	-	0.222	0.133	0.740	-	0.298	0.679

Note. Absolute standard errors in parentheses, +, * and ** denote significance at 10, 5 and 1 percent level, respectively.

Columns (3) to (5) of Table 4 present models with country fixed effects that mitigate a potential omitted variable problem.[16] In particular, the fixed effects approximate various omitted factors such as organization of schools, curriculum, school autonomy and centralization of exams. In the specific context of Table 4, they also capture that student achievement is highest in rich and well-educated countries. However, including country fixed effects in addition to year effects in column (3) does not change the point estimates of GDP and adult education attainment much compared to column (2), although the standard errors are twice as large as in the model with only time fixed effects (column (2)). On the other hand, the effect of government consumption is reduced and becomes insignificant.

[16] The number of observations is reduced in the fixed effects models because only countries with at least two observations can contribute to the identification.

In these fixed effects models, our variable of interest becomes sometimes insignificant so that identification of effects needs to be discussed. The relative large standard errors in the models with country fixed effects in columns (3) to (5) may indicate that the within country variations in student achievement and government size are too small for identification. Notably, the OLS R^2 is as high as 0.94. However, it may equally be that it is the time-specific effects that complicate identification, e.g. in column (3). The purpose of the scaling of the test scores described above is to make the scores comparable over time. In consequence, with time-specific effects, the model in essence draws inference on which other countries that participated on a particular test and year. Our motivation for including year effects is that, on average, all the independent variables have positive trends. Indeed, while the p-value of joint significance of the time-specific effects is 0.02 in the model in column (3) in Table 4, the p-value is only 0.13 when a simple trend is added. For this reason, we replace in column (4) the time-specific effects with such a time trend. The coefficient of the trend variable is negative, indicating a positive trend in the other independent determinants, as expected. The OLS R^2 appears only marginally lowered, while the within-R^2 is clearly reduced. Interestingly, the effect of government consumption spending is significantly negative in this specification. We conclude that it is not country-specific factors that make the effect of government size insignificant in column (3), but the handling of the variation over time.

What kind of within-country variation in student achievement and government consumption is driving the results? Is it country-specific trends, or fluctuations around the trends? Figure 2 above suggests that some countries exhibit a trend-like development in student achievement. To investigate this question, column (5) in Table 4 expands the model with country-specific trends. Then the effect of government consumption increases to about the same magnitude as in the model without country-specific factors in column (2), and is highly significant. Thus, it seems like it is variations around trends that account for the association between government consumption and student achievement. The result is independent of whether the model includes time-specific effects, and the difference between the models in columns (3) and (5) is not related to the fact that the number of countries that contribute to identification is necessarily smaller in the latter model.

In columns (6) to (8) of Table 4 we estimate the same models restricting our sample to OECD countries, to ease comparison with Table 5 where an OECD-specific measure of the welfare state is employed. We define the subsample of OECD countries by membership in the year 2000, but test later the robustness of our results for post-communist period effects. In the OECD sample, the effect of government consumption is significant in the model without country fixed effects (column (6)), with its size appearing independent of whether country fixed effects and country-specific trends are included in the model (columns (7) and (8)). The effect of -1 is remarkably similar to that in column (4) - the preferred model for the whole

world sample. In columns (7) and (8), however the standard errors are relatively large. Nevertheless, the variation that drives these results appears to differ between OECD countries and non-OECD countries. While the variation across country-specific trends aids identifying a strong effect for the whole sample, inclusion of trends in the OECD sample does not influence the effect of government size.

In sum, we identify negative correlations between government's public goods creation, measured by its consumption spending, and student achievement. This evidence is in accordance with our hypothesis that a more generous welfare system generates disincentives for educational investment. The result indicates that when government consumption spending as a share of GDP increases by 0.1 log-points, student achievement is reduced by about 0.1 "core country" standard deviations.

5.2. Social expenditures

Table 5 presents results for government social expenditures, measured as share of GDP, in 29 OECD countries, resulting in a sample of 124 observations. Column (1) shows that the unconditional correlation between welfare transfers to households and student achievement is negative and significant at 5 percent level. Inclusion of co-variates even increases the effect of social expenditures both in terms of magnitude and statistical significance (column (2)).[17] Also within OECD countries, there are positive impacts of GDP and adult education attainment. Column (3) of Table 5 adds country fixed effects. The effect of social expenditures is still significant at 5 percent level. The magnitude of the coefficient is larger than those when country fixed effects are excluded, but of similar size compared to when government consumption spending in Table 4 is employed. Column (4) shows that the results are not sensitive to whether the model includes time-specific fixed effects or a time trend common to all countries.

However, when including country-specific time trends, the effect of social expenditures completely disappears (column (5) of Table 5). We conclude that it is country-specific trends that drive the results, which indicates that in the OECD sample there are some systematic medium-term changes in government policy that students and parents react on.

In the last part of Table 5 we distinguish between different components of social expenditures relating to specific social policy areas such as health, family care, labor market, pension system, etc. In column (6) we replace total social expenditures with all its various components. All components have a negative sign as expected, except 'health care spending', 'other spending', and 'family allowances'. Notably, the spending category 'other spending' is

[17] The sample is smaller in column (2) than in column (1) of Table 5 because adult educational attainment is not available for Luxembourg.

of a rather 'kitchen-sink' nature so that its estimate is not easy to interpret. The effect of family allowances is in fact significant at 10 percent level, which may indicate that relaxing parents' budget constraints in the poorest families may have a positive effect on the average achievement level of students.

Table 5. The effect of social expenditures on student achievement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Gov. social expenditures, percent of GDP (log)	-0.488*	-0.670**	-0.997*	-0.899+	0.505	-	-	-
	(0.223)	(0.230)	(0.464)	(0.457)	(1.070)			
GDP per capita (log)	-	0.717*	0.523	1.508+	6.500**	-0.686	0.546	-0.008
		(0.315)	(0.895)	(0.863)	(2.355)	(1.262)	(0.872)	(1.000)
Percentage secondary school attained among adults (log)	-	1.225**	0.548	0.433	0.572	-0.026	0.671	0.444
		(0.275)	(0.627)	(0.604)	(1.167)	(1.255)	(0.6216	(0.704)
Population size (log)	-	-0.028	4.448+	2.865	-11.63	-2.474	3.653	4.268
		(0.61)	(2.398)	(2.371)	(9.178)	(4.845)	(2.323)	(3.132)
Trend	-	-	-	-0.031	-	-	-	-
				(0.022)				
Pension spending (log)	-	-	-	-	-	-2.470*	-0.986**	-
						(1.033)	(0.333)	
Active labor market policy spending (log)	-	-	-	-	-	-0.392	-	-0.576**
						(0.298)		(0.191)
Unemployment spending (log)	-	-	-	-	-	-0.317	-	-
						(0.210)		
Family allowances (log)	-	-	-	-	-	0.824+	-	-
						(0.422)		
Health care spending (log)	-	-	-	-	-	0.208	-	-
						(0.814)		
Housing spending (log)	-	-	-	-	-	-0.042	-	-
						(0.198)		
Other spending (log)	-	-	-	-	-	0.912	-	-
						(0.667)		
Country fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	No	Yes	Yes	No	Yes	Yes	Yes	Yes
Country specific trends	No	No	No	No	Yes	No	No	No
Observations	124	121	121	121	118	80	121	113
No of countries	29	28	28	28	26	19	28	28
R^2	0.038	0.339	0.861	0.835	0.932	0.825	0.868	0.863
R^2 (within)	-	-	0.258	0.118	0.665	0.571	0.293	0.376

Note. Absolute standard errors in parentheses, +, * and ** denote significance at 10, 5 and 1 percent level, respectively.

The different expenditure components are positively correlated, which may contribute to the heterogeneous and mainly insignificant effects in column (6). Thus, we have run regressions

including each of the components separately. In all cases, the effects are negative, except for family allowances. Columns (7) and (8) in Table 5 report the single two cases where the separate effects are significant at least at 10 percent level. The effects of spending on active labor market policies and pension payments are both performance lowering. Since the former constitutes only a small part of total social expenditures, the negative effect of social expenditures in columns (1) to (4) seem to a large part be driven by pension spending.[18]

Taken all together, we identify a negative effect of social spending in OECD countries for student achievement, in support of our hypothesis. Among the different types of welfare transfers, it is pension benefits that contribute most to this effect. This result is not surprising given that in most developed countries pensions systems are highly redistributive in nature: on the one hand, they guarantee an income-independent minimum rent to every contributor, on the other hand, the place a cap on the maximum rent, equalizing rent incomes in the non-active elderly population.

5.3. Tax progressivity

Table 6 uses the same models as Table 4, but replaces the government size variable with a 10-point scale index of the top marginal income tax rate, adjusted for the income bracket, our measure of progressivity of the income tax system. For this welfare state generosity measure, there is a negative correlation with student achievement for the whole sample (column (1)), but the effect disappears when we include the control variables (column (2)). However, in the models with country fixed effects in columns (3) and (4), the effect is significant at five and 10 percent level, respectively. When the index increases by one standard deviation, which is about 2.5 point on the index, student achievement is reduced by 0.21 adjusted "core" country standard deviations.

How the effect of the variable of interest changes when we alter the model specification varies greatly between the government size, social transfers and tax progressivity models, which indicates that the variables have very different features.[19] Nevertheless the main result for both variables is that they tend to reduce student achievement. The quantitative effects of the adjusted top tax rate are difficult to gauge as using an index variable makes quantitative predictions difficult.

[18] We are unable to exclude the possibility that more public expenditures on pension may equally proxy for a large body of civil servants. In this case, the prospects of becoming a civil servant with high job security and generous retirement options may equally lower effort in mandatory schooling.

[19] The correlation coefficient between government consumption spending (log) and the top marginal tax rate index is -0.44, and for the social spending (log) in OECD countries -0.30. Please note that financing of government activities also occurs through corporate taxation and indirect taxes on e.g. consumption goods.

Table 6. The effect of tax progressivity on student achievement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Income tax rate progressivity	-0.141*	-0.009	-0.086*	-0.074+	-0.074	-0.133**	-0.025	0.018
	(0.058)	(0.059)	(0.042)	(0.039)	(0.057)	(0.046)	(0.050)	(0.066)
GDP per capita (log)	-	1.297**	1.113*	1.031+	2.118+	0.001	1.399+	4.794**
		(0.226)	(0.558)	(0.523)	(1.197)	(0.258)	(0.772)	(1.534)
Percentage secondary school attained among adults (log)	-	0.945**	-0.601	-0.51	1.013	1.403**	-0.067	0.637
		(0.336)	(0.496)	(0.449)	(1.043)	(0.285)	(0.671)	(1.258)
Population size (log)	-	0.109	0.654	-0.632	0.857	0.005	4.553	-1.245
		(0.085)	(1.245)	(1.178)	(10.023)	(0.061)	(2.789)	(10.999)
trend	-	-	-	-0.006	-	-	-	-
				(0.017)				
Country fixed effects	No	No	Yes	Yes	Yes	No	Yes	Yes
Time fixed effects	No	Yes	Yes	No	Yes	Yes	Yes	Yes
County specific trends	No	No	No	No	Yes	No	No	Yes
Observations	206	180	180	180	180	116	116	116
No of countries	64	56	56	56	56	28	28	28
Sample	All	All	All	All	All	OECD	OECD	OECD
R^2	0.029	0.387	0.958	0.952	0.983	0.280	0.855	0.939
R^2 (within)	-	-	0.188	0.066	0.674	-	0.290	0.700

Note. Absolute standard errors in parentheses, +, * and ** denote significance at 10, 5 and 1 percent level, respectively.

Regarding OECD countries, there is a strong negative effect of tax progressivity (column (6)) when the model does not include country fixed effects, but the effect disappears when country fixed effects are added. This finding is similar to the pattern observed for government consumption spending in the OECD countries (columns (6) to (8) of Table 4). In Table 6, it most possibly suggests that the conditional within country variation of tax progressivity is too small to identify a statistically significant effect.

To sum up, we find that a more income redistributive state as measured by the progressivity of the income tax system is negatively correlated with student performance in international tests of Mathematics and Science. Again, this result holds for the world sample, and in tendency for the OECD sample. It is in strong support of our hypothesis.

5.4. Generosity of the welfare system

The sizes of government consumption and welfare spending, expressed in percentage of GDP, are commonly viewed as proxies for the generosity of redistributive activities by the government. However, in principle, generosity of social transfers can be more directly assessed when values per recipient of social benefits in place of per capita numbers are employed. However, precise information on number of recipients is not easily available. Thus,

we analyze the effects of generosity by estimating models with those selected components of social expenditures for which appropriate population shares serving as proxies for number of recipients are available. We employ either the share of elderly in the population or the share of unemployed in the active population. Indeed, omission of beneficiary measures might have biased our previous results as the spending estimates might capture population composition effects. Simple correlations of the spending measures with the number of their specific beneficiaries are large.[20]

In Table 7 we present results for models with measures of the number of recipients included. Taken all together, the results are not sensitive to inclusion of proxies for the number of beneficiaries. Pension spending and active labor market policy spending still exert a student performance lowering impact when the share of elderly and the unemployment rate, respectively, are included in the model (columns (1) and (2)), while the effects of unemployment and health care spending remain insignificant (columns (3) and (4)). The similarity of the coefficients on the spending variables with the original models reported in Table 5 suggests that the bias from using spending measured per GDP (conditional on population size) in place of per recipient is rather small. Regarding pension spending in column (1), the significance level is reduced to 5 percent when the share of elderly is included. While the share of the population above 60 years of age is insignificant, the test of joint significance clearly suggests that both variables are jointly related to student achievement. The effect of active labor market policies spending is equally lowered in significance (now at 5 percent level) when the unemployment rate is included (column (2)), while this time the test of joint significance clearly suggests that only one of the variables is related to student achievement.

5.5. Robustness analyses

The student test scores from the 1980s are not average results for jointly conducted Mathematics and Science tests as those achievement tests in the 'post-communist period' (1990s and beyond), but separate tests on the two subjects. Another reason for splitting the sample into two time-periods is that many argue that test designs and test procedures have improved over time. Therefore, the dependent variable may incorporate a larger measurement error in the 1980s than in later periods. Table 8 presents results for regressions on the subsample for the post-communist period 1990-2003. Columns (1)–(3) in the table show that the effects of all our measures of government generosity - government consumption spending, social transfers, and income tax progressivity - are in fact larger in this subsample compared

[20] The correlation coefficients between unemployment spending and unemployment rate is 0.51 and between pension spending and the share of the population above the age of 60 is 0.86. The correlation between active labor market policy spending and the unemployment rate is only 0.17.

to the full sample that includes the pre-1990 tests, although the effect of government size is still insignificant in the model with time-specific fixed effects.

Table 7. Generosity of the welfare state: OECD countries

	(1)	(2)	(3)	(4)
GDP per capita (log)	0.766	-0.156	0.975	1.029
	(0.919)	(1.340)	(1.211)	(0.963)
Percentage secondary school attained among adults (log)	0.707	0.747	1.578+	0.357
	(0.658)	(0.832)	(0.846)	(0.668)
Population size (log)	2.969	2.919	5.581+	2.619
	(2.437)	(3.523)	(2.885)	(2.518)
Pension spending (log)	-1.057*	-	-	-
	(0.502)			
Active labor market policy spending (log)	-	-0.544*	-	-
		(0.209)		
Unemployment spending (log)	-	-	-0.112	-
			(0.182)	
Health care spending (log)	-	-	-	0.104
				(0.590)
Share of elderly (log)	-0.783	-	-	-2.048*
	(0.964)			(0.868)
Unemployment rates	-	-0.027	0.000	-
		(0.032)	(0.037)	
Country fixed effects	yes	yes	yes	Yes
Time fixed effects	yes	yes	yes	Yes
Observations	110	110	113	110
Countries	27	28	27	27
R^2	0.8632	0.8687	0.8479	0.8544
R^2 (within)	0.3328	0.3995	0.3079	0.2902
F-test of joint significance (p-value)	5.5078	3.458	0.2524	3.1086
	0.006	0.0372	0.7776	0.051

Note. Absolute t-values in parentheses, +, * and ** denote significance at 10, 5 and 1 percent level, respectively.

It might also be argued that government size and the generosity of the welfare state are proxies for educational expenditures, albeit some empirical evidence both from analyses within and between countries indicates that educational expenditures have at most a minor effect on student performance (e.g. Hanushek and Luque 2003).[21] Similarly, welfare effects might also proxy school (or teacher) quality – more generous governments in term of public

[21] For contrasting evidence exploiting variation across Swiss states, see Fischer (2007). She finds that educational spending exerts a decisive impact through teachers' wages.

goods creation and social transfers may also have a preference for creating a high quality public education system, equalizing incomes and earning prospects of the future generations.[22] If, however, educational expenditures or education quality have a positive effect on student achievement, their exclusion will positively bias the effect of government size in our previous models since the variables are likely to be positively correlated.[23]

Table 8. Post-communist period and school quality measures

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Gov. consumption expenditures, percent of GDP (log)	-0.856	-	-	-0.924	-	-	-0.587	-	-
	(0.569)			(0.637)			(0.550)		
Gov. social expenditures, percent of GDP (log)	-	-1.360*	-	-	-1.496**	-	-	-1.394**	-
		(0.537)			(0.563)			(0.511)	
Income tax rate progressivity	-	-	-0.093+	-	-	-0.116*	-	-	-0.132**
			(0.054)			(0.046)			(0.049)
GDP per capita (log)	0.554	0.659	0.412	1.579*	0.152	2.005**	1.347*	1.216	0.901
	(0.577)	(1.008)	(0.676)	(0.616)	(1.024)	(0.697)	(0.542)	(1.030)	(0.618)
Percentage secondary school attained among adults (log)	-0.498	0.658	-1.039	-1.168+	-0.168	-1.426+	1.154*	0.571	-0.166
	(0.717)	(1.077)	(0.878)	(0.646)	(0.868)	(0.768)	(0.533)	(0.794)	(0.725)
Population size (log)	2.836+	4.303	1.579	1.178	6.258	0.906	1.647	2.266	0.628
	(1.557)	(4.117)	(1.771)	(1.516)	(4.087)	(1.466)	(1.190)	(3.189)	(1.363)
Primary education expenditures per pupil, percent of GDP (log)	-	-	-	0.412	0.433	0.579*	-	-	-
				(0.295)	(0.341)	(0.287)			
Pupil-teacher ratio in primary education	-	-	-	-	-	-	-0.005	0.048	-0.003
							(0.032)	(0.031)	(0.033)
Sample	All	OECD	All	All	OECD	All	All	OECD	All
Time period 1990 - 2003	yes	yes	yes	no	no	no	no	no	no
Country fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Time fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	177	102	163	166	104	152	185	103	160
Countries	58	28	56	52	27	50	58	27	55
R^2	0.9617	0.8791	0.9626	0.9558	0.8740	0.9609	0.9484	0.8833	0.9645
R^2 (within)	0.2251	0.3290	0.1744	0.2377	0.3010	0.3161	0.2462	0.3398	0.2094
F-test (social spending, school quality)				1.5582	3.749	4.3742	0.5928	5.2155	3.9535
(p-value)				0.2156	0.029	0.0155	0.5545	0.0081	0.0226

Note. Absolute t-values in parentheses, + * and ** denote significance at 10, 5 and 1 percent level, respectively.

[22] In principle, the quality of teaching may also be influenced by the government creating incentives for teachers. That teachers may influence student effort is discussed in the model section.

[23] The correlation coefficients of per pupil spending in primary education with our government consumption and social spending exceed well 0.5, while those with pupil-teacher ratio in primary education are -0.76 and -0.29, respectively.

In columns (4) to (9) of Table 8 we add to our model educational expenditures per pupil in primary schools as a percentage of GDP and pupil–teacher ratios in primary schools from the World Bank education database.[24] The effect of educational expenditures and pupil–teacher ratio are insignificant in all but one specification. Most important, the effects of government consumption, social expenditures, and the progressivity of the income tax system remain qualitatively unchanged when these measures of resource use and school quality in primary education are accounted for. Overall, we find no indication that the generosity of the welfare system and government public goods' creation proxies previously unobserved educational expenses or school quality.

Lastly, we investigate whether the choice of functional form of the model is important. One may argue that it is not short-term fluctuations in the independent variables that are important, but the development in the medium or long term. We have carried out identical regressions as reported in Tables 4-6 using 5-year moving averages of the independent variable in place of current values. The findings for government size in Table 4 appear partly sensitive to the choice of time window, although a robust and large performance lowering effect at the 1 percent level remains if country-specific time trends are included. The effect of social expenditures analyzed for the OECD countries (analogously to Table 5) appears insignificant throughout, albeit their coefficients prevail in size and direction. In contrast, the negative association between active labor market policies spending and pension benefit spending is strongly corroborated.[25] Estimation of 5-year moving averages corroborates the results for progressivity of the tax system for OECD countries (analogously to Table 6), while the coefficients for the whole country sample are now smaller and insignificant, albeit all with negative signs. We have also investigated whether the results for government spending are sensitive to using the log of the spending shares. The analogous results for Table 4 are similar and show, again, the importance of country-specific time trends to identify the effect of government size in the world sample. In contrast, the coefficients for social spending in the OECD become now insignificant, suggesting a model misspecification. Results for social spending components are, again, comparable to the original Table 5, while we now also observe family allowances to be significantly and positively contributing.

[24] For secondary education, the number of observations was insufficient. The inclusion of interpolated values lets the samples in Table 8 increase by 19-31 percent.

[25] Significant at least at the 5 percent level. In addition, housing subsidies appear now conducive to student performance, (at the 5 percent level) in a similar manner as family allowances in the current value model.

6. Conclusion

The recent publications of international comparative student achievement tests such as PISA and TIMSS have spurred the debate on quality of public education in many countries. Indeed, empirical analyses suggest that it is quality of the educational outcome rather than quantity, expressed in terms of years of education, that matters to economic growth. While most of the discussion has been centered around educational resource use and institutions, analyses of macro incentives implicit in economic policy are limited. At the same time, there is an ongoing debate whether the welfare state has excessively grown over the last decennials leading to hampered macro-economic growth through bureaucratic waste and high income taxation. Indeed, recent happiness research suggests that government consumption spending reduces welfare in society, both at the individual as well as at the aggregate level (Bjørnskov, Dreher and Fischer, 2007 and 2008). This paper aims to contribute to these discussions by empirically linking the quality-of-education-question to the size-of-welfare-state-debate.

This paper is among the first analyses which provide an investigation into the effects of the size of the welfare state on investment in human capital during mandatory education. First, we develop a simple economic model that demonstrates how individuals' optimal investment in human capital declines in the generosity of the welfare system. In our empirical analysis, we test the impact of a variety of welfare state measures on student achievement in Mathematics and Science at the lower secondary education level We employ a panel of up to 72 participating countries during the period 1980-2003, making the test scores comparable across testing institutions, test years and ways to assess student performance. In fixed effects panel models that account for unobserved country heterogeneity, we identify a student performance deteriorating impact of a more generous welfare state, measured by either general government consumption spending, the progressivity of the income tax system, and, for OECD countries, more narrowly by social welfare spending on households and persons. An investigation into the single components of social spending reveals the dominance of redistributive pension systems and active labor market policies for creating disincentives to human capital investment.

Overall, this paper contributes to the branch of empirical (and theoretical) literature which suggests that cuts in the welfare state and government consumption spending might have beneficial effects for society – in both OECD and non-OECD countries. However, reality often is more complex and a differentiated view is advisable. The policy implications from our empirical results are limited by, e.g., the fact that findings in form of econometric 'point estimates' always must be interpreted as marginal, 'local' changes. Thus, our results cannot be interpreted as if it were optimal in terms of student performance to cut government spending down to zero, abolish state pensions systems or cease to redistribute income via the tax

system. Indeed, in the OECD country sample, different components of social welfare impact educational investment in different ways. Arguably, our results even suggest in tendency that student effort would increase when means-tested family allowance programs targeted at poor households with children were extended. However, we also have to acknowledge that our findings are for high- and middle-high income countries only - leaving the question unanswered to what extent least developed countries are affected by such redistributive government involvement in the economy.[26] For the least developed countries which are stuck in poverty traps and to which additional economic growth may be most beneficial, empirical research on the welfare state effects is still lacking, and different policies may appear as optimal.

Acknowledgement

We thank Fredrik Carlsen, Urs Fischbacher, Stefan Hupfeld, Thomas Lange, Bjarne Strøm, Ludger Wößmann, and seminar participants at the CESifo Area Conference on Employment and Social Protection, 2008, Economics Departments at the University of Konstanz and the Norwegian University of Science and Technology for helpful comments and suggestions. Justina Fischer acknowledges a EU Marie Curie experienced researcher fellowship (RTN ENABLE) while visiting the Stockholm School of Economics.

[26] Indeed, also Bjørnskov et al. (2007) argue that their well-being lowering influence of size of government sector may be driven by exclusion of least developed countries, possible disguising an inverted U-shaped relation.

References

Agell, J., H. Ohlsson and P. S. Thoursie (2006): ”Growth effects of government expenditure and taxation in rich countries: A comment”. *European Economic Review* 50, 211-218.

Barro, R. J., and J.-W. Lee (2001a): “Schooling quality n a cross-section of countries.” *Economica* 68, 465-488.

Barro, R. J., and J.-W. Lee (2001b): “International data on educational attainment: updates and implications.” *Oxford Economic Papers* 53, 541-563.

Becker, G.S. (1964, 1993, 3rd ed.). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago, University of Chicago Press.

Bjørnskov, C., Dreher, A. and Fischer, JAV (2007): "The bigger the better? Evidence of the effect of government size on life satisfaction around the world". *Public Choice* 127(3), 267-292.

Bjørnskov, C., Dreher, A. and Fischer, JAV, (2008). "Cross-country determinants of life satisfaction: exploring different determinants across groups in society". *Social Choice and Welfare* 30(1), 119-173.

Charles, K. K. and M.-C. Luoh (2003): “Gender differences in completed schooling”. *Review of Economics and Statistics* 85, 559-577.

Ehrlich, I. and J. G. Zhong (1998): “Social security and the real economy: An inquiry into some neglected issues”. *American Economic Review* 88, 151-157.

Ehrlich, I., and J. Kim (2007): "Social security and demographic trends: Theory and evidence from the international experience", *Review of Economic Dynamics* 10, 55-77.

Fischer, Justina A.V., (2007). "The Impact of Direct Democracy on Public Education: Evidence for Swiss Students in Reading, Mathematics and Natural Science," Working Paper Series in Economics and Finance 688, Stockholm School of Economics.

Fölster, S. and M. Henrekson (2001): “Growth effects of government expenditure and taxation in rich countries”. *European Economic Review* 45, 1501-1520.

Glomm, G., and B. Ravikumar (1992): "Public versus private investment in human capital: Endogenous growth and income inequality". *Journal of Political Economy* 100, 818-834.

Gwartney, J. and R. Lawson. (2002), *Economic Freedom of the World: 2002 Annual Report*. Vancouver: Fraser Institute.

Hanushek, E. A. and D. D. Kimko (2000): "Schooling, labor-force quality, and the growth of nations". *American Economic Review* 90, 1184-1208.

Hanushek, E. A. and J. A. Luque (2003): "Efficiency and equity in schools around the world". *Economics of Education Review* 22, 481-502.

Hanushek, E. A. and L. Wössmann (2007): The role of school improvement in economic development". NBER Working Paper 12832.

Hogan, V. and I. Walker (2007): "Education choice under uncertainty: Implications for public policy". *Labour Economics* 14, 894-912.

Jacobs, B. (2007): "Real options and human capital investment". *Labour Economics* 14, 913-925.

Jamison, E. A., D. T. Jamison and E. A. Hanushek (2007): "The effects of education quality on income growth and mortality decline". *Economics of Education* 26, 771-788.

Kneller, R., M. Bleaney and N. Gemmell (1999): "Fiscal policy and growth: Evidence from OECD countries": *Journal of Public Economics*, 74, 171-190.

Laibson, D. (1997): "Golden eggs and hyperbolic discounting". *Quarterly Journal of Economics* 112, 443-477.

Lee, J.-W., and R. J. Barro (1997): "Schooling Quality in a Cross-Section of Countries" *Economica* 68, 465-488. Data on http://www.nber.org/pub/barro.lee/

O'Donoghue, T., and M. Rabin (1999): "Doing it now or later". *American Economic Review* 89, 103-124.

OECD (2007): "$ocial €xpenditure 1980-2003: Interpretative Guide of SOCX", Paris: OECD 2007.

Postlethwaite, T. N., and D. E. Wiley (1992): "The IEA study of science II: Science achievement in twenty-three countries". Pegamon Press, New York

Poutvaara, P. (2007): "Social security incentives, human capital investment and mobility of labor". *Journal of Public Economics* 91, 1299-1325.

Sinn, H.-W. (1995): "A Theory of the Welfare State, *Scandinavian Journal of Economics* 97, 495- 526.

Travers, K. J., and I. Westbury, eds. (1989): "The IEA Study of Mathematics I: Analysis of Mathematics Curricula". Pergamon Press, New York.

Warm, T.A. (1989). "Weighted Maximum Likelihood Estimation of Ability in Item Response Theory". *Psychometrika*, 54, 427-450.

Wössmann, L. (2003a): "Specifying human capital". *Journal of Economic Surveys* 17, 239-270.

Wössmann, L. (2003b): "Schooling Resources, Educational Institutions and Student Performance: the International Evidence". *Oxford Bulletin of Economics & Statistics* 65, 117-170.

Zhang, J. and J. Zhang (2004): "How does social security affect economic growth? Evidence from cross-country data". *Journal of Population Economics* 17, 473-500.

CESifo Working Paper Series

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2321 Rainald Borck and Matthias Wrede, Subsidies for Intracity and Intercity Commuting, June 2008

2322 Patricia Apps and Ray Rees, Testing the Pareto Efficiency of Household Resource Allocations, June 2008

2323 Amihai Glazer, Vesa Kanniainen and Panu Poutvaara, Firms' Ethics, Consumer Boycotts, and Signalling, June 2008

2324 Claudia M. Buch, Jörg Döpke and Kerstin Stahn, Great Moderation at the Firm Level? Unconditional vs. Conditional Output Volatility, June 2008

2325 Helmuth Cremer, Philippe De Donder, Dario Maldonado and Pierre Pestieau, Forced Saving, Redistribution and Nonlinear Social Security Schemes, June 2008

2326 M. Hashem Pesaran and Paolo Zaffaroni, Optimal Asset Allocation with Factor Models for Large Portfolios, June 2008

2327 Harald Badinger and Peter Egger, Horizontal versus Vertical Interdependence in Multinational Activity, June 2008

2328 Jan K. Brueckner and Harris Selod, A Theory of Urban Squatting and Land-Tenure Formalization in Developing Countries, June 2008

2329 Paolo M. Panteghini, Corporate Debt, Hybrid Securities and the Effective Tax Rate, June 2008

2330 Guglielmo Maria Caporale, Juncal Cuñado and Luis A. Gil-Alana, Modelling Long-Run Trends and Cycles in Financial Time Series Data, June 2008

2331 Avi Ben-Bassat and Momi Dahan, Social Identity and Voter Turnout, June 2008

2332 Martin R. West and Ludger Wößmann, "Every Catholic Child in a Catholic School": Historical Resistance to State Schooling, Contemporary Private Competition, and Student Achievement across Countries, June 2008

2333 Erkki Koskela and Panu Poutvaara, Outsourcing and Labor Taxation in Dual Labor Markets, June 2008

2334 Philippe Choné and Laurent Linnemer, Optimal Litigation Strategies with Signaling and Screening, June 2008

2335 Albert Solé-Ollé and Pilar Sorribas-Navarro, Does Partisan Alignment Affect the Electoral Reward of Intergovernmental Transfers?, June 2008

2336 Antonio Cabrales and Piero Gottardi, Markets for Information: Of Inefficient Firewalls and Efficient Monopolies, June 2008

2337 Sumon Majumdar and Sharun W. Mukand, The Leader as Catalyst – on Leadership and the Mechanics of Institutional Change, June 2008

2338 Ulrich Hange, Tax Competition, Elastic Labor Supply, and Growth, June 2008

2339 Guy Laroque and Bernard Salanié, Does Fertility Respond to Financial Incentives?, June 2008

2340 Adriano Paggiaro, Enrico Rettore and Ugo Trivellato, The Effect of Extending the Duration of Eligibility in an Italian Labour Market Programme for Dismissed Workers, June 2008

2341 Helmut Seitz, Minimum Standards, Fixed Costs and Taxing Autonomy of Subnational Governments, June 2008

2342 Robert S. Chirinko, Leo de Haan and Elmer Sterken, Asset Price Shocks, Real Expenditures, and Financial Structure: A Multi-Country Analysis, July 2008

2343 Wolfgang Leininger, Evolutionarily Stable Preferences in Contests, July 2008

2344 Hartmut Egger and Udo Kreickemeier, Fairness, Trade, and Inequality, July 2008

2345 Ngo Van Long and Bodhisattva Sengupta, Yardstick Competition, Corruption, and Electoral Incentives, July 2008

2346 Florian Baumann, Employment Protection: The Case of Limited Enforceability, July 2008

2347 Alessandro Balestrino, Cinzia Ciardi and Claudio Mammini, On the Causes and Consequences of Divorce, July 2008

2348 Dirk Schindler and Benjamin Weigert, Insuring Educational Risk: Opportunities versus Income, July 2008

2349 Lammertjan Dam and Ben J. Heijdra, The Environmental and Macroeconomic Effects of Socially Responsible Investment, July 2008

2350 Avner Greif, Contract Enforcement and Institutions among the Maghribi Traders: Refuting Edwards and Ogilvie, July 2008

2351 Helmuth Cremer, Philippe De Donder, Dario Maldonado and Pierre Pestieau, Habit Formation and Labor Supply, July 2008

2352 Francesco Menoncin and Paolo M. Panteghini, The Johansson-Samuelson Theorem in General Equilibrium: A Rebuttal, July 2008

2353 Michael Kaganovich and Itzhak Zilcha, Alternative Social Security Systems and Growth, July 2008

2354 Keith Blackburn, Kyriakos C. Neanidis and M. Emranul Haque, Corruption, Seigniorage and Growth: Theory and Evidence, July 2008

2355 Edward Castronova, A Test of the Law of Demand in a Virtual World: Exploring the Petri Dish Approach to Social Science, July 2008

2356 Harald Badinger and Peter Egger, GM Estimation of Higher-Order Spatial Autoregressive Processes in Cross-Section Models with Heteroskedastic Disturbances, July 2008

2357 Wolfgang Buchholz and Jan Schumacher, Discounting the Long-Distant Future: A Simple Explanation for the Weitzman-Gollier-Puzzle, July 2008

2358 Luca Anderlini, Leonardo Felli and Alessandro Riboni, Statute Law or Case Law?, July 2008

2359 Guglielmo Maria Caporale, Davide Ciferri and Alessandro Girardi, Are the Baltic Countries Ready to Adopt the Euro? A Generalised Purchasing Power Parity Approach, July 2008

2360 Erkki Koskela and Ronnie Schöb, Outsourcing of Unionized Firms and the Impacts of Labour Market Policy Reforms, July 2008

2361 Francisco Alvarez-Cuadrado and Ngo Van Long, A Permanent Income Version of the Relative Income Hypothesis, July 2008

2362 Gabrielle Demange, Robert Fenge and Silke Uebelmesser, Financing Higher Education and Labor Mobility, July 2008

2363 Alessandra Casarico and Alessandro Sommacal, Labor Income Taxation, Human Capital and Growth: The Role of Child Care, August 2008

2364 Antonis Adam, Manthos D. Delis and Pantelis Kammas, Fiscal Decentralization and Public Sector Efficiency: Evidence from OECD Countries, August 2008

2365 Stefan Voigt, The (Economic) Effects of Lay Participation in Courts – A Cross-Country Analysis, August 2008

2366 Tobias König and Andreas Wagener, (Post-)Materialist Attitudes and the Mix of Capital and Labour Taxation, August 2008

2367 Ximing Wu, Andreas Savvides and Thanasis Stengos, The Global Joint Distribution of Income and Health, August 2008

2368 Alejandro Donado and Klaus Wälde, Trade Unions Go Global!, August 2008

2369 Hans Gersbach and Hans Haller, Exit and Power in General Equilibrium, August 2008

2370 Jan P.A.M. Jacobs and Jan-Egbert Sturm, The Information Content of KOF Indicators on Swiss Current Account Data Revisions, August 2008

2371 Oliver Hülsewig, Johannes Mayr and Timo Wollmershäuser, Forecasting Euro Area Real GDP: Optimal Pooling of Information, August 2008

2372 Tigran Poghosyan and Jakob de Haan, Determinants of Cross-Border Bank Acquisitions in Transition Economies: A Latent Class Analysis, August 2008

2373 David Anthoff and Richard S.J. Tol, On International Equity Weights and National Decision Making on Climate Change, August 2008

2374 Florian Englmaier and Arno Schmöller, Reserve Price Formation in Online Auctions, August 2008

2375 Karl Farmer, Birgit Friedl and Andreas Rainer, Effects of Unilateral Climate Policy on Terms of Trade, Capital Accumulation, and Welfare in a World Economy, August 2008

2376 Monika Bütler, Stefan Staubli and Maria Grazia Zito, The Role of the Annuity's Value on the Decision (Not) to Annuitize: Evidence from a Large Policy Change, August 2008

2377 Inmaculada Martínez-Zarzoso, The Impact of Urbanization on CO_2 Emissions: Evidence from Developing Countries, August 2008

2378 Brian Roberson and Dmitriy Kvasov, The Non-Constant-Sum Colonel Blotto Game, August 2008

2379 Ian Dew-Becker, How Much Sunlight Does it Take to Disinfect a Boardroom? A Short History of Executive Compensation Regulation, August 2008

2380 Cécile Aubert, Oliver Falck and Stephan Heblich, Subsidizing National Champions: An Evolutionary Perspective, August 2008

2381 Sebastian Buhai, Miguel Portela, Coen Teulings and Aico van Vuuren, Returns to Tenure or Seniority?, August 2008

2382 Erkki Koskela and Jan König, Flexible Outsourcing, Profit Sharing and Equilibrium Unemployment, August 2008

2383 Torberg Falch and Justina AV Fischer, Does a Generous Welfare State Crowd out Student Achievement? Panel Data Evidence from International Student Tests, September 2008