Education, equality, and efficiency – An analysis of Swedish school reforms during the 1990s

by

Anders Björklund
Per-Anders Edin
Peter Fredriksson
Alan Krueger

January 28, 2004

Table of contents

Preface 5

1 Introduction 7
1.1 The Swedish schooling system around 1990 8
1.2 What happened during the 1990s? 10
1.3 Education and growth 13
1.4 Questions for our report 14
References 16

2 Education, equality and efficiency 18
2.1 Market failures 18
2.1.1 Static externalities 18
2.1.2 Growth externalities 20
2.1.3 Credit constraints 21
2.1.4 Internality: A paternalistic argument for intervention 22
2.1.5 Negative externalities from sorting 22
2.2 Pre-existing distortions 23
2.3 Targeting 24
2.3.1 Means targeted vs. universal programs 24
2.3.2 Targeted for younger vs. older recipients 25
2.4 Concluding remarks 26
References 27

3 Education, skills and earnings – the Swedish record 31
3.1 Inputs 31
3.2 Development of formal schooling 35
3.3 The skills provided by the Swedish schooling system 36
3.3.1 Skills among those still in school 36
3.3.2 Skills in the adult population 37
3.4 Wage inequality and the returns to schooling 42
3.5 The dispersion of wages and skills 46
3.6 Summary and conclusions 47
References 49

4 Resources, decentralization, and student achievement 51
4.1 How important are resources for outcomes? 52
4.1.1 Class size 52

4.1.2 Teacher qualifications ... 56
4.1.3 Computers, IT and new teaching techniques ... 57
4.2 School inputs during the 1990s and early 2000s ... 59
4.2.1 The aggregate development ... 59
4.2.2 The distribution of resources ... 61
4.3 Resources and achievement: Swedish evidence ... 68
4.3.1 The effect of resources on achievement ... 68
4.3.2 Resources and the between school variance in achievement ... 71
4.4 Summary and conclusions ... 73
References ... 74

5 Teacher supply ... 77
5.1 The incentives to become a teacher ... 78
5.2 The supply of skills to the teacher profession ... 83
5.3 Teacher mobility ... 90
5.4 Summary and conclusions ... 95
References ... 97

6 The consequences of school choice ... 99
6.1 The Swedish system in an international perspective ... 100
6.2 International evidence on school competition ... 100
6.3 Is school choice a big deal? ... 105
6.4 The effects of school choice: segregation and costs ... 107
6.4.1 The effect of school choice on segregation ... 107
6.4.2 The effect of school choice on costs ... 109
6.5 The effects of school choice: student achievement ... 111
6.5.1 Previous Swedish studies ... 111
6.5.2 A reanalysis of the data ... 113
6.6 Summary and conclusions ... 118
References ... 120

7 Quantitative tests as an evaluation device ... 122
7.1 Problems and limitations with standardized tests ... 124
7.2 National tests in Sweden ... 126
7.3 The predictive power of the Swedish national tests ... 126
7.4 Conclusions ... 128
References ... 131

8 Family background and earnings: What have education and education policy to do with it? ... 141

8.1 Intergenerational earnings mobility in Sweden, the United States and other countries ... 142
8.2 What have education and education policy to do with it? ... 144
8.2.1 Theory ... 144
8.2.2 Empirical results ... 146
8.3 Lessons from some specific reforms ... 149
8.4 A backsliding during the 1990s? ... 150
8.5 Conclusions ... 152
References ... 154

9 Conclusions ... 156
9.1 Has education policy equalized outcomes and opportunities? ... 156
9.2 Is there an efficiency price for these egalitarian outcomes? ... 158
9.3 How did the development during the 1990s affect equality and efficiency? ... 159
9.4 What is the best use of scarce resources? ... 161
9.5 A call for evaluation ... 164
References ... 169

Preface

This report is an English version of the 2003 Report form the SNS Welfare Policy Group (*Välfärdspolitiska rådet*). This version is slightly more technical than the Swedish version and contains more details on some of the analyses. The initiative to the theme of the report originated from SNS, while the research presented here to a large extent represents ongoing research projects of the individual authors.

We are all empirical labor economists. Therefore, we are inclined to focus on the measurable aspects of what schools produce, in part because making sense of the data is our expertise. We also firmly believe that statistical evaluation of school performance can help guide policy and help to quantify the precision or imprecision of our knowledge of specific school reforms. Hence, outcomes such as student achievement as measured by the results on standardized tests will figure prominently in this report. This quantitative focus is not driven by the belief that outcomes that are more difficult to measure – such as democratic values – are unimportant. Nevertheless, we think that student knowledge is mainly what schools should produce. One may, of course, quibble about whether test results accurately measure "knowledge". Still, they clearly have some informative value. The opposite position is easily falsified by the fact that adult labor market success is predicted by the performance on standardized tests in primary school.

This report is a joint product. But since we are economists we believe in the virtues of specialization. Therefore, the work in chapters 2-8 has been divided among the four of us. The division has been as follows: Anders Björklund has been primarily responsible for the material in chapters 7 and 8; Per-Anders Edin for the material in chapter 3; Peter Fredriksson has been mainly responsible for the analysis in chapters 4, 5, and 6; Alan Krueger, finally, has written chapter 2 and contributed with the reviews of the international literature in chapters 4 and 6.

The writing of the report has been funded by The Swedish Council for Working Life and Social Research (FAS), The Swedish National Agency for School Improvement, and firms and agencies within the SNS-group (see the Swedish version for a complete list). The underlying research projects are funded by The Swedish Council for Working Life and Social Research (FAS), The Institute for Labour Market Policy Evaluation (IFAU), and Jan Wallanders and Tom Hedelius Research Foundation. We are grateful for useful comments

from Göran Arvidsson, Robert Erikson, Bertil Holmlund, Stephen Machin, Erik Mellander, Roope Uusitalo, Björn Öckert, Thomas Östros, and our reference group at SNS. We thank Lalaina Hirvonen, Louise Kennerberg, Caroline Runeson, and Krister Sund for expert research assistance.

1 Introduction

The Swedish school system is in crises. That is often the impression one has gotten from Swedish media reports about education policy in the last 10-15 years. During the same period the Swedish school system has undergone rapid changes. This notion of a school system in severe crises is the starting point for this report. We will examine Swedish education policy with special focus on the turbulent reform period during the 1990s. We focus on two very established goals for education policy: equality and efficiency.

Ever since the early introduction of the compulsory school in 1842, egalitarian goals have been important ones in Swedish education policy. A reading of a more recent policy document by the present government – see Regeringens skrivelse 2001/02: 188 – reveals that even the egalitarian goals for education policy are multifaceted. In general, though, one can distinguish between two main egalitarian goals, namely equality of outcomes and equality of opportunity.

The ambition to influence the distribution of outcomes, e.g., cognitive skills, has been evident in many ways. First of all, the compulsory schooling system has gradually become more comprehensive so that all pupils are kept together in one class with a similar curriculum. Further, extra resources have been allocated to pupils with special needs, such as handicapped pupils and immigrant children.

The ambition to equalize opportunities has generally been interpreted as an ambition to reduce the importance of pupils' family background for their subsequent educational attainment. In fact, in their thorough report to government Erikson and Jonsson (1993) note that politicians have also stressed efficiency arguments in favor of a policy that may weaken the direct link between family background and educational attainment. One popular expression has been to "mobilize the reserve of talents" among children with a family background without educational tradition. Economists would rather talk about policies that eliminate "credit constraints" that low-income families face when their children are contemplating longer education.

Through the 1980s, Sweden appeared to have been quite successful in terms of overall economic equality. At least according to readily available measures like hourly wages and annual disposable household income, Sweden generally

ranked high in cross-country comparisons of equality.[1] Comparisons of equality based on long-run measures of earnings and income are more complicated, but the available evidence suggests the same cross-country patterns as found in point-in-time data.[2] However, education policy's contribution to these egalitarian outcomes is more of an unsettled issue.

The period since 1990 has been turbulent in many respects. At the macro-economic level, the decade started with the most severe economic downturn since the 1930s. Unemployment rose from two percent to almost ten percent in just three years. As a consequence, public budgets deteriorated when tax revenue fell and expenditures to support the unemployed rose. Before the end of the decade, these deficits had been eliminated, partly by reductions of expenditures that also affected Swedish schools.

During the 1990s, Sweden also implemented "the tax reform of the century"[3], entered the European Union, settled successfully for low inflation, and deregulated many markets. Further, both hourly earnings and disposable income inequality started to rise.

The 1990s was a turbulent decade for Swedish education policy too. Although, some policy initiatives, like an expansion of adult "second-chance" education, were motivated by traditional egalitarian arguments, many changes occurred that were unexpected for those who followed the Swedish discussion during previous decades. Education policy is our focus in this report. Before we summarize the reforms in Swedish education policy during the 1990s, we offer a brief description of the Swedish schooling system around 1990.

1.1 The Swedish schooling system around 1990

Around 1990 Sweden had an extensive public daycare system for ages 1-6. Public daycare was heavily subsidized; the fees covered some 15-20 percent of the average cost per child. The public daycare system had been built up quite quickly from the late 1960s. In 1989, the capacity was around 50 slots per 100 children aged 0-6.

[1] See e.g. Björklund and Freeman (1997) and Edin and Topel (1997) in the SNS-NBER study of the Swedish economy.
[2] See Aaberge et al. (2002).
[3] This was a major tax reform that, among other things, lowered the highest marginal income tax rate from over 80 to 50 percent.

Since the mid 1960s until today, Sweden has had nine years of tuition-free compulsory education starting at age 7. It is a comprehensive education, so all children follow basically the same curriculum, which is determined by the central government.

Through 1990, upper-secondary school was voluntary and offered several programs, ranging from vocational training to programs that prepared for further studies at the university level. At this time, some 80 percent of a cohort continued from compulsory school to any of the study tracks at the upper-secondary level.

The municipalities have for a long time run daycare, primary, and secondary education. Nonetheless, the system was a strongly centralized one through the 1980s. The central government decided the basic goals, the curriculum, and provided earmarked money for schools. A national body was responsible for evaluation of the schools. So there was little leeway for the individual municipality to deviate from the national standard, although the rules allowed the municipality to "top up" the resources with local funding.

Swedish youth could typically apply for university at age 19 after having completed three years of high school. Swedish universities are, with a few exceptions, public, and run by a central agency. They were not allowed to charge tuition. All students, who were admitted to a university or college and completed their courses at an acceptable speed, were eligible for subsidized student loans and a smaller stipend (around $300 per month). An important purpose of the public financial support system has been to eliminate any credit constraints that prospective students might face. Nonetheless, the financial support system has been universal, so even students from wealthy families have been eligible for the loans and the stipend. This rule reflects the preference for universalism in Swedish education and social policy.

By tradition, prospective university students apply for both a specific university and a specific field of study. So the applicant must choose between fields of study like law, medicine, engineering, business administration and social work etc. Whether this makes young people wait for a few years before they move on from high school to university is hard to say. Anyhow, Swedish university students are relatively old by international standards. The admittance rules also to some extent favor older applicants.

1.2 What happened during the 1990s?

A number of quite radical changes of the Swedish schooling system took place during the 1990s. The changes in governance of the system are probably the most interesting ones, since they in many respects represent a radical ideological change. But in addition several important changes took place as regards enrolment at different levels of the system and the resources allocated to different types of education. Some of these changes were motivated by more traditional egalitarian arguments. Still, they raise many important issues about tradeoffs in education policy.

Governance

The changes in governance of Swedish schools during the 1990s can be described by key words like "decentralization", "goal steering", "accountability", "parental choice" and "competition". In this respect, it seems like Sweden has followed the same route as many other OECD countries. Indeed, Levin (1998) talks about a "policy epidemic" in the OECD countries. Nonetheless, the fast and radical restructuring of Swedish education during the 1990s seems to have changed the system to one of the most decentralized ones in OECD (Lindblad et al. 2002).

A major step towards decentralization was taken in 1990 when the authority to run primary and secondary education was transferred to the municipalities. After this reform, the municipalities got full financial responsibility for the schools offering such education. Although the central government continued with a system that redistributed financial resources from rich to poor municipalities, the ear marked money for schools gradually disappeared and had been eliminated completely by 1993. Thus the scope for differences in educational expenditure across municipalities increased.

The wage-setting for teachers was also decentralized in the mid 1990s. Since then, school-level factors may affect wages to a greater extent. Many school managers have used this opportunity to more "individual" wage setting of teachers. During the decade a long-predicted shortage of trained teachers developed, and schools that wanted licensed teachers needed to offer higher starting wages to fill their vacancies. This development came after an initial wage increase for teachers in 1990. The latter increase was a visible price that the government paid to make teachers' unions accept the decentralization of primary and secondary education to the municipalities.

School choice was introduced in 1992. The reform requires municipalities to satisfy parental preferences regarding school choices subject to space limitations. However, residing close to a school (the residence principle, *närhetsprincipen*) is the main principle for allocating students to schools. So if students residing close to a particular school fill the available slots, the other parents' preferences vis-à-vis this school are given little weight. In 2000, however, Stockholm city introduced a major deviation from the residence principle. For upper-secondary education, they then introduced a system where admission is based on student achievement (i.e., compulsory school grades) only.

Also in 1992, municipalities were required to fund independent, privately run, schools. In 2002, almost six percent of students at primary and lower-secondary level attend a private school, a sharp rise from less than one percent around 1990. There is complete freedom to choose between a private and public school provided that a private alternative exists. Private schools exist in about half of the municipalities and especially in big city areas.

A parallel development to decentralization and the introduction of school choice, was to rely more on goal steering than before. The basic principle has been school-based self-evaluation. The schools thus decide on how to evaluate, and whether they fulfilled, the general goals or not. As a guide for this self-evaluation, the schools can use the national tests that were made available by the National Agency of Education in grades 5 and 9. The actual use of these tests has varied among schools. Some schools and municipalities have started to publish league tables based on these tests or grade point averages, others have not.

It is natural to ask what motivated these rather radical policy changes. We are not political scientists – not even political economists – so we do not claim any expertise in explaining political behavior. But there is no doubt that these changes took place during a period of much critique of public schools in the media; see, e.g., Bergström (1998). The famous Lindbeck commission (1993), which delivered its report in the midst of the deep economic downturn in 1993, also had a critical tone about Swedish schools. Among others things, they claimed that more resources for smaller classes would not improve school achievement. So they recommended more home work and larger classes.

Resources

The financial crisis in the public sector had consequences also for the resource development in daycare and schools. Starting with daycare, the child-staff ratios have increased markedly since the late 1980s. During most of the 1990s, fees were also raised considerably. However, the fees were drastically reduced in 2002 when a decision by the parliament in practice enforced a ceiling on the fees. The magnitude of the reduction varied among municipalities depending on the previous structure, but proportionate reductions above 50 percent were very common. By 2001 the fees were had become quite high, so the reduction implied a big absolute improvement in living standards for families with small children.

The pupil-teacher ratio increased in compulsory schools as well. After having decreased gradually over a long period of time, the ratio increased from close to 11 in 1991 to over 13 in 1997. Expenditure per pupil in compulsory schools fell quite markedly from 1990 to 1995, followed by a slight recovery.

At the same time as resources per pupil declined, the 1990s was the decade when computers were introduced in Swedish schools at a large scale. Teaching techniques changed as well, partly because the computers offered new opportunities in this respect. In many schools, the traditional concept of a "school class" lost its meaning due to new teaching styles. It is fair to say that there was no overall consensus in the school community behind these changes. Thus, these changes contributed to the turbulence in Swedish schools during the decade.

School enrolment

By 1990 public daycare slots more or less met the demand so the expansion of slots tapered off during the decade. One change was that 6-year olds were brought closer to primary school so in practice Sweden got a Kindergarten-type of school start at age six.

School enrolment rates did not change in compulsory school. However, some important changes took place in upper-secondary school. In 1991 a reform added one year of mainly theoretical studies to vocational programs. A completed curriculum at a vocational program now implies that the student fulfils the so-called general requirement for entering university studies. This in turns means that some of the university programs are available for these students. The enrolment rate at upper-secondary school increased during 1990s

and in the second part of the decade around 95 percent of a cohort participated in such studies.

During the second half of the 1980s, university enrolment rates had started to increase somewhat after having fallen radically in the early 1970s, see Fredriksson (1997). When the youth labor market deteriorated in the beginning of the 1990s, enrolment rates increased rapidly. This expansion could not have taken place without political decisions to expand the number of slots at the public universities. Indeed, one motivation for this expansion was that university education was considered a much better alternative than being unemployed or participating in labor market programs for the unemployed. Much of the expansion took place at new regional colleges rather than at the older and more established universities. One political motivation for this change would facilitate the recruitment of new students to higher education, especially students with working class background.

Other types of adult education also expanded in the 1990s. During 1990-93, when employment fell rapidly, labor market training (typically with a big class-room training component) became the most common labor market program. By the mid-1990s, this measure had to a large extent been replaced by more work-related programs. Such temporary programs were inter alia used to help unemployed persons to renew their benefit entitlement when the typical period of 60 weeks with UI-benefits expired.

During the second half of the decade, adult education expanded enormously as a consequence of the so-called adult education initiative (*kunskapslyftet*). This program gave unemployed individuals the opportunity to keep their unemployment benefits while taking part in education at the primary and secondary levels. In contrast to the short courses provided in labor market training, this new initiative helped low-skilled persons to raise their overall education level and eventually get qualified for university studies.

1.3 Education and growth

So far we have emphasized the equality dimension of education policy. But Swedish politics is not so possessed by equality that growth and efficiency considerations have not been part pf the discussion about education policy. Indeed, Sweden's growth performance in a cross-national perspective has been

discussed quite intensively during the last 10-15 years.[4] In particular, the discussion has focused on the interpretation of Sweden's drop in the GDP per capita league from a top position around 3-4 in 1970 to a more mediocre rank around 15-18 in 1995.

Much of the growth discussion has centred on the classical issue whether high taxes and public spending hinder economic growth. But education policy has also been part of this discussion. The critics of Swedish education have mainly focused on the low estimated earnings returns to post-compulsory education.[5] These numbers were particularly low in the early 1980s when both wage compression and high marginal taxes contributed to low private after-tax return to additional schooling. Some data also suggested that Sweden's labor force was poorly educated compared to other countries[6]. Swedish manpower training was also heavily criticized since most evaluation studies from the 1990s suggested that this program had poor results.

The defenders of Swedish education policy got some comfort from the International Adult Literacy Study that was published in the mid 1990s. The results from this innovative study showed that Swedish adults did very well in terms of the literacy and numeracy skills that were tested in this study. Not only did Swedes perform well on average, but the lower tail of the distribution did remarkably well in a cross-national perspective; see e.g. Nickell and Layard (1999) and Björklund et al. (1998).

1.4 Questions for our report

Needless to say, Swedish education policy through the 1980s and the changes since then raise a large number of interesting issues that we would like to address. But space and time prevent us from discussing all of them.

We start out in chapter 2 with a discussion about the basic theoretical arguments in favor of a public education policy and how these arguments are related to efficiency and equality considerations. In chapter 3 we examine some basic empirical evidence about Swedish education policy. We report how Swedish pupils have achieved in international comparisons and how the skills

[4] See ,e.g., the interchange between Korpi (2000) och Henrekson (2001). See also Lindbeck (1997) and Freeman, Topel and Swedenborg (1997).
[5] See, e.g., Andersen et al. (1997).
[6] As shown in Edin et al. (1994) this is a fallacy, however.

of the Swedish adult labor force compare to those of adults in other countries. We also report estimates of the private returns to schooling.

The next three chapters focus on the experience of the 1990s. Chapter 4 shows that the decentralization reform changed the allocation of school resources among Swedish municipalities. This change in turn is exploited by us to study the impact of changes in school resources on pupils' achievement. In chapter 5 we examine the supply of teachers to Swedish schools. We emphasize that many Swedish teachers will retire in the next decade and that the incentives to become a teacher have deteriorated over time. Then, in chapter 6, we investigate whether the competition induced new independent schools in combination with free school choice has improved productivity by raising achievement in all schools.

A system with free school choice requires that parents have good information about the quality of schools. In chapter 7 we discuss what role quantitative measures like grades and test scores can play to guide parents in their school choice. We also examine how well grades and test results in compulsory school predict outcomes at adult age like final education and earnings.

In chapter 8 we examine whether Swedish ambition to equalize educational and labor market performance among individuals from different family backgrounds. In particular, we investigate whether the 1990s was a backsliding in this regard.

Finally, we summarize our main findings in chapter 9 where we also make suggestions about future evaluations of Swedish education policy.

References

Aaberge, R., A. Björklund, M. Jäntti, M. Palme, P.J. Pedersen, N. Smith, and T. Wennemo (2002), Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States, *Review of Income and Wealth,* 48, 443-469.

Bergström, H. (1998), Om konsten att lyfta den svenska skolan, artiklar och föredrag om tillståndet i och utvecklingen av den svenska skolan, Bertil Ohlin institutet.

Björklund, A. and R. Freeman (1997), Generating Equality and Eliminating Poverty, the Swedish Way, in R. Freeman, R. Topel, and B. Swedenborg (eds.), *The Welfare State in Transition*, NBER and the University of Chicago Press.

Björklund, A. P-G. Edebalk, R. Ohlsson, and L. Söderström (1998), *Välfärdspolitik i kristid – håller arbetslinjen?* Välfärdspolitiska rådets rapport 1998, SNS Stockholm.

Edin, P-A., P. Fredriksson, and B. Holmlund (1994), Utbildningsnivå och utbildningsavkastning i Sverige, in A. Björklund and A. Forslund (eds.), *Studier av svensk utbildning*, Ekonomiska Rådets årsbok 1993, Fritzes, Stockholm.

Edin P-A. and R. Topel, Wage Policy and Restructuring: The Swedish Labor Market since 1960, in R. Freeman, R. Topel, and B. Swedenborg (eds.), *The Welfare State in Transition*, NBER and the University of Chicago Press.

Erikson, R. and J. Jonsson (1993), *Ursprung och utbildning*, SOU 1993:85.

Fredriksson, P. (1997), Economic Incentives and the Demand for Higher Education, *Scandinavian Journal of Economics*, 99, 129-142.

Freeman, R., R. Topel, and B. Swedenborg (1997), *The Welfare State in Transition*, NBER and the University of Chicago Press.

Gylfason, T. et al. (1997), *I otakt med omvärlden?* Konjunkturrådetes rapport 1997, SNS Stockholm.

Henrekson, M. (2001), Swedish Economic Growth: A Favorable View of Reform, *Challenge*, 44, 38-58.

Korpi, W. (2000), Welfare States, Economic Growth, and Scholarly Objectivity, *Challenge*, 43, 49-66

Levin, B. (1998), An Epidemic of Education policy: (what) can we learn from Each Other?, *Comparative Education*, 34, 131-141.

Lindbeck, A. (1997), The Swedish Experiment, *Journal of Economic Literature*, 35, 1273-1319.

Lindbeck, A., T. Persson, A. Sandmo, and B. Swedenborg (1993), *Nya villkor för ekonomi och politik*, SOU 1993:16.

Lindblad, S., L. Lundahl, J. Lindgren, and G. Zackari (2002), Educating for the New Sweden, *Scandinavian Journal of Educational Research*, 46, 283-303.

Nickell S. and R. Layard (1999), Labor Market Institutions and Economic Performance, in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics* 3C, North-Holland, Amsterdam.

2 Education, equality and efficiency

In his classic 1975 book, *Equality and Efficiency: The Big Tradeoff*, Arthur Okun (p. 81) noted that, "A vigorous social effort to narrow the educational financing gap can improve both equality and efficiency." Okun's concern was that unequal access to credit restricts many talented children from poor families from attending college in the United States, which decreases both economic efficiency and equality. This market failure, he argued, created a win-win situation where there was no tradeoff between efficiency and equality. In these situations, Okun emphasized that society should make the most of interventions to enhance equity and efficiency. In this chapter, we will review and evaluate several explanations that have been put forward to explain why countries may want to subsidize human capital accumulation to improve efficiency. The particular tax policies and social objectives of Sweden would make educational policy all the more important in this regard. Moreover, even if the rationales discussed in this chapter do not create a win-win situation like Okun envisioned, it is nonetheless the case that they are likely to produce a smaller tradeoff between equality and efficiency for appropriately selected human capital policies than for many other forms of redistribution. And in any event improving the efficiency of education is in society's interest even if all of the benefits of education are private benefits that do not entail externalities.

Human capital policy can influence the distribution of income in two important ways. First, a targeted human capital policy can increase access to higher education and higher quality education for children from lower income families, and thus raise their economic opportunities and aspirations. A second, and more subtle effect, is that if human capital policy results in there being fewer unskilled workers, then the economic circumstances of the remaining unskilled workers would be improved because they would face less (domestic) competition.

2.1 Market failures

2.1.1 Static externalities

It is well known that if there are externalities from education, individuals will invest less than the optimal amount in their human capital absent government subsidies. We divide potential externalities into two types: those that lead to a

one-time increase in welfare, which we call static externalities, and those that affect the growth trajectory, which we call growth externalities.

We begin by reviewing arguments and evidence on possible external effects from education involving civic participation, crime, unemployment, participation in the political process, and health, and then consider growth externalities in the next section.

(1) More educated voters make the democratic process work better. First, people with more education are more likely to be informed and more likely to participate in democracy. Second, more informed citizens are likely – though certainly not guaranteed – to make better decisions. No less a devotee of free markets than Milton Friedman (1982) cited just such an externality: "A stable and democratic society is impossible without a minimum degree of literacy and knowledge on the part of most citizens and without widespread acceptance of some common set of values." For this reason, Friedman supported compulsory minimum schooling ages.

(2) Available evidence suggests a link between crime, education and inequality (e.g., Ehrlich, 1973, Freeman, 1983 and 1995, and Imrohoroglu et al., 2001). Other things being equal, the incentive for those with limited market opportunities to commit property crimes rises as inequality increases. From the criminal's perspective, the potential gain from crime is higher if inequality is higher, and the opportunity cost is lower. Society can devote more resources to crime prevention and incarceration, or to reducing inequality. Education raises the market opportunities of potential criminals, and is therefore expected to reduce crime.

(3) A wealth of studies show that unemployment and education are negatively associated (e.g., Ashenfelter and Ham, 1979). It is not obvious that the link is causal, but there is also experimental US evidence suggesting that education has a negative effect on unemployment (Eberwein et al., 1997). Therefore, there are externalities associated with education that work through the public budget: higher education increases tax revenues and reduces unemployment expenditure. The externality arises since individuals do not take these effects into account in their job search and job acceptance decision (e.g., Fredriksson and Holmlund, 2001).

(4) Negative externalities can arise from the policy domain from an unequal distribution of income and education. Benabou (2000), for example, develops a model in which the progressivity of educational funding and taxation is endogenous. He shows that the political influence of the wealthy interacts with income inequality to block efficient progressive policies, or impose inefficient regressive ones. When inequality is high, the wealthy are more likely to block *efficiency enhancing* programs that would improve educational opportunities for the less well off.

(5) Education has been linked to improvements in health. Although some authors have suggested that the correlation between health habits and education is spurious (e.g., Fuchs, 1982, argues that differential discount rates affect both educational attainment and smoking), other research suggests a causal impact running from education to health (e.g., Lleras-Muney, 2002). In a country with publicly provided health care, education will therefore generate external benefits.

2.1.2 Growth externalities

Nelson and Phelps (1966) and Romer (1990) model the level of education as generating positive externalities for economic growth. Romer argues for more government subsidies for scientists and engineers. The available empirical support for these model is mixed (see Krueger and Lindahl, 2001; Heckman and Klenow, 1998; Acemoglu and Angrist, 2000; and Bishop, 1996), however.

Persson and Tabellini (1994) develop a model of economic growth in which inequality negatively influences growth through the political process. In their model, high initial inequality leads to political demands for redistribution which, in turn, is detrimental for growth. A growing body of cross-country and cross-state studies has estimated the relationship between initial inequality and subsequent GDP growth.[7] Although attributing causality is difficult in these studies, the correlation between inequality and growth is negative, conditional on variables like initial GDP per capita and average education. Two-stage least squares estimates that instrument for inequality with variables such as initial literacy and infant mortality also show an inverse relationship between GDP growth and inequality.

[7] Early papers are Persson and Tabellini (1994) and Alesina and Rodrik (1994). See Benabou (1996) for a survey of the empirical and theoretical literatures.

2.1.3 Credit constraints

In a world with perfect credit markets, children from all families would invest in educational resources up to the point that their marginal return equals their discount rate, and all families would have equal access to credit. The evidence suggests that education decisions are not made in such a world, however. Children from poor families *behave* as if they have higher discount rates. The most plausible explanations for this phenomenon are that poor families are credit constrained (i.e., cannot borrow at the same rate as everyone else), or that they discount future benefits of human capital investments at a greater than market rate because they are impatient, have a greater disutility of schooling, or fail to appreciate the benefits of education. Credit constraints have received the most attention in the literature because students cannot easily use the return on their future human capital as collateral. This may be a reason for discount rates to vary. Poor families face different borrowing costs than rich ones. (See Carneiro and Heckman, 2002, for a skeptical view of the role credit constraints should play in human capital policy.)

The following five observations are consistent with the view that low-income families face credit constraints when it comes to education. First, Behrman and Taubman (1990) find that the timing of parental income matters for children's educational attainment. Using data from the PSID, they find that father's income earned when children are teenagers has a stronger effect on children's educational attainment than income earned later on. Second, Shea (2000) looks at the effect on children's human capital of differences in parental income emanating from noncompetitive factors, such as employment in a high-paying union job or industry. Wage differences for these reasons arguably are independent of parents' ability. He finds that family income matters for children's human capital investment in a sample of low-income families, but not for the broader population. He concludes that these findings are "consistent with models in which credit market imperfections constrain low income households to make suboptimal investments in their children." Third, Ellwood and Kane (2000) find that when the return to college education increased in the 1980s, four-year college enrollment increased for children from all quartiles of the income distribution, except the bottom one. Fourth, results surveyed in chapter 8 indicate stronger family income effects on children's outcomes in the United States than in Sweden. Although much of this correlation is due to lower private returns to education in Sweden than in the United States, some is

probably due to the fact that Sweden provides much more generous educational subsidies than the United Sates, so credit constraints are less of an issue for low-income families in Sweden. Fifth, the reaction of college enrollment to changes in tuition, especially at the two-year-college level, are substantially larger than the reaction of college enrollment to equivalent, present-value changes in the payoff to education (see Kane, 1999).

2.1.4 Internality: A paternalistic argument for intervention

It is common to assume that individuals make educational decisions by implicitly comparing the costs and benefits associated with further investment. If the marginal benefits of further study exceed the marginal costs, then an individual would invest the time and money in further study. As is standard in economics, this assumes rational decision making. But a great deal of evidence suggests that individuals, and youth in particular, do not always make decisions that are in their best interest or rational. For example, individuals often tend to discount future benefits at an irrationally high interest rate (see Warner and Pleeter, 2001). That youth are particularly prone to impatience, impulsiveness, and irrational risk taking is not surprising, and this was even commented on by Adam Smith in *The Wealth of Nations*: "The contempt of risk and the presumptuous hope of success are in no period of life more active than at the age at which young people chose their professions."

Situations like these give rise to what Matthew Rabin (1998) in another context has called an *internality*. It may be in students' best interest to stay in school longer, but because of short sightedness they drop out. Like externalities, internalities cause the economy to operate at less than peak efficiency. The loss to society due to internalities could be quite large. For cigarette smoking, for example, Gruber and Koszegi (2001) contend that the societal loss due internalities greatly exceeds the loss due to externalities (e.g., second-hand smoke). We suspect a major reason for public intervention in the education field involves the implicit belief that internalities are important. Indeed, because education is expected to improve individuals' decision making, one can argue that improving education reduces internalities in a number of domains.

2.1.5 Negative externalities from sorting

We would be remiss if we did not mention that education can produce negative static externalities as well as positive ones. In particular, if education served

only as a sorting mechanism that did not enhance individuals' abilities, then encouraging education or requiring certification for certain jobs could result in wasteful investment (Spence, 1974). For example, if education only serves to sort individuals by their inherent ability, and the minimum school leaving age is increased, then other individuals will ratchet up their schooling attainment to distinguish themselves from those with lower ability. This additional schooling is costly and does not increase productivity (because it was assumed that education only serves to sort individuals, not raise productive abilities). Alison Wolf (2002) argues that this is the case in England; see Card (2002) for a critique.

In practice, the importance of the sorting externality has been very difficult to assess. Many implications of the sorting model are similar to a human capital model, so it has proved difficult to distinguish between the two models, and it is likely that education serves both to enhance ability and to sort individuals. In principle, the strongest evidence is from international comparisons that look at how differences in educational attainment across countries relate to GDP, or how increases in education over long periods of time relate to GDP growth. Cross-country evidence is always difficult to interpret, however, because there are only a relatively small number of countries and many potential influences on GDP, and because there are difficulties measuring educational attainment in many countries. Nonetheless, we interpret the bulk of the evidence as indicating that increases in education are associated with higher living standards because education raises individuals' productive capacities and generates more positive than negative externalities (see Krueger and Lindahl, 2001, Cohen and Soto, 2001, and Heckman and Klenow, 1998; for a different view, see Benhabib and Spiegel, 1994.)

2.2 Pre-existing distortions

Wage compression due to union policies and high marginal tax rates caused by a progressive income tax would create incentives for suboptimal private investment in education and training.[8] The reason is that part of the return to

[8] In the special case of education raising the value of leisure as much as it raises a worker's productivity, then taxes would not distort educational investment. Additionally, if there is a flat tax and direct educational expenses are deductible (or negligible), the educational investment decision is not distorted; see Jacobs (2000) and Bevia and Iturbe-Orxmaetxe (2002) for a discussion of these and related issues.

investment is taxed away, so investment decisions are distorted. Edin and Holmlund (1995) calculate that the after-tax (pre-subsidy) internal rate of return from completing a university education as opposed to leaving school after the upper secondary level is low by international standards, only 6.6 percent. In this "second best" world (that is, second best compared to an idealized, frictionless economy without distortions), human capital policy can help improve efficiency.

Subsidized education is one way to provide optimal incentives for human capital acquisition while still maintaining compressed wages and progressive taxes to meet other social objectives. Indeed, Edin and Holmlund calculate that, after taking account of the subsidy for university attendance, the after-tax internal rate of return from completing university education in Sweden is 11 percent, about as high as the pre-tax return in the United States.

In addition, because of generous social welfare benefits for those who are unemployed or employed at low earnings, raising the earnings and employability of otherwise low-income workers has a particularly high payoff to society in Sweden. This creates another kind of externality that is likely to be substantially larger in Sweden than in most other countries.

2.3 Targeting

Another issue concerns whether subsidies should be targeted or universal. Targeted subsidies vary with the family background or age of the recipient, while universal subsidies are available to all.

2.3.1 Means targeted vs. universal programs

If the rationale for intervention in the education market is that there are credit constraints mainly on the poor, then the case for means tested subsidies is stronger.[9] Means testing also increases the redistributive effect of education.

On the other hand, means tested subsidies provide a disincentive for wealth or income accumulation; in essence, subsidies that are greater for lower income families act like a tax. A literature in the US suggests that families respond to wealth-related college tuition grants by adjusting their asset accumulation behavior (e.g., Feldstein, 1995). Combined with high existing marginal tax rates, this is an argument for universal programs, such as are common in

[9] Means-tested subsidies are subsidies that decline with the income or wealth of an individual.

Sweden. Moreover, the compressed after-tax distribution of income in Sweden is a reason why credit constraints among the poor are likely to be less severe in Sweden than in most other countries. Participation in means tested programs could also carry a stigma, and may even be a negative signal. And yet another argument for universal programs is that the political support for the program would be deeper.

The fact that the dispersion in achievement test scores for adults is lower in Sweden than in other countries, combined with the fact that the mean level is comparatively high, suggests that human capital policy historically has been successful at raising the achievement at the bottom of the distribution, by overcoming credit constraints and lower educational aspirations among the poor.[10]

2.3.2 Targeted for younger vs. older recipients

One issue involves the returns: are they higher for the young or old. The usual argument among labor economists is that people should invest in education while they are young, because the opportunity cost of education is lower and because they have a longer period (i.e., remaining work years) over which to amortize the costs of education. If there are liquidity constraints or if the return to education unexpectedly increases, then it may make sense for some workers to make their investment at older ages. (The reason why liquidity constraints could have this effect is that individuals would have to work while they are young, to some extent, in order to finance their education later on.)

Some interpret the available evidence as indicating that human capital programs that invest in disadvantaged young children have the highest return because there are critical points for learning, and it is too late for older disadvantaged children (Heckman, 2000). Randomized evaluations of the Perry Preschool Program and ABCEDARIAN program do indeed find high payoffs to investment in preschool education for disadvantaged, primarily African American children. Jacobson et al. (2003b), however, provide evidence that "you can teach an old dog new tricks": they find that returns are sizable for older workers, at least for a subset of workers who opt for training.

In another study, Jacobson et al. (2003a) find that returns are very low for older displaced workers in some fields of study, such as history, and reasonably high in others, such as nursing assistance. One inference from this line of

[10] In chapter 3 we summarize the evidence on the skills in the adult population in more detail.

research is that some forms of education may only be of consumption value, rather than of investment value, to older individuals who return to school.

Jacobson et al. (2003b) also make the important point that even if there are critical stages of development, *on the margin*, returns could be as high or higher for older workers because of declining marginal returns to investment. Therefore, the comparative payoff to targeted investments at different ages will depend upon the context, the particular training programs, the course of study, and the amount of investment to that point. Generalizing from evidence on US adult training programs, even though it is often based on randomized field experiments, is also difficult because the US has such a large disadvantaged population with low skills that may have had inadequate human capital investments at critical junctures. Finding the right allocation at which the return from marginal investment in the education of older and younger workers is equalized is a difficult challenge. In the next section, we will consider evidence on second chance and adult education programs in Sweden.

2.4 Concluding remarks

In this chapter we have discussed alternative arguments for public investments in education. The arguments for public intervention are, for instance, based on (various) positive externalities associated with education and capital market imperfections. In addition, education policy may correct existing distortions. Moreover, we have made the point that there is probably a smaller trade-off between equality and efficiency when it comes to human capital policy than for many other forms of redistribution. For instance, there may also be an efficiency case for policies targeted at disadvantaged groups.

Note that what is relevant is how human capital policy affects equality and efficiency *on the margin*. While hardly anyone would argue against requiring citizens to achieve a minimum level of schooling in an economically advanced democracy, there would be little support for setting that minimum level at the university graduation level or higher.

One must recognize that all advanced countries already provide a great deal of public subsidy for education. Nevertheless, the equality and efficiency arguments outlined in this chapter underscore the importance of monitoring and evaluating the education system to make sure that it is performing efficiently. If the education system can be made to work better, then both equality and efficiency can be improved.

References

Acemoglu, D. and J. Angrist (2000), How Large are the Social Returns to Education? Evidence from Compulsory Attendance Laws, *NBER Macro Annual*, 15.

Alesina, A. and D. Rodrik (1994), Distributive Policies and Economic Growth, *Quarterly Journal of Economics*, 109, 465-490.

Ashenfelter, O and J. Ham (1979), Education, Unemployment, and Earnings, *Journal of Political Economy*, 87, S99-S116.

Behrman, J. and P. Taubman (1990), The Intergenerational Correlation Between Children's Adult Earnings and Their Parent's Income: Results from the Michigan Panel Survey of Income Dynamics. *Review of Income and Wealth*, 36, 115-127.

Benabou, R. (1996), Inequality and Growth, *NBER Macroeconomics Annual 1996,* MIT Press, Cambridge, MA, 11-74.

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

Benhabib, J. and M. Spiegel (1994), The Role of Human Capital and Political Instability in Economic Development, *Journal of Monetary Economics*, 34, 143-173.

Bevia, C. and I. Iturbe-Orxmaetxe (2002), Redistribution and Subsidies for Higher Education, *Scandinavian Journal of Economics*, 104, 321-340.

Bishop, J. (1996), Signalling Incentives and School Organization in France, the Netherlands, Britain and the United States: Lessons for Education Economics, in E. Hanushek (ed.) *Incentives and Education Reform*, National Academy Press, Washington DC.

Card, D. (2003), Comment on Alison Wolf, *Milken Institute Review.*

Carneiro, P. and J. Heckman (2002), The Evidence on Credit Constraints in Post-secondary Schooling, *Economic Journal*, 112, 705-734.

Cohen, D. and M. Soto (2001), Growth and Human Capital: Good Data, Good Results, CEPR Discussion Paper no. 3025.

Eberwein, C., J. Ham and R. LaLonde (1997), The Impact of Being Offered and Receiving Classroom Training on the Employment Histories of Disadvantaged Women: Evidence from Experimental Data, *Review of Economic Studies*, 64, 655-682.

Edin, P-A. and B. Holmlund (1995), The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy? in R. Freeman and L. Katz (eds.) *Differences and Changes in Wage Structures*, University of Chicago Press, Chicago, 307-344.

Ehrlich, I. (1973). Participation in Illegitimate Activities: A Theoretical and Empirical Investigation, *Journal of Political Economy*, 81, 521-565.

Ellwood, D. and T. Kane (2000), Who is Getting a College Education? Family Background and the Growing Gaps in Enrollment, in S. Danziger and J. Waldfogel (eds.), *Securing the Future*, Russell Sage, New York.

Feldstein, M. (1995), College Scholarship Rules and Private Saving, *American Economic Review*, 85, 552-566.

Fredriksson, P. and B. Holmlund (2001), Optimal Unemployment Insurance in Search Equilibrium, *Journal of Labor Economics*, 19, 370-399.

Freeman, R. (1983), Crime and the Labor Market, in J. Wilson (ed.), *Crime and Public Policy,* Institute for Contemporary Studies, San Francisco, 89-106.

Freeman, R. (1995), The Labor Market, in J. Wilson and J. Petersilia (eds.), *Crime*, Institute for Contemporary Studies, San Francisco, 171-192.

Friedman, M. (1982), *Capitalism and Freedom*, University of Chicago Press, Chicago.

Fuchs, V. (1982), Time Preference and Health: An Exploratory Study, in V. Fuchs (ed.), *Economic Aspects of Health*, University of Chicago Press, Chicago, 93-120.

Gruber, J. and B. Koszegi (2001), Is Addiction 'Rational'? Theory and Evidence, *Quarterly Journal of Economics*, 116, 1261-1303.

Heckman, J. (2000), Policies to Foster Human Capital, *Research in Economics*, 54, 3-56.

Heckman, J. and P. Klenow (1998), Human Capital Policy, in M. Boskin (ed.), *Capital Formation*, Hoover Economic Growth Conference, Hoover Institution.

Imrohoroglu, A., A. Merlo and P. Rupert (2001), What Accounts for the Decline in Crime? Penn Institute for Economic Research Working Paper no. 01-012.

Jacobs, B. (2000), A Note on Taxation and Human Capital Accumulation, manuscript, University of Amsterdam.

Jacobson, L., R. Lalonde and D. Sullivan (2003a), The Returns from Classroom Training for Displaced Workers, *Journal of Econometrics*, forthcoming.

Jacobson, L., R. Lalonde and D. Sullivan (2003b), Can You Teach Old Dogs New Tricks? The Impact of Community College Retraining on Older Displaced Workers, manuscript, University of Chicago.

Kane, T. (1999), *The Price of Admission. Rethinking How Americans Pay for College*, Brookings Institute Press, Washington, DC.

Krueger, A. and M. Lindahl (2001), Education and Growth: Why and for Whom?, *Journal of Economic Literature*, 39, 1101-1136.

Lleras-Muney, A. (2002), The Relationship Between Education and Adult Mortality in the United States, NBER Working Paper no. 8986.

Nelson, R. and E. Phelps (1966), Investment in Humans, Technological Diffusion, and Economic Growth, *American Economic Review*, 56, 69-75.

Okun, A. (1975), *Equality and Efficiency: The Big Tradeoff.* Brookings Institution, Washington DC.

Persson, T. and G. Tabellini (1994), Is Inequality Harmful for Growth?, *American Economic Review*, 84, 600-621.

Rabin, M. (1998), Psychology and Economics, *Journal of Economic Literature*, 36, 11-46.

Romer, P. (1990), Endogenous Technological Change, *Journal of Political Economy*, 98, 71-102.

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

Smith, A. (2003), *The Wealth of Nations*, Bantam Books, New York.

Spence, M. (1974), *Market Signaling: Informational Transfer in Hiring and Related Processes,* Harvard University Press, Cambridge MA.

Warner, J. and S. Pleeter (2001), The Personal Discount Rate: Evidence from Military Downsizing Programs, *American Economic Review*, 91, 33-53.

Wolf, A. (2002), Does Education Matter: Myths About Education and Economic Growth, Penguin Books, London.

3 Education, skills and earnings – the Swedish record

What has the Swedish education system accomplished? To what extent has education policy produced more equality in skills? What effects have a compressed distribution of skills had on wage inequality? These questions are discussed in this chapter. We will give brief account of some key characteristics of the Swedish education system and what it has produced. The main focus will be on the current situation in Sweden compared to other (mainly OECD) countries. We will also provide information on the historical development of some aspects of education in Sweden.

The chapter contains five main sections. We start in section 3.1 with a discussion of the resources devoted to education in Sweden. We then turn to three sections on various forms of output from the education system: the formal level of schooling (section 3.2), the skills provided by the schooling system (section 3.3), and the returns to schooling and wage inequality (section 3.4). Finally, in section 3.5 we discuss the important issue of the relationship between the distribution of skills and the distribution of wages.

3.1 Inputs

We start by noting that Sweden spends a large amount of resources on the education system. In 2001, the overall expenditures of the education system amounted to 7.7 percent of GDP. Using somewhat older figures, Table 3.1 shows that Sweden is a country that relative to GDP spends substantially more on education compared to most other countries in the OECD area. The table also reveals that Swedish spending has been fairly constant over the 1990s – there was a drop in resources up to the mid 1990s, but spending has returned to previous levels in the late 1990s. At the same time, most other OECD countries, including the United States, have reduced their spending relative to GDP. Average OECD education expenditure dropped from 6.1 to 5.6 percent between 1992 and 1999.

In Table 3.1 we also show a measure of expenditure per student for the entire education system (excluding pre-primary education). We have calculated these numbers from expenditure data at different levels of education, using a fixed set of weights, so that the differences across countries and over time will not be affected by changes in the composition of students across levels. The

pattern in these numbers is similar to the pattern for the expenditure level. Sweden spends about 34 percent of GDP per capita on an "average student", and this number has not changed since the early 1990s. Most other OECD countries spend somewhat less than Sweden, and spending has declined over the 1990s from 30 to 28 percent of GDP per capita.

Table 3.1 Total education expenditure as a percentage of GDP and expenditure per student as a percentage of GDP per capita, 1995 and 1999

Country	*Total expenditure*		*Expenditure/student*	
	1992	1999	1992	1999
Sweden	6.8	6.7	34	34
United States	7.0	6.5	35	32
OECD average	6.1	5.6	30	28

Source: OECD (1995, 2002).
Note: Expenditure per student is a weighted average of expenditure per student at different levels of education. The weights are common across countries; one half for primary and lower secondary education, and one quarter each for upper secondary and tertiary education.

A break down of education expenditures by level of education is given in Table 3.2. Like in most other countries, primary and lower secondary education (grades 1–9 in Sweden) is the largest source of expenditure (3 percent of GDP). In an international comparison, Sweden spends a lot on primary and lower education, as well as tertiary education. Regarding upper secondary, Sweden is more like an average OECD country.

Table 3.2 Total education expenditure as a percentage of GDP by level of education, 1999

Country	*Primary and lower secondary education*	*Upper secondary education*	*Tertiary education*
Sweden	3.0	1.4	1.7
United States	NA	NA	2.3
OECD average	2.3	1.3	1.3

Source: OECD (2002).

The picture of Sweden as a high-spending country is modified if we instead focus on spending on teacher's salaries. Sweden clearly deviates from other

OECD countries in this respect. The share of total education expenditure devoted to teacher salaries is 48 percent in Sweden compared to the OECD average which is 65 percent (OECD, 2002, Table B6.3). Expenditures on teacher salaries in Sweden stood at 3.2 percent of GDP in 1999; the corresponding figure for the OECD average was 3.6 percent. In terms of expenditures directly related to teaching, the resources in Swedish schools are fairly low compared to many other countries. The main reason for low teaching expenditures in Sweden is that teacher salaries are comparatively low (Gustafsson and Myrberg, 2002).

Table 3.3 shows a breakdown of expenditure per student across countries in 1992 and 1999. To make the comparison across countries (i.e. the United States), we have a somewhat different categorization of education levels here.[11] In terms of expenditure per student 1999, Sweden actually spends about the same on upper secondary education as the OECD average (25 percent of GDP per capita). In primary education we spend more and in tertiary education substantially more than the average in OECD.

Table 3.3 Education expenditure per student as a percentage of GDP by level of education, 1992 and 1999.

Country	*Primary education*		*Secondary education*		*Tertiary education*	
	1992	1999	1992	1999	1992	1999
Sweden	29	24	33	25	43	61
United States	24	20	26	24	61	57
OECD average	21	19	27	25	49	44

Source: OECD (1995, 2002).

The most interesting feature of Table 3.3, though, is the changes over time within each country. In both the United States and the average OECD country, it seems like the overall reduction in resources for education (see Table 3.1) have been symmetrically distributed across different levels of education. All three levels of education have experienced reduced expenditures per student. The Swedish experience clearly deviates from this pattern. Within a virtually

[11] Primary education consists of Swedish grades 1–6, while grades 7–9 are grouped with high school education (”gymnasium”) in secondary education.

constant level of overall resources relative to GDP, there has been a substantial shift in resources across levels of education. Resources for tertiary education have expanded dramatically, from 43 to 61 percent of GDP per capita, while primary and secondary education both have experienced substantial reductions in resources per student.

Another notable feature of the Swedish education system is that we spend quite a lot of resources on relatively old students. This is evident in Table 3.4, where we show enrollment rates by age in Sweden compared with the OECD average. Looking at enrollment rates for 30–39 year olds, we find that it is about 10 percentage points higher in Sweden than in the United States and the OECD average. There are two key aspects of this: adult education and high university enrollment age. Over the 1990s adult education expanded substantially, especially with the introduction of the Adult Education Initiative ("Kunskapslyftet"), see e.g. Ekström (2003).

Sweden has also traditionally had a high age of enrollment at university. The last column of Table 3.4 shows the median age of entry into tertiary education. While the median age of entry is around 19 or 20 years in most countries, Swedish entrants to tertiary education are 22.7 years old. Along with Iceland and New Zealand, Sweden actually has the highest median age of entry in the OECD area.

Table 3.4 Overall enrolment rates by age and median age of entry to tertiary education, 2000

Country	*Age 0-4*	*5-14*	*15-19*	*20-29*	*30-39*	*Age 40-*	*Median tertiary entry age*
Sweden	70.5	97.8	86.4	33.4	15.0	3.4	22.7
United States	49.9	99.3	73.9	21.2	5.4	1.5	19.4
OECD average	63.8	97.9	77.3	21.4	4.9	1.3	20.2

Source: OECD (2002).
Note: The younger ages include enrollment in child care.

An interesting question is how these high ages of entry affect calculations of the internal rates of returns to higher education. Using the information provided in Björklund and Kjellström (2002) (their estimates for 1991 and a discount rate of 2 percent) we find that increasing the age of entry from 19 to 22 years reduces the internal rate of return to a university degree by 7.5 percent. Thus,

we are paying a non-negligible price for the fact that young Swedes are not able (or not willing) to make the transition directly from high-school to university.

3.2 Development of formal schooling

At the start of the 20th century, higher education was a rare thing. In 1930, less than two percent of the adult population had upper secondary education or more (Table 3.5). A large majority of the population only had some primary education of varying length and intensity. Over the next 70 years, the situation has changed rapidly. In 1970, the expansion had become evident for upper secondary schooling, while the expansion of tertiary schooling was still limited. The effects of the reforms of the university system in the 1960s show up more clearly in the data for 2000. Even if the expansion of the Swedish university system occurred long after the US expansion, it still predates the expansion in many European countries (Edin et al., 1994).

Table 3.5 The level of education in the Swedish population aged 20 – 60, percent

	1930	*1970*	*2000*
Primary, lower secondary education and below	92.8	59.0	17.8
Upper secondary education	0.6	29.6	50.8
Tertiary education	1.2	7.4	30.5
No information	5.4	4.0	0.9

Source: Computations from the 1930 Census micro file (Bång, 2001) and LINDA.

An international perspective on the educational attainment of the Swedish population is provided by Table 3.6. It is clear that the level of schooling still is lower than in the US, but that it is higher than in most other OECD countries. Another way of summarizing the current situation is to calculate the expected years of schooling in each country based on current enrollment rates. Such an exercise suggests that Sweden will overtake the US in the future if the current enrollment patterns prevail; see OECD (2002).

Table 3.6 Educational attainment of the adult population, 2001

Country	*Primary and lower secondary education*	*Upper secondary education*	*Tertiary education*
Sweden	19	49	32
United States	13	50	37
OECD average	34	41	21

Source: OECD (2002).

3.3 The skills provided by the Swedish schooling system

A simple comparison of years of schooling across countries, or over time, is of course rather uninteresting unless we have reasons to believe that student actually learn something in school. In the end we are interested in how the schooling system in Sweden affects the skills of the student population and the labor force. Over the last decades a number of comparative international studies of various forms of skills have been conducted, both for students at different levels of schooling and for the adult population as a whole. Comparing skills across countries is admittedly a very difficult task. Still these studies have been carefully executed and convey interesting information. We will use some of these studies to illustrate how the skills of various groups of Swedes compare with those in other countries. We will start with a comparison of the skills of students still in school, and then turn to a comparison of the skills in the entire labor force.

3.3.1 Skills among those still in school

Since the first International Study of Achievement in Mathematics conducted in 1964 (Husén, 1967), a number of international studies of student achievement have been done. These cover students in different ages and subjects – mainly reading, mathematics and science. Here we will not attempt to survey this large literature in any detail. Instead we will summarize our main impressions from this literature in the form of four main observations.

A first observation is that the youngest Swedish children (age 9–10) do very well on various literacy and science tests. In the most recent study in reading, PIRLS 2001, Swedish students actually came out at the top of 36 countries. The interpretation of these results may be a bit muddled by the fact that there were small differences in at what age and in what grade the tests were administered across countries. However, accounting for age and grade at test

across countries does not change the main picture: Swedish students in primary school come out at the top in reading in an international comparison.

If we look at school children in lower secondary schooling (age 13-15), the Swedish results are more modest. In most comparisons Swedish children do about average or above average in math and science. In terms of literacy skills Swedish students still perform well. In the PISA 2000 study, Swedish 15-year olds did significantly better than the average on general literacy tests as well as tests for mathematical and scientific literacy (OECD, 2002).

The final group of students consists of finishing year students in upper secondary education. In the 1995 TIMSS, Sweden comes out as one of the top countries on both mathematics and science if we look at the average of all upper secondary students. If we instead focus on "specialists" – students taking advanced courses in these subjects – Swedish finishing students keep a top position in science (physics) but get mathematics results that are close to the international average.[12]

The overall impression from the international comparisons of student achievement is that the average Swedish student performs well. This is most evident in reading tests for the youngest children, but also students in upper secondary schooling do very well in mathematics and science.

A final observation on these tests is that Sweden also differs from most other countries in terms of differences in achievement across schools. In the PISA 2000, the overall variance of student performance in reading literacy was somewhat lower (92 percent) in Sweden compared to the OECD average. The striking aspect of the Swedish case, though, was that the share of the variance that was attributed to between school variation was very low – 9.7 percent. Except for Iceland, the Swedish between school variance was the lowest among the OECD countries, where the country average was 35 percent.

3.3.2 Skills in the adult population

The most comprehensive comparative study of adult literacy skills is found in the International Adult Literacy Survey (IALS), see OECD and Statistics Canada (1995). This study provides tests of three measures of literacy (prose, document and quantitative) for representative samples of the adult population in several OECD countries.

[12] The TIMSS results are available at: http://isc.bc.edu/timss1995i/TIMSSPDF/C_Hilite.pdf

The average skill levels in countries participating in the first two waves of the IALS are reported in Table 3.7. As observed by others (e.g. Björklund et al., 1998), the results show that the skills of the population are very high in Sweden. The average score is the highest among the 16 countries participating in the initial rounds of the survey. The dispersion of skills in Sweden is on the lower side, but not among the lowest.

Table 3.6 Skills in the adult population according to the IALS

Country	*Mean*	*Standard deviation*
Sweden	312	51
Norway	300	41
Finland	299	43
Denmark	296	40
Netherlands	294	43
Canada	292	60
Germany	292	47
Czech Republic	289	47
United States	285	65
Hungary	263	48
Switzerland	283	54
Italy	252	62
Poland	243	64
Slovenia	240	61
Chile	215	58

Source: Leuven et al. (2002).

Another way of comparing the results is to see where individuals in various points of the Swedish skill distribution would end up in the US skill distribution. In Table 3.8 we report the quantitative test score for individuals at the 10th, 50th and 90th percentile in the Swedish test score distribution. The final column shows at what percentiles of the US test score distribution these score are located. Consequently, an individual with a score of 230 would end up in the 10th percentile of the Swedish distribution and in the 30th percentile of the US distribution. The table shows that there are substantial differences in test scores throughout the skill distribution, but that these differences seem to be more pronounced at the lower part of the distribution.

Table 3.7 A comparison of the Swedish and US skill distribution, quantitative skills IALS

Swedish percentile	*Score*	*US percentile*
10th	230	30th
50th	303	69th
90th	370	95th

Source: Calculation from the 1994 IALS.

Has the Swedish population become more skilled over time? One way of trying to look at how schooling (and other factors) has changed the skills of the population over time is to look at skills across cohorts. These cohort skills will of course be affected by both true cohort effects and ageing. It turns out, however, that ageing effects are relatively unimportant, at least in Swedish data, with some qualification for the oldest cohort (Nathanaelsson, 2003). Also, when we compare cohort skills across countries, it seems fairly innocuous to assume that the effect of ageing on skills is similar across countries. Thus, we can interpret different trends in skills over cohorts as trends driven by cohort specific effects due to, e.g., education.

In Figure 3.1 we give the overall picture of skills across cohorts for the entire adult population in 1994 in five countries. There are (at least) three interesting points to be noted. First, there is a general tendency that skills are higher in younger cohorts. Second, the "growth rate" of skills across cohorts in Sweden seems to be declining. Third, Sweden comes out at the top in all five cohorts, but Finland is closing the gap rapidly.

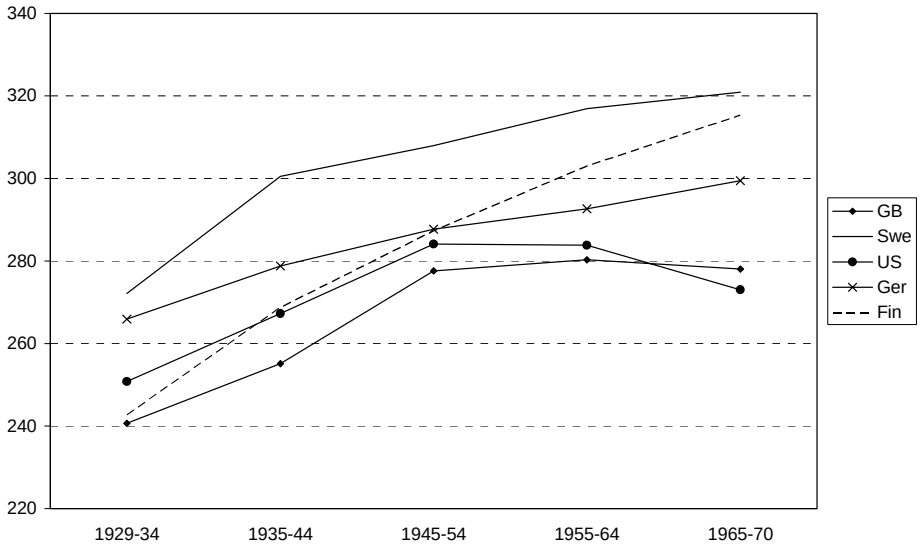

Figure 3.1 Skills by cohort.
Source: Nathanaelsson (2003).

In the next set of graphs we break down the analysis by level of education. In Figure 3.2.a we show average skills by cohort for individuals with tertiary education. It is evident from this graph that the differences in trends and levels across countries are smaller for the highly educated. Once again, Sweden comes out at the top with Finland closing in. Also, we see that skills are higher in younger cohorts in all countries, except for the US and the UK.

Figure 3.2.b shows the corresponding trends in skills for individuals with only primary or lower secondary schooling. Here, the differences across countries are larger. This can partly be explained by differences in the compulsory schooling systems across countries, but also by the fact that the test is constructed to be more precise in the lower part of the skill distribution. The figure shows that the US is something of an outlier in terms of the level of skills. The US is also the country that deviates mostly from the pattern of increasing skills over cohorts. The development in Sweden and Germany is almost identical, with Finland catching up over time.

a. University educated

b. Primary and lower secondary schooling

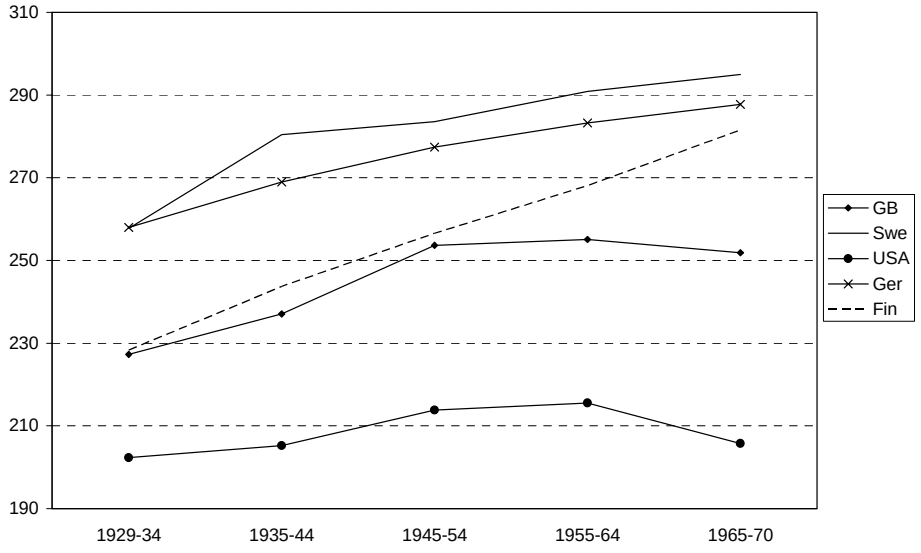

Figure 3.2 Skills by cohort and level of education.
Source: Nathanaelsson (2003).

The graphs presented here suggest that the skill formation process in Sweden is improving over time in a way to keep the skills of the population in the higher part of the international skill distribution, even if some countries seem to be catching up. This seems to be true both among those with university education and those with only compulsory education. To the extent that the trends in skills reflects changes over time in the quality of education in different countries, our results do not indicate any major deterioration of the Swedish schooling system over time.

3.4 Wage inequality and the returns to schooling

In this section we will provide some basic facts about the Swedish wage structure with a focus on the monetary returns to schooling. The starting point for this discussion is that monetary incentives are key determinants of the individual's decision to invest in formal education. We are not arguing that this is the only determinant of the demand for education, but we strongly argue that monetary incentives do have an effect at the margin.[13] We also feel that the recurring debate of whether the returns to investing in education are too low in Sweden motivates a short survey of the evidence.

It is well known that from an international perspective Sweden has a relatively compresses wage distribution. This was particularly evident in the early 1980s when wage dispersion stood at its minimum. During the last two decades, there has been a clear tendency to increasing wage dispersion in Sweden as well as in many other countries, but Sweden is still a country with a comparatively compressed wage distribution.

One of the key components in the returns to education is the wage differential between education groups – the wage premium. In Figure 3.5 we give the long run perspective on the education wage premium. The estimates are based on a simple Mincer-specification with a linear years-of-schooling variable.[14] Since the 1930 data do not contain hourly earnings, we also report estimates based on annual earnings for males for the 1930–1991 period.

[13] For Swedish evidence, see Fredriksson (1997).

[14] We are not claiming that these estimates necessarily reflect the causal effect of education on earnings; the estimates are probably affected by selection (e.g. Card, 1999). Still they provide a measure of wage differentials across education groups that are comparable over time and across countries. It is not obvious that these comparisons are affected by the selection problem.

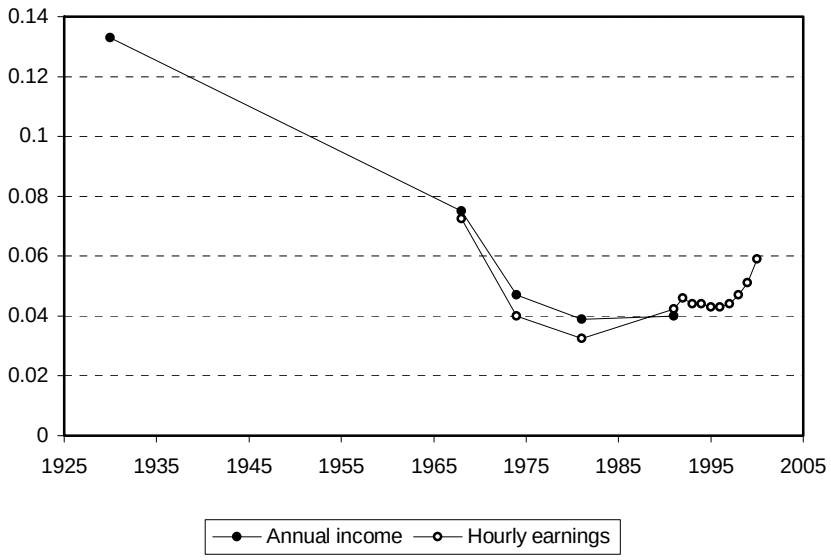

Figure 3.5 The premium to a year of schooling, 1930–2000.
Source: Bång (2001), Björklund and Kjellström (1994), and calculations from LINDA.
Note: The annual income estimates are based on males only. The hourly earnings estimates are based on survey data 1968–1991, and on register data 1992–2000.

In the early part of the 20th century educated labor was in very short supply (see Table 3.5), a fact that is also reflected in a very high wage premium. We have very little information on what happened between 1930 and 1968, but from 1968 the development of the education wage premium closely mimics the overall changes in the wage distribution. There is a dramatic reduction of the wage premium in the 1970s followed by a slight recovery up to 1991.

The estimates for the post-1991 period are based on data from other sources than the previous – register data instead of survey data. Therefore the levels of these estimates may not be completely comparable. This can also be part of the explanation for the break in the trend between 1991 and 1992. Over the 1990s, though, there is a clear pattern of increasing education wage premiums, in particular at the end of the period.[15] At the end of our observation period, the

[15] This upturn is not as marked in survey data; see Le Grand et al. (2001). During this period the coverage of private sector workers increased in our register data. Reweighing the estimates to account for this does not affect the main pattern though.

average wage premium associated with one additional year of schooling has increased to around 6 percent.

The wage premium associated with education is a key component of the returns to education, but not the only one. The private returns to investing in education are also affected by various other factors like tuition fees, student support and taxes. While tuition fees are not present in Sweden, the effects of the tax system and student support (grants and subsidized loans) are non-negligible. When incorporating these two factors in a calculation of the internal rate of return to higher education, it seems that it has been fairly stable over the 1990s. Edin and Holmlund (1995) compute an internal rate of return to university education for 1991 of about 11 percent when taxes and student support are included. We have replicated their calculation for the year 2000 and get an internal rate of return of about 10 percent. The slight decrease in the return is due to less generous student subsidies (loans and grants) in 2000. Note also that these calculations are based on survey data (the LNU), where the upturn of the wage premium is less pronounced than in the register data.

An international comparison of the returns to schooling is provided by OECD (2002). Their computations of internal rates of returns are not fully comparable with those above, e.g. they are based on annual earnings instead of hourly wages and they also include differences in unemployment across education groups. Still they provide a comparable measure of returns across countries.

In Table 3.8 we report the internal rates of returns for both upper secondary and tertiary education for males and females separately. Starting with the returns to tertiary education in the last two columns, we see that the Swedish numbers are fairly close to those reported above – around 11 percent. It is also clear that Sweden comes out in the middle of the pack when comparing returns across countries. The Swedish returns are substantially below those in some countries, like the US and UK, but are comparable to or above those in other countries.

An interesting feature of Table 3.8 is that it demonstrates how much internal rates of returns differ across levels of education within countries. The Swedish return to upper secondary education is strikingly low (for females it is not even computable) compared both to the returns in other countries and to the return to tertiary education. It seems reasonable to argue that at least part of this low returns is associated with the massive expansion of upper secondary education that has taken place in Sweden over the last decades.

Table 3.8 Internal private rates of returns to upper secondary and tertiary education, 1997-2000

Country	*Upper secondary, Males*	*Upper secondary, Females*	*Tertiary, Males*	*Tertiary, Females*
Canada	13.6	12.7	8.1	9.4
Denmark	11.3	10.5	13.9	10.1
France	14.8	19.2	12.2	11.7
Germany	10.8	6.9	9.0	8.3
Italy	11.2	NA	6.5	NA
Japan	6.4	8.5	7.5	6.7
Netherlands	7.9	8.4	12.0	12.3
Sweden	6.4	0.0*	11.4	10.8
United Kingdom	15.1	NA	17.3	15.2
United States	16.4	11.8	14.9	14.7

Source: OECD (2002).
Notes: The earnings difference between females with lower and upper secondary schooling in Sweden is not large enough to produce a positive rate of return.

The low return to upper secondary education also raises the question of the return to adult education, since most of adult education is at the upper secondary level. There are issues about selection into adult education which are difficult to handle in empirical work. Still, studies that try to handle these selection problems suggest that the average returns to adult secondary education are even lower than the returns presented above. Ekström (2003) estimates the earnings premium for various subgroups around 10 years after they started adult education. She finds that there is no positive return for Swedish born participants – for males there is even a significant negative return. A somewhat more positive picture emerges for immigrants, where the positive effects are close to significant for females. Similar results are obtained by Stenberg (2003) for the Adult Education Initiative. He finds that participants in this initiative obtain lower incomes than comparable participants in labor market training.[16]

[16] Note here that available evaluations of labor market training suggest that this form of education has non-positive returns; see e.g. Calmfors et al. (2001).

3.5 The dispersion of wages and skills

Combining the evidence on the distribution of skills across countries with the pattern of wage inequality across countries leads naturally to the question of the relationship between the two: Are differences in wage dispersion driven by differences in the supply of skills? The simple answer to this question is yes, but we don't really now how much of the differences that can be explained. To be able to answer this question we need to pin down the magnitude of two effects. First, there is an obvious direct effect from skill dispersion to wage dispersion. At a given price of skills, countries with a higher dispersion of skills will have more wage dispersion. The size of this effect depends on the price of skills in the country, which in turn is affected by the particular wage setting institutions and other factors. The second effect is an indirect effect that works through the price of skills. The net supply of skills will affect the price of skills. We will summarize some evidence on these two mechanisms in turn.

Starting with the direct effect, Devroye and Freeman (2002) present a simple variance decomposition exercise that suggests that only a small part of cross-country differences in wage inequality is driven by differences in skill dispersion. Applying the US price of skills to the Swedish skill distribution would explain only 12.5 percent of the difference in the standard deviation of earnings across countries. They conclude that the differences in wage inequality between the US and European countries are mainly driven by different wage setting institutions.

Similar evidence is reported by Blau and Kahn (2001) using a decomposition method that allows for different effects at different part of the wage distribution. An interesting aspect of their results is that the contribution of measured characteristics (including skills) is higher at the lower end of the wage distribution. Once again comparing Sweden and the US, differences in age, schooling and skills account for 26 percent of the 50-10 log wage differential.

Both these studies only measure the direct effect of differences in the distribution of skills across countries. Leuven et al. (2002) attempt to include also the indirect effect, i.e., the effect of the net supply of skills on the price of skills. To do this they apply the methodology of Katz and Murphy (1992). The bottom line of their study is that they find much stronger effects from country differences in skills. They find that about one third of the variation across countries in the relative wages of skill groups can be attributed to the net supply of skill groups. Their analysis does an even better job in explaining

relative wages in the lower parts of the wage distribution, where about 60 percent of the variation is explained. Even though we cannot make strong statements about the exact magnitude of the contribution of skill differences to differences in wage inequality across countries based on this limited number of studies, it seems clear that the distribution of skills is one of the factors that produce differences in wage inequality across countries.

3.6 Summary and conclusions

In terms of overall resources devoted to education, Sweden appears to be an ambitious country. The level of ambition is not as impressive if we measure resources as expenditures on teachers. Using this measure, Sweden actually spends less resources on education compared to the OECD average.

During the 1990s, the total expenditure per student was roughly constant, but the structure of expenditure changed drastically. We have witnessed reductions in resources per student for primary and secondary education, while resources per student in tertiary education have increased substantially.

The Swedish population appears to be highly educated in an international comparison. The average level of formal schooling is high and available direct measures of skills tell a similar story. Sweden shows very good results in international comparisons of basic literacy skills.

The dispersion of skills in the population is relatively small in Sweden. Even if wage setting institutions, i.e. union wage policy, are important, the dispersion of skills is a partial explanation to the compressed wage structure in Sweden. The returns to higher education in Sweden are comparable to those in many other OECD countries. The returns to upper secondary school, on the other hand, are very low by international standards. Whether these returns are too low is an open question. Today, a very large share of each cohort enters upper secondary school in Sweden and differences in the supply of education may explain part of these differences.

The private returns to adult education appear to be even lower than the returns to upper secondary education. It is hard justify massive subsidies to adult education on the basis of increased productivity at the individual level. These low returns provide one argument for reconsidering the current allocation of resources. Another reason is that Swedish student are relatively old in an international comparison – the public returns to education is unambiguously decreasing when education is taken at higher ages. We will

return to a discussion of the allocation of resources to different parts of the education system in the final chapter.

References

Blau, F.D. and L. Kahn (2001), Do Cognitive Test Scores Explain Higher US Wage Inequality? NBER Working Paper No. 8210.

Björklund, A. and C. Kjellström (1994), Avkastningen på utbildning i Sverige 1968 till 1991, in R. Erikson and J. Jonsson (eds.), *Sorteringen i skolan*, Carlssons förlag, Stockholm.

Björklund, A. and C. Kjellström (2002), Estimating the Return to Investment in Education: How Useful is the Standard Mincer Equation, *Economics of Education Review*, 21, 195-210.

Björklund, A., P-G. Edebalk, R. Ohlsson and L. Söderström (1998). *Välfärdspolitik i kristid – håller arbetslinjen?*, Välfärdspolitiska rådets rapport 1998, SNS Stockholm.

Bång, J. (2001), The Returns to Education – using data retrieved from the Swedish National Census of 1930, mimeo, Department of Economics, Uppsala University.

Calmfors, L., A. Forslund and M. Hemström (2001), Does Active Labour Market Policy Work? Lessons from the Swedish Experiences, *Swedish Economic Policy Review*, 8, 61-124.

Card, D. (1999), The Causal Effect of Schooling on Earnings, in O. Ashenfelter and D. Card (red.), *Handbook of Labor Economics* vol. 3A, North-Holland, Amsterdam.

Devroye, D. and R.B. Freeman (2002), Does Inequality in Skills Explain Inequality of Earnings across Advanced Countries?, Discussion Paper no. 552, Centre for Economic Performance.

Edin, P-A., P. Fredriksson, and B. Holmlund (1994), Utbildningsnivå och utbildningsavkastning, in A. Björklund and A. Forslund (ed.), *Studier av svensk utbildning*, Ekonomiska rådets årsbok 1993, Fritzes, Stockholm.

Edin, P-A and B. Holmlund (1995), The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy? in R.B. Freeman and L.F. Katz (eds.), *Differences and Changes in Wage Structures*, 307-343, University of Chicago Press, Chicago and London.

Ekström, E. (2003), Earnings Effects of Adult Secondary Education in Sweden, in *Essays on Inequality and Education,* PhD thesis, Department of Economics, Uppsala University.

Fredriksson, P. (1997), Economic Incentives and the Demand for Higher Education, *Scandinavian Journal of Economics* 99, 129-42.

Gustafsson, J-E. and E. Myrberg (2002), *Ekonomiska resursers betydelse för pedagogiska resultat – en kunskapsöversikt*, Skolverket, Stockholm.

Husén, T. (1967), International study of achievement in mathematics: a comparison of twelve countries, Almqvist & Wicksell, Stockholm.

Le Grand, C., R. Szulkin and M. Tåhlin (2001), Lönestrukturens förändring i Sverige, in J. Fritzell, Michael Gähler and O. Lundberg (eds.), *Välfärd och arbete i arbetslöshetens årtionde*, SOU 2001:53, Fritzes, Stockholm.

Leuven, E, H. Oosterbeek, and E. van Ophem (2002), Explaining International Differences in Male Wage Inequality by Differences in Demand and Supply of Skill, *Economic Journal*, forthcoming.

Nathanelsson, K. (2003), The Development of the Supply of Skills across Countries and the Evolution of Wage Inequality, mimeo, Department of Economics, Uppsala University.

OECD (1995), *Education at a Glance*, Organization for Economic Cooperation and Development, Paris.

OECD (2002), *Education at a Glance*, Organization for Economic Cooperation and Development, Paris.

OECD and Statistics Canada (1995), *Literacy, Economy and Society. Results of the First International Adult Literacy Survey*, OECD and Statistics Canada, Paris.

Stenberg, A. (2003), An Evaluation of the Adult Education Initiative Relative to Labor Market Training, PhD thesis, Department of Economics, Umeå University.

4 Resources, decentralization, and student achievement

The 1990s was a very eventful decade for Swedish youth education. Perhaps most importantly, this decade saw an overall decline in education expenditures relative to GDP. Expenditure per student in compulsory schools, as a share of GDP per capita, declined from 34 percent in 1991 to 24 percent in 1999; see OECD (1994) and Table 3.3. Education expenditure declined even more at the secondary level: from 41 percent in 1991 to 25 percent in 1999. These developments took place during a time period when the typical OECD country experienced no major changes in resources. Sweden now spends about as much on primary and secondary education relative to GDP per capita as the average OECD country.

Part of the decrease in resources should certainly be attributed to the severe economic slump that hit Sweden in the beginning of the 1990s. In the beginning of the 1990s, authority over schools was also decentralized to the local level. Formally this change took place in 1991. But it was not until 1993 that the local authority could decide on the allocation of spending over its various responsibilities. Between 1991 and 1993 the system with ear-marked money granted from the central government was still running, albeit in a less strict form.

The main question in this chapter is whether and how resource changes affect student achievement. To answer this question we begin in section 4.1 by reviewing the received literature on the effect of resources on student performance. Section 4.2 documents the evolution of resources over the 1990s and early 2000s. We document the aggregate development and examine whether the distribution of resources changed after decentralization. In section 4.3 we investigate the effect of the resource changes during the 1990s. We look at the average effect, but we also pay particular attention to pupils from disadvantaged backgrounds. Throughout this chapter we focus on the developments in compulsory schools. The main reason for this focus is that we think that resource changes have the greatest potential effect early on in life. So, if

we cannot find an effect for students in compulsory school, we probably cannot find it in upper secondary school either.[17]

4.1 How important are resources for outcomes?

This section briefly reviews the literature on the effect of various forms of resources on student performance. We focus on three types of resources: class size, teacher qualifications, and computer aided instruction. In addition to setting the stage for this chapter, this section also serves as a background for chapter 5 on teachers.

4.1.1 Class size

The pupil-teacher ratio is a major determinant of the cost of schooling, as teacher salaries are a major component of school budgets. Most summaries of the literature on the effect of class size (or the effect of the pupil-teacher ratio if the unit of observation is larger than a class) and student achievement begin with the admonition that the literature is mixed. We are no exception: the findings in the literature are mixed. Indeed, the quantitative reviews of the literature itself are mixed, with some (e.g., Hedges, et al. 1994; Krueger, 2003) finding positive impacts of smaller classes on student outcomes, and others (most prominently, Hanushek, 1997) finding no beneficial effects of smaller classes. The fact that the literature does not find ubiquitous evidence of a beneficial effect of smaller classes, however, does not necessarily mean that teachers can be furloughed without consequence, or that reducing class size would not help students. The effect of class size may be difficult to distinguish from background factors and other school inputs, and observational studies have the further problem that school authorities may deliberately assign weaker students to smaller classes, hoping to reverse their disadvantages, creating bias due to reverse-causality.

Despite these problems, we believe a reasonable interpretation of the available research is that it is likely that smaller classes have a beneficial effect on student achievement, especially if the population under consideration includes young children or disadvantaged children. In addition to our reading of the comprehensive literature summaries, we reach this conclusion by putting

[17] As we have shown above, the resource developments were similar in compulsory and upper secondary school. Teacher density, however, has not fallen as much in upper secondary schools as in compulsory schools.

weight on the experimental evidence, on the quasi-experimental evidence, and on work specifically for Sweden described here. Nevertheless, reducing class size remains a costly solution to raising educational achievement, and the benefits may not be worth the costs, especially for older students or students from advantaged families. But policymakers should not consider raising class size a free lunch: there are most likely costs in terms of reduced educational achievement.

Interested readers are advised to consult the literature reviews (e.g., Hanushek, Hedges et al., Krueger, and Gustafsson, 2003) for a fuller account of the literature, and the debate over methodological issues and interpretation. Here we just give a brief flavor for the issues. In perhaps the most influential work in the economics of education, Hanushek (1986, p. 1162) argues that: "The results are startlingly consistent in finding no strong evidence that teacher-student ratios, teacher education or teacher experience have an expected positive effect on student achievement." This conclusion was challenged by Hedges, et al. (1994), who point out that Hanushek did not take into account the magnitude of the estimates; instead, he simply counted whether the effects were positive or negative. When they take into account the magnitude of the coefficients, using formal meta-analysis techniques, Hedges, et al. find a large impact of school resources such as smaller class sizes. Krueger (2003) also questions Hanushek's methods, noting that Hanushek extracted multiple estimates from some studies – as many as 24 in two cases – and only one estimate from many others. If each study in the literature is accorded equal weight, Krueger finds that the literature as a whole, points in the direction of a beneficial effect of smaller classes: 57 percent more studies find a positive effect than find a negative effect. Moreover, Krueger notes that if studies published in more highly ranked journals are accorded more weight, the evidence of a beneficial effect is even stronger. Krueger also questions the appropriateness of many of the specifications underlying the estimates that Hanushek included in his analysis. A large number of studies, for example, estimated the effect of the pupil-teacher ratio while holding the amount of expenditures per student constant, thereby offsetting smaller class sizes by reductions in other inputs. Hanushek (2003) has argued in response that the studies that he considers to be the best in the literature, those using a "value added" specification (where the outcome is measured in terms of improvement from one year to the next), tend to find no consistent effect of class size. Whether the value added specification does yield more reliable estimates is a

topic of much controversy, however. Todd and Wolpin (2003), for example, question the appropriateness of the value added specification. We return to this issue in the context of Sweden shortly.

A number of analysts have placed a great deal of weight on the Tennessee STAR experiment, the only large-scale randomized experiment conducted on the effect of smaller classes. This experiment involved students in the first year of school through the fourth year of school. Within each participating school, three assignment groups were used: small class (15 students on average); regular size class (22 students on average); and regular size class with a teacher aide. After four years, all students returned to regular size classes. Because random assignment was used to assign students and teachers to their classes, biases due to reverse causality and omitted variables are eliminated, although one can be concerned that attrition was high in the experiment. Analyses of the STAR experiment consistently find that students assigned to the smaller classes scored higher on standardized achievement tests and were more likely to take a college entrance exam (see Finn and Achilles, 1990, Nye et al., 1994, Krueger, 1999, and Krueger and Whitmore, 2001).[18] On average the gains were in the order of 0.2 standard deviations while the pupils were assigned to smaller classes. The gains were also larger for African American students and low-income students than for other students.

Angrist and Lavy (1999) examine the effect of class size on student achievement in Israel by exploiting a natural experiment created by "Maimonides' rule," a cap on class size. Maimonides' rule requires schools to maintain class size below 40 students per class. Angrist and Lavy note that the application of such a class size cap creates a sea-saw pattern in the relationship between class size and the number of students in a grade level. When enrolment approaches the cap, class size rises, and then it drops discretely once the cap is exceeded. Variability from this source is close to random, as there is little rhyme or reason for the jumps in class size other than the vagaries of the number of students enrolled in a grade level. They exploit these discontinuous movements to estimate the effect of class size on the test scores of Israeli 4th and 5th graders in 1991 and 3rd graders in 1992. They provide convincing evidence that the up-and-down movements in class size induced by Maimonides' rule is mirrored in test scores: reductions in class size are found to induce a significant

[18] The assignment of a teacher aid, however, usually was found to have little impact on achievement.

increase in reading and math scores for 5th graders and a smaller increase in reading scores for 4th graders. Results for 3rd graders are statistically insignificant, but they argue this could be due to shortcomings of the test administered in 1992. Similar to findings from the STAR experiment, their estimates further suggest that the achievement gains from small classes are largest among students from disadvantaged backgrounds.

Whether findings for students in Israel or Tennessee are relevant for Sweden is an open question. The socio-economic characteristics of students and their families are clearly different. We therefore think it is important to also consider studies that focus on Sweden. In the last 20 years, we know of only one study that has made a serious attempt to examine the impact of class-size on student achievement in Sweden, Lindahl (2001).

Lindahl collected and analyzed longitudinal data on 556 pupils from 16 schools in Stockholm in 1998 and 1999. The students were given identical math tests in the spring of the 5th grade, and in the fall and spring of the 6th grade. These data allow him to estimate value added models, where he can look at the gains in scores specifically over the school year. Moreover, he can use changes in achievement over the summer months to make an adjustment for the "pure" effect of family background and other non-school factors on student learning. The typical value added model in the previous literature ignores the fact that family background affects the trajectory as well as the level of student achievement, independently of what goes on in school. Lindahl can address this omitted influence by using the change in achievement over the summer months as a control for the family influence on the achievement trajectory.

Class size in his study was measured by the number of pupils taught together in the math class. In sixth grade, the average math class had 19.9 students, and the range went from 6 to 25. In fifth grade, the average class had 22.9 students, and the range went from 3 to 32. When Lindahl estimates level regressions (e.g., 6th grade test scores as outcome) or fall-to-fall change (from 5th to 6th grade) value added regressions, he finds no significant relationship between class size and achievement, or a positive relationship. But when he estimates value-added models using growth over the school year less changes over the summer months as the outcome variable (or controlling for changes over the summer as an explanatory variable), he finds a strong inverse relationship between class size and test scores. That is, Lindahl's results suggest that students who are in a smaller math class for a year make larger gains that school year than students in larger classes. The order of magnitude is similar to

the estimates from STAR. A reduction of class size by seven students increases the percentile rank by 2.6 to 6.9 percentile ranks according to his preferred set of estimates; this corresponds to a gain of 0.09 to 0.24 standard deviations. Lindahl does not find a larger gain for students from lower socioeconomic status families, however.

Rarely do researchers move beyond the simple question of whether class size matters to ask how much it matters, or whether the gain is worth the cost. Krueger and Lindahl (2002) use Lindahl's estimates to conduct a hypothetical cost-benefit analysis for Sweden of the effect of reducing class size from an average of 19.9 in grades 1-6 to an average of 13.3. Assuming a discount rate of 3 percent and future productivity growth of 1 percent per annum, they estimate that every SEK spent on class size reductions would eventually pay back 1.8 SEK in higher income for students. Put differently, the internal rate of return that equates the benefits and costs of class size reduction in Sweden is about 5 percent. Although this calculation involves several assumptions, it is worth noting that the internal rate of return is remarkably similar to the 6 percent figure Krueger (2003) finds for the United States based on the STAR experiment.

4.1.2 Teacher qualifications

Teachers matter, there can be little doubt. A large literature finds "teacher effects" – that is, particular teachers tend to have students who consistently score higher or make larger gains from one year to the next. Although part of the measured teacher effects is probably due to the fact that the same teachers tend to have high achieving (or high improving) students from one year to the next, part is also very likely due to the teachers' unique contributions. But knowing what it is about teachers that matter, or predicting which teachers will be more successful than others from their observable characteristics and credentials, is a much more difficult matter. Literature reviews tend to find very little systematic evidence of an impact of teachers' educational attainment on student outcomes (e.g., Hanushek, 1986)

Studies do find that teachers who score higher on standardized tests tend to have more successful students, especially at the high school level and especially when the tests are related to the subjects the teachers are teaching (e.g., Strauss and Sawyer, 1986, and Ferguson, 1991). Gustafsson (2003) is quite optimistic about the potential for upgrading teacher competencies, concluding, "Given the strengths of effects associated with teacher competence

it would seem that investments in teacher competences would have a higher likelihood of paying off in terms of student achievement than would other investments." He acknowledges, however, that very little research has been done on the effects of teacher in-service training.

In general, the literature on the importance of teacher suffers from a similar problem as the literature on class size. Teachers having certain characteristics might be systematically allocated to students. A few studies do not suffer from this problem. Since teachers were randomly assigned to classes in STAR, these data can be used to examine the effect on student performance of observed teacher characteristics. The estimates from STAR suggest that teacher experience tends to have, at best, a very modest positive effect on student achievement; the educational attainment of the teacher had no effect at all (Krueger, 1999). Another study by Rockoff (2003), which also does not suffer from problems caused by the systematic allocation of teachers to students, obtains a similar conclusion, although teacher experience was more important than in STAR.

A very interesting question involves teacher continuing training. New Information Technology offers many opportunities for teacher training, especially because of the flexibility that self-paced computer instruction affords. Yet there has been little research done to examine the modes of teacher training that are most successful. This would seem to us to offer many research possibilities, especially because randomized experiments could be designed to assess the impact of various teacher training programs.

4.1.3 Computers, IT and new teaching techniques

For all the enthusiasm about using computers and new Information Technology in the classroom, there is little evidence that such methods have raised student achievement, even when state-of-the-art, scientifically-based instructional programs are considered (see Kirkpatrick and Cuban, 1998). Wenglinsky (1998) finds perhaps the most optimistic results for the efficacy of computer-aided instruction, but, even taken at face value, his results are puzzling: the frequency of computer use is inversely related to math test scores in his data. He emphasizes that computers raise math achievement when used for applications and simulations, but reduce it when used for drills and practice. One could raise serious concerns as to why teachers chose particular methods of computer instruction, however. Maybe teachers in classes with more advanced students are more likely to use computers for applications and

simulations. No one would be surprised if teachers tended to use traditional drill and practice exercises (the digital equivalent of flash cards) for eighth grade students when those students were low achievers, or if they used applications of integral calculus when their students were high achievers. Yet one would be reluctant to say the causality runs from the computer-aided instruction method to student achievement.

To avoid these problems, Angrist and Lavy (2002) analyze a natural experiment in Israel. The program Tomorrow-98 distributed computers to some elementary and middle schools in Israel. By 1996, 10 percent of elementary schools and 40 percent of middle schools applied for and received computers under this program. Yet Angrist and Lavy find no evidence that the computers actually raised pupils' test scores. But it is unclear what instructional packages were being used or whether the teachers had the expertise to use the packages and computers in this setting.

Two independent randomized evaluations have been done in the United States of the computerized instruction program known as *Fast ForWord*. *Fast ForWord*, which was developed by neuroscientists to aid reading, is the leading edge of scientifically-based computer technology in schools, and one of the more expensive programs available, so it affords a test of state-of-the art applications of computers. In both Borman and Rachuba (2001) and Rouse and Krueger (2003), low-achieving students in schools in two urban school districts in the US were randomly assigned to train on *Fast ForWord* or to serve in a control group. Both studies found no effect of *Fast ForWord* on students' reading achievement scores, although Krueger and Rouse found a marginally significant effect on one pre-reading exam. The findings by these two research teams were strikingly different from those of the Corporation that markets the program, even when a common test was used to assess performance. This disparity raises another, often overlooked, issue: the potential for conflict of interests when for-profit educational companies advertise the scientific basis for their products.

So what does one make of this literature? Computer-aided instruction and IT still hold great promise for productivity increases in education. To date, however, that promise has probably not been realized. We believe the situation calls for heightened monitoring and evaluation of computer-aided instruction and other IT innovations in schools, particularly in light of the conflicts of interest that arise between the business interests of the firms that develop and market the equipment and the interests of the students who use them.

4.2 School inputs during the 1990s and early 2000s

The purpose of this section is to document the evolution of inputs. We pay particular attention to teaching inputs and focus on the 1990s. We start by depicting the resource changes in the aggregate and then move on to consider the distribution of resources.

4.2.1 The aggregate development

Figure 4.1 shows the evolution of the median teacher/student ratio (i.e. the number of teachers per student) in compulsory schools during the school years 1990/91—2002/03. These numbers are calculated as the median over municipalities weighted by the size of the student population. Hence they have the interpretation of the resource development facing the median student.

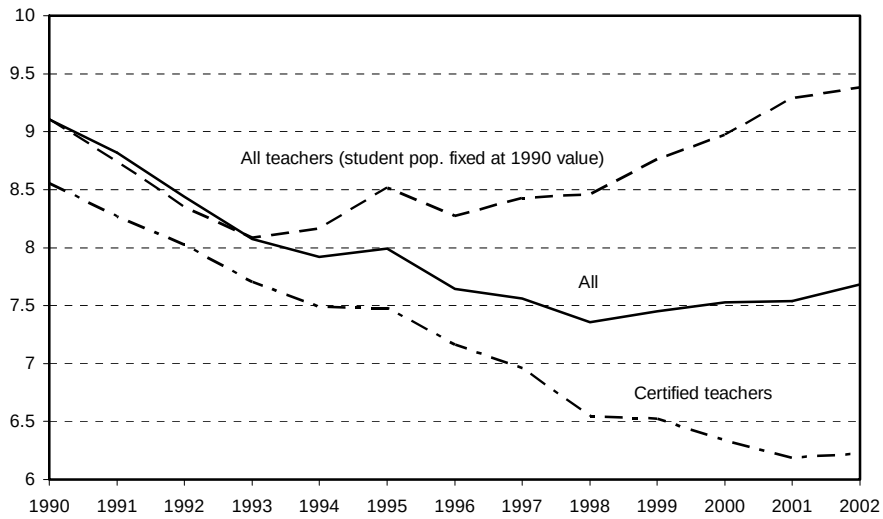

Figure 4.1 Median teacher/student ratios, percent, 1990/91 – 2002/03.
Source: Calculations using the Teacher register.
Notes: The figure shows the population weighted median over municipalities. Certified teachers refer to those with a degree from teacher education. During 1990/91—1997/98 the number of teachers have been converted to full time equivalents using 25.3 stipulated teaching hours per week as the measure of full-time teaching load. For later years there is a (reliable) measure of employment intensity directly available in the data. In doing these calculations, we have imposed the restriction that individuals holding more than one position can work no more than 120 % of full time.

The 1990s saw a continuous decrease in teacher density (solid line). From 1990/91 to 1999/00 the median teacher/student ratio decreased by 1.7 percentage points – from 9.1 to 7.4 percent. In the beginning of the 2000s teacher density rebounded somewhat. The decline during the 1990s is driven by the fact that the number of teachers has not kept pace with the increase in the student population. Between 1990/91 and 1999/00 teacher density declined by 0.4 percentage points holding the student population fixed at the 1990 value.

The literature on the effect of resources on outcomes is generally about class size. It is somewhat unfortunate, therefore, that there is no recent information on average class sizes available. The most recent information dates back to the beginning of the 1990s. In 1991/92, the average class consisted of 21.8 students. Let us make a rough translation of the resource development shown in Figure 4.1 into changes in class size. The numbers shown in Figure 4.1 implies that the number of students per teacher increased by 18 percent between 1991/92 to 1999/00. Thus, multiplying 21.8 with 1.18 gives a predicted class size of 25.8 students, i.e., an increase of four students per class. In 2002/03 the predicted class size, calculated analogously, equals 25 students.

In section 4.1 we surveyed the evidence pertaining to teacher quality. We documented that teachers are important, but that it was hard to explain the teacher effect by their educational attainment. Nevertheless, we think it is interesting to examine the evolution of the number of certified teachers per student; see the lower dotted line.[19] The figure illustrates that the increase in the teacher/student ratio starting in 1998/99 was accomplished via an increase in density of non-certified teachers (i.e. teachers without pedagogical training). In 2002/03 almost 19 percent of teachers do not hold a certification. The density of qualified teachers has declined rather dramatically – from 8.6 percent in 1990/91 to 6.2 percent in 2002/03 – during the period.

Unlike teacher inputs, computers/IT investments soared during the 1990s. This is shown in Table 4.1, which reports the number of students per computer in compulsory school since 1993. Following massive computer investments – starting in 1994 and continuing through 1999 – financed by "KK-stiftelsen",

[19] One aspect of teacher quality is to present the material in an interesting way. Pedagogical training presumably contributes to this quality, but it is not obvious that having a Masters rather a Bachelors degree does. We will analyze the development of teacher supply in closer detail in Chapter 5.

the student/computer ratio decreased from 38 in 1993 to 8 in 2001. Relative to the other countries of the OECD, Sweden ranks second (after Canada) in terms of computer intensity in primary school; Statistics Sweden (2002).

Table 4.1 The number of students per computer in compusory schools

	1993	*1994*	*1995*	*1996*	*1997*
# students per computer	38	19	13	10	8

Source: Söderlund (2000) and National Agency for Education (2001).

What should one make of the resource development during the 1990s, given previous research? The number of students per teacher has clearly increased, which, according to our reading of the received literature, is detrimental to student achievement. The number of students per computer has decreased, but there is little in the literature to suggest that this will benefit academic performance. All in all, then, the resource development during the 1990s has probably been detrimental to student achievement.

4.2.2 The distribution of resources

Has the distribution of resources across the country been affected by decentralization? A priori the effect of decentralization on the distribution of resources can go either way: on the one hand, the fact that there is greater leeway for municipalities to allocate expenditure on different items suggests an increase in the spread of the distribution; on the other hand, the fact that municipalities are held responsible for the development in compulsory schools to a greater extent, suggests that they pay closer attention to what goes on in other municipalities. The net effect of decentralization on the distribution of resources is a priori ambiguous.

While it is difficult to have a clear prior about the spread of the distribution, it is natural to expect some mobility within the distribution as a consequence of the reform. After all, the decentralization implied that there was a regime shift since the rule for allocating expenditure over municipalities was changed. We will examine whether this was the case. And in the next section we will exploit the regime shift in an attempt to estimate the effect of resources on student achievement.

Before starting the analysis, it is important to date the crucial year of the reform. Prior to 1991/92, the allocation of resources was determined through a strict system with ear-marked money. The central government – via the re-

gional schooling authorities (*länsskolnämnderna*) – more or less determined the amount of resources at the school level. For instance, the number of teachers at the school level was effectively determined by a central government grant. The municipalities had little freedom to allocate expenditure on different items within, e.g., the compulsory schooling system. This was changed in 1991/92. The municipalities still received a grant for, e.g., compulsory schools from the central government. However, the municipalities could freely allocate the money across schools and various items within the compulsory schooling system. Thus, the system is still one of ear-marked money but it contained more degrees of freedom for the municipalities. One should also note that the reform was implemented such that the real amount of resources for each municipality was the same as in the old system. The new system survived only one additional year. As of 1 January 1993, the money, previously ear-marked for education, was incorporated into an overall equalization grant. From this year and onwards, the municipalities, in principle, can freely allocate resources over its different responsibilities. Arguably, this is the big change in terms of the allocation of resources to compulsory (and upper secondary) schools over municipalities.

Then, let us start by looking at the overall spread of the distribution. Figure 4.2 shows percentile ratios of the distribution of the teacher/student ratio over municipalities. The solid line shows the 90/10 ratio, while the dotted lines "decompose" this ratio into 90/50 ratio and the 50/10 ratio. The spread of the distribution stood at a low in 1993/94. Since then there is an upward trend in the 90/10 ratio. The widening of the distribution appears to have taken place mainly at the bottom of the distribution. The 90/50 ratio is largely constant throughout the period; but there is an upward trend in the 50/10 ratio.

An interesting question when it comes to the distribution of schooling inputs is the association with mean income in the municipality. Has the relationship between resources and mean income changed over the years? Figure 4.3 reports the coefficients (solid line) from regressions of the (log of the) teacher/pupil ratio on the (log of) average municipal income; the dotted lines are 95 percent confidence bands. We run these regressions separately for each year from 1988 to 2000 and weight the regressions by the number of students in each municipality. Since we apply the weighting procedure, the regression coefficient captures the relationship between resources and income for the typical student.

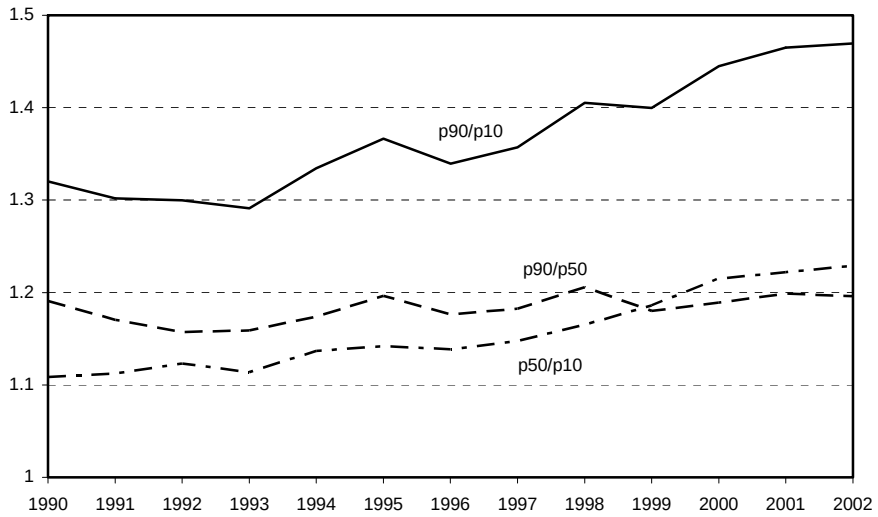

Figure 4.2 Percentile ratios of the teacher/student distribution, 1990/91-2002/03.
Source: Calculations using the Teacher register.
Notes: These numbers have been obtained by first calculating the population-weighted percentiles over municipalities and then forming the percentile ratios.

The allocation of resources appears to have been redistributive throughout. On average, a one percent increase in mean income is associated with a decrease in the teacher/student ratio by 0.1 percent. It is interesting to see the shift in the relationship between resources and income occurring in 1993. Thus, also when looking at the data it appears that the crucial regime shift took place in that year. As shown in the figure, the coefficient on income turns more negative in 1993. Since this time point marks the end of the system with ear-marked money it is difficult to have an idea about why this shift took place. In the longer run – i.e., comparing 2000 with the pre-1993 period – it seems that decentralization had little effect on the relationship between schooling inputs and income across municipalities.

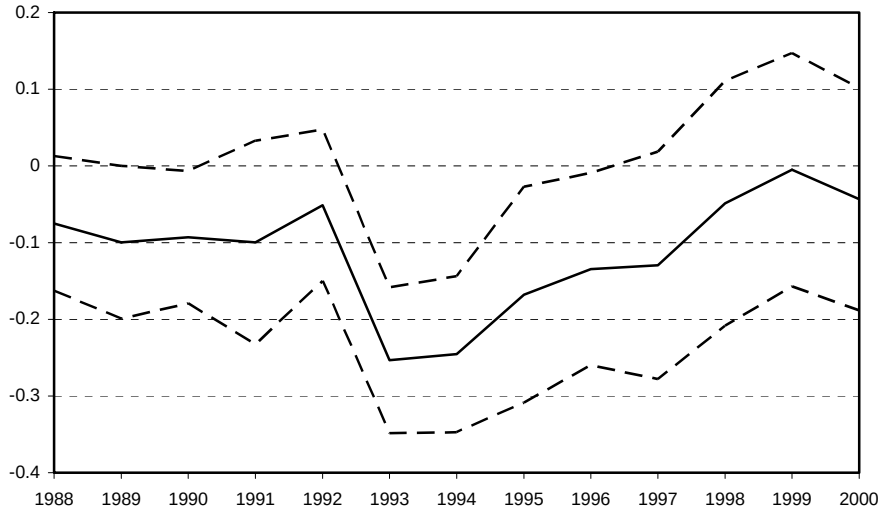

Figure 4.3 The relationship between teacher density and average income, 1988-2000.
Sources: Unpublished statistics and Official statistics (various issues of SM IF 20), Statistics Sweden.
Notes: The solid line shows the estimated coefficients from regressions of the log of teacher density on log income. The dotted lines show the estimates ± two standard errors. The regressions are run separately for each year and are weighted by the number of students in the municipality. The information on teacher density has kindly been supplied by Inge Göransson at Statistics Sweden. Income refers to the average of total income from gainful employment among individuals aged 16 and above.

Figure 4.3 suggests that the allocation of resources over municipalities changed in 1993. Let us try to substantiate this change further. We argued earlier that we should expect some mobility in the distribution of resources as a consequence of decentralization. Figure 4.4 examines this issue by reporting the rank correlation in teacher densities over time. The solid line shows the correlation with the rank in the preceding year (e.g., the entry for 1991 shows the correlation between 1991/92 and 1990/91). The dotted line shows the correlation with the rank three years prior to the year in question (e.g., 1991/92 with 1988/89).

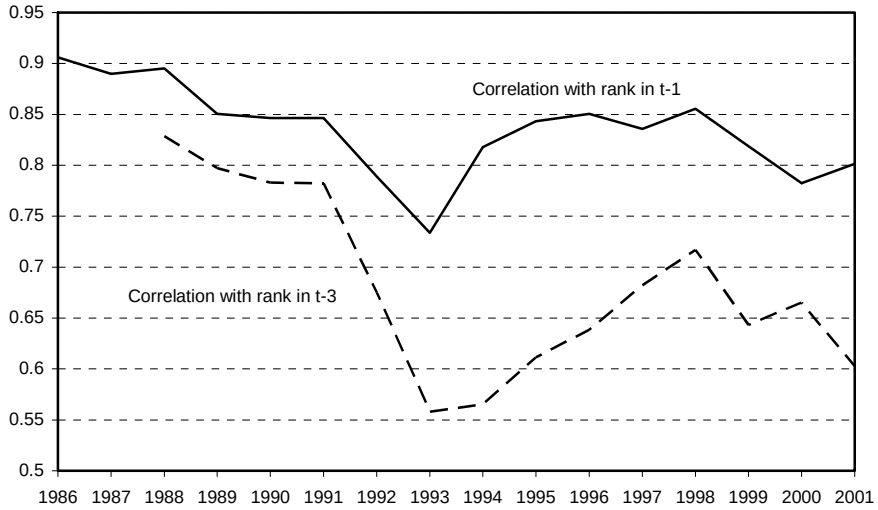

Figure 4.4 The rank correlation in the current and lagged teacher/student ratio, 1986-2001.
Source: Unpublished statistics, Statistics Sweden.
Notes: The solid line shows the correlation with the closest preceding year. The entry for, e.g., 1991 thus shows the correlation between the ranks in 1991/92 and the ranks in 1990/91. The dotted line shows the correlation with the rank lagged three years. The data has kindly been supplied by Inge Göransson at Statistics Sweden.

It is clear that the reform shifted municipalities around in the distribution of resources. The first solid line shows that the correlation is markedly lower for years involving 1993, i.e., 1992/93 and 1993/94. This is illustrated even more sharply by the dotted line in the figure. The correlations are substantially lower for years that come from different resource allocation systems. After the reform, the correlation between the preceding teacher density picks up in 1995/96 (which involves the correlation between 1995/96 and 1992/93) and continues to rise for a couple of years. Towards the end of the period, however, it seems that the stability of the distribution has declined somewhat. This is perhaps what one should expect given that the state of the municipal budget has become more important for the determination of schooling expenditures.

So far we have not taken into account that some of the change in the distribution of resources may be due to changes in the characteristics of students. Now, the characteristics of students change slowly over time so it is unlikely that they will account for the sudden shifts in Figure 4.3 or Figure 4.4.

Nevertheless, to give an accurate picture of the reform induced change in the distribution of resources, the characteristics of students and the municipality should be taken into account.

To calculate the change in the distribution of resources we proceed as follows. We take two years – one prior to the reform and one after the reform. As shown in Figure 4.4, the last year that was unaffected by the reform is 1991/92; we use the school year 2000/01 to characterize the situation after the reform. Then we estimate a regression relating resources in 1991/92 to a set of characteristics that we know are important determinants of resources prior to the reform. This regression represents an estimate of the resource allocation formula prior to the reform.[20] Then we use the estimated allocation formula to predict the amount of resources in 2000/01 given the characteristics of the municipality in 2000/01. The *deviation* of actual expenditures in 2000/01 from this prediction is a measure of the regime shift that takes the evolution of characteristics into account.[21]

Figure 4.5 presents the results generated by the above procedure. As the measure of resources we use the log. of teacher density. Since we are interested only in the distribution of resources, we have normalized the data such that they have same mean in the two years. The solid line depicts the deviation of the actual amount of resources in 1991/92 from the level predicted by the regression; the dotted line shows the same thing but for 2000/01. The regression relates teacher density to a set of characteristics of 9th graders in the municipality (e.g. the fraction of females, the fraction of foreign-born, and the share with university educated parents) and municipal characteristics (e.g.

[20] It would perhaps have been preferable to have information on the actual resource allocation formula. However, it is impossible to construct this formula, in part because resource allocation was subject to the discretion of the regional schooling authorities (*länsskolnämnderna*); see Du Rietz et al. (1987) for an excellent description of the system prevailing prior to decentralization.

[21] In slightly more formal terms we think of resources (R) in municipality m and year t = 1991/92, 2000/01 as being determined by a set of characteristics (x): $R_{mt} = \beta_t x_{mt} + \varepsilon_{mt}$. It is straightforward to decompose resources in 2000/01 into

$$R_{m(2000/01)} = \beta_{1991/92} x_{m(2000/01)} + (\beta_{2000/01} - \beta_{1991/92}) x_{m(2000/01)} + \varepsilon_{m(2000/01)} .$$

The first component reflects the level of expenditure had the previous regime prevailed. The sum of the second and third components is the deviation of resources from the predicted level had the previous regime prevailed.

average income, the average size of schools, and the density of students in the municipality).[22]

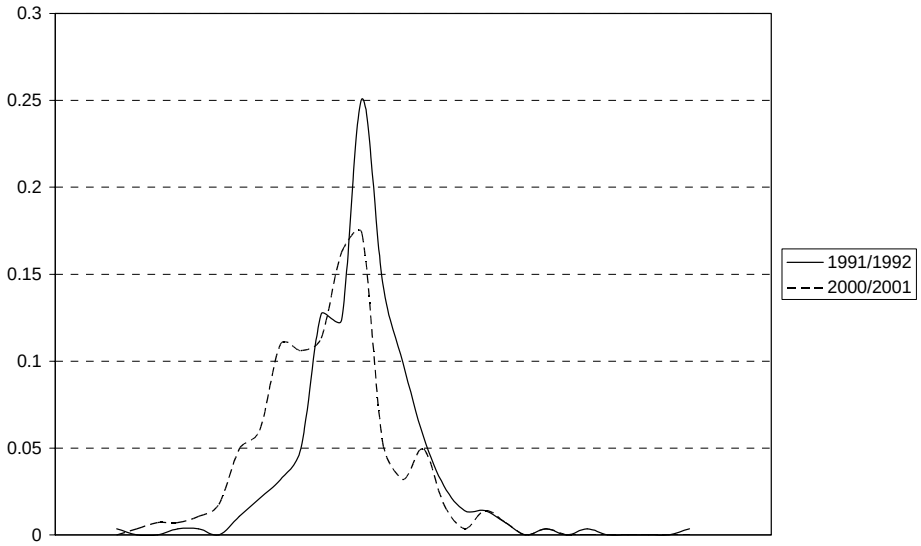

Figure 4.5 Deviations of resources from estimated resource allocation formula.
Sources: Calculations using the Teacher register, the Grade-9 register, and official statistics published by Statistics Sweden and the National Agency for Education.
Notes: The figure shows the densities of the deviations of log teacher density from the predicted level. Predicted teacher density is generated from a regression relating the log of teacher density in 1991/92 to the characteristics of 9th graders in the municipality and a set of municipality characteristics. The student characteristics include: the fraction of females, the fraction of foreign-born, the fraction with two foreign-born parents, the fraction of students having immigrated in the preceding five years, the fraction of students with at least one high-school educated parent, the fraction of students with at least one university educated parent. The municipal characteristics include: average income, the average size of schools, the number of individuals aged 7-15, and the (geographical) size of the area. All municipal characteristics are measured in logs.

[22] According to the regulations, the most important form of central government support for schooling costs were so called "base resources". The number of base resources was determined by something like "Maimonides' rule"; see Angrist and Lavy (1999) for a clever application of this rule. A school having 25 students in a particular grade at the primary level would be given one base resource, while a school having 26 students in a grade was given two base resources. The variations induced by this rule were to some extent mitigated by particular clauses and the regional schooling authorities. The size of each base resource was fairly closely linked to regional costs. Extra resources were granted to municipalities with a high fraction of students in need of extra teaching in Swedish and teaching in their mother tongue. Additional support was also given to rural municipalities; see Du Rietz et al. (1987). We think that the characteristics we include in the regression capture these various aspects of the resource allocation system well.

As shown by Figure 4.5, the reform induced a shift of the distribution. The meaning of this shift is, again, that decentralization changed the allocation of teacher inputs over municipalities. Taking the evolution of the observed characteristics into account, the distribution of resources is clearly different before and after the reform.

4.3 Resources and achievement: Swedish evidence

As argued in section 4.1, it is notoriously difficult to estimate the "true" effect of resources on student performance. The difficulties stem from compensatory behavior on the part of educational authorities. Credible evidence presumably requires some kind of experimental – either explicit or quasi – variation. In this section we make use of a variation that is potentially exogenous – the change in schooling expenditure induced by the decentralization reform. We thus ask the question: Did the change in schooling inputs induced by the reform have an effect on student performance?

Although the reform clearly represents an exogenous shift in the distribution of resources over municipalities, it is not evident that the induced resource change is exogenous to student achievement. Exogeneity in the latter respect might not hold if the local governments target students that are weaker in the unobserved sense differently than the central government. With this caveat in mind, let us proceed to the evidence.

4.3.1 The effect of resources on achievement

The basic strategy is to calculate difference-in-differences estimates of the effect of the reform. As the outcome measure we use the percentile ranked grades (for 9th graders). The difference-in-differences approach amounts to comparing the average performance of students leaving compulsory schools in a particular municipality in 2001 to those leaving compulsory schools in 1992 in the same municipalities and to relate this change in performance to the change in resources. We also standardize for a set of student and municipal characteristics. Therefore, the variation we are utilizing to estimate the effect of resources is akin to the shift shown in Figure 4.5.

The principal data set used for estimating the effects of the resource change consists of pooled data on individual 9th graders leaving compulsory school in the spring of 1992 and 2001. These data contain the grade point averages of these students along with their individual and family background character-

istics. We match information about the (resident) municipality onto these data. The key variables at the municipal level are measures of schooling inputs such as the teacher/student ratio and teaching expenditures per student.

Table 4.2 presents a set of regression results. The dependent variable is the percentile ranked grade point average. The regression standardizes for gender, age, whether the individual is a first or second generation immigrant, whether (s)he immigrated during the five years directly preceding graduation, and the educational attainment of the parents. The regressions also include the (log. of) average municipal income, the average size of schools in the municipality, the size of the student population, as well as municipality and time fixed effects. The fact that we include municipality fixed effects implies that we are using the variation within a municipality over time to estimate the effects. The time effects control for resource changes that are common across municipalities (e.g., the resource change induced by the downturn of the cycle).

We begin with a specification which links student performance to teaching expenditure per student. The bulk of the variation in teaching expenditure is driven by teacher wages and the number of teachers per student. It is potentially important to allow for an effect of wages on student outcomes since higher wages may attract abler teachers which, in turn, have positive effects on student achievement. The specification in column (1) is based on a decomposition of teaching expenditure per student into the number of teachers per student and a remainder, i.e., mostly wages. It turns out that only the number of teachers is important; variations in "wages" have no effects. The estimate on the teacher/student ratio implies that a ten percent increase in the number of teachers per student improves student performance by three quarters of a percentile point.

Column (2) asks whether it matters whether teachers are certified or not. Since the estimates are not statistically different from each other the answer to this question is no. The results in column (1) and (2) thus lead to the parsimonious specification in column (3), which relates student outcomes only to the log. of teacher density.[23] Multiplying the estimate reported in column (3) with the actual decline in teacher density between 1990/91 and 2002/03

[23] The careful reader will have noticed that the sample sizes are different in columns (1) and the remaining columns. With the sample in column (1) we estimate the effect of the log of teacher density to be 7.01, which is virtually identical to the estimate presented in column (3) with a larger sample.

(around 17 log. points) implies that the actual resource change lowered student performance by 1.2 percentile points.

Table 4.2 The relationship between student achievement and resources. Dependent variable: percentile ranked grade point average.

	(1)	*(2)*	*(3)*	*(4)*
ln(# teachers per student)	7.45		7.05	6.09
	(2.43)		(2.36)	(2.06)
ln(teaching expenditure per student)-	1.04			
ln(# teachers per student)	(0.34)			
ln(# certified teachers per student)		6.84		
		(2.30)		
Share of non-certified teachers		4.22		
		(0.94)		
ln(# teachers per student) *				28.92
(immigrated within last 5 years)				(4.34)
ln(# teachers per student) * (no				3.84
parent with high-school ed. or more)				(2.07)
# individuals	192,017	196,952	196,952	196,952
# municipalities	265	283	283	283
Adjusted R-square	0.1847	0.1848	0.1848	0.1850

Notes: The *t*-ratios in parentheses are based on standard errors that allow for arbitrary correlation between individuals residing in the same municipality. The regressions are based on pooled individual data from the Grade-9 register in 1992 and 2001. The share of non-certified teachers in column (2) is measured as ln(1+(# non-certified teachers)/(# certified teachers)). The teacher/student ratio is measured in 1991/92 and 2000/01 and are based on the Teacher register. Teaching expenditures pertain to 1992 and 2001 and have been downloaded from the website of the National Agency for Education. The regressions control for gender, age, age squared, whether the individual was foreign born, whether the individual immigrated within the last five years, whether both parents are foreign born, whether there is at least one parent with upper secondary education, whether there is at least one parent with a university education, a time fixed effect, municipality fixed effects, the log of the average income in the municipality, the log of average school size, and the log of the number of pupils in the municipality.

Alternatively, we can link these estimates to the ones in Krueger (1999) and Lindahl (2001). The average class size in 1991/92 was around 22 students and the median teacher/student ratio stood at 8.8 percent. Suppose that class size is reduced by seven students, roughly the magnitude in STAR. This corresponds to an increase of teacher density by 4.1 percentage points (around 59 log.

points).[24] Evaluated at this point, the estimates suggest that the effect of reducing class size by seven students is a gain in achievement of 4.2 percentile ranks. This gain corresponds to an increase by 0.14 standard deviations, which is within the range of the estimates reported in Lindahl (2001).

An important question when it comes to resource changes concerns their distributional effects. Column (4) reports estimates that address this issue. In addition to the main effect of teacher density we have interacted teacher density with indicators reflecting disadvantageous family background. The regression includes an interaction term with low parental education and an interaction term with recent immigrant status.[25] These estimates clearly indicate that students from disadvantageous family backgrounds are more susceptible to variations in resources. For instance, a ten percent reduction of teacher density lowers the performance of recent immigrants by 3.5 percentile points. This result concurs with the conclusions in Krueger (1999). Notice also that Lindahl (2001) found that class size mattered more for immigrant than native children. However, the class size effects in Lindahl did not vary by socio-economic status in general.

4.3.2 Resources and the between school variance in achievement

We have documented above that resource changes do affect student outcomes. Also, we have found that students from disadvantaged background are particularly susceptible to variations in resources. This suggests that the variance in outcomes has increased as a consequence of the decline of teacher density during the 1990s.

Another way of examining whether the decline in resources has affected the variance in outcomes is to look at the evolution of the between school variance over time. This is what is done in Figure 4.6. The top solid line shows the between school variance in percentile ranked grade point averages in compulsory schools. The bottom line shows the same thing for percentile ranks standardized for student characteristics. By comparing these two lines we get an idea about whether changes in segregation in terms of student characteristics are important for the changes in the raw between school variance.

[24] We are well aware that we are almost outside the range of the data. There were only two municipalities in 1991/92, and none in 2000/01, that had a teacher density of the implied magnitude.

[25] We also tried interaction terms with immigrant status and having foreign born parents without finding any statistically significant effects.

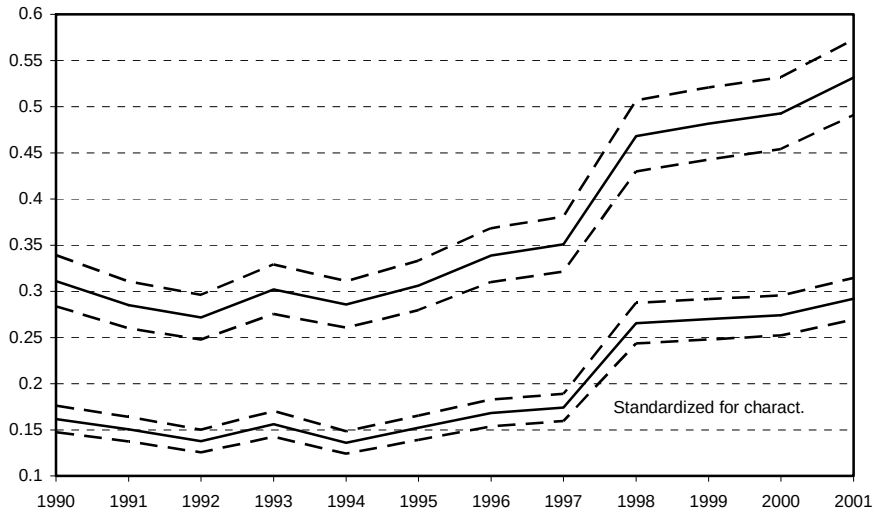

Figure 4.6 The evolution of the between school variance in outcomes, 1990-2001.
Source: Calculations based on the Grade-9 register 1990-2001.
Notes: The dotted lines are confidence bands. The upper solid line shows the "raw" between school variance. The lower solid line shows the between school variance after standardizing for student characteristics. To be consistent over time we have excluded students with no grade point average.

The most striking feature in the figure is the upward jump in the between school variance between 1997 and 1998. This jump coincides with the introduction of the new grading system. The question is if this jump is just a mechanical effect reflecting the new grading system. Now, remember that we are converting the grades to percentile ranks. Due to this transformation, the total variance (the sum of the within and between school variance) is the same over time. Since we are normalizing the overall variance, the jump cannot be mechanical. However, it is difficult to argue that the sudden change is related to an increase in the variance of resources or segregation – these processes are presumably smooth. One possible interpretation of the change is that the new grading system has increased the scope for school policies in terms of determining what constitutes a "pass" or a "pass with distinction". In the late 1990s, it matters more what school you attend because the school interpretation of absolute performance varies more than the school interpretation of relative performance.

What about the trends in the between school variance? There is some evidence of an increase in the variance. The variance (standardized for student characteristics) increased by a third from 1992 to 1997 (from 0.152 to 0.201). Also there is some evidence suggesting that sorting on characteristics has become more important since the mid 1990s. The slope of the top line appears to be decidedly more positive than the slope of the bottom line after 1995. This was not the case during the first half of the 1990s.

4.4 Summary and conclusions

The 1990s saw a decline of teacher density in compulsory schools, from 9.1 percent in 1990/91 to 7.4 percent in 1999/2000. Since the turn of the century the number of teacher per pupil has rebounded somewhat. The number of certified teachers, however, declined precipitously throughout the period.

The reduction in inputs is of course driven by the downturn of the cycle to a large extent. It is impossible to tell if decentralization also contributed to the resource decline. However, decentralization and the subsequent abolition of ear-marked central government money for schools did shift the distribution of resource inputs over municipalities.

Our literature review and the new evidence we have presented also suggest that the resource decline matters. Increases in teacher density appear to improve student outcomes on average. The size of the average effect concurs with the estimates reported in Lindahl (2001). The effect is larger for students from disadvantaged family backgrounds – recent immigrants in particular – which is in agreement with the results in the previous literature. The fact that the effect varies by family background characteristics implies that the resource decline contributes to the upward trend in the between school variance in student outcomes that we observe in the data.

So, should we increase the number of teachers in schools on the basis of the evidence presented here? If we take the social rate of return calculation presented in Krueger and Lindahl (2002) literally the answer is a qualified yes. A rate of return of five percent must be considered high among the possible public investment projects. The qualification is of course that their calculation involves some assumptions regarding the unknowns of the cost-benefit calculation. In addition, it is not without problems to argue in favor of reductions in class size, given the predicted future shortages of teachers.

References

Angrist, J. and V. Lavy (1999), Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement, *Quarterly Journal of Economics*, 114, 533-575.

Angrist, J. and V. Lavy (2002), New Evidence on Classroom Computers and Pupil Learning, *Economic Journal*, 112, 735-765.

Borman, G. and L. Rachuba (2001), Evaluation of the Scientific Learning Corporation's Fast ForWord Computer-Based Training Program in the Baltimore City Public Schools, Report Prepared for the Abell Foundation.

Du Rietz, L., U. Lundgren, and O. Wennås (1987), Ansvarsfördelning och styrning på skolområdet, DsU 1987:1, Utbildningsdepartementet, Stockholm.

Ferguson, R. (1991), Paying for Public Education: New Evidence on How and Why Money Matters, *Harvard Journal of Legislation*, 28, 465-498.

Finn, J. and C. Achilles (1990), Answers and Questions about Class Size: A Statewide experiment", *American Educational Research Journal*, 27, 557-577.

Gustafsson, J-E. (2003), What do We Know about Effects of School Resources on Educational Results?, *Swedish Economic Policy Review*, forthcoming.

Gustafsson, J-E. and E. Myrberg (2002), Vad betyder ekonomiska resurser för pedagogiska resultat? En kunskapsöversikt, Skolverket, Stockholm.

Hanushek, E. (1986), The Economics of Schooling: Production and Efficiency in Public Schools", *Journal of Economic Literature*, 24, 1141-1177.

Hanushek, E. (1997), Assessing the Effects of School Resources on Student Performance: An Update, *Educational Evaluation and Policy Analysis*, 19, 141-164.

Hanushek, E. (2003), The Failure of Input-based Schooling Policies, *Economic Journal*, 113, 64-98.

Hedges, L., R. Laine, and R. Greenwald (1994), Does Money Matter? A Meta-Analysis of Studies of the Effects of Differential School Inputs on Student Outcomes, *Education Researcher*, 23, 5-14.

Kirkpatrick, H. and L. Cuban (1998), Computers Make Kids Smarter – Right?, *Technos Quarterly for Education and Technology*, 7, http://www.technos.net/tq_07/2cuban.htm.

Krueger, A. (1999), Experimental Estimates of Educational Production Functions, *Quarterly Journal of Economics*, 115, 1239-1285.

Krueger, A. and M. Lindahl (2002), *The School´s Need for Resources – A Report on the Importance of Small Classes*, ESO, Ds 2002:12, Stockholm.

Krueger, A. and D. Whitmore (2001), The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project STAR, *The Economic Journal*, 111, 1-28.

Krueger, A. (2003), Economic Considerations and Class Size, *Economic Journal*, 113, 34-63.

Lindahl, M. (2001), Home versus School Learning: A New Approach to Estimating the Effect of Class Size on Achievement, Discussion Paper No. 261, IZA.

Nye, B., J. Zaharias, B.D. Fulton, et al. (1994), The Lasting Benefits Study: A Continuing Analysis of the Effect of Small Class Size in Kindergarten through Third Grade on Student Achievement Test Scores in Subsequent Grade Levels, Seventh grade technical report, Center of Excellence for Research in Basic Skills, Tennessee State University.

OECD (1994), *Education at a Glance*, Organization for Economic Cooperation and Development, Paris.

OECD (2002), *Education at a Glance*, Organization for Economic Cooperation and Development, Paris.

Rouse, C. and A. Krueger (2003), Putting Computerized Instruction to the Test: A Randomized Evaluation of a 'Scientifically-based' Reading Program, manuscript, Princeton University, April.

National Agency for Education (2001) *Skolans datorer 2001*, Skolverkets rapport nr 208.

Statistics Sweden (2002), *Swedish Education in International Statistics.*

Strauss, R. and E. Sawyer (1986), Some Evidence on Teacher and Student Competencies, *Economics of Education Review*, 5, 41-48.

Söderlund, A. (2000), *Det långa mötet, IT och skolan*, Doktorsavhandling, Institutionen för lärarutbildning, Luleås tekniska universitet.

Todd, P. and K. Wolpin (2003), On the Specification and Estimation of the Production Function for Cognitive Achievement, *Economic Journal*, 113, 3-33.

Wenglinsky, H. (1998), Does it compute? The Relationship between Educational Technology and Student Achievement in Mathematics, Princeton, NJ: Educational Testing Service. (ERIC Document Reproduction Service No. ED425191).

5 Teacher supply

In chapter 4 we reviewed the literature on the importance of teacher qualifications for student achievement. The quality of teachers is clearly one of the most important inputs for achievement, although there is some controversy in the literature about what the crucial characteristics of teachers are. The purpose of this chapter is to present some evidence pertaining to change in the quality of teacher supply to Swedish schools.

Given the decrease in the public sector wage premium (see le Grand et al, 2001), one might expect that the monetary incentives for becoming a teacher have declined over time. Also, working conditions have probably become worse in Swedish schools. At least one is inclined to infer this from the development of teacher density presented in chapter four and media reports. These developments most likely contribute to the predicted teacher shortages in the years to come; see National Agency for Education (2003) on the demand for teachers until 2020.

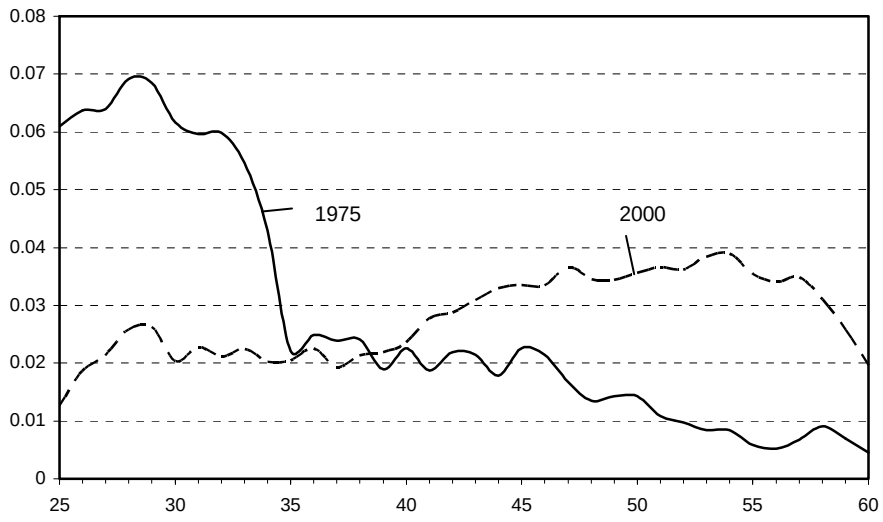

Figure 5.1 The age distribution of certified teachers in 1975 and 2000.
Source: Calculations based on LINDA.

Figure 5.1 presents some evidence relating to the demographic development in the teacher profession over the past 25 years. The figure suggests that

teacher supply has been reduced rather drastically over the time period. In the past 25 years, the median age of certified teachers has increased from 32 years-of-age in 1975 to 46 years-of-age in 2000. Moreover, the supply of certified teachers seems to have been reduced relative to individuals with a similar educational attainment. The median university graduate was aged 33 in 1975; in 2000 the median age had increased to 40. In this chapter we will direct attention to the underlying forces yielding a picture such as the one below.

5.1 The incentives to become a teacher

We begin by looking at the development of the incentives for becoming a teacher. In Figure 5.2 we plot the evolution of teacher relative wages over the past eighty years. Relative to the wages of production workers, teacher wages have declined precipitously since the beginning of the 1940s.[26] Since 1945, the decline in the relative wage amounts to almost 50 percent. The evolution of the relative wage is less dramatic over the past thirty years. Still there are no signs of a turning tide during this period.

The decline of teacher relative wages is not a unique Swedish phenomenon. For the countries where there are data, it is universally true that teacher relative wages have declined since the mid 1960s. However, the international evidence suggests that the decline in Sweden was particularly sharp; see Lakdawalla, 2001. Today, experienced Swedish teachers are paid less than their Nordic and OECD counterparts (OECD, 2002). For instance, experienced Swedish teachers in lower secondary school earned 82 percent relative to the average OECD counterpart. The relative wage gap is lower in primary schools and it is lower when it comes to starting wages.

Although Figure 5.2 is informative about the long-run development, it is susceptible to a number of objections. For instance the share of females in teaching has increased over time. Also, the development of the teacher relative wage shown in the figure may be the result of a decline in the overall university wage premium.

Of course there are other potential comparison occupations. Nurses and teachers have similar gender composition and education level. Over the past

[26] We thank Anders Nilsson for supplying the pre 1975 data on teacher wages.

thirty years, teacher wages have declined relative to nurses as well: the relative wage decreased by 20 percent between 1970 and 2001.[27]

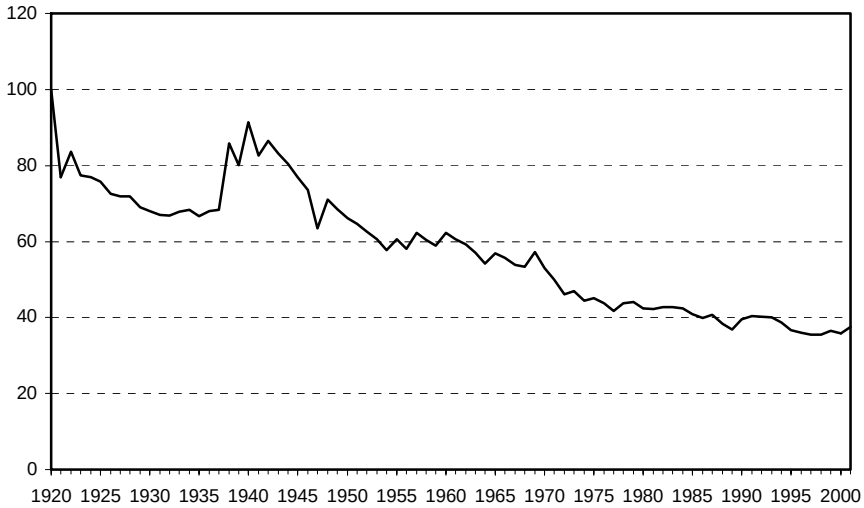

Figure 5.2 Index of teacher wages relative to production worker wages, 1920-2001, 1920=100.
Sources: Nilsson (1984); Swedish Official Statistics, Wages, (*SOS Löner*) 1975-1989; Statistical Yearbook of Sweden, various issues; Swedish Teachers' Union wage statistics.
Notes: During the period 1920-89, teacher wages refer to the wage in the highest wage category for teachers in primary schools (*folkskola*) or teacher in grades 4 to 6 of the comprehensive school. We have used information on the growth rate of average teacher wages to impute comparable wages from 1990 to 2001. Anders Nilsson has kindly supplied data on teacher wages for the period 1920-74; see Nilsson (1984) for more information about the definition of this series. We thank Sune Johansson at the Swedish Teachers' Union (*Lärarförbundet*) for supplying information on average teacher wages. Production worker wages in general refer to mining and manufacturing, although the included industries vary somewhat over time.

Nevertheless, comparisons with single occupations do not answer the most relevant question: How have the incentives to opt for teaching, given a decision to opt for university education, evolved over time? Figure 5.3 examines this question.

Figure 5.3 presents the standardized (annual) earnings premium for male and female teachers relative to those with the same educational attainment; see

[27] See Statistics Sweden's wage statistics (*Kommunal personal* 1970 and *Statistical Yearbook of Salaries and Wages* 2001) for information on the wages of nurses.

Figure 5.3.[28] The calculations are based on the LINDA data base (see Edin and Fredriksson, 2000).

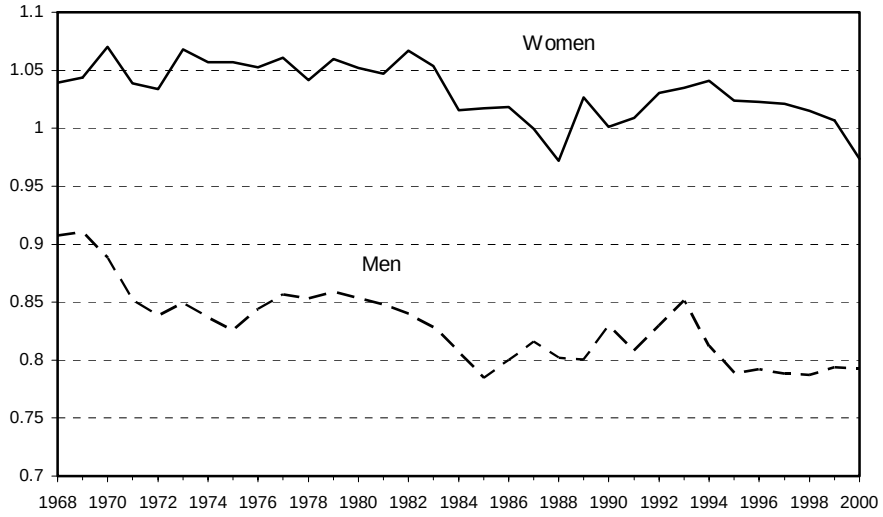

Figure 5.3 Teacher earnings relative to those with the same educational attainment, 1968-2000.
Source: LINDA 1968-2000.
Notes: When making this calculation we have standardized with respect to age and educational attainment (short and long university education).

Female teachers (the vast majority of teachers) saw no earnings decline relative to other females with a similar educational attainment during the 1970s.[29] There is a downward jump for female teachers around the mid 1980s, however, when their relative earnings were reduced by five percent.

[28] We would have preferred to use wage data, but data for a representative comparison group is only available from 1995 and onwards. The evolution of the gender specific relative teacher wage during 1995-2000 is consistent with evolution shown in Figure 5.3.
[29] During the time period the share of females among those with a teacher education increased from 66 percent in 1968 to 78 percent in 2000.

It is noteworthy that female teachers have earned more than females in other professions for most of the period.[30] Things are just the opposite for male teachers. In the late 1960s male teachers earned ten percent less than men with a similar educational attainment. There is a downward trend in the relative earnings of male teachers. At the turn of the century their relative earnings have fallen to less than 80 percent of the earnings of men in other professions.

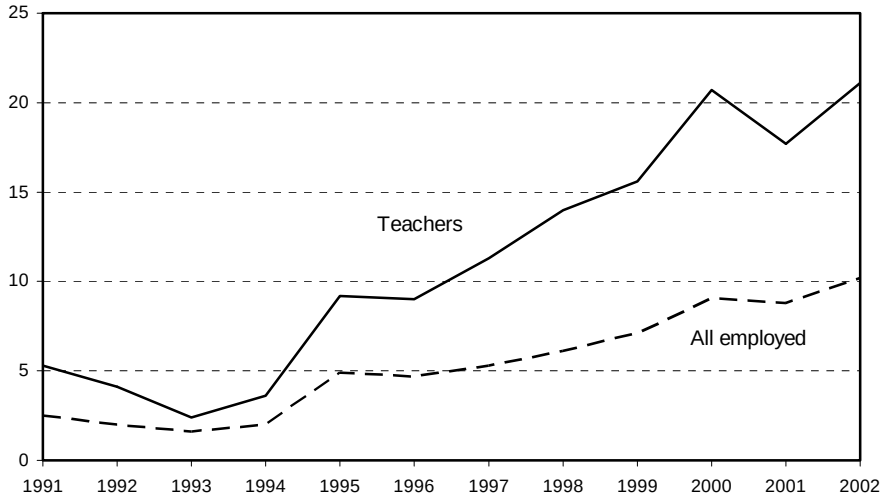

Figure 5.4 Percent of the employed reporting health problems due to mental stress.
Source: Work related health problems 1991-2002, Swedish Official Statistics, Statistics Sweden.
Note: The figures pertain to teachers at the compulsory, upper secondary, and university level.

So let us turn to the non-monetary incentives to opt for the teacher profession. In Figure 6.4, we present the percentage of employed teachers suffering from problems due to mental stress at the workplace. For comparative purposes we also report this number for all employed. During the first half of the 1990s there is a small difference between teachers and the average employee. From 1995 and onwards the gap between teachers and other employees

[30] Comparing relative wages and relative earnings for female teachers it seems that female teachers are less inclined to work part time than females in other professions. No such differences exist for males.

widens rather drastically. In 2002, 21 percent of teachers report suffering from mental stress and 10 percent of all employees report having the same problem.

It is relevant to ask whether the differential change is related to the gender and age composition of the teacher profession. Unfortunately, we do not have the information to adjust these figures for potential differences in gender and age. It seems highly unlikely, however, that the differences shown in Figure 6.4 would disappear when standardized with respect to age and gender.

These numbers are so interesting that they deserve closer investigation. Table 5.1 reports the shares of teachers, white collar employees, and all employed suffering from mental stress and work related health problems in general (excluding accidents) for the years 1991 and 2002. The table shows that there is a trend increase in mental stress and health problems in all categories.[31] Since we are interested in the development of the relative attractiveness of the working conditions in teaching, we eliminate the trend in the comparison groups. The result is shown in the last two rows of Table 5.1.[32] The next to last row shows that there is a significant relative increase of the mental stress in teaching. Whether work related health problems in general have increased depends on the comparison group. The most relevant comparison group is arguably the average white collar worker and, therefore, we conclude that there is no relative increase in worker related health problems as a whole.

So let us sum up this subsection. The relative wages of teachers have declined for a long time. There is nothing special about the 1990s in this sense. During the 1990s, the earnings of teachers have declined by 2-3 percent relative to individuals with a similar educational attainment. It seems that working conditions have become substantially worse in Swedish schools during the 1990s. We think that the most likely candidate for explaining this development is the reduction of teacher density described in chapter 4. Indeed, in a review of the literature of teachers' perceptions about the effects of class size, Granström (1998) concludes that larger classes are associated with higher levels of mental stress for teachers.

[31] There are, of course, many potential reasons for this increase. The list of candidates includes: an ageing population, greater job insecurity, public sector cut-backs, and a greater propensity to report health problems.

[32] In technical terms these are difference-in-differences estimates.

Table 5.1 Mental stress and work related health problems for different categories in 1991 and 2002, percent of the employed in each category.

		Mental Stress	*Work related health problems (other than accidents)*
All employees	2002	10.2	24.6
	1991	2.5	18.0
	Change	7.7	6.6
		(0.2)	(0.3)
White collar workers	2002	13.1	23.8
	1991	3.1	13.6
	Change	10.1	10.2
		(0.3)	(0.5)
Teachers	2002	21.1	29.7
	1991	5.3	19.7
	Change	15.8	10.0
		(1.3)	(1.6)
	Change relative to all employees	8.1	3.4
		(0.5)	(0.4)
	Change relative to white collar workers	5.7	-0.2
		(0.5)	(0.4)

Source: Work related health problems 1991, 2002, Swedish Official Statistics, Statistics Sweden.
Notes: Standard deviations in parentheses.

5.2 The supply of skills to the teacher profession

To our knowledge, there is no research about the response of teacher supply in Sweden to variations in incentives. The international evidence, however, suggests that teacher supply do respond to monetary incentives. Wages influence the decision to become a teacher as well as the decision to remain a teacher; see Dolton (1990) and Murnane et al. (1991). There is also Swedish evidence suggesting that the overall demand for university education responds to monetary incentives; see Fredriksson (1997).

There is even less evidence on the impact of working conditions on teacher supply. However, after reviewing the literature on resources on student achievement Gustafsson (2003) suggests that "[w]hile the direct effects of class size on student achievement may be too weak to justify class size reductions, the indirect effect via the influence on teacher competence may provide a justification for class size reduction." The premise for this argument is that

class size is important for the working conditions of teachers and that supply responds to variations in working conditions.

This section asks the question: Has the development documented in the previous section had any consequences for the supply of skills to the teaching profession? We present two sets of evidence pertaining to this issue. First, we look at the number of applicants per slot and the evolution of admittance grades to teacher education. Second, we examine whether there are changes over time in the performance on standardized tests for those who later on opted for teacher education.

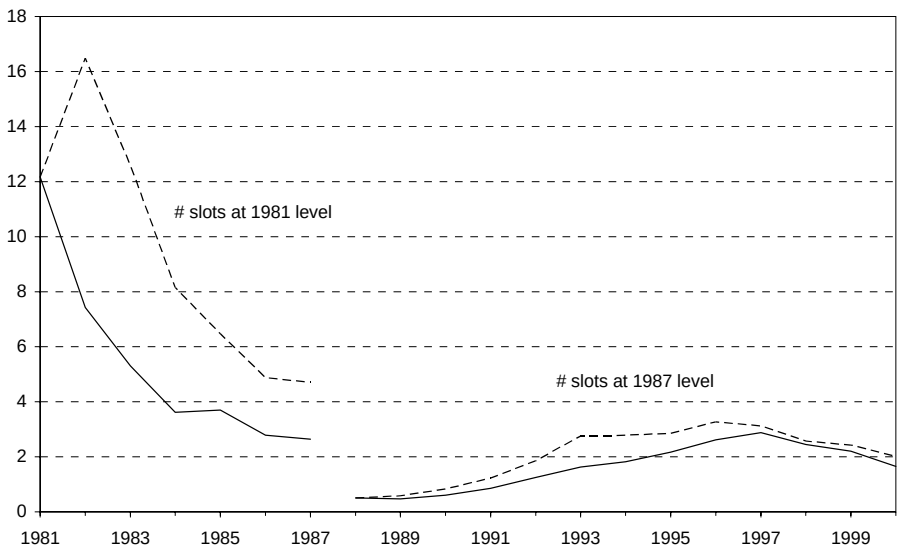

Figure 5.5 The number of qualified first-hand applicants per slot (solid line), 1981-2000.

Sources: Application statistics (Antagningsstatistik), various issues, UHÄ and VHS.

Notes: The dashed line shows the number of applicants per slot when the number of slots is held constant. Teacher education was reformed in 1988. During 1981-87 the numbers refer to "Lågstadielärarlinjen"; for 1988-2000 they refer to "Grundskollärarlinjen 1-7". In 1981 there is no information on the number of qualified first-hand applicants. The 1981 value has been imputed using information on the total number of qualified applicants in 1981 and the average share of qualified first-hand applicants in the total during 1982-87.

Figure 5.5 presents the number of qualified applicants per slot.[33] At this stage, we restrict attention to the education of teachers for the lowest level of elementary school (From 1988 and onwards this is equivalent to the primary level.). The time series break in the figure is due to a reform of teacher education in 1988; a new reform was launched in 2001 which is the reason for not extending the graph beyond 2000.

Figure 5.5 shows a drastic reduction in the popularity of teacher education during the 1980s. The decline during the 1980s is not due to an expansion in the number of slots (see dashed line). In fact, the number of slots was reduced for the most part of the 1980s. During the 1990s, there appears to be a slight rebound in the demand for teacher education at the primary level. However, it is possible that this rise is driven by the business cycle; as unemployment rose in the beginning of the 1990s, the overall demand for university education presumably increased.

To eliminate the variations due to the overall popularity of university education it is natural to examine the *relative* demand for teacher education. Figure 5.6 shows the number of qualified first-hand applicants to the sector for pedagogical training[34] (*sektorn för undervisningsyrken*) relative to the overall number of qualified first-hand applicants. According to Figure 5.6, the relative demand for teacher education indeed increased during the first half of the 1990s after a decade of decline during the 1980s. Since the mid 1990s, however, the relative popularity has declined yet again. In the beginning of the 2000s, relative demand is even lower than around 1990.

Thus, there is a secular decline in the popularity of teacher education during the past two decades. What has happened to the formal qualifications of those opting for teaching? Figure 5.7 sheds light on this question by presenting the admittance grades for those at the margin of entering teacher education. Teacher education was reformed in 1988. From 1988 and onwards there are two tracks for teachers at the primary level. They can either choose Swedish/Social Science or Math/Science as their major subject. Before 1988 there was only one track. From 1988 and onwards we show the marginal entry grade for those majoring in Swedish/Social Science. We do not show the evolution of entry grades for those majoring in Math/Sciences for a very simple reason:

[33] Applicants to programs, such as teacher education, are admitted at the central level.
[34] This is a wider concept than the one that we analyzed in Figure 5.5. We choose this measure in order to be better able to compare the periods before and after the reforms in 1988 and 2001.

everyone that applied was admitted! Figure 5.7 illustrates that the formal qualifications of the *marginal* entrant declined along with the demand for teacher education.

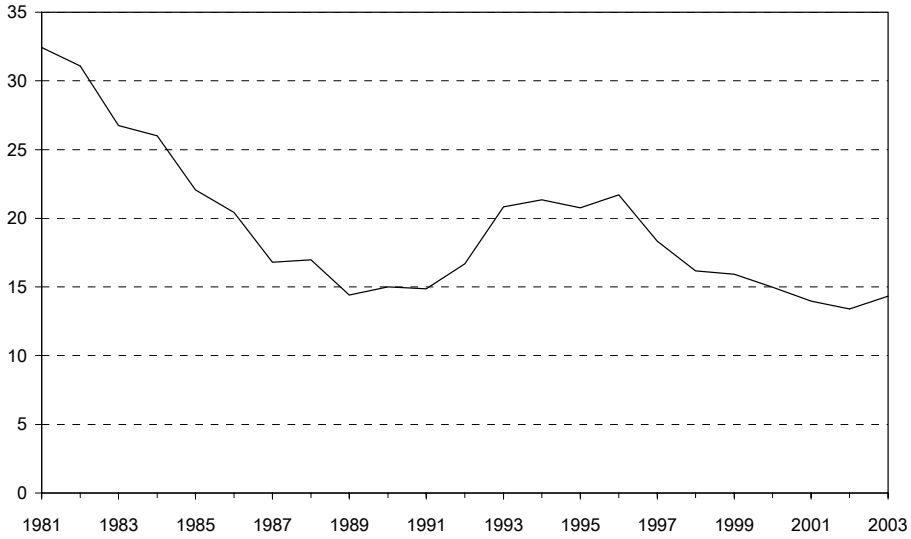

Figure 5.6 The relative number of qualified first-hand applicants to pedagogical education, 1981-2003.
Source: Application statistics (*Antagningsstatistik*), various issues, UHÄ and VHS.
Notes: Number of qualified applicants relative to the overall number of qualified first-hand applicants to university education. The 1981 value has been imputed using an analogous procedure as in Figure 5.5. We wish to thank Linn Brohmé at VHS for supplying data for 2001-2003.

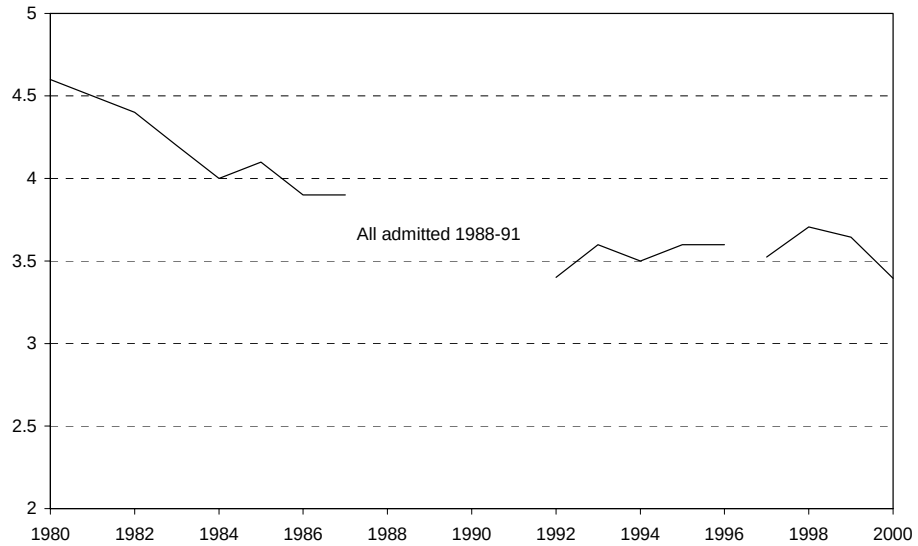

Figure 5.7 Entry grades for the marginal entrant to primary teacher education, 1980-2000.

Sources: Application statistics (*Antagningsstatistik*), various issues, National Agency for Education reports (Skolverkets rapporter nr. 135, 157, 173, 192), and Grade 9 register 1998.

Notes: During 1980-1987 we show the marginal entry grade for those admitted to "Lågstadielärarlinjen". During 1988-2000 we show the marginal entry grade for those entering "Grundskollärarlinjen 1-7" with a major in Swedish/Social Sciences. The grading system was reformed in 1997. We have standardized the new grading system such that it has mean 3 and a unit variance. To estimate the standard deviation we used the Grade 9 register excluding 2.8 percent of the lower tail of grade point average distribution. This number corresponds to the share not attending upper secondary school.

So let us have a look at the evolution of average teacher quality over the longer run. The Departments of Education in Göteborg and Stockholm have conducted tests of a random sample of 6th graders for the cohorts born in 1948, 1953, 1967, and 1972; see Härnqvist (1998) for a description of the data. Among other things these data contain the scores on verbal, inductive, and spatial ability tests. The three ability tests have been used in identical forms for these cohorts. To the data, information on the educational attainment and the type of education later on in life has been matched.[35] Using the matched data

[35] We are very grateful to Jan-Eric Gustafsson and Åsa Berndtsson for letting us access and helping us with the data.

sets we can examine the relative performance on tests at age 13 for those who later on decided to take teacher education.

These cohorts made their career choice at very different time points. The majority of those born in 1948 presumably made their career choice around 1970 while those born in 1972 made their choice in the mid 1990s. The question is if the trend decline in teacher relative wages has affected the relative position in the ability distribution of those who enter teacher education. To calculate this relative position we use the sum of the scores from the three tests mentioned above.[36] We refer to the sum of scores as ability scores for short.

Table 5.2 presents the results of these calculations; see Nickell and Quintini (2002) and Evans et al. (2002) for similar calculations for the UK and US respectively. For comparative purposes we also report the rank for the average university graduate. For both teachers and other university educated individuals we focus on those with at least three years of tertiary education.

Let us begin with the column for university education. In the cohort born in 1948, the average college educated individual had a percentile rank of 72. As we move on to later cohorts, there is a decrease in the average rank. The average individual born in 1972 had a rank that was 4.5 points lower than the corresponding individual in the 1948 cohort. To some extent this decline is unsurprising given that the share proceeding to tertiary education increased by almost 7 percentage points between the 1948 and 1972 cohort. If there is a positive association between the performance at the test and the probability of going on to tertiary education, the relative performance must decrease along with the increase in the probability of going on to higher education. However, it is not only the increase in the number of university graduates that explains the decline in the relative rank of university graduates. The share going on to university education is about the same in the 1948 and 1967 cohorts but the relative rank of university graduates is significantly lower in the 1967 cohort.

[36] We have also conducted the analysis for each separate test. The changes in performance are practically uniform across the different tests.

Table 5.2 Average percentile ranked ability scores for individuals with teacher and university education.

Birth Cohort	*Teachers*	*University*
1948	68.9	72.0
	(..)	(..)
	[..]	[..]
	{..}	{..}
Percent of sample	4.7	15.4
1953	67.5	70.7
Test for mean equal to 1948	(-0.85)	(-1.48)
	[..]	[..]
	{..}	{..}
Percent of sample	3.6	13.5
1967	63.2	69.6
Test for mean equal to 1948	(-2.85)	(-2.69)
Test for mean equal to 1953	[-1.89]	[-1.13]
	{..}	{..}
Percent of sample	2.4	15.2
1972	62.3	67.5
Test for mean equal to 1948	(-3.93)	(-5.35)
Test for mean equal to 1953	[-2.68]	[-3.47]
Test for mean equal to 1967	{-0.42}	{-2.23}
Percent of sample	4.3	22.2

Sources: Calculations based on UGU-data.
Notes: The ability scores are the sum of the scores on verbal, inductive, and spatial ability tests. Numbers within parentheses are (equal variance) t-tests of the equality of the "current" rank and the 1948 rank; numbers within brackets (braces) are analogous but with the 1953 (1967) rank as the norm. Teachers are individuals with a pedagogical degree of at least 3 years at the tertiary level. University refers to individuals who have at least three years of tertiary education. Number of individuals with scores on all three tests: 10560, 9372, 8098, and 7938 for the 1948, 1953, 1967, and 1972 cohorts respectively. See Härnqvist (1998) for more details on the study population and the sampling procedure. We thank Jan-Eric Gustafsson and Åsa Berndtsson for making the 1967 and 1972 data available.

Consider instead the relative performance of teachers. Between 1948 and 1972, there is a significant reduction of the relative performance of teachers by 6.6 percentile points. To put this decline into perspective we have calculated the average difference between individuals of varying educational attainment for the cohort born 1972. On this metric, the decline in the relative performance of teachers almost corresponds to the average difference between individuals having a 3-year and a 2-year upper secondary education.[37] This is a substantial decline both in its own right but also when compared to the average individual with a university degree given that there is nothing "mechanical" about the decline. Thus those opting for teacher education in the early 1990s appear to be less able than those entering teaching in the late 1960s.[38]

In this section we have presented a collection of evidence pertaining to the size and quality of the inflow into teacher education. We think that this evidence tell a consistent story. The decline in the quality, as measured by the scores on standardized tests, broadly mirrors the decline in the teacher relative wage from the late 1960s to the early 1990s. Interestingly, Nickell and Quintini (2002) obtain a similar conclusion, at least for male teachers in Britain; in a similar vein, Evanset al. (2002) note that, over time, it has become less likely that females in the top decile of the ability distribution opt for the teaching profession in the US. The relative attractiveness of teaching declined substantially during the 1980s. There appears to have been a slight rebound during the early 1990s, at least when we look at the inflow into teacher education. This rebound, however, seems to have subsided during the second half of the 1990s.

5.3 Teacher mobility

In the previous section we focused on the size and characteristics of the inflow into teacher education. Of course, there are several adjustment margins in response to variations incentives. Here we will look at the outflow from

[37] The gap between these two groups was 8.2 percentile points. In passing, note that the gap between individuals having a 2-year tertiary and a 3-year upper secondary education was identical.

[38] We have also done these calculations separately by gender. The reduction in the rank is smaller for male teachers than their female counterparts. The average percentile rank among male teacher, within the male ability distribution, was 67.8 in the 1948 cohort. By the 1972 cohort it had declined to 64.5. Female teachers had an average rank of 70.4 among females born in 1948; in the 1972 cohort their rank had decreased to 61.4.

teaching at the compulsory and upper secondary level. This is perhaps a more promising avenue for tracing responses in teacher supply to variations in working conditions.

In Figure 5.8 we graph the outflow from the teaching profession for individuals under the age of 60.[39] We have done a separate calculation for teachers initially employed in compulsory and upper secondary school. There is a trend increase in the probability of leaving the profession altogether. This goes for compulsory and upper secondary school, although the secular increase is more pronounced for teachers in compulsory school. The probability of leaving teaching within the next year for those originally employed in compulsory schools stood at 5 percent in 1986. By 1990, the probability of leaving had risen to almost 7 percent. In 2001, the outflow rate had increased to over 9 percent.

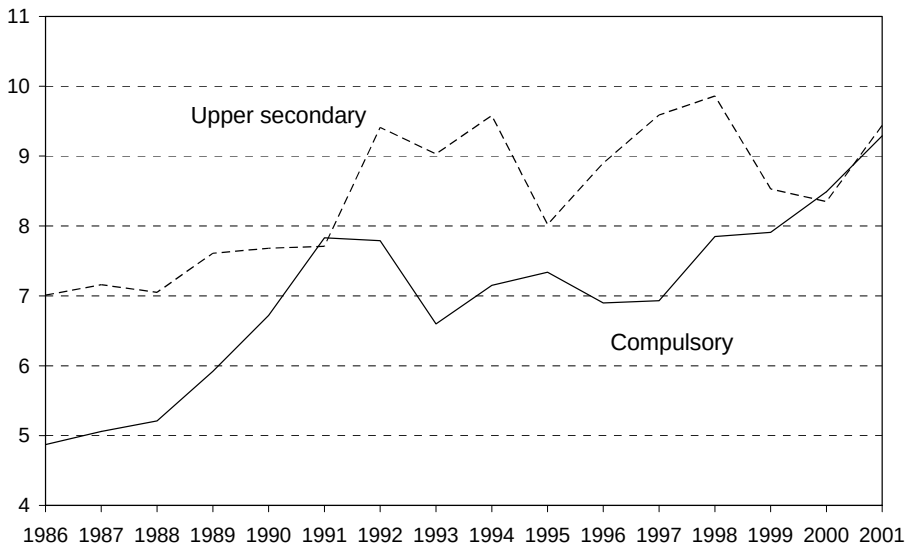

Figure 5.8 The probability of leaving the teacher profession, percent.
Source: Calculations based on the Teacher register.
Notes: The figure shows the outflow rate out of teaching between *t* and *t*+1 for those that were teachers in compulsory (upper secondary) school and less than age 60 in *t*. In the data teachers may hold several positions. We have restricted the population to those holding only one position and excluded those that hold administrative positions.

[39] The population includes all individuals holding a teaching position. The implies that the population includes individuals on temporary as well as permanent contracts and certified as well as non-certified teachers.

The total outflow (i.e. out of the profession and to other levels) from a particular schooling level exhibits a similar upward trend. However, there is an upward jump in the total outflow around 1998 which is when enrolment in adult education peaked. Therefore we take a closer look at the destination choices for those that left youth education for other forms of teaching. We are particularly concerned with the outflow from upper secondary schools to adult teaching. Is the introduction of the Adult Education Initiative (AEI) visible in the data? As already mentioned the AEI was introduced in 1997. Total enrolment in adult education peaked in 1998 and remained high during 1999 and 2000. Concomitantly, there is a peak in the outflow rate to adult education from upper secondary school; see Figure 5.9. Thus, the development in the adult education market feeds on to youth education. Teachers may be leaving youth education to enjoy more pleasant working conditions.[40]

There is some evidence that the competition for teachers has increased. For one thing the establishment of privately run schools has implied that teachers in public school have more alternatives available than previously. In the data there is an upward trend in the probability of leaving public schools for private schools. This development is basically the mirror image of the expansion of private schools. Also, a system of individual wage setting was introduced during the second half of the 1990s. So employers are able to attract teachers by offering higher wages. This may increase the mobility across schools within a municipality and also mobility across municipalities. Unfortunately, we cannot examine whether school mobility has increased because we only have the school identifier since the mid 1990s. Nevertheless, the teacher mobility between municipalities has increased; the probability of changing municipality is twice as large at the turn of the century in comparison to the mid 1990s.

[40] Another possible explanation is that teacher wages in adult education have grown faster than wages in regular upper secondary school. It seems that this is not the case. In 1996, the earnings of those working in adult education were higher than the earnings of teachers in youth education. In 1998, this relationship had been overturned and by 2000 the earnings of teachers in youth education had grown even more. We have used the Teacher register and LOUISE to do these calculations.

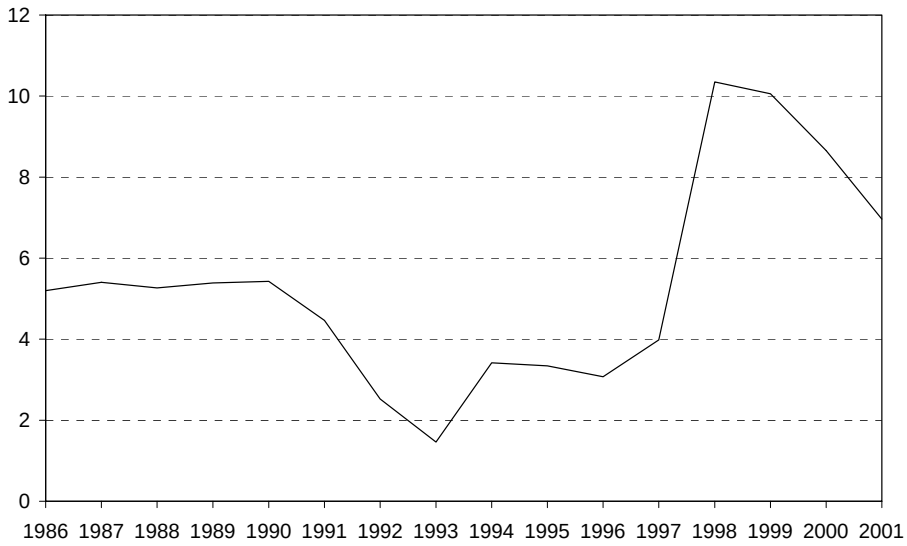

Figure 5.9 The outflow rate from upper secondary school to adult teaching, percent.
Source: Calculations based on the Teacher register.
Notes: The figure shows the outflow rate to adult education between *t* and *t*+1 for those that were teachers in upper secondary school and less than age 60 in *t*. We have restricted the population to those holding only one position and excluded those that hold administrative positions.

Above we documented a trend increase in the probability of leaving the teaching profession. What are the characteristics of those that are leaving? Table 5.3 presents the results of (logit) regressions where the probability of leaving teaching is related to a set of characteristics of the individual and the position. We look at the outflow over a three year horizon for those initially employed in compulsory schools. We examine the determinants of this outflow at two points in time. Column (1) presents a regression where the base year is 1992 and column (2) the results where the base year is 1999.

The regressions for each time point look as one might expect. Certified teachers and teachers on permanent contracts are less likely to leave the profession for instance. The more interesting question is if there is a change in the importance of any characteristic between the two time points. Comparing columns (1) and (2) there are some changes. However, the major changes refer to teaching fields ("other" fields is the reference category) and there is no obvious association between the quality of the teacher and the teaching field. When it comes to teacher quality the most interesting characteristics are teacher certification and teaching experience (as approximated by age). There appears

to be no major changes in the importance of these characteristics between the beginning and the end of the 1990s. Having said this, we should also note that this does not necessarily imply that there has been no change in the relationship between teacher quality and the outflow rate. Remember that the literature suggests that it is difficult to explain teacher quality (as measured by their contribution to student performance) by observed indicators of teacher quality.

Table 5.3 The relationship between the probability of leaving teaching and characteristics

	(1) *1992*	*(2)* *1999*
Female	.016	.024
(=1 if "yes")	(4.88)	(7.07)
Age	-.112	-.126
	(77.3)	(82.2)
Age squared/100	.145	.163
	(84.8)	(87.3)
Certified teacher	-.158	-.145
(=1 if "yes")	(22.8)	(27.2)
Science	-.018	-.017
(=1 if "yes")	(1.68)	(2.45)
Social Science	-.202	-.040
(=1 if "yes")	(3.48)	(5.35)
Language	-.202	-.040
(=1 if "yes")	(3.48)	(5.35)
Public employer	-.108	-.095
(=1 if "yes")	(7.54)	(11.2)
Permanent contract	-.225	-.151
(=1 if "yes")	(34.3)	(32.3)
# individuals	81,521	77,737

Notes: The table shows the marginal effect of each characteristic on the probability of leaving within 3 years; absolute value of z-statistics based on robust standard errors in parentheses. The sample is restricted to those with only one position and who were below age 60 in the base year (1992 or 1999). The regressions are based on the Teacher register data.

To summarize this section: the major finding is that there is an upward trend in the probability of leaving the teaching profession. There is also evidence suggesting that the build-up of adult education has implied that teachers leave youth education for adult teaching. We think that the changes we have documented here are more related to changes in working conditions in teaching than changes in monetary incentives. After all, teacher relative wages have not changed that much since the mid 1980s.

5.4 Summary and conclusions

Teacher relative wages have declined for a long time. To some extent this decline is due to a decrease in the wage differentials between different education categories. But there is also a decline relative to those with a similar education level as teachers: for female teachers, relative earnings decreased by 6 percent between 1968 and 2000; for male teachers, the decline amounted to almost 13 percent during the same time period. The 1990s featured fairly minor changes in relative wages.

Working conditions are of course also important. During the 1990s, working conditions in teaching appear to have become substantially worse, at least as evidenced by the relative increase in health problems caused by mental stress. This increase is most likely related to cut-backs in resources and a consequent increase in class size.

In sum, it seems that the teaching profession has become less attractive over time. This will affect the number of teachers as well as the composition of the pool of teachers. We have shown that the relative demand for teacher education has declined. In addition, it seems that high ability individuals have become less inclined to opt for the teaching profession over time.

The retirement of the baby boomers of the 1940s implies future shortages that will hit schools particularly hard. How should one improve the attractiveness of teaching? The national and local governments, at least to some extent, control several instruments that influence the attractiveness of teaching and, hence, the pool of teachers: the dimension of teacher education, teacher wages, and school resources in general are a few, but important, examples.

The number of slots in teacher education seems to be the main instrument used. This may be an ineffective policy without due consideration to the evolution of incentives – pecuniary as well as non-pecuniary. This is clearly illustrated by the demand for teacher education in math and sciences where the pool of qualified applicants is not sufficient to fill the available slots.

Prospective teachers in math and sciences apparently have better options outside teaching.

Are there any developments on the teacher labor market that imply that the future shortage will not be as severe as predicted? One possibility is the move towards individualized wage setting. Clearly, some wage flexibility is required to equalize demand and supply. With more flexibility we will probably observe pay differentiation across, e.g., subjects and regions to a greater extent. It seems to us that such differences are necessary outcomes.

The analysis in this chapter has shown that basic economics applies also to the teacher labor market. If school authorities decide to change the amount of resources going to schools, this will feed on to teacher quality. The quality of teachers is, in turn, important for student performance. In addition, one can argue that teacher skill becomes even more important in larger classes. It is these considerations that policy makers must take into account when they decide on the allocation of scarce resources.

References

Dolton, P. (1990), The Economics of UK Teacher Supply: The Graduate's Decision, *Economic Journal*, 100, 95-104.

Edin, P-A. and P. Fredriksson (2000), LINDA – Longitudinal INdividual DAta for Sweden, Working Paper 2000:19, Department of Economics, Uppsala University.

Corcoran, S., W. Evans, and R. Schwab (2002), Changing Labor Market Opportunities for Women and the Quality of Teachers 1957-1992, NBER Working Paper No. 9180.

Fredriksson, P. (1997), Economic Incentives and the Demand for Higher Education, *Scandinavian Journal of Economics*, 99, 129-142.

Granström, K. (1998), Stora och små undervisningsgrupper. Forskning om klasstorlekens betydelse för elevers och lärares arbetssituation, FOG-report no. 37, Department of Behavioral Sciences, Linköping University.

Gustafsson, J-E. (2003), What do we Know about Effects of School Resources on Educational Results?, *Swedish Economic Policy Review*, forthcoming.

Härnqvist, K. (1998), Evaluation through Follow-up. A Longitudinal Program for Studying Education and Career Development, Report No. 1998:01, Department of Education, Göteborg University.

Lakdawalla, D. (2001), The Declining Quality of Teachers, NBER Working Paper No. 8263.

Lazear, E. (2003), Teacher Incentives, Manuscript, *Swedish Economic Policy Review*, forthcoming.

Le Grand, C., R. Szulkin, and M. Tåhlin (2001), Lönestrukturens förändring i Sverige, in J. Fritzell, Michael Gähler and O. Lundberg (eds.), *Välfärd och arbete i arbetslöshetens årtionde*, SOU 2001:53, Fritzes, Stockholm.

Murnane, R., J. Singer, J. Willett, J. Kemple, and R. Olsen (1991), *Who Will Teach? Policies that Matter*, Harvard University Press, Cambridge.

National Agency for Education (2003), Behov av och tillgång på pedagogisk personal – Skolverkets bedömning 2003.

Nickell, S. and G. Quintini (2002), The Consequences of the Decline in Public Sector Pay in Britain: A Little Bit of Evidence, *The Economic Journal*, 112, F107-F118.

Nilsson, A. (1984) Studiefinansiering och social rekrytering till högre studier 1920-1976, PhD Thesis, Economic History Society, vol. XLI, Lund.

OECD (2002), *Education at a Glance*, Organization for Economic Cooperation and Development, Paris.

6 The consequences of school choice

Following the wake of decentralization, another school reform was implemented the 1st of July 1992. After the reform, municipalities were obliged to provide funding for independently run schools. Moreover, parents were given the right to choose among all schools – public as well as private.[41] As a general rule, the choice is made subject to the availability of slots. In case of excess demand, educational slots in public schools are allocated on the basis of residence while slots in private schools are allocated on a first-in—first-serve basis. An important deviation from this general rule is that Stockholm city, in 2000, started to admit students to upper secondary schools solely on the basis of performance.

Arguably, this reform has increased the (potential and actual) competition between schools. Will it also promote the efficiency of schools? The Swedish evidence on this question is fairly limited but it has received a great deal of attention recently.

The recent interest started with a report by Bergström and Sandström (2001). Their analysis suggested that competition from independent schools increased the performance of public schools. Their conclusion was cast in doubt by Wibe (2002) who argued that the results are not robust to reasonable alternative specifications of the outcomes of interest. The debate between Bergström and Sandström, on the one hand, and Wibe, on the other hand, has been heated. The contenders have accused each other of being "scientific humbugs" and producing research of "deficient standards". For us outside observers, it seems that the debate is fueled by a non-negligible amount of ideological drive. To be more constructive, we therefore devote part of this chapter to examining the consequences of school choice for various outcomes, the most important being student outcomes.

The remainder of this chapter is outlined as follows. We first describe the main features of the Swedish reform. Then we present the international evidence on the effects of school competition. In section 6.3, we move on to present some descriptive evidence on the prevalence of school choice. Sections 6.4 and 6.5 examine the consequences of school choice. Section 6.4 deals with

[41] We will use "independent schools" and "private schools" interchangeably. The reader should remember, however, that independent schools are really publicly funded, but privately run, schools. Thus, they are more similar to US charter schools than US private schools.

the effects on segregation and costs, while section 6.5 examines the effect on achievement.

6.1 The Swedish system in an international perspective

In the United States, the public typically does not offer funding to support school choice. If public funds are made available it is usually in the form of a voucher enabling students from disadvantaged backgrounds to opt for a private school that charges tuition. The programs are small scale and targeted at specific groups; the voucher program in Milwaukee is a prominent example.[42]

In contrast, the Swedish reform entailed universal school choice. Therefore, the Swedish system is large scale and more similar to the systems operating in Chile and New Zeeland than the US system. The Swedish reform also opened up the possibility of establishing privately run (or independent) schools. These are not private schools in the US sense since they are not allowed to charge tuition.[43] Rather, they are financed through public funds. The grant from the local government takes the form of a grant per pupil and it is set in negotiations with the local government. The guiding principle is that the per pupil grant for private schools should be allocated on the same basis as the allocation of funds between public schools in the municipality.[44] Also, independent schools are not allowed to select students although it may be difficult to enforce this requirement in practice.

The establishment of independent schools is determined in an application procedure at the National Agency for Education (NAE). The municipality concerned is allowed to express its opinion on the application but it has no veto. If the application is approved by the NAE, the independent school is eligible for public funding.

6.2 International evidence on school competition

The effect of private school attendance, and competition more generally, on student achievement is a contentious subject in many countries. The research

[42] Rouse (1998a) evaluates the Milwaukee program.
[43] Prior to 1997 they were allowed to charge a small amount of tuition.
[44] This implies that the principle for determining the grant per pupil varies by municipality. The rules for determining this grant has been changed on several occasions. Initially, there was a nationally determined floor of 85 percent of the cost per pupil in local public schools. There was less than full compensation since private schools were exempt from certain administrative duties and VAT.

community has not reached a consensus as to the impact of school choice. Moreover, the particular organization of a country's public school system, as well as other factors -- such as the extent of residential mobility, discrimination in the housing market, and the extent of separation of church and state -- are likely to condition the extent and direction of any differences in outcomes produced by public and private schools. Here we mainly address a narrow question: What has the research found regarding the effect on achievement of going to a private rather than a public school? We also cover the more difficult question of the effect of *competition* among schools on student achievement.

We begin by considering charter schools. Charter schools in the United States represent a rapidly growing form of school choice and competition. Charter schools are schools that are formed by parents or other groups as an alternative to neighborhood public schools. Charter schools are frequently administered by private contractors. Charter schools receive public funding, and do not charge tuition; in this sense, they are public schools. Charter schools, however, are exempt from many of the rules and regulations that govern other public schools. For example, charter schools are often exempt from teacher certification, and curriculum and testing requirements that apply to public schools. Private schools differ from charter schools primarily because they can charge any amount they like for tuition. In this respect, American charter schools come closer than American private schools to independent schools in Sweden, as independent schools in Sweden cannot charge tuition and receive public support.

The literature on charter schools is exceedingly thin, in part because they are a relatively recent innovation. Probably the most careful study is by Bettinger (1999), who provides an analysis of the effect of charter schools in Michigan. Bettinger estimates a variety of models where charter and traditional public schools are matched to each other, and he also implements an instrumental variables strategy in which the exogenous variation in the occurrence of charter schools is derived from proximity to a state university, as the Michigan charter school law allowed state universities to approve charter schools. He examines both the level and changes in test scores of the charter students versus other public school students, and he examines the effect of competition from a charter school on the performance of students in nearby public schools. He concludes "When charter schools are compared to public schools with similar pre-charter characteristics, pupils in charter schools score no higher, on

average, and may even be doing worse." He also concludes "charters have had little effect on student achievement in neighboring public schools."

The literature on private schools is much more extensive and contentious; see Neal, 2002 and Ladd, 2002 for a recent debate over private school vouchers. Early studies such as Coleman et al. (1982) examined differences in performance between those who attend public and private high schools, without modeling why some students elected to attend private schools.[45] The next wave of studies sought to exploit exogenous instruments – that is, reasons why students might attend private school that are unrelated to their ability – for private school education. Many studies (e.g., Neal, 1997; Evans and Schwab, 1995) used religious affiliation or distance from a Catholic school as an exogenous variable to generate variability in private school attendance. Recent work by Altonji et al. (2000) cast doubt on these identification strategies, however. They show, for example, that religious affiliation is related to performance by those who attend public schools. They also cast doubt on the plausibility of the assumption that distance from a Catholic school is unrelated to student achievement: parents of students who live nearer to a private school have greater educational expectations for their children and higher income, for example. These attributes are likely to affect their children's achievement irrespective of whether they attend a private school.

The next wave of studies examined public voucher programs, which, in part, used lotteries to select students. These lotteries were not intended to facilitate research, but instead to allocate scarce slots. As a result, complete information on the composition of the application pools used to make admission decisions is typically unavailable in these studies, and proxies typically are used. (We consider these studies natural experiments because they were not explicitly designed as randomized field experiments.) In the leading research of this genre, Rouse (1998a) finds that students in the choice schools outperformed those in the public schools in terms of gains on their math tests, but not reading tests. Rouse (1998b), however, finds that students who attended Milwaukee public schools in the so-called P-5 program, a program enabling some public schools to reduce class sizes to levels comparable to those in the private schools, "have similar math test scores gains to those in the choice schools, and students in the P-5 schools outperform students in the choice schools in reading." She concludes that "one *potential* explanation for these results is that

[45] For a critique, see Goldberger and Cain (1982).

students perform well with smaller classes" regardless of whether they are in public or private schools.

The most recent wave of research, initiated by Paul Peterson of Harvard and his collaborators, involves actual randomized experiments conducted in three cities: New York, Dayton and Washington, DC. The New York City experiment had the largest sample size and lowest attrition rate of the three experiments, and is the only one for which data have been made available (by Mathematica Policy Research). In these experiments, low-income applicants for a privately funded voucher program were randomly assigned to a treatment group that was offered a voucher of around $1,400 a year for up to three years, and a control group that was not offered a voucher. At the conclusion of the experiment, there was no statistically discernable difference in average test scores between those offered and those not offered a voucher for the sample as a whole (Howell and Peterson, 2002). When participating students were broken down into racial groups, however, Howell and Peterson (2002) claim that private school attendance significantly raised test scores for African American students in New York, and possibly in Dayton. Krueger and Zhu (2003) reanalyze the data from New York, and their findings raise doubts that the offer of vouchers did raise scores for African American students in that experiment. Regardless of the controversy over the impacts for racial subgroups, the bottom line of all three experiments was that test scores were not significantly different between those offered and those not offered a school voucher when the broadest set of students was considered.

Outside the United States, the evidence is mixed as well. Angrist et al. (2002) study a voucher program in Colombia that provided vouchers to 125,000 children from poor neighborhoods that covered roughly half of the cost of private secondary school attendance. Many of the vouchers were allocated in a lottery, and Angrist et al. compare lottery winners and losers. Their results are generally positive: those who won vouchers in the lottery were less likely to repeat grades and scored an impressive 0.20 standard deviations higher on standardized tests than those who did not win. Hsieh and Urquiola (2002), on the other hand, find much less positive results for school vouchers in Chile. Although private school enrollment increased by 20 percentage points after Chile introduced a nationwide school choice program, they find no improvement in achievement or grade repetition. When they look across 300 municipalities, moreover, they find evidence of an increase in sorting by ability, as the higher achieving students were more likely to switch to private

school. Hsieh and Urquiola make the important point that it is difficult to estimate the effects of school competition *per se* on student achievement if there is a change in sorting: higher (lower) achieving students could disproportionately leave the public schools for private schools, for example, causing the average of the remaining students to fall (rise) just because of a change in the composition of students.

Even if attending a private school does not produce significantly better achievement outcomes for students who switch from public to private schools, competition among schools could lead to better outcomes for students in all schools. According to the Tiebout model of local public goods, families vote with their feet by choosing which municipality to settle in, and an important consideration in family decisions involves the quality of schools. This introduces an element of competition even in the absence of private schools. In her influential study, Hoxby (2000) examines the effect of the concentration of public schools across 316 metropolitan areas on student achievement in public schools and on private school enrollment. Concentration is measured by the Herfindahl index of school districts' shares within metropolitan areas. Presumably, metropolitan areas with less enrolment concentrated in one school district have greater choice in terms of educational quality in the area.

To estimate these effects, Hoxby implements an instrumental variables strategy, which utilizes the variability in district concentration associated with rivers. She argues that this is a valid instrument since rivers form natural boundaries in a geographic area. The estimates imply that more concentration is associated with lower achievement and higher private school enrollment. Moreover, these models indicate that expenditures per student are greater in areas where public school concentration is greater. She concludes, "Tiebout choice among public-school districts raises school productivity."

Hoxby's study is often cited as evidence that greater private school choice will produce desirable educational results. We think it is an open question if this inference is valid, in general, and it is unclear how her results should be interpreted in the context of Sweden. First, Hoxby finds that Tiebout competition already exists among many public schools. It is possible that competition among public schools was already sufficient in Sweden prior to the choice reform and that opening competition up to more schools will have only a minor effect. Second, Hoxby's results do not address the question of how private school competition would affect public school performance. It is possible that the sorting of students will change in a way that offsets any gain from enhanced

competition; or maybe sorting will change in a way that increases efficiency. Only a direct examination of school choice in the Swedish context can shed light on these issues.

Based on this literature review, we reach the tentative conclusion that the injection of more private schools is unlikely to be a panacea. Private schools could perform better or worse than public schools, but there is little reason to expect big differences, or to generalize the results across countries with very different educational institutions or demographics. The academic benefits of vouchers are often exaggerated by their advocates, and the negative effects are often exaggerated by their critics.

6.3 Is school choice a big deal?

It is difficult to establish whether school choice is an important phenomenon for a couple of reasons. First, there are no data on the amount of mobility between municipalities that is caused by variations in school quality. Although this is likely to be a minor issue, recent research has shown that movers are attracted to municipalities that are generous in terms of teaching expenditures per student; see Dahlberg and Fredriksson (2001). Second, it is difficult to get a grip on the fraction of families that choose to reside close to the public school that they prefer.

Until recently there was not much information on families that chose another public school than that dictated by the residence principle. What has been readily available in the data is the fraction of students attending independent schools.

Before the school choice reform in 1992 there were around 90 independent compulsory schools. By the school year 2002/2003, the number of schools had increased to 539 and 5.7 percent of children attended private schools. Private schools are to a large degree an urban phenomenon. This is shown in Figure 6.1, which plots the evolution of the private school share across different municipalities grouped by population density. In the most urbanized areas (4th quartile) the private school share increased by 5.7 percentage points – from 2 percent to 7.7 percent – between 1992 and 2001. For the least densely populated areas (1st quartile) the private school share increased from 0.1 to 1.2 percent during the same time period.

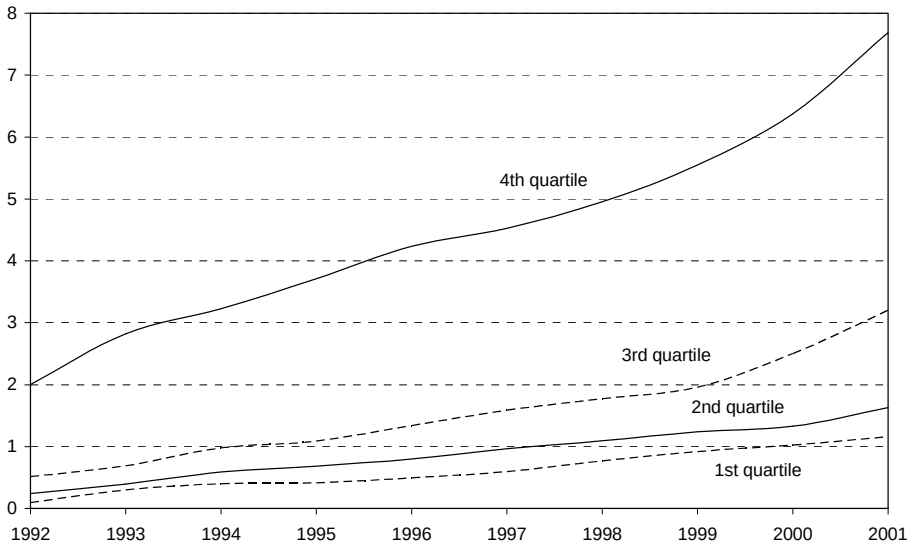

Figure 6.1 The share of students attending an independent compulsory school by population density, percent.
Source: National Agency for Education website
Notes: The private school share is the share of students (as of 15 October) in the municipality attending an independent compulsory school. The numbers have been generated by sorting all municipalities into quartiles based on population density. The 4th quartile refers to the most densely populated areas and the 1st quartile to the least densely populated areas.

As of 1992, the majority of independent schools had a special pedagogical profile, such as Steiner schools. The growth in private schools, since the introduction of school choice, has to a large extent been concentrated to schools with a general profile. In the school year 2002/2003, 44 percent of private schools had a general or a subject profile, 17 percent were confessional or ethnic, 32 percent had a pedagogical profile, and 7 percent had other profiles.[46]

As mentioned above, there has not been much information available on the choice between public schools. A recent study by the National Agency for Education to some extent fills this void; see National Agency for Education (2003). Among other things, this study asks parents about their choice of school. As one would expect, choice between public schools is a bigger deal

[46] These numbers are taken from the National Agency for Education website.

than the choice of a private alternative. It turns out that choosing a public school different from the one dictated by the residence principle is twice as common as choosing a private alternative. Around a quarter of parents also report having chosen the closest public school, although it is slightly difficult to interpret this number.

6.4 The effects of school choice: segregation and costs

Here we start examining the consequences of school choice. We begin with two intermediate outcomes: segregation and costs. In section 5.5, we examine what we are ultimately interested in, namely student achievement.

6.4.1 The effect of school choice on segregation

Does school choice imply less mixing of children from different backgrounds? This may be a particularly important question depending on the presence and nature of so called peer effects. There is fairly convincing evidence suggesting that the behavior and performance of your peers influence your own behavior; see, e.g., Sacerdote (2001). However, we think that the nature and quantitative importance of peer effects in the classroom is still an unresolved issue. Nevertheless, if the result of school choice is that students from poor backgrounds interact to a greater extent with other students from poor backgrounds it is a potential source of concern.

The general question we are interested in is: Does the existence of independent schools increase sorting on observed characteristics? Since the existence of independent schools is largely an urban phenomenon, and, e.g., the highly educated and immigrants live in urban areas to a greater extent than the remainder of the population, it does not make much sense to analyze the question by comparing across municipalities. Therefore, we ask whether the probability of attending an independent school varies with observed family characteristics given that there is one in the municipality (i.e. we conduct a within municipality analysis).

To answer this question we run individual level (logit) regressions relating private school attendance to individual characteristics. The most important of these characteristics are parental education and immigrant background. We run these regressions for those that completed compulsory schools in the spring of 2001. The results confirm previous research (e.g. Hsieh and Lindahl, 2003) in the sense that private school attendance is positively related to parental education and immigrant background. Students with parents that have a

university education are 4.5 percentage points more likely to attend a private school than students with parents that have compulsory education. Foreign-born students are 3.3 percentage points more likely to attend a private school than native-born students of Swedish ancestry. Interestingly, however, the relationship between student background and private school attendance is substantially weakened (and reduced to statistical insignificance) when considering enrollment in private schools with a general profile (remember that schools with a general profile constitute the largest category of independent school). Then there is literally no association between private school attendance and immigrant background. Students with university-educated parents are 1.4 percentage points more likely to attend general purpose private schools (the standard error is 1.4). This suggests, therefore, that much of the relationship between private school enrollment and student background comes from schools with a particular profile, be it a subject, pedagogical, confessional or ethnic profile.

In the previous analysis we did not have the information to examine whether choice between public schools affect segregation. We attempt to assess this issue using information from Stockholm upper secondary schools. In 2000, a new admittance system was introduced in the Stockholm municipality. The new system admits students on the basis of their performance (their grade point average) in compulsory school. Thus we can examine the effect on segregation by comparing the situation in (private and public) schools prior to the reform to the situation after the reform. However, one should probably be a little bit cautious when inferring from the Stockholm experience to the system operating in the rest of the country. The effect on segregation may be substantially different in a system with entrance to public and private schools based on student performance than in a system without explicit sorting on performance.

The statistical office in Stockholm has evaluated the reform by comparing sorting on observed characteristics in Stockholm upper secondary schools in 2001 and 1999; see USK (2002). Thus, this evidence is based on a before-and-after calculation and the magnitudes may be biased because of time effects, i.e., there may be an underlying trend in the data that is wrongfully attributed to the reform. Nevertheless, it seems to us that the magnitudes are too large to come only from a time effect.

One pattern that emerges from the evaluation is that mobility has increased in the sense that the probability of going to school in the resident neighborhood has declined. Moreover, segregation across schools has increased in all relevant

dimensions: the between school variance in immigrant status, parental income and parental education increased between 1999 and 2001. So, for instance, the index of dissimilarity (Duncan and Duncan, 1955) increased by around 9 points for family income (from 0.23 to 0.32). The meaning of this increase is that an additional nine percent of students would have to change schools in order for there to be equalization of family income across schools. Whether the increase in sorting on characteristics will affect the outcomes for the least well-off students depends on the nature of peer effects and it is a question that we cannot answer at this stage.

6.4.2 The effect of school choice on costs

The main argument for introducing school choice is that the force of competition increases the productivity of schools: faced with competition, schools produce the same amount of knowledge at lower cost. A recent (and highly publicized) study conducted by the Swedish Confederation of Trade Unions (see Fransson and Wennemo, 2003) at first glance delivered a fatal blow to this argument. The main result was that total costs per pupil increases along with the independent school share. A percentage point increase in the independent school share raised total costs per pupil in private schools by SEK 25,000 according to their analysis. In this section we investigate whether this conclusion holds up to reasonable alternative specifications.

The conclusion of Fransson and Wennemo (2003) is based on a cross section regression for 2001. The cross section analysis asks whether costs are higher in municipalities with a greater share of students in independent schools. We would argue that this is not the right question to ask.[47] The relevant question is whether *increases* in the private school share *raises* total cost.

Table 6.1 illustrates the fallacies of the cross-section regression rather effectively. Column (1) reports the coefficient on the independent school share in the 2001 cross-section. The regression also standardizes for a set of observed characteristics that we know are related to costs; see chapter 4. The coefficient on the private school share is statistically significant; total costs in public schools are 0.6 percent higher when the private school share is one percentage

[47] In addition, one might want to hold student performance constant, since proponents of school competition argue that it reduces the cost of achieving a given level of student achievement.

point higher.[48] Converted into monetary amounts this estimate implies that total costs increase by SEK 350 per pupil in the municipality. We can of course run an analogous regression using 1992 data. At this time point there were some private schools but essentially the choice reform had not yet been implemented. It turns out that the coefficient is again significant (at least at the 10 percent level of significance): a percentage point increase in the private school share is related to an increase in costs by 1.1 percent; see column (2). We illustrate the problems associated with the cross-section analysis even more effectively in column (3) where we show the regression of total costs in 1992 on the private school share in 2001. As such, this is of course a non-sense regression; the only thing that will be picked up is a spurious relationship. Again, the coefficient is well-determined and positive, but it does not reflect a causal relationship. This clearly indicates that unobserved characteristics about the municipalities drive the cross-section results.

Table 6.1 The relationship between total costs and the indepedent school share (percent). Dependent variable: logarithm of total cost per pupil

	2001	*1992*	*1992 total costs 2001 private school share*	*Change between 1992 and 2001*
	(1)	*(2)*	*(3)*	*(4)*
Share of pupils in independent schools	0.57	1.09	0.42	0.14
	(3.12)	(1.84)	(2.41)	(0.72)
Adjusted R-square	0.304	0.486	0.490	0.794
# Municipalities	263	263	263	526

Notes: *t*-ratios in parentheses. In addition to the percent of students in independent schools the regressions include (the log of) the average size of schools, (the log of) the density of pupils in the municipality, the log of average municipal income among 16-64 year olds, the share of students that are foreign born, the share of students with two foreign born parents, the share of students that have immigrated within the last five years, the share of students with at least one high-school educated parent, the share of students with at least one parent with a university education.

[48] Fransson and Wennemo (2003) used total costs including expenditure to students in private schools. Here we use total costs in public schools because this is what is available in 1992. Using total costs in all schools in 2001 gives a lower and somewhat less precise coefficient: the estimate is 0.49 (t-ratio: 2.78).

So let us instead examine if the *changes* between 1992 and 2001 are due to changes in the private school share; see column (4).[49] The fixed effects regression suggests that costs increase by 0.1 percent in response to an increase in the private school share by one percentage point, but the coefficient is not significant. One should not be surprised by this result, because causality may run in both directions here. Arguably, the incentives are geared such that it is potentially more attractive to start an independent school in regions where the costs in public school are high.

In sum, the conclusion from this simple robustness check is that there is no support for the conclusion that private school choice increases total costs. At the same time, there is no support for the conclusion that competition lowers costs either. Thus if there are beneficial effects of school competition on school productivity it must be because competition improves student achievement. It is to this issue we turn next.

6.5 The effects of school choice: student achievement

The prime difficulty in estimating the causal effect of school choice on student achievement is that alternatives are not established by chance. In particular, it is reasonable to suspect that independent schools enter where parents demand them. The demand for alternatives is likely to be higher when parents are unhappy with the performance of existing schools. The number of independent schools may thus be endogenous with respect to the achievement in public schools. To avoid this problem some exogenous source of variation (an instrument) is needed. Ideally, this source of variation should come from the supply side. Sources of variation that come from the demand side are unlikely to be exogenous to student performance.

6.5.1 Previous Swedish studies

The analysis of school competition in Sweden started with Bergström and Sandström (2001) and Sandström and Bergström (2002). They used a single cross-section consisting of 9th graders that finished public compulsory schools in 1998. The key independent variable was the share of students in the municipality that attended an independent compulsory school. The outcome

[49] The definition of costs changes somewhat in 1995. The estimate based on the change between 1995 and 2001 is equivalent to that using 1992 and 2001 data.

was the performance on tests (*Nationella prov*) in Math, English, and Swedish. The data consist of a sample of students from about 30 municipalities. So a potential problem may be that there are only around 30 variations in the data.

Bergström and Sandström adopt a selection cum instrumental variables approach. They want to correct for selection since public school attendance is not determined by chance. They apply an instrumental variables approach since the independent school share is not exogenous to student performance. Both of these steps require, we would argue, rather arbitrary exclusion restrictions and it is crucial that all characteristics relevant for student performance are included in the regression.[50]

With one important exception, Ahlin (2003) uses the same kind of data as Bergström and Sandström. The crucial exception is that Ahlin has two observations for each individual on performance. In addition to the test conducted in 9th grade, the individuals in the data set were also tested in 6th grade. She can use this information to control for each individual's prior performance. So, the fact that the private school share may be a function of previous achievement does not cause concern since she can include the latter among the control variables. The issues revolving around the selection into public schools are avoided by simply including all students (public as well as private) in the regressions. This appears to be sensible thing to do in its own right since the prospective benefits of school competition should accrue in all schools. And even if one would only be interested in the effects in public schools it is difficult to get at these since there is sorting of students into public and private schools. Some of the methodological quibbles you might raise against the work by Bergström and Sandström are thus not present in Ahlin.

Despite this, the results are very similar. An increase in the independent school share by ten percentage points improves math scores by 0.19 of a standard deviation in Sandström and Bergström and by 0.17 of a standard deviation in Ahlin. Translated to the effect on the percentile rank, the Ahlin results suggest that the percentile rank increases by six points. In none of the studies there is an effect on the performance in Swedish and English. Moreover, in both studies there appears to be a slight downward bias in the plain

[50] Even if one can produce a sensible argument suggesting that a proposed instrument can be excluded from the outcome equation, this argument is contingent on all relevant characteristics being included in the regression. If not, the proposed instrument may be invalid because it is correlated with the omitted characteristics.

OLS estimate of the coefficient on the independent school share, but endogeneity does not appear to be a big issue. Also, there is no evidence suggesting that disadvantaged students gain less from competition. Our purpose next is to reanalyze the data. This reanalysis is inspired by the work of Runeson (2003) and the preliminary work reported in Hsieh and Lindahl (2003).

6.5.2 A reanalysis of the data

We have access to repeated cross sections of test score and grade data for the time period 1998-2001. Between 1998 and 2000, the test score data were constructed in a similar fashion, i.e., the data includes observations from about 30 municipalities. In 2001, the data were obtained by random sampling at the school level. Around 30 municipalities appear at least twice in the test score data during these years. The grade data, on the other hand, refer to the entire population of students.

The nice feature of using repeated cross-sections is that we can control for unobserved differences across municipalities that potentially biases the results. Thus, our estimates are robust to the fact that the level of student performance in the municipality may influence private school entry.[51]

Given a choice between test score data and grade data of equivalent quality we would always prefer the test data. One reason is that the test results, at least in principle, are collected in standardized fashion across the country. Another reason is that the test scores are generically more informative since they contain more variation. The grading system, on the other hand, is not standardized across country.[52] This is not to say that grades are uninformative about student performance. We are just saying that the metric may vary across the country.

But the collection of the test score data raises some concerns. First, the data during 1998-2000 only give 30 variations in the key independent variable – the independent school share. Second, not all students are present to do the test. Non-attendance is most likely higher among poorly performing students. If non-attendance is related to competition, then the estimate on the independent school share will be biased because individuals with a test score form a selected sample. Third, independent schools are not required to report test

[51] However, if private school entry is influenced by the trends in achievement we are in trouble.
[52] Chapter 4 shows, for instance, that the variation in grading standards seems to have increased with the introduction of the new grading system.

results. For our current purposes this is a minor issue, but it seriously compromises any attempt to estimate the gains of private school attendance. The grade data do not suffer from these problems since they contain the population of individuals.

There are two tests in Math (tests A and B) and two tests in English (tests B1 and B2) where we know the actual score obtained on the test. Test A in Math tests the ability to understand mathematical symbols while test B consists of short mathematical problems. Test B1 in English is a reading test while test B2 is a listening comprehension test. There is one reading comprehension test in Swedish (test A) where we, unfortunately, know only the grade on the test. These five tests are given in a comparable fashion over time.

In our regression analysis the basic set of controls includes student and family characteristics (gender, immigrant status, and parental education), private school attendance, and municipality characteristics (the share of immigrants in the population, the share of low-educated in the population, and income). Importantly, the regressions also include municipality fixed effects. These take care of observed and unobserved characteristics about the municipality that stay constant over time. We follow previous studies by measuring competition from independent schools by the share of students in such schools. This key variable of interest thus only varies at the municipal level and, therefore, it is vital to have an estimator that is robust to unobserved municipality characteristics. There is also an issue of the dating of the key independent variable. Our estimates will be utilizing the relationship between the (standardized) *change* in outcomes and the (standardized) *change* in private school enrollment. Introducing the independent school share dated contemporaneously implies that the feed-back from competition onto performance is assumed to be instantaneous. This mechanism appears to be less plausible – one would also think that the history of competitive pressure is (perhaps more) relevant. We attempt to deal with this problem by introducing average private school enrollment in t, t-1, and t-2.[53]

[53] Throughout the present analysis we choose the dating convention that t is a school year. Thus grades awarded in the spring of 2001 (pertaining to the school year 2000/01) are related to private school enrolment in October 2000. We have also experimented with introducing the independent school share in t, t-1, and t-2 freely. Most of the time enrolment in t mattered mostly but in a few instances we obtained the result that one of the lags was more important. Since it is difficult to interpret such variations we settled on imposing equality of the coefficients. There are also other issues here. For instance, one may quibble with the use of the share of student in

Table 6.2.a-c presents the regression results. Table 6.2.a pertains to Math, 6.2.b to English, and 6.2.c to Swedish. As we move along the columns, from left to right, we present regressions for test results, final grades for the tested population, final grades for the municipalities in the sample, and final grades in the entire population. For each outcome we present a basic specification, which only includes a main effect of independent school enrollment, and an extended specification which also includes interaction terms between private school enrollment and characteristics indicating whether the students are disadvantaged. The estimates on the interaction terms capture the extent to which foreign born students and students with low-educated parents are differentially affected by competition from independent schools.[54] The tables also report estimates on the indicator for attending an independent school. We should emphasize that these estimates only have descriptive value; they should not be interpreted as the effect of attending a private school. It may well be that private school attendance has beneficial effects on students, but it is equally plausible that selection into private schools is driving the estimates.

Before probing deeper into the estimates let us reveal our prior on what we think is robust evidence on this issue. First, we have the prior that if competition from private schools is truly beneficial for students the effects should be roughly similar across subjects. Second, about the same information should be contained in test results and grades. The evidence with respect to grades may be weaker because different standards are used across the country, but at least these two measures should not move in opposite directions. Third, test results and grades probably have the greatest correspondence in Math. In Swedish one might expect a greater deviation since, e.g., the ability to write essays (which is not explicitly tested) is a vital component of the subject.

To keep a long story short it seems that the evidence based on the sample of municipalities does not meet with these standards. The case in point is the results for Math. When the estimates jump from being significantly positive (test scores) to being significantly negative (final grades) within the same population of students, the results are simply not credible. We think that the

---

compulsory schools as a whole. Would it not be more sensible to have the share of students in 9th grade when we measure performance? Perhaps, but this formulation would exclude the possibility that the school as a whole changes, say, its way of teaching in response to a change in competition.

[54] The indicator variable for low educated parents is defined to equal unity if both parents have only a comprehensive education.

key reason for this frailty is that there are only around 30 municipalities that can be observed over time.

Table 6.2.a The relationship between independent schools and students' 9th grade performance. Repeated cross-sections 1998-2001 (dependent variable is percentile ranked).

	Math									
	Test score						*Final grade*			
	Basic		*Problem solving*		*Test score pop.*		*Sampled municip.*		*Entire pop.*	
Independent school share	1.01	1.02	1.11	1.13	-.56	-.55	-.25	-.24	.40	.42
	(1.81)	(1.85)	(2.33)	(2.44)	(2.18)	(2.61)	(1.16)	(1.13)	(3.66)	(3.80)
…interacted with foreign born		-.18		-.19		-.14		-.19		-.13
		(1.46)		(1.21)		(1.24)		(1.83)		(1.98)
…interacted with low-ed. parents		-.40		-.58		-.36		-.45		-.10
		(1.37)		(2.62)		(1.47)		(2.43)		(2.82)
Attending independent school	2.46	1.93	6.62	5.61	.91	.02	5.56	5.25	5.09	4.86
	(1.27)	(1.03)	(2.48)	(2.06)	(0.82)	(0.02)	(3.08)	(2.31)	(4.14)	(4.08)
…interacted with foreign born		3.20		4.50		6.08		3.32		3.29
		(0.76)		(1.98)		(4.01)		(1.26)		(3.13)
…interacted with low-ed. parents		1.23		5.94		2.50		-2.73		-1.99
		(0.21)		(1.53)		(0.64)		(0.64)		(1.60)
# individuals	49298	49298	49298	49298	49298	49298	57268	57268	385054	385054

Notes: *t*-ratios in parentheses. Standard errors allow for correlation between individuals residing in the same municipality. "Low-ed. parents" means both parents have a comprehensive degree. The regressions include municipality fixed effects and indicator variables for gender, immigrant status, recent immigrant status (entered within 5 years prior to graduation), no parent is native-born, at least one parent has an upper secondary degree, at least one parent has university education, the share of the population with low education, immigrant density, and mean income in the municipality. Results for samples are weighted by the inverse of the sampling weights.

The results based on the entire population of students seem to be a lot more robust over subjects. What do the estimates mean? It is customary to evaluate at the typical variation observed in the data. The independent school share has grown by one percentage point between 1998 and 2001; the standard deviation is 1.2 percentage points.[55] The estimates thus mean that students gain half a

[55] This is the population-weighted average of the growth within municipalities.

percentile rank in response to an increase in the independent school share by one standard deviation (i.e. 1.2 percentage points). The estimated effect in Math is slightly smaller than the one estimated by Ahlin (2003).[56] Another robust result is that disadvantaged students gain significantly less from increases in competition. In fact, if the equations are estimated freely for each of the two disadvantaged groups, their performance is unrelated to the independent school share. All across the board, students attending private schools score higher than their public school counterparts. As emphasized earlier this may just be self-selection of the abler students into private schools.

Table 6.2.b The relationship between independent schools and students' 9th grade performance. Repeated cross-sections 1998-2001 (dependent variable is percentile ranked).

	English									
	Test score						*Final grade*			
	Reading comprehension		*Listening comprehension*		*Test score pop.*		*Sampled municipalities*		*Entire pop.*	
Independent school share	2.06	2.07	1.64	1.65	0.41	0.41	1.05	1.07	.41	.46
	(4.19)	(4.59)	(2.68)	(2.72)	(1.37)	(1.37)	(2.56)	(2.82)	(3.30)	(3.60)
…interacted with foreign born		-.37		-.26		-.04		-.33		-.33
		(1.95)		(1.64)		(0.02)		(2.33)		(2.49)
…interacted with low-ed. parents		-.77		-.74		-.34		-.56		-.36
		(4.20)		(3.78)		(2.07)		(4.82)		(9.08)
Attending independent school	9.36	8.32	8.57	7.60	6.14	6.18	10.52	9.40	7.41	6.69
	(6.13)	(4.37)	(8.32)	(6.36)	(2.86)	(2.43)	(7.16)	(6.55)	(5.01)	(5.04)
…interacted with foreign born		4.31		2.38		1.69		8.49		6.58
		(3.20)		(0.69)		(0.88)		(3.29)		(6.48)
…interacted with low-ed. parents		4.57		-6.50		-3.98		-.24		-.79
		(0.80)		(1.00)		(0.75)		(0.07)		(0.57)
# individuals	47725	47725	47725	47725	47725	47725	57268	57268	385054	385054

Notes: See Table 6.2.a

[56] Ahlin evaluates her estimates at an increase of the independent school share by ten percentage points. In her case the gain in Math is 0.17 standard deviations. The estimates in Table 6.1, evaluated at this point, imply that the average gain is 0.14 standard deviations.

Table 6.2.c The relationship between independent schools and students' 9th grade performance. Repeated cross-sections 1998-2001 (dependent variable is percentile ranked).

	Swedish							
	Test grade				*Final grade*			
	Reading comprehension		*Test score pop.*		*Sampled municipalities*		*Entire pop.*	
Independent school share	1.96	1.96	.91	.91	.89	.90	.36	.40
	(2.57)	(2.66)	(1.22)	(1.27)	(1.34)	(1.42)	(2.66)	(2.97)
…interacted with foreign born		-.34		-.32		-.54		-.32
		(2.06)		(1.85)		(3.71)		(3.40)
…interacted with low-ed. parents		-.42		-.31		-.26		-.17
		(3.08)		(1.54)		(1.45)		(4.37)
Attending independent school	7.73	7.68	4.35	4.83	5.71	5.68	5.39	5.15
	(5.33)	(6.24)	(4.35)	(3.97)	(4.86)	(4.22)	(4.99)	(4.94)
…interacted with foreign born		1.96		-2.81		1.67		1.89
		(0.40)		(0.75)		(0.84)		(2.30)
…interacted with low-ed. parents		-3.55		-2.71		-3.45		-.19
		(1.69)		(0.93)		(1.22)		(0.15)
#individuals	51828	51828	51828	51828	57268	57268	385054	385054

Notes: See Table 6.2.a

In sum, there is no evidence suggesting that students are hurt by competition from private schools. However, competition from independent schools is no panacea either. The gains we estimate for native-born students without low-educated parents are relatively small. Perhaps this is not too surprising. The possibility of choosing a school by moving to its take-up area has always been there. The school choice reform also introduced a choice between public schools; the extra competitive pressure added by allowing a choice of an independent school is probably relatively minor.

6.6 Summary and conclusions

What are the major findings in this chapter? The international evidence on the effects of private school attendance and competition between schools is mixed. There are some estimates suggesting positive effects, but more often than not

there is no relationship between student performance and private school attendance and school competition respectively.

With respect to the Swedish experience we have four findings that we wish to emphasize. First, there is some evidence suggesting that independent schools contribute to the segregation across schools. Immigrants and children with high-educated parents are more likely to attend an independent school. Second, the independent school share is unrelated to total costs per pupil. In particular, independent schools do not contribute to increasing costs as argued in a recent report. Third, there is a positive association between the 9th grade performance of the average native-born student and the independent school share. Evaluated at a reasonable point this gain appears to be relatively small. Fourth, there is no association (neither positive nor negative) between achievement and the independent school share for foreign born students and children with low-educated parents.

School choice is a contentious issue. It seems to us that proponents as well as critics exaggerate the prospective benefits and costs. The results presented here suggest that independent schools do make a difference, but it is unlikely that they make a major difference.

In this chapter we have only examined one component of the choice reform: the introduction of the possibility for parents to opt for a publicly funded independent school. However, the reform also introduced the possibility of choosing a public school different from the one dictated by the residence principle. However, little is know about the effects of this possibility. This is unfortunate, since this aspect of the reform is probably at least as important as the choice between public and independent school.

References

Ahlin, Å. (2003), Does School Competition Matter? Effects of a Large-Scale School Choice Reform on Student Performance, Working Paper 2003:2, Department of Economics, Uppsala University.

Altonji, J., T. Elder, and C. Taber (2000), Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools, NBER Working Paper 7831.

Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer (2002), Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment, *American Economic Review*, 92, 1535-1558.

Bergström, F. and M. Sandström (2001), Konkurrens bildar skola – en ESO-rapport om friskolornas betydelse för de kommunala skolorna, DS 2001:12.

Bettinger, E. (1999), The Effects of Charter Schools on Charter Students and Public Schools, National Center for the Study of Privatization in Education, Occasional Paper No. 4, Teachers College, Columbia University.

Dahlberg, M. and P. Fredriksson (2001), Migration and Local Public Services, Working Paper 2001:12, Department of Economics, Uppsala University.

Coleman, J., T. Hoffer, and S. Kilgore (1982), *High School Achievement: Public Catholic and Private Schools Compared,* Basic Books, Inc., New York.

Duncan, O. and B. Duncan (1955), A Methodological Analysis of Segregation Indexes, *American Sociological Review*, 20, 210-217.

Evans, W. and R. Schwab (1995), Finishing High School and Starting College: Do Catholic Schools Make a Difference, *Quarterly Journal of Economics*, 110, 941-974.

Fransson, A. and I. Wennemo (2003), Valfrihetens pris – En analys av grundskolan 1995-2001, Näringspolitiska enheten, LO.

Goldberger, A., and G. Cain (1982), The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report, *Sociology of Education*, LV, 103-122.

Howell, W. and P. Peterson (2002), *The Education Gap: Vouchers and Urban Schools*, Brookings Institute Press, Washington, DC.

Hoxby, C. (2000), Does Competition Among Public Schools Benefit Students and Taxpayers?, *American Economic Review*, 90, 1209-1238.

Hsieh, C-T. and M. Lindahl (2003), Did School Choice in Sweden Improve Academic Achievement?, manuscript, University of Amsterdam.

Hsieh, C-T. and M. Urquiola (2002), When Schools Compete, How do They Compete? An Assessment of Chile's Nationwide School Voucher Program, manuscript, Princeton University.

Krueger, A. and P. Zhu (2003), Another Look at the New York City School Voucher Experiment, *American Behavioral Scientist*, forthcoming.

Ladd, H. (2002), School Vouchers: A Critical View, *Journal of Economic Perspectives*, 16, 3-24.

National Agency for Education (2003), Valfrihet och dess effekter inom skolområdet, Skolverkets rapport nr. 230.

Neal, D. (1997), The Effects of Catholic Secondary Schooling on Educational Achievement, *Journal of Labor Economics*, 15, 98-123.

Neal, D. (2002), How Vouchers Could Change the Market for Education, *Journal of Economic Perspectives*, 16, 25-44.

Rouse, C. (1998a), Private School Vouchers And Student Achievement: An Evaluation of The Milwaukee Parental Choice Program, *The Quarterly Journal of Economics*, 113, 553-602.

Rouse, C. (1998b), Schools and Student Achievement: More Evidence from the Milwaukee Parental Choice Program, *Economic Policy Review*, Federal Reserve Bank of New York, 4, 61-76.

Runeson, C. (2003), Do Students Benefit from Private School Competition?, masters thesis, Department of Economics, Uppsala University.

Sacerdote, B. (2001), Peer Effects with Random Assignment: Results for Dartmouth Roommates, *Quarterly Journal of Economics*, 116, 778-794.

Sandström, M. and F. Bergström (2002), School Vouchers in Practice: Competition Won't Hurt You!, Working Paper No. 578, The Research Institute of Industrial Economics.

USK (2002), Förändrad elevsammansättning på gymnasieskolan i samband med ändrad intagningsprincip, Utrednings- och Statistikkontoret, Stockholms stad.

Wibe, S. (2002), Leder konkurrensen från friskolor till högre kvalitet i undervisningen? En kritisk granskning av Bergström och Sandströms artikel, *Ekonomisk Debatt*, 30, 244-255.

7 Quantitative tests as an evaluation device

A precondition for school choice and free entry of independent schools to raise overall productivity is that parents are reasonably well informed about what school is best for their children. This in turn requires that schools provide information about their activities and that parents can evaluate this information. Most likely, parents collect information about different schools in a number of ways. Other parents who have, or have had, children in the schools under consideration constitute one obvious source of information. Personal visits at the schools could also be useful.

In this chapter we discuss how useful quantitative information like grade point averages (GPA) and results on tests like those used in the national tests for Swedish schools can be for parents' school choice. There is no doubt that the availability of such information at the school level has increased in recent years. The National Agency for Education has developed a special data base (called *SIRIS*), which is available at its web page and contains such information at the municipality level and to some extent also for single schools. Many schools have also started to report such information at their own web pages. Some schools even report league tables over own and competing schools GPAs and test results. Thus it is important to know how to evaluate such information.

Quantitative tests can also be used for other purposes in education policy. For example, it has been suggested in the public policy discussion that resources shall be allocated to schools based on their performance on such tests.[57] It has also been suggested that teacher salaries could have a bonus component based on changes in test scores.[58] Such systems obviously require testing instruments that reliably can measure changes in performance for small units like a school class.

Tests can also be used for more general evaluations of schools. When the National Agency for Education last year got a more focused evaluative task, the education minister Thomas Östros wrote: "Every school shall do annual self-evaluations of their own results and compare them with the basic national goals.…All schools' evaluations shall contain common and comparable

[57] See Storesletten and Zilibotti (1999) for a discussion of such proposals.
[58] See Lazear (2003) for a discussion.

measures of the results and their quality". (Dagens Nyheter, March 3, 2003). It is hard to see how this goal could be achieved if national tests and other quantitative tests are not used more frequently than before.

Quantitative tests as an evaluation device are, however, quite often controversial. Our impression is that teachers often are critical to such tests. They emphasize that most quantitative tests measure only limited skills, and that one has to take into account that different pupils develop with different speed. These arguments are important to consider when interpreting the information that tests provide. But it is also likely that many teachers have negative attitudes to tests because tests are associated with one of the more unpleasant aspects of the teacher profession, namely to grade and thus also rank the pupils. It is also likely that many teachers feel that tests directly or indirectly also evaluate themselves.

Quantitative tests are controversial also within the academic community. There is a visible divide between quantitatively and qualitatively oriented education researchers and the view on the informative value of quantitative tests is often the dividing line between these two groups of researchers. We find this divide very unfortunate because obviously good qualitative research based on in-depth interviews and participation in daily work is a complement to – not a substitute – to statistical analysis based on quantitative tests.

For many reasons it is obviously important to know what information there is in data like grades and tests for evaluations of individual schools and the education system in general. The purpose of this chapter is to contribute to the discussion about the role of such data in Swedish education policy. The structure of the chapter reflects our own mixed priors on this issue. On the one hand, it is easy to misuse quantitative test information and go too far in using it. In section 7.1 we discuss such risks more generally, drawing on recent mainly US studies. On the other hand, we believe that many tests contain useful information about pupils' performance and that this information can be collected at reasonable costs. We report a statistical analysis that examines the predictive power of Swedish grades and national tests. In particular, we look at the relationship between grades and test results in compulsory school and subsequent educational level and labor earnings at adult age. Finally, in section 7.4, we summarize our results and discuss how tests can be used more frequently in future evaluations of Swedish schools.

7.1 Problems and limitations with standardized tests

There is a huge international literature on the usefulness of standardized tests for accountability purposes. Koretz (2002), who offers a historical overview of these issues, identifies long-run waves in the political interest in using standardized tests for such purposes. The recent years seem to be a period of rapidly rising interest, in particular in the United States, and not surprisingly a number of recent research reports shed new light on the problems. While some research suggests significant benefits with such test-based evaluation schemes, the literature also clearly demonstrates a number of problems of limitations that must be taken seriously.

All skills and knowledge cannot be easily measured

This problem is in a sense illustrated by the fact that national tests over the years only have been done in a limited number of subjects. Further, some of the goals of Swedish public education policy are rather broad and extend beyond measurable skills and knowledge. In its most recent policy document, for example, the government writes: "The school has an important up-bringing role to play, not least to teach and consolidate society's basic values". (Free translation from Regeringens skrivelse 2001/2002: 188, page 4.). It is not easy to measure the extent to which such goals are reached successfully.

Teaching towards the test

If schools and teachers expect to be evaluated according to test performance and the outcome of the evaluations will have real consequences, they obviously get incentives to "teach towards the test". Because the domain of the tests is limited to some specific skills and knowledge, other important goals for education policy may be neglected.

That this is a real problem in an accountability system that relies on tests has been clearly documented in recent research. Jacob (2002) investigates the experience in Chicago schools, where standardized tests got a more central role for evaluations of schools in 1996-1997. He found statistically significant and substantial improvements in those subjects for which tests were introduced: reading and mathematics. But the gains were concentrated to those specific skills which were tested, not to other dimensions of reading and math. Further, he could observe that teaching in other fields that were not tested (like physics and social science) was reduced. He also found evidence of other disincentive

effects, like increasing placement of pupils in special education classes that are exempt from the tests.

Cheating by teachers and pupils

Cheating should in this context be interpreted in a broad sense and include activities like making extra efforts to make sure that the best students show up at the days of tests but ignoring whether weaker students show up. Also, the behavior of school administrators is relevant as well the behavior or teachers.

Such problems have also been subject to recent research. Jacob and Levitt (2003) develop a statistical technique to infer cheating from combinations of strange answers to tests used in Chicago's public schools. They could convincingly show that cheating had taken place in at least 4-5 percent of the tests. More important they showed that cheating became more likely when the test results started to be used for evaluation purposes. Another striking result was that a bad result in one year increased the likelihood of cheating the next year.

Of course, cheating can be counteracted by deliberate actions of various types and Jacob and Levitt (2002, 2003) have many constructive suggestions in this regard. Nonetheless, their analysis demonstrates an unpleasant problem that has to be taken into account if tests are to be used as a more frequent evaluation device.

Low precision at the school class level

It is easy to forget that the statistical precision of mean performance on a specific standardized test can be rather low when the mean is defined for a single class or a small school. Obviously several temporary factors can affect the performance of a single pupil at a specific test occasion. The pupil might be temporarily ill, or something might disturb the concentration before or during the test. When calculating the mean for a class or school, such temporary factors specific to a single pupil will cancel out. But some temporary disturbing factors could also be shared by most pupils in a class and then the problem will remain.

Even more important in this respect is the fact that *changes* is test scores are more appealing to use to evaluate teachers and schools than levels. It is a well-known experience from empirical research that the statistical precision in estimates of changes is lower than the precision in levels. Recent work by Kane and Staiger (2001, 2002) show quite convincingly that changes in test scores

for units as large as schools to a substantial extent can be due to temporary and irrelevant factors rather than to real ones. Thus the allocation of resources and bonus payments based on changes can yield quite arbitrary outcomes. In addition, the identification of particularly successful (or unsuccessful) schools based on test-score changes can be misleading.

Just as in the case of cheating, this problem with low precision can be reduced in many different ways. Kane and Staiger (2001, 2002) offer several constructive suggestions to achieve better precision. Nonetheless, it is important to stress that the precision problem is a real one when small units like classes and small schools are to be evaluated on a regular basis.

7.2 National tests in Sweden

National tests have a long history in Sweden, see Ljung (2000). The first ones were done in 1944. The background to these tests is that it used to be the case that admittance to secondary school (*realskolan*) was based on tests in Swedish and Math, which were done at one occasion. This test was regarded as mentally quite demanding for the children and not very reliable. Thus it was suggested that admittance instead could be based on GPA, which more reliably could reflect performance during a longer period of time. But GPA can only be used if grades are comparable across classes and schools. Therefore national tests were introduced to make grading comparable for the whole country.

Ever since 1944 such national tests have been used in Swedish compulsory school and their main purpose has been to make grades comparable. In general, they have been done in Swedish, English and Math but sometimes also in other subjects. Since 1997 national tests are voluntary in fifth grade and compulsory in ninth grade. In both grades the tests are done in Swedish, English and Math.

In upper secondary school national tests are of more recent origin. They were introduced in 1966 after a major reform of upper secondary school. They have been done in many subjects and have generally been compulsory.

7.3 The predictive power of the Swedish national tests

Although we recognize that misuse – or even too much use – of quantitative tests can deteriorate their informative value, it is useful to know how well the actual Swedish national tests have been able to predict future outcomes. Because our focus is on education's ability to affect equality in the labor market, annual labor earnings are a natural final outcome to study. It is informative, though, to know whether an association between test scores and

earnings is mediated by educational attainment, or if there also is an association with earnings after controls for educational attainment. Thus we examine the predictive power in two ways: first we examine how well tests predict educational attainment at adult age. We look specifically at low education and high education. Second, we run log earnings equations with seven levels of educational attainment as well as test scores as explanatory variables.

It is also of special interest to compare the predictive power of grades per se and the tests per se. When teachers grade their pupils, they have the test results at their disposal. However, they are also free to use additional information from the students' daily achievement in the class room; this information might be "quantitative" or "qualitative" in nature. Thus it is interesting to know whether grades have stronger predictive value than the tests. To make this comparison we use dummy variables for grades as predictors and compare their predictive value with the predictive value of the national tests. In doing so, we divide the test results into five levels with the same fraction at each level as the recommended fraction according to the relative grade system (7 percent get 1 and 5 respectively, 24 percent get 2 and 4 respectively, and 38 percent get 3).

We did this analysis on the UGU-data for the cohorts 1948 and 1953. We used the national tests in Math, Swedish and English in sixth grade and the grades in the same subjects. We compared the predictive value for each subject separately. The detailed results are reported in a set of tables in an appendix. The main results are the following:

- Both tests and grades have separately strong predictive power in the sense that the coefficients are strongly significantly different from zero. The predictive power is particularly strong for educational achievement, but by conventional statistical standards both grades and tests significantly predict earnings differentials conditional upon educational attainment.
- Test results contribute to the explanatory power of the educational attainment and earnings even conditional upon grades. This result shows that the tests have additional predictive information that is not contained in the grades.
- Tests and grades predict high educational attainment – college degree – as well as low attainment – compulsory school only.
- The predictive power of the three subjects was about the same. Thus the results do not support the common claim that math skills are particularly important.

- In the 1953 data we could do separate analyses for one group of pupils who were graded according to the older grading system based on letters and the relative grading system that was used until the late 1990s. There were no marked differences in the explanatory power of the two grading systems.

7.4 Conclusions

We have discussed the value of standardized tests like those used in the Swedish national tests for the purpose of evaluating Swedish education policy. Our discussion has illustrated that there are trade-offs involved in the use of such information. On the one hand, such tests measure skills and knowledge that are strongly related to educational achievement and labor earnings at adult age. Even controlling for grade levels, there is additional explanatory value in the national tests. Thus one cannot simply dismiss the value of tests by claiming that they do not contain useful information.

On the other hand, there are many risks involved in using standardized tests to evaluate small units as classes, small schools and in particular single teachers. The obvious reason is that the presence of tests can affect the behavior of teachers and schools in a non-productive way. We have discussed risks like "teaching towards the test", rather than towards broader skills that cannot be evaluated by standardized tests. We have also discussed the risk of "teacher cheating" when the tests are used for accountability purposes. That we have stressed such potential negative behavioral effects of tests does not imply that we dismiss the possibility of productive incentive effects, namely that evaluations make teaching more efficient. Indeed, we believe that there is a good potential for such effects and two recent studies on Israeli data (Lavy 2002, 2003) suggest that they can be substantial. Our main message is that any accountability system using standardized tests must be implemented with much care to be successful.

Is it possible to productively increase the use of standardized tests in Swedish education policy? We believe that it is, and we conclude by discussing some guidelines for such a more ambitious testing policy.

As a general remark, we first want to stress the important role of a central evaluation body in a decentralized system. Of course, one could argue that it is useful if individual municipalities take own initiatives and implement new tests of different types. Other municipalities could then learn from the experience. But there is a risk that such approaches will not be evaluated properly if they are part of local politicians' own political agenda.

One good step towards better use of the available information was taken in 1998, when the results on national tests in grade 9 were sent to Statistics Sweden so that the data could be merged with other useful information. Some of our own assessments of the impact of resources and independent schools were possible thanks to such merged data. From the school year 2002/2003 onwards, the results on national tests will be collected for the whole country. The lack of data before 1998 has severely limited our and other researchers' possibilities to evaluate the school reforms during the 1990s. We have been forced to rely only on grade data for that period. That this useful national test information was not saved and reported to a central body before 1998 probably reflects a negative attitude to quantitative analysis in Swedish education administration and politics.

It is also a natural step that the national tests in fifth grade are made compulsory and that the results are saved and collected by Statistics Sweden in the same manner as the ninth grade test results. In this way, it will be possible to perform analyses of *changes* in test performance at the level of the individual, a substantial improvement over the present situation.

Next, we recommend that (compulsory) national tests are done earlier than in fifth grade (age 11) as is the case today. There are strong reasons to believe that the very first years in school, and the years in pre-school, are very important for children's development. To be able to analyze the contribution of schools and school resources to the development for very young children, it would be very useful to have "starting values" when they start school.

The issue about the time of the first national tests is related to the issue when grades are to be given the first time. This is, to say the least, at sensitive issue in Sweden politics that shows up in almost every election campaign. Today, and for more than a decade, Swedish pupils are graded the first time after the fall semester of grade 8, that is, at age 14. Before that parents and pupils get more informal information about school achievement. This system implies that in ninth grade the national tests can be used as guidance for teachers' grading. The grades in turn are used for admittance to alternative high schools and various high school tracks. Thus the grades are important.

The national tests, in turn, are important for the grades. Although the teacher has some discretion to deviate from the result on the national test, this test is generally considered the most important one during the year. Thus the pupils also have incentives to do well on the national tests and as a

consequence the tests should be more informative than they otherwise would have been.

The situation is different for the national tests in fifth grade. These tests are not connected to any grades so one could suspect that pupils take the tests less seriously. Therefore one could fear that these tests are less informative and less useful for overall evaluations of education policy. We have no empirical evidence that this problem is a substantial one, but it is worthwhile considering whether the lack of grades make tests less informative. If that is the case, one cost for starting with grades as late as age 14 would be less reliable evaluations of education policy. Because we have no empirical evidence, we can only speculate on this point. It seems to us, however, that alternative starting ages for grading is a strong candidate for experimental evaluation where some municipalities are randomly selected to start grading earlier than other ones.

Finally, we question whether teachers should grade their own pupils' national tests as is the case today. Of course, such a system is easy to administer and thus inexpensive. But there are many risks involved, in particular today when many schools use their performance on these tests for marketing purposes. A simple system that requires only slightly more administration than today's system is to have teachers grade the test of each others' students. Even better, but possibly more expensive, is to rely on central grading of the tests.

References

Härnqvist, K. (2000), Evaluation through Follow-up, in C.-G. Janson (ed.), *Seven Swedish Longitudinal Studies in the Behavorial Sciences*, Forskningsrådsnämnden, Stockholm.

Jacob B. (2002), Accountability, Incentives and Behavior: the Impact of High-stakes Testing in the Chicago Public Schools, NBER Working Paper no. 8968.

Jacob B. and S. Levitt (2002), Catching Cheating Teachers: the Results of an Unusual Experiment in Implementing Theory, NBER Working Paper no. 9414.

Jacob B. and S. Levitt (2003), Rotten Apples: an Investigation of the Prevalence and Predictors of Teacher Cheating, *Quarterly Journal of Economics*, 118: 843-877.

Kane T. and D. Staiger (2001), Improving School Accountability Measures, NBER Working Paper no. 8156.

Kane T. and D. Staiger (2002), The Promise and Pitfalls of Using Imprecise School Accountability Measures, *Journal of Economic Perspectives*, 16, 91–114.

Koretz D.M. (2002), Limitations and the Use of Achievement tests as Measures of Educators' Productivity, *Journal of Human Resources*, 38, 752-777.

Lavy V. (2002), Evaluating the Effect of Teachers' Group Performance Incentives on Students' Achievements, *Journal of Political Economy*, 110, 1286-1318.

Lavy V. (2003), Paying for Performance: The Effect of Individual Financial Incentives on Teachers' Productivity and Students Scholastic Outcomes, manuscript, Hebrew University, Jerusalem.

Lazear E. (2003), Teacher Incentives, *Swedish Economic Policy Review*, 10, forthcoming.

Ljung B-O (2000), Standardproven – 53 år i skolans tjänst, Rapport 13 från PRIM-gruppen, Lärarhögskolan i Stockholm.

Storesletten K. and F. Zilibotti. (1999), Utbildning, utbildningspolitik och tillväxt, in L. Calmfors and M. Persson (eds.), *Tillväxt och ekonomisk politik*, Studentlitteratur, Lund.

Table A.1 Regression estimates, 1948 cohort of UGU-data. Dependent variable: log annual earnings 1993

	Math		*Swedish*		*English*	
Male	0.377	0.370	0.387	0.386	0.403	0.404
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Grade B	0.129	0.080	-0.205	-0.213	0.095	0.046
	(0.042)	(0.045)	(0.218)	(0.219)	(0.040)	(0.043)
Grade Ba	0.205	0.126	-0.140	-0.175	0.196	0.122
	(0.042)	(0.048)	(0.218)	(0.219)	(0.040)	(0.047)
Grade AB	0.282	0.184	-0.056	-0.108	0.249	0.171
	(0.043)	(0.052)	(0.218)	(0.220)	(0.041)	(0.051)
Grade a	0.339	0.223	-0.076	-0.130	0.264	0.152
	(0.047)	(0.061)	(0.219)	(0.222)	(0.050)	(0.063)
Grade A	0.469	0.344	0.015	-0.037	0.387	0.252
	(0.177)	(0.184)	(0.294)	(0.298)	(0.345)	(0.348)
Test 2	-	0.106	-	0.031	-	0.098
		(0.037)		(0.037)		(0.036)
Test 3	-	0.129	-	0.064	-	0.123
		(0.040)		(0.040)		(0.039
Test 4	-	0.157	-	0.073	-	0.113
		(0.047)		(0.046)		(0.045)
Test 5	-	0.177	-	0.070	-	0.181
		(0.060)		(0.058)		(0.059)
Adjusted R-sq.	0.151	0.152	0.145	0.145	0.149	0.150
# observations	7657	7657	7657	7657	7657	7657

Notes: Constants and coefficients for six educational levels not reported. Standard errors within parentheses.

Table A.2 Regression estimates, linear probability models. 1948 cohort of UGU-data. Dependent variable: University degree.

	Math		*Swedish*		*English*	
Male	0.019	0.004	0.056	0.047	0.072	0.072
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Grade B	0.000	-0.010	-0.069	-0.066	0.010	-0.010
	(0.021)	(0.023)	(0.114)	(0.112)	(0.020)	(0.022)
Grade Ba	0.072	0.028	-0.017	-0.053	0.104	0.045
	(0.021)	(0.025)	(0.113)	(0.112)	(0.020)	(0.024)
Grade AB	0.255	0.143	0.167	0.044	0.295	0.173
	(0.022)	(0.027)	(0.113)	(0.112)	(0.021)	(0.026)
Grade a	0.481	0.274	0.437	0.187	0.570	0.335
	(0.024)	(0.031)	(0.113)	(0.114)	(0.024)	(0.032)
Grade A	0.768	0.491	0.655	0.340	0.744	0.450
	(0.091)	(0.094)	(0.153)	(0.153)	(0.179)	(0.179)
Test 2	-	0.005	-	-0.009	-	0.022
		(0.019)		(0.019)		(0.018)
Test 3	-	0.050	-	0.070	-	0.068
		(0.020)		(0.020)		(0.020)
Test 4	-	0.156	-	0.186	-	0.144
		(0.024)		(0.023)		(0.023)
Test 5	-	0.285	-	0.350	-	0.306
		(0.030)		(0.029)		(0.030)
Adjusted R-sq.	0.162	0.177	0.159	0.184	0.166	0.180
# observations	7657	7657	7657	7657	7657	7657

Notes: Constants not reported. Standard errors within parentheses.

Table A.3 Regression estimates, linear probability models. 1948 cohort of UGU-data. Dependent variable: less than or equal to compulsory school.

	Math		*Swedish*		*English*	
Male	0.041	0.047	0.023	0.028	0.014	0.014
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Grade B	-0.074	-0.039	-0.154	-0.130	-0.039	0.021
	(0.018)	(0.020)	(0.096)	(0.096)	(0.018)	(0.018)
Grade Ba	-0.160	-0.101	-0.024	-0.155	-0.140	-0.029
	(0.018)	(0.021)	(0.096)	(0.096)	(0.017)	(0.020)
Grade AB	-0.225	-0.014	-0.034	-0.205	-0.211	-0.070
	(0.018)	(0.023)	(0.096)	(0.096)	(0.018)	(0.022)
Grade a	-0.266	-0.163	-0.383	-0.228	-0.239	-0.080
	(0.020)	(0.026)	(0.096)	(0.097)	(0.021)	(0.028)
Grade A	-0.280	0.170	-0.392	-0.237	-0.239	-0.075
	(0.077)	(0.080)	(0.129)	(0.130)	(0.151)	(0.152)
Test 2	-	-0.074	-	-0.078	-	-0.105
		(0.016)		(0.016)		(0.016)
Test 3	-	-0.092	-	-0.149	-	-0.163
		(0.018)		(0.018)		(0.017)
Test 4	-	-0.131	-	-0.184	-	-0.195
		(0.020)		(0.120)		(0.019)
Test 5	-	-0.142	-	-0.188	-	-0.209
		(0.027)		(0.025)		(0.025)
Adjusted R-sq.	0.058	0.063	0.058	0.070	0.060	0.073
# observations	7657	7657	7657	7657	7657	7657

Notes: Constants not reported. Standard errors within parentheses.

Table A.4 Regression estimates, 1953 cohort of UGU-data. Dependent variable: log annual earnings 1993. Old system, letter grades.

	Math		*Swedish*		*English*	
Male	0.401	0.392	0.410	0.411	0.411	0.410
	(0.041)	(0.042)	(0.043)	(0.042)	(0.043)	(0.042)
Grade B	0.103	0.073	-0.454	-0.500	-0.094	-0.117
	(0.123)	(0.132)	(0.739)	(0.744)	(0.141)	(0.153)
Grade Ba	0.015	-0.081	-0.467	-0.560	-0.075	-0.134
	(0.120)	(0.139)	(0.737)	(0.747)	(0.142)	(0.161)
Grade AB	0.053	-0.088	-0.410	-0.538	-0.035	-0.096
	(0.123)	(0.150)	(0.738)	(0.749)	(0.145)	(0.170)
Grade a	0.180	0.012	-0.399	-0.550	-0.046	-0.092
	(0.132)	(0.165)	(0.740)	(0.754)	(0.157)	(0.191)
Grade A	0.540	0.343	-0.044	0.209	NO OBS	NO OBS
	(0.388)	(0.410)	(0.904)	(0.917)		
Test 2	-	0.053	-	0.059	-	0.033
		(0.095)		(0.118)		(0.094)
Test 3	-	0.155	-	0.111	-	0.101
		(0.105)		(0.125)		(0.107)
Test 4	-	0.174	-	0.145	-	0.061
		(0.121)		(0.1369		(0.119)
Test 5	-	0.221	-	0.178	-	0.052
		(0.151)		(0.162)		(0.150)
Adjusted R-sq.	0.098	0.098	0.093	0.092	0.093	0.092
# observations	1302	1302	1302	1302	1302	1302

Notes: Constants and coefficients for six educational levels not reported. Standard errors within parentheses.

Table A.5 Regression estimates, linear probability models. 1953 cohort of UGU-data. Dependent variable: University degree. Old system, letter grades.

	Math		*Swedish*		*English*	
Male	-0.007	-0.009	0.023	0.022	0.028	0.027
	(0.017)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
Grade B	0.040	0.029	0.009	-0.008	0.052	0.038
	(0.053)	(0.056)	(0.318)	(0.316)	(0.060)	(0.065)
Grade Ba	0.059	0.038	0.070	0.028	0.087	0.036
	(0.051)	(0.059)	(0.317)	(0.317)	(0.060)	(0.068)
Grade AB	0.162	0.144	0.166	0.064	0.176	0.095
	(0.051)	(0.063)	(0.317)	(0.318)	(0.060)	(0.071)
Grade a	0.353	0.278	0.346	0.143	0.408	0.285
	(0.064)	(0.069)	(0.318)	(0.320)	(0.065)	(0.080)
Grade A	0.253	0.073	0.999	0.712	NO OBS	NO OBS
	(0.164)	(0.174)	(0.387)	(0.389)		
Test 2	-	0.020	-	0.017	-	0.020
		(0.040)		(0.050)		(0.040)
Test 3	-	0.032	-	0.044	-	0.087
		(0.045)		(0.053)		(0.046)
Test 4	-	0.005	-	0.124	-	0.085
		(0.050)		(0.058)		(0.050)
Test 5	-	0.187	-	0.287	-	0.163
		(0.064)		(0.068)		(0.063)
Adjusted R-sq.	0.096	0.105	0.082	0.102	0.091	0.096
# observations	1302	1302	1302	1302	1302	1302

Notes: Constants not reported. Standard errors within parentheses.

Table A.6 Regression estimates, linear probability models. 1953 cohort of UGU-data. Dependent variable: less than or equal to compulsory school. Old system, letter grades.

	Math		*Swedish*		*English*	
Male	0.100	0.101	0.061	0.062	0.059	0.059
	(0.023)	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)
Grade B	-0.122	-0.157	-0.496	-0.404	-0.199	-0.181
	(0.0070)	(0.075)	(0.420)	(0.421)	(0.080)	(0.086)
Grade Ba	-0.223	-0.261	-0.633	-0.488	-0.303	-0.235
	(0.067)	(0.079)	(0.418)	(0.422)	(0.079)	(0.091)
Grade AB	-0.312	-0.345	-0.776	-0.580	-0.418	-0.295
	(0.068)	(0.084)	(0.418)	(0.423)	(0.080)	(0.095)
Grade a	-0.464	-0.472	-0.0896	-0.653	-0.500	-0.353
	(0.072)	(0.092)	(0.420)	(0.425)	(0.086)	(0.106)
Grade A	-0.545	-0.513	-0.997	-0.724	NO OBS	NO OBS
	(0.218)	(0.232)	(0.511)	(0.518)		
Test 2	-	0.068	-	-0.126	-	-0.028
		(0.054)		(0.067)		(0.053)
Test 3	-	0.049	-	-0.152	-	-0.111
		(0.060)		(0.070)		(0.061)
Test 4	-	.052	-	-0.219	-	-0.153
		(0.069)		(0.077)		(0.067)
Test 5	-	-0.016	-	-0.273	-	-0.181
		(0.086)		(0.090)		(0.084)
Adjusted R-sq.	0.079	0.078	0.078	0.083	0.073	0.077
# observations	1302	1302	1302	1302	1302	1302

Notes: Constants not reported. Standard errors within parentheses.

Table A.7 Regression estimates, 1953 cohort of UGU-data. Dependent variable: log annual earnings 1993. New system, figure (relative) grades.

	Math		*Swedish*		*English*	
Male	0.434	0.430	0.443	0.441	0.443	0.443
	(0.020)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)
Grade 2	0.019	0.039	0.113	0.080	0.015	0.042
	(0.054)	(0.060)	(0.063)	(0.067)	(0.053)	(0.059)
Grade 3	0.066	0.067	0.124	0.057	0.054	0.067
	(0.052)	(0.065)	(0.062)	(0.070)	(0.052)	(0.063)
Grade 4	0.109	0.077	0.151	0.078	0.042	0.041
	(0.054)	(0.070)	(0.064)	(0.076)	(0.053)	(0.069)
Grade 5	0.176	0.122	0.117	0.054	0.086	0.087
	(0.060)	(0.082)	(0.070)	(0.087)	(0.060)	(0.082)
Test 2	-	-0.044	-	0.049	-	-0.057
		(0.051)		(0.045)		(0.049)
Test 3	-	-0.002	-	0.098	-	-0.010
		(0.056)		(0.048)		(0.055)
Test 4	-	0.038	-	0.084	-	-0.010
		(0.078)		(0.056)		(0.061)
Test 5	-	0.058	-	0.074	-	-0.016
		(0.078)		(0.072)		(0.076)
Adjusted R-sq.	0.120	0.120	0.118	0.118	0.118	0.118
# observations	5776	5776	5776	5776	5776	5776

Notes: Constants and coefficients for six educational levels not reported. Standard errors within parentheses.

Table A.8 Regression estimates, linear probability models. 1953 cohort of UGU-data. Dependent variable: University degree. New system, figure (relative) grades.

	Math		*Swedish*		*English*	
Male	0.009	0.000	0.073	0.067	0.062	0.061
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Grade 2	0.031	0.014	0.044	0.024	0.030	0.009
	(0.025)	(0.028)	(0.029)	(0.031)	(0.025)	(0.028)
Grade 3	0.078	0.042	0.110	0.063	0.114	0.055
	(0.024)	(0.030)	(0.028)	(0.032)	(0.024)	(0.029)
Grade 4	0.236	0.141	0.277	0.184	0.254	0.141
	(0.024)	(0.032)	(0.029)	(0.035)	(0.024)	(0.032)
Grade 5	0.498	0.329	0.545	0.386	0.475	0.277
	(0.026)	(0.037)	(0.032)	(0.040)	(0.027)	(0.038)
Test 2	-	0.024	-	0.027	-	0.031
		(0.023)		(0.020)		(0023)
Test 3	-	0.043	-	0.059	-	0.087
		(0.026)		(0.022)		(0.026)
Test 4	-	0.130	-	0.118	-	0.136
		(0.029)		(0.025)		(0.028)
Test 5	-	0.218	-	0.211	-	0.253
		(0.035)		(0.033)		(0.035)
Adjusted R-sq.	0.147	0.155	0.143	0.150	0.121	0.130
# observations	5776	5776	5776	5776	5776	5776

Notes: Constants not reported. Standard errors within parentheses.

Table A.9 Regression estimates, linear probability models. 1953 cohort of UGU-data. Dependent variable: less than or equal to compulsory school. New system, figure (relative) grades.

	Math		*Swedish*		*English*	
Male	0.079	0.087	0.036	0.042	0.040	0.040
	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Grade 2	-0.100	-0.032	-0.083	-0.017	-0.101	-0.072
	(0.026)	(0.029)	(0.031)	(0.032)	(0.026)	(0.028)
Grade 3	-0.230	-0.099	-0.247	-0.122	0.268	-0.194
	(0.025)	(0.031)	(0.029)	(0.034)	(0.025)	(0.030)
Grade 4	-0.343	-0.172	-0.346	-0.180	-0.338	-0.234
	(0.025)	(0.034)	(0.030)	(0.036)	(0.025)	(0.033)
Grade 5	-0.388	-0.195	-0.384	-0.201	-0.401	-0.270
	(0.027)	(0.039)	(0.033)	(0.041)	(0.028)	(0.039)
Test 2	-	-0.110	-	-0.105	-	0.046
		(0.024)		(0.021)		(0.024)
Test 3	-	-0.184	-	-0.172	-	-0.115
		(0.027)		(0.023)		(0.026)
Test 4	-	-0.022	-	-0.206	-	-0.123
		(0.030)		(0.026)		(0.029)
Test 5	-	-0.239	-	-0.215	-	-0.160
		(0.037)		(0.034)		(0.036)
Adjusted R-sq.	0.089	0.098	0.086	0.096	0.086	0.091
# observations	5776	5776	5776	5776	5776	5776

Notes: Constants not reported. Standard errors within parentheses.

8 Family background and earnings: What have education and education policy to do with it?

Equality of opportunity has always been an important goal in Swedish politics. Although philosophers might argue that equality of opportunity is a complicated concept, it is fair to say that it has been interpreted in a rather pragmatic way in the public discussion: a strong association between the socio-economic status of parents and their offspring has been considered as a violation of the equality-of-opportunity norm. Most empirical research that has claimed to address the issue of equality of opportunity has also used such a framework. Sociologists have for long investigated the association between parents' and offspring's social class, in particular the association between father's and son's class. Economists, who more recently have entered this field of research, have rather used earnings or income. Most likely, politicians and the general public would consider both social class and earnings relevant outcomes.

There are more dimensions of equality of opportunity than the intergenerational one. The strive for gender equality in the labor market can also be interpreted in terms of equality of opportunity. The same applies to the ambition to integrate immigrants in the labor market. Yet another example is regional policy that has aimed at equalizing outcomes among people born in different parts of the country.

Because we study Sweden's education policy, we find it natural to focus on intergenerational mobility. The ambition to promote such mobility has been inherent in most parts of Swedish education policy. Going through the education system by age of pupils, one can trace this goal all the way from the expansion of daycare, to the comprehensive compulsory-school reform, to the centralized governance of education through the 1980s, to the tuition-free university system with universal financial student support, and to the expansion of the second-chance adult education during the 1990s. Thus we would ideally like to evaluate the overall impact of Swedish education policy on intergenerational mobility, but we would also like to know how effective various policies have been.

In this chapter we address the following questions. We first ask whether Sweden has been successful in its ambition to promote intergenerational

mobility. We thus ask the question: Does Sweden have a comparatively weak association between parental and offspring's socio-economic status. We survey a recent literature on the association between family background and lifetime earnings and conclude that Sweden (and its neighbor Nordic countries) really appears to have higher intergenerational earnings mobility than other countries, most notably the United States and United Kingdom. Then we ask whether this favorable result can be attributed to education policy. Our major finding is that it is mainly due to lower returns to schooling in Sweden. Finally, we turn to the consequences of the reforms during the 1990s. We find that the association between school achievement, measured by the grade-point-average at age 16, and family background was remarkably stable during the turbulent decade with school reforms, cuts in school budget, and high unemployment.

8.1 Intergenerational earnings mobility in Sweden, the United States and other countries

Although education in several regards can be considered valuable per se, and hence also a goal in itself, one of the major functions of education is to generate future earnings (or income). Thus, we start out with a survey of recent research on parental-offspring relationships in earnings. We interpret a strong (weak) such association as low (high) intergenerational earnings mobility.

This literature has focused on a very simple statistical regression model that relates the logarithm of offspring's earnings to the logarithm of parents' earnings.[59] The coefficient of parent's earnings in such an equation is interpreted as the elasticity of offspring's earnings with respect to parents' earnings. The elasticity provides an answer to questions like, if parents' earnings are 50 percent above the average in his generation, what percentage above the average is the offspring's earnings predicted to be in the own generation. But if the variances in the logarithmic earnings variables are about the same in the parents' and offspring's generations, the elasticity will approximately equal the correlation between log earnings in the two generations. The correlation in turn provides the answer to a slightly different question: if parent's earnings are a standard deviation unit higher than average, how many standard deviation units will the offspring's earnings deviate from the average in the next generation.

[59] See Solon (1999) for a survey.

Further, this literature has focused on long-run earnings. The reason is that annual earnings are affected by transitory factors, so the intergenerational association of annual earnings would be misleadingly low. Most research has focused on fathers and sons. This is a most unfortunate gender bias in the literature. But, due to mothers' intermittent labor force behavior up to the 1960s, it is quite problematic to measure their long-run earnings in a reliable way. Finally, we note that these intergenerational associations should not be given a causal explanation. They only measure the degree of association between outcomes in two generations.

Table 1 reports a number of recent estimates of intergenerational father-son elasticities. Swedish estimates range from 0.13 to 0.28. Although the standard errors of these estimates are non-trivial, the discrepancy also reflects alternative approaches to measuring long-run earnings and different sample criteria. The two Finnish estimates are basically in the same ballpark, whereas the Norwegian study got even lower estimates.

United States and United Kingdom have the highest estimates, well above 0.4. A recent study by Mazumder (2002) suggests that these US estimates might be too low. The two German studies provide ambiguous results, but a higher weight on Wiegand's study would indicate a somewhat higher elasticity than in the Nordic countries. Interestingly, the estimate from the Canadian study is closer to the ones for the Nordic countries than to the US and UK ones.

That family background factors are less important determinants of long-run earnings in Sweden than in the United States is also corroborated by recent comparative results regarding brother correlations in long-run earnings. A correlation among siblings is a most useful measure for our purposes. All childhood factors that influence long-run earnings as an adult and that are shared by siblings make siblings' outcomes more equal. Thus, the higher the correlation among sibling's earnings during adulthood the more important such factors are. The measure's virtue is that it not only captures family factors ("nature" as well as "nurture"), but also neighborhood conditions shared by siblings. School quality and school peers typify such neighborhood factors. Further, the correlation has a straightforward statistical interpretation, namely as the fraction of the variation in the outcome (long-run earnings in our application) that is attributed to the factors shared by siblings.

Björklund et al. (2002) estimated brother correlations in long-run earnings for the United States, Denmark, Finland, Norway and Sweden. They used as similar variables and sample criteria as possible from their data sets. Their US

estimate exceeds 0.4, whereas the estimates for the Nordic countries were in the range 0.15-0.25. Norway had the lowest numbers, a result that corroborates the low father-son correlation reported in Table 1. The Nordic-US differences were significantly different from zero.

Table 8.1 Estimated father-son earnings elasticities for various countries.

Country	*Study*	*Elasticity*	*Comment*
Sverige	Gustafsson (1994)	.14	Gustafsson emphasizes that his only measures father's income in a single year and that the estimate might be too low because of this.
	Björklund and Jäntti (1997)	.28	
	Österberg (2000)	.13	
	Björklund and Chadwick (2003)	.25	
Finland	Jäntti and Österbacka (1996)	.22	
	Österbacka (2001)	.13	
Norway	Bratberg et al. (2002)	.12	.12 refers to the 1960-cohort and .17 to the 1950-cohort
		.17	
Canada	Corak and Heisz (1999)	.23	
UK	Atkinson et al. (1983)	.42	
	Dearden et al. (1997)	.57	
Germany	Couch and Dunn (1997)	.11	Couch and Dunn use very young sons. Because earnings at young age poorly reflect long-run earnings, the young sons could explain the low estimate compared to Wiegand.
	Wiegand (1997)	.34	
US	Solon (1992)	.43	
	Zimmerman (1992)	.45	

8.2 What have education and education policy to do with it?

8.2.1 Theory

The natural question that follows from this short survey is what education and education have to do with it. To examine the role of education and education policy in the intergenerational earnings mobility process, it useful to start with a simple theoretical framework that gives a role for the family and a role for

education policy. Gary Solon (2003) has recently elaborated on a classical model of Becker and Tomes (1979, 1986) to highlight factors that could help explain cross-national patterns of intergenerational earnings mobility. The model is a highly stylized one, and considers a family's decision to invest in the human capital of their child. The human capital of the child that can be influenced by parents should be interpreted broadly as both education and health.

The parents face a budget constraint, so they can use their lifetime earnings to either consume themselves or invest in their child's human capital. Because they are assumed to get utility both from own consumption and from their child's utility, they invest in their child's human capital until the marginal utility of the investment equals the marginal utility of their own consumption. The marginal benefit of investing in the child's human capital depends on several factors.

First, there is a technology that translates parental investment into their child's human capital. The efficacy of this technology determines how much human-capital output is received from a certain amount of input. The input to this human-capital production process comes from the family, but also from public education policy. Because there is diminishing marginal returns to investment in human capital, the more input that the public provides, the lower the marginal return will be for parent's additional contributions. Thus, public investments will, at least partly, crowd out parental investments. Public investments, in turn, need not be the same for all children. Solon's model allows for a degree of progressivity in public investment; the more the ratio of public investment to parent's income declines with parental income, the more progressive policy is.

The second factor that determines parent's marginal benefits of investing in the child's human capital is the rate of return to the human capital in the labor market. The more the labor market pays for the human-capital attributes that investments generate, the more valuable the investments will be.

The human capital that the child brings to the labor market has two sources, namely the human capital that is generated by the investments and the human capital that is mechanically transmitted to the child without any investments. For example, genetically inherited human-capital attributes are received without any investment expenditure.

The earnings of the child will thus depend on (i) parental investments in human capital, (ii) public investments in human capital, (iii) the efficacy of the

investments, (iv) the mechanically transmitted human capital attributes, and (v) the earnings return to human capital. Finally, Solon shows that within the framework of this model, the intergenerational elasticity is greater as:

- mechanical heritability is greater;
- human-capital investment is more productive;
- the earnings return to human capital is greater; and
- public investment in children's human capital is less progressive.

It is hard to believe that the first and second mechanisms should be stronger in the United States than in Sweden. The main candidates to explain the US-Sweden differences are the third and fourth ones.

8.2.2 Empirical results

Could the difference between the US and Swedish intergenerational elasticities be driven by the differences in the return to schooling? Or is schooling per se more equally inherited in Sweden than in the United States? To investigate whether these are reasonable explanations, we perform a simple decomposition analysis of the elasticities for the two countries.[60]

The starting point is that we can model father's and child's earnings as a function of a simple indicator of education. We follow the labor-economics tradition and use years of schooling. Consider the following equations for father's (*f*) and child's (*c*) long-run earnings:

$$Y_f = \alpha_f + \beta_f X_f + e_f \tag{8.1}$$

$$Y_c = \alpha_c + \beta_c X_c + e_c \tag{8.2}$$

where Y_f and Y_c are the long-run log earnings measures for fathers and children used to estimate the intergenerational elasticities, X_f and X_c are years of schooling with associated returns β_f and β_c, α_f and α_c are intercepts, and e_f and e_c are errors terms.

It follows that the intergenerational elasticity equals:

[60] This exercise is inspired by Österbacka (2001).

$$\begin{aligned}\mathrm{Cov}(Y_f, Y_c)\big/\sigma^2 &= [\beta_f \beta_c \,\mathrm{Cov}(X_f, X_c) + \beta_f \,\mathrm{Cov}(X_f, e_c)]\big/\sigma^2 \\ &\quad + [\beta_c \,\mathrm{Cov}(e_f, X_c) + \mathrm{Cov}(e_f, e_c)]\big/\sigma^2\end{aligned} \tag{8.3}$$

where σ^2 equals the variance of father's earnings.

By estimating (8.1) and (8.2) and the four covariances in (8.3), one can compute the four components of the elasticity on the right-hand side of (8.3). Then one can address a set of counterfactual questions: what would Sweden's elasticity be if Sweden had US returns to schooling and what it would be if it had US covariances? The corresponding questions can be asked about the US.

We report the results of such an exercise in Table 2. We used a large data set of Swedish sons born 1951 to 1963 and their fathers defined as resident fathers in the 1970 census. We used annual earnings from employment as our outcome variable. The sons' outcome is measured in 1993 and fathers' earnings as the average of annual earnings in 1970 and 1975. Further, we transformed information about fathers' and sons educational level and field into years of education. For this purpose we used the 1970 census for fathers and the 1996 education register for sons.

The estimate of the intergenerational income elasticity is 0.211, the components of which are reported in the first row in Table 2. As seen in the table, our estimated β's are quite large; they are 0.089 for fathers in 1970-75, and 0.075 for sons in 1993. In particular the latter estimate is quite high by Swedish standards.

Next, we did the estimations for the United States using data from the Panel Study of Income Dynamics (PSID), the most frequently used US data source for labor market studies in general and intergenerational studies in particular. We defined an analysis sample as close as possible to the Swedish one. In this case a pair of a father and a son was defined as those who lived together in the same family in the first PSID survey in 1968. We restricted the sons to those who were born 1951 to 1963, the reason being that we measure sons' earnings in 1993 and we want them to be at least 30 years of age when we observe their earnings. Further, we employed the average of fathers' earnings in 1970 and 1975 as our measure of fathers' long-run earnings. We report the estimates in row 2 of Table 2. The intergenerational elasticity is higher in the United States than in Sweden, 0.343 vs. 0.211. The former estimate is slightly lower than in previous studies reported in Table 1, but for cross-national comparability

purposes we could not use the same sample and variables as in previous studies.[61] The difference between the two countries is quite marked though. Further, one can see that the estimated earnings returns for both fathers and sons are higher in the US.

Finally, we use the estimated equations to address the counterfactual questions: What would Sweden's intergenerational elasticity be if Sweden would have had the US returns? And what would the US intergenerational elasticity be if the returns would have been like in Sweden. The results from this exercise are quite striking. The Swedish counterfactual estimate would be as high as 0.329, which is very close to 0.343 for the United States. And the US estimate would fall to 0.247, which is not far from the estimated 0.211 for Sweden.

An alternative explanation to the Sweden-US differential in intergenerational elasticities could have been differences in the covariance between fathers' and sons' years of education. As can be seen in the third column of the table, the US covariance is indeed higher, 3.315 vs. 2.733 for Sweden. But that difference is not big enough to explain much of the difference. For example, applying the US covariance on Swedish data (but sticking to the Swedish returns) would only raise the Swedish estimate from 0.211 to 0.220.

Although this exercise is completely mechanical (and is not the outcome of a sophisticated behavioral model), it clearly demonstrates the importance of the differences in the earnings returns to education as a crucial factor behind the cross-national differences in intergenerational earnings elasticities.

Table 2 Actual and counterfactual components of the intergenerational income elasticity.

	$\frac{\mathrm{Cov}(Y_f, Y_c)}{\sigma^2}$	$\frac{\beta_f \beta_c \mathrm{Cov}(X_f, X_c)}{\sigma^2}$	$\frac{\beta_f \mathrm{Cov}(X_f, e_c)}{\sigma^2}$	$\frac{\beta_c \mathrm{Cov}(e_f, X_c)}{\sigma^2}$	$\frac{\mathrm{Cov}(e_f, e_c)}{\sigma^2}$
The intergen. elasticity					

[61] Most likely, the estimates would be slightly higher and closer to those in Table 1 if fathers' earnings would have been measured for more years. Using only two years – although the use of the time interval five years rather than two consecutive years probably helps – does not reduce the transitory component of earnings as much as if long-run earnings is estimated for more years.

Swedish estimates	0.211	$\frac{.089 \times .075 \times 2.733}{.249}$ = .073	$\frac{.089 \times .037}{.249}$ = .013	$\frac{.075 \times .119}{.249}$ = .037	$\frac{.022}{.249}$ = .088
US estimates	0.343	$\frac{.107 \times .136 \times 3.315}{.423}$ = .114	$\frac{.107 \times .253}{.423}$ = .064	$\frac{.136 \times .155}{.423}$ = .050	.115
US with Swedish returns	.247	.052	.053	.027	.115
Sweden with US returns	.329	.160	.016	.065	.088

Note: The components are explained in text.

8.3 Lessons from some specific reforms

So far we have focused on the overall intergenerational earnings relationship and asked what education in general has to do with it. But we are also interested in specific parts of education policy and learn about their impact on intergenerational associations. Although, we have found that the earnings return to education to a large extent explain the lower intergenerational elasticity in Sweden compared to the United States, it is natural to ask what specific parts of the Swedish educational system has done for intergenerational mobility. Interesting parts of Swedish educational policy in this respect are: the daycare system, the financial support for college students and the comprehensive school reform.

Our reading of the literature on these issues is that the empirical evidence is meager. This, in turn, is not due to the fact that those researchers who have addressed these issues have done a poor job, but rather that the information that is required for convincing empirical results is not there (or has not been found yet).

There is one interesting exception though. Meghir and Palme (2003) have recently exploited the variation is schooling that was generated by the comprehensive school reform that was implemented in Sweden during the 1950s and 1960s. The reason that the consequences of this reform can be analyzed in a compelling way is that it deliberately was implemented in different municipalities at different points in time. Although, the design was not purely experimental, the "quasi-experimental" variation in the data proved useful for Meghir and Palme. They found that those who lived in the

experimental municipalities and thus got longer compulsory education in a comprehensive school system continued with post-compulsory education to greater extent than those who lived in the other municipalities and got shorter compulsory education. Moreover, the impact on further post-compulsory education was bigger for those with a working-class family background. The study consequently supports the view that this reform was conducive to intergenerational mobility.

8.4 A backsliding during the 1990s?

It is obviously too early to see if the policy reforms and budget cuts during the 1990s created a backsliding so that intergenerational earnings mobility decreased. The students who were affected by these reforms were born from the mid-1970s onwards and have only very recently entered the labor market; some have not even finished their education. To address this issue, we instead look at the association between family background and school performance. We measure school performance by the grade-point-average at age 16 for graduates from the compulsory primary school. Considering what happened during the 1990s, it is natural to investigate whether the association between such an outcome and parental earnings increased or not. But it is also natural to consider a broader relationship between school performance and pupils' backgrounds. Due to the decentralization of school responsibility to municipalities and the introduction of school choice, it could be that municipality- and school-specific factors became more important. The sibling correlation is a useful measure that captures such broader neighborhood conditions in addition to family background factors. As noted above, the sibling correlation in an outcome like grades measures the fraction of the total variation in grades that is attributable to family and neighborhood conditions that are shared by siblings.

We report such results in Figure 1. The sibling correlations are estimated for biological full siblings who are born within three calendar years. Thus, they are quite closely spaced and their family and neighborhood conditions were likely quite similar as well. The table shows that the correlations were very close to 0.50 during the whole period. One can possibly discern a decline from the first years, but no significant increase over the period. Note also that the magnitude of the sibling correlations around 0.50 implies that half of the grade variation is explained by factors that siblings share. One can always ask whether such a

number suggests that "the glass is half-full or half-empty", but in any case siblings seem to share quite a lot of characteristics that affect their grades.

The second measure in the figure is the correlation between the grade-point-average and father's long-run earnings. This statistic has also been quite stable during the period. In particular, one cannot see any decline. The magnitude of the correlation is in the range 0.20-0.22. This magnitude implies that a one standard deviation move in the distribution of father's long-run earnings is associated with approximately a 0.20 standard deviation move in the grade distribution. Starting from median in the earnings distribution, this implies that a move to the 84th percentile is associated with a move from the median to the 58th percentile in the grade distribution. There is also another interesting interpretation of the two correlations in the table. The fraction of grade variation that can be explained – in a statistical sense, not necessarily causal sense – by father's earnings is only slightly above 0.04, to be compared to sibling correlation around 0.50. Remember that the sibling correlation also measures the fraction of the variation that is due to all factors that siblings share. Thus, other factors than parental earnings explain the bulk of what siblings have in common.

The overall conclusion then is that there is no backsliding in intergenerational mobility during the 1990s as long as we focus on school performance. The same result applies to the big-city areas where privatization went further and parents had more options for their children's school choice. Nonetheless, a complete analysis of the evolution of intergenerational mobility during the 1990s must also consider the consequences of rising earnings inequality and rising return to schooling. As we emphasized in section 8.3, the earnings return to education is also a driving force behind overall intergenerational earnings mobility.

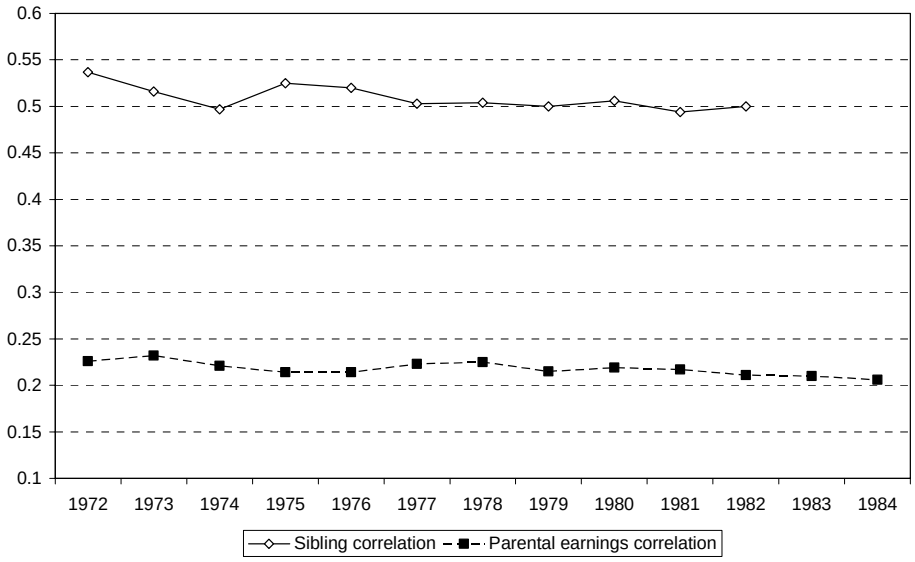

Figure 8.1 Family background and grade-point-average at age 16. Sibling correlations and correlations between grade-point-average and father's earnings.
Source: Björklund et al. (2003).
Notes: 1. The sibling correlations are estimated for full biological siblings who were born within three calendar years. The estimate for 1972 refers to siblings born 1972-74 and so on. 2. Standard errors are small, around 0.010 for sibling correlations, and 0.007 for the correlation with father's earnings.

8.5 Conclusions

We want to emphasize three major conclusions from this chapter. The first is that the association between parental and offspring's earnings that has been found in Sweden (and Finland and Norway) is weak compared to other countries and in particular compared to the United States. These results are based on long-run earnings at adult age for both generations. Thus, the results have been obtained for generations who went to school from approximately the late 1950s to the early 1970s. To the extent that these patterns are caused by education policy, they must be caused by the educational systems during these periods of time.

The second conclusion is that the Sweden-US differentials in these intergenerational associations to a large extent seem to be driven by higher earnings return to education in the United States. The intuition behind this explanation is that parental background in both countries has a quite similar impact on years of education, but these educational differences translate into a

higher earnings differential in the United States. In our concluding chapter, we return to the policy implications of this finding.

Finally, we examined whether the school reforms and other turbulent events during the 1990s has created a backsliding so that the association between family background and educational achievement will become stronger. For obvious reasons, it is yet too early to study the long-run consequences of these events on intergenerational associations of long-run earnings or final educational attainment. Our analysis of the evolution of the association between family background and grade point average at age 16 reveals that there is not much support for such a concern. This finding, as well, we return to in our concluding chapter.

References

Atkinson A., A. Maynard, and C. Trinder (1983), *Parents and Children: Incomes in two Generations*. Heinemann, London.

Becker G. and N. Tomes (1979), An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility, *Journal of Political Economy*, 87, 1153-1189.

Becker G. and N. Tomes (1986), Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, 4, S21-S39.

Björklund A. and M. Jäntti (1997), Intergenerational Income Mobility in Sweden Compared to the United States, *American Economic Review*, 87, 1009-1018.

Björklund, A and Chadwick (2003), Intergenerational Earnings Mobility in Permanent and Separated Families, *Economics Letters*, 80, 239-246.

Björklund A., T. Eriksson, M. Jäntti, O. Raaum, and E. Österbacka (2002), Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden compared to the United States, *Journal of Population Economics*, 15, 757-772.

Björklund A., M. Lindahl, and K. Sund (2003), Family Background and School Performance during a Turbulent Era of School Reforms, *Swedish Economic Policy Review*, forthcoming.

Bratberg E., O. Anti Nilsen, and K. Vaage (2002), Assessing Changes in Intergenerational Earnings Mobility, Working paper No. 26/02 Department of Economics, University of Bergen.

Corak, M. and A. Heisz, The Intergenerational Earnings Mobility of Canadian Men, *Journal of Human Resources*, 34, 504-533.

Couch K and T. Dunn (1997), Intergenerational Correlations in Labor Market Status: A Comparison of the United States and Germany, *Journal of Human Resources*, 32, 210-232.

Dearden L., S. Machin, and H. Reed (1997), Intergenerational Mobility in Britain, *Economic Journal*, 110, 47-64.

Erikson R. and J. Jonsson (1993), *Ursprung och utbildning*, SOU 1993:85.

Gustafsson B. (1994), The Degree and Pattern of Income Immobility in Sweden, *Review of Income and Wealth*, 40, 277-302.

Mazumder B. (2002), Earnings Mobility in the US: a New Look at the Intergenerational Mobility, mimeo.

Meghir C. and M. Palme (2003), Ability, Parental Background and Education Policy: Empirical Evidence from a Social Experiment, Working paper 03/05 Institute for Fiscal Studies.

Solon, G. (1992), Intergenerational Income Mobility in the United States, *American Economic Review*, 82, 393-408.

Solon, G. (1999), Intergenerational Mobility in the Labor Market, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Volume 3A, North Holland.

Solon G. (2002), Cross-country Differences in Intergenerational Earnings Mobility, *Journal of Economic Perspectives*, 16, 59-66.

Solon G (2003), A Model of Intergenerational Mobility over Time and Place, in M Corak (ed.), *Generational Income Mobility in North America and Europe*, Cambridge University Press, forthcoming.

Wiegand, J. (1997), Intergenerational Earnings Mobility in Germany, mimeo.

Zimmermann D. (1992), Regression Toward Mediocrity in Economic Stature, *American Economic Review*, 82, 409-436.

Österbacka E. (2001) Family Background and Economic Status in Finland, *Scandinavian Journal of Economics* 103, 467-484.

Österberg T (2000), Intergenerational Income Mobility in Sweden: What do Tax Data Show? *Review of Income and Wealth* 46, 421-436.

9 Conclusions

Although not comprehensive, our analyses have covered much ground. In this concluding chapter we return to the general issues that we addressed in the introduction and discuss the policy implications that follow from our analyses. One major issue is: How successful was the "traditional" Swedish educational policy, with its strong focus on egalitarian outcomes, when it was in place prior to the market-oriented reforms of the 1990s? What outcomes were equalized, and how much? Did equalization take place at the cost of efficiency? Our conclusions on these matters are discussed in sections 9.1 and 9.2. Then in section 9.3 we turn to our verdict regarding the market-oriented reforms during the 1990s. Next in section 9.4 we address the question of how to make the best use of scarce resources for Swedish education. The predicted shortage of teachers is an important part of this discussion.

Many of our conclusions are quite cautious for the obvious reason that we do not consider the available empirical evidence sufficiently convincing to render a strong verdict. One reason for this is simply that convincing causal analyses of many important educational policy questions are intrinsically difficult to conduct. Another reason is that relevant data are missing. After our examination of Swedish and international research we are convinced that evaluations of Swedish educational policies can be improved in several respects. Therefore, we conclude in section 9.5 with some recommendations for future policy evaluations. We stress that the governance of the Swedish schools as well as the overall evaluation of Swedish schools would be improved by more information from tests like those produced by the national tests. We also emphasize that Sweden has much to learn from the US tradition of running randomized experiments to evaluate new policy initiatives.

9.1 Has education policy equalized outcomes and opportunities?

If education policy successfully has equalized earnings, one would like to see direct causal links running from education policy to the skill distribution, and from the skill distribution to the earnings distribution. International comparisons show that Swedish pupils overall have done very well on standardized tests. The variation among students, however, is not particularly low, but the variation among schools is. Going one step further in the causal chain, we have

looked at the data from the International Adult Literacy Survey, a literacy test that was conducted in 1994. Swedish adults perform very well in this test. This holds for all birth cohorts (cohorts born from 1929 until 1970) and all education levels. In general, Swedes in the lower part of the skill distribution did very well compared to their fellows in other countries, in particular relative to those in the United States. These patterns do not directly prove anything about the contribution of education to the Swedish skill distribution, but they are hard to reconcile with the notion that the Swedish school system was in crises during the 1980s when the cohorts born through 1970 completed their compulsory education.

Finally, we have discussed the results from some recent attempts to estimate how much of the differentials in the skill distributions between Sweden (and other European countries) and the United States that could be due to the differences in the skill distributions. To us, the bottom line of these estimates is that a significant part of the lower earnings variation in Sweden can be attributed to a more equal skill distribution. Educational policy is probably important for this result but other factors may also contribute.

Another goal for Swedish education policy has been to reduce the impact of family background on educational achievement and thus increase equality of opportunity. A consistent result in the literature is that the correlation between the earnings of fathers and sons is higher in the United States than in Sweden and the other Nordic countries, indicating higher intergenerational mobility in the Nordic countries than in the United States. Equality of opportunity thus appears to be higher in Sweden than in the US.

Next we asked whether the Sweden-US differential in intergenerational mobility could be attributed to education and education policy. If, for example, subsidies to Swedish college students have successfully eliminated credit constraints facing prospective students from poorer family background, one would expect that higher intergenerational mobility in school attainment explains the intergenerational earnings mobility differentials. Our examination of US and Swedish intergenerational data showed, however, that the family associations in school attainment were not markedly different between the two countries. Instead, the weaker Swedish family associations in earnings could mainly be attributed to the lower earnings returns to schooling in Sweden. This outcome is intuitively quite reasonable: if children inherit (in a broad sense) educational attainment from their family background (also in a broad sense), the impact of this inheritance on earnings differentials becomes larger the

higher the earnings returns to education is. The fact that the difference in the returns to education mainly contributes to the difference in intergenerational earnings mobility does not necessarily mean that educational policies are unimportant. The earnings return itself is presumably a function of the kind of policy that has been pursued.

Nevertheless, this analysis suggests a potential trade-off for education policy. If high earnings returns to schooling are needed to attract students to higher education, there will also be larger earnings differentials by family background.

9.2 Is there an efficiency price for these egalitarian outcomes?

A major concern in Swedish public policy discussions has been that the private rates of return to higher education have been inefficiently low due to, e.g., union wage compression and progressive income taxes. This is a relevant question to address. The fact that Swedish adults rank very high on literacy and numeracy tests does not rule out that higher private incentives would have enhanced productivity and economic growth.

The private rate of return to university studies was quite low in the 1980s, a period when gross earnings returns were low and extremely high marginal taxes took much of the gross return from individual take-home pay. Even when the subsidy component of the student loans was taken into account, the estimated returns were low. As in many other countries, however, market forces have driven the gross returns upwards since then. Further, the tax reform in 1990-91 reduced the marginal tax rates to a maximum of around 50 percent. In addition, if the net value of the student loans is taken into consideration, the returns appear more substantial. Indeed, an international comparison of private rates of returns in the late 1990s for ten OECD countries shows that the Swedish returns were average at this point in time.

The Swedish returns are, however, substantially affected by the subsidies imparted by the student loans (and the tax free grant around $300 per month). It has been argued in the public policy discussion that this is a concern. The argument is that the returns are contingent on being a student, rather than to studying hard, picking the most rewarding fields of study, and using the skills efficiently in the labor market. This is an interesting argument, but we do not know of any evidence that can shed light on its validity. Rather, one could refer

to a recent Dutch study of students at University of Amsterdam; see Leuven et al. (2003). In this study, a randomly selected student group was offered financial rewards ranging from 200–600 Euros if they passed their exams at a quick (but feasible) speed. It turned out that this group of students did not complete their studies faster than the control group that was not offered such rewards.

However, there are other potential efficiency problems in Swedish education. One possible problem is that Swedish university students are old by international standards. This is probably the result of restricted entry in combination with generous admittance rules for older students. In addition, the possibility of using adult education to improve one's grades was introduced in 1997, further delaying the transition to university education. We have presented some crude calculations of the efficiency cost of an unnecessary delay in the completion of university studies; these costs are not trivial.

9.3 How did the development during the 1990s affect equality and efficiency?

The reforms involving decentralization and school choice during the 1990s are far reaching both from a Swedish historical perspective and from an international perspective. For the obvious reason that the pupils who were affected by these reforms have not yet entered the labor market, our analyses of these issues have employed school achievement measures like grades and national test scores.

The crucial productivity issue is whether the competition among schools that was introduced by these reforms has been productivity enhancing. Previous (mainly US-based) research suggests that private schools *per se* are not more efficient than public ones. Neither does the international literature support the notion that the presence of private schools raises productivity in nearby public schools.

The Swedish evidence is of course most pertinent for the productivity issue. A study, published a few years ago, attracted much attention by arguing that the presence of so called independent schools in a municipality is associated with better school achievement in the public schools in the same municipality, at least in Math. Our own analysis went one step further by looking at data covering several years. In our analysis we asked the following question: Did student achievement in municipalities with a rising share of students in independent schools increase more than achievement in other municipalities?

Our results suggest that this was the case. We found that the final grade in Math, English and Swedish at age 16 improved somewhat more where independent schools started than elsewhere, and the same result applied for the smaller samples where we used test scores.

Although we believe that our analysis has exploited the available non-experimental data in the best possible way, we have stressed that these questions are very difficult to answer. A crucial issue is why there are differential growth rates in the fraction of students in independent schools. Our estimates rely on the assumption that the factors that caused independent schools to sprout up more quickly in some areas than in others do not otherwise affect the growth rate of student performance. That is, we must make the rather strong assumption that observed differences in growth rates of performance associated with differential usage of independent schools can be attributed to the independent schools, rather than to unmeasured factors that are related to the diffusion of independent schools *per se*.

The main equality issue related to the reforms is whether students from wealthy and high-educated families have benefited more from the reforms than students from other family backgrounds. We have seen rising segregation of students by family background at the school level and that between-municipality differences in resources have increased. Further, we have presented new evidence suggesting that lower teacher density is particularly harmful for immigrant students. Thus the decline in resources over the decade might have been more harmful for students from poor family background than for students from more well-to-do family backgrounds.

Despite signs of increasing segregation and increasing dispersion of student performance, the relationship between family background and school performance (measured as the grade point average at age 16) has been remarkably stable in Sweden from 1988-2000. The association between the grade point average and parental income has been very stable. This is also true for the sibling correlation in grades. Thus, the composite importance of families, neighborhoods, and schools has not increased during the 1990s.

So, finally, where does this evidence lead us in terms of a final verdict of the reforms? We are inclined to stress that the evidence suggests that the positive (productivity) effects are not as obvious and as big as many advocates of school reforms have argued. But neither are the negative effects in terms of more unequal outcomes with respect to family background as evident as many skeptics have feared. At present, it seems unrealistic to argue that Sweden

should go back to the centralized system without much room for school choice that prevailed in the late 1980s. We think it is more prudent to find new ways for schools, parents and school politicians to evaluate the present system and to learn from the experience. In particular, we think it would be very unwise to, yet again, introduce a series of major school reforms without solid evidence about the consequences of such reforms. Our final recommendations in section 9.5 provide some examples on how to better evaluate teaching techniques, schools, and the present school system.

9.4 What is the best use of scarce resources?

Resource allocation at the school level

Economics is basically about getting the most out of scarce resources. Thus many observers of our profession would perhaps expect that the typical education economist is busy solving sophisticated optimization problems such as: How should a certain school budget be allocated among alternative uses in order to maximize (some well-defined measure of) student performance? Such an exercise would require information on the parameters in an "educational production function", showing how inputs like teachers, school buildings and equipment can be combined to produce output in terms of pupil attainment. At present, the research is far from providing such detailed knowledge. And it seems highly unlikely that future research will be able to fill the knowledge gaps.

Nonetheless, we are convinced that recent research on the importance of school resources offers some useful lessons for politicians. Our survey of recent research confirms two points that at first glance seem obvious: both teacher quality and class size (or the number of students per teacher) matter significantly for student performance. Ten years ago, a reading of the international literature on class size would rather suggest that class size hardly matters at all. This view was also widespread in our country despite the lack of studies on Swedish data. Today the consensus has changed, mainly due to a number of recent studies that have been able to circumvent the severe selection problems involved in studying the causal impact of class size.

Moreover, a recent Swedish study, which also extensively controls for omitted variables, estimates class-size effects in the same ballpark as the recent international studies. The results in this study are consistent with our own estimates. We have exploited the drastic differences in the teacher-student

ratios among Swedish municipalities that evolved because of the decentralization of school authority to the local level. Our new analysis shows that changes in student achievement at age 16 are positively associated with changes in the teacher-student ratio. These two Swedish studies are also in agreement with the international evidence in the sense that the impact on school achievement of smaller classes is stronger for pupils from disadvantaged family background.

A cost-benefit analysis based on these new estimates – neglecting all potential benefits except the earnings impact associated with enhanced cognitive skills – suggests that the internal rate of return on investments in smaller classes is around five percent. Although there are a lot of caveats to this calculation, it is clear that saving in school budgets by raising class size is no free lunch.

So what are reasonable guidelines for politicians and school leaders to use their scarce school budgets in the future? We think that this question must be considered in the light of the predicted future shortage of teachers. Our analysis of the teacher labor market showed that teachers' relative wages have declined for a long time and working conditions appear to have deteriorated during the 1990s. Given these facts it is no surprise that it has become more difficult to attract new generations of teachers to the profession. The main instrument used to affect teacher supply seems to have been the number study slots at teachers colleges. Our view is that this policy will not succeed. To increase the supply of teachers, and the quality of the teacher pool, it is important that the teaching profession becomes more attractive.

In surveys teachers often respond that smaller classes would make their job more attractive. So a possible policy recommendation could be to pursue smaller classes; such a policy might fulfill the dual objective of increasing the attractiveness of the teaching profession and improving student performance. In the present situation, however, such a policy could be counterproductive if it is phased in too quickly. The problem we have in mind is that smaller classes would raise the demand for teachers at the same time as the supply of teachers is falling. In turn, this might have severe consequences in schools that are deemed unattractive by teachers. In a situation with excess demand for teachers, it is possible that good teachers will leave troubled neighborhoods in order take a position in a more pleasant work environment.

So what should one do? This is a hard question. Because of decentralization it has become more difficult for the central government to influence the amount

of resources going into Swedish schools. At the same time we think that decentralization and individual wage setting may have injected a self-correcting mechanism that did not exist in the old system. Market forces will imply that teachers in unattractive areas and unattractive subjects will be offered a more favorable "compensation package" than others. The exact contents of this package – higher salaries, smaller classes, or other working conditions – might well differ across regions, subjects, and different parts of the school system. We simply do not have enough information to advocate one single measure to increase the attractiveness of the teaching profession.

Allocation among parts of the educational system

A fundamental issue, in terms of efficiency as well as equity, concerns the allocation of the total education budget among alternative uses like pre-school (or daycare), primary and secondary schools, post-secondary education, and adult "second-chance" education. At the theoretical level one can identify some key factors such as the evolution of individual learning ability, the "social discount rate", and the functioning of credit markets. But this is a long way from concrete proposals based on compelling evidence.

An interesting question in the larger resource allocation problem concerns pre-school. The Swedish government strongly emphasizes that this is a schooling activity rather than just the provision of child care. In this report we have not presented any Swedish evidence on the contribution of pre-school attendance to individual school performance, let alone labor market performance. We want to stress that this neglect is not due to lack of interest. It is simply due to the fact there are no data to base credible estimates on.[62]

With respect to the allocation of money between youth and adult education we think that there was too much money allocated to adult education during the second half of the 1990s. One basis for this skepticism is that some recent studies suggest markedly lower earnings returns on Swedish adult education than on other education. Another reason is that adult education at this point in time was subsidized at an unprecedented rate. Individuals may have been attracted to adult education by the fact that they received student pay at the

[62] The available evidence is based on simple regressions where individual school performance is related to time in pre-school and parental occupation, education, and income; see Söderström et al. (1999). There is a positive association between pre-school time and student achievement, but we are reluctant to say that this association reflects a causal relationship. Selectivity issues are likely to be a big concern here.

level of the UI-benefit, which in Sweden replaces up to 80 percent of foregone earnings. Finally, adult education competes with youth education in attracting qualified teachers. Since skilled teachers appear to be in short supply, this competition could have been detrimental for youth education.

There are also reasons to express doubts about another recent phenomenon within adult education. Prospective university students can improve on their grades by taking relatively easy courses within adult education. It seems to us that this is a bad use of scarce resources in general, particularly in a situation where Swedish university students are relatively old, there is a shortage of qualified teachers in youth education, and one can enter the university via the Swedish equivalent to the Scholastic Aptitude Test (*Högskoleprovet*). We think that the incentives for improving grades within adult education should be drastically reduced.

Having raised these critical points about adult education we wish to emphasize two things. First, we do not believe that individual learning capacity declines so fast that you cannot "teach old dogs new tricks". Indeed, some recent US evidence that we have referred to suggest that certain fields of adult education can have a substantial marginal return. Second, we think there is a role for second-chance education in an efficient and fair education system. We just think that the scale and the rate of subsidization of adult education have exceeded the optimal levels during recent years.

9.5 A call for evaluation

More frequent and better tests

The governance of Swedish schools would benefit from better and more frequent tests such as the national tests and standardized ("IQ-") tests like those done by the Department of Education in Gothenburg. Our empirical analysis of the latter data set demonstrated that both the national tests and the standardized tests at age 13 had substantial predictive power for adult earnings even after controlling for grades.

The use of quantitative tests is controversial – to a greater extent in Sweden than in the US. We acknowledge that tests can be misused. Problems may arise if school authorities use such test information to allocate resources to specific schools or to determine bonus payments to individual teachers. Such resource allocation systems may cause the informational content of the tests to be destroyed by excessive "teaching towards the test" or teacher cheating.

The Swedish school system does not feature high-stakes testing. But Swedish schools increasingly use results from national tests for marketing purposes. And the new requirement that "all schools' self-evaluations shall contain common and comparable measures of the results" implies that schools must continue to openly report their national test results. To avoid problems of a similar nature as in the US we think that school authorities should consider grading the tests centrally. In general, there is need for centralized evaluations of a decentralized school system for it to operate appropriately.

But test data also contain indispensable information for more general evaluations of the school system. The fact that test data can be misused should not prevent intelligent use of these data. Our report offers several analyses where changes in municipalities' school policies are related to changes in student performance. Such evaluations are, however, limited by the fact that national tests only are done in fifth and ninth grade. Further, until recently, data on test results have only been collected for the test in the ninth grade and a sample of municipalities. These facts have hindered the possibilities to gain valuable experience about Swedish education policies, in general, and the reforms during the 1990s in particular.

Recently it has been decided that the population of test results will be collected for the national test conducted in the ninth grade. All schools must participate in this test, but participation in the test in the fifth grade is voluntary. Even if participation in the fifth grade test would be obligatory, we think that it is far too late to get the first indication of potential problem at this point time. Presumably, it is easier to correct possible problems the sooner they are detected. We therefore recommend that tests are given at an earlier age, possibly already in the first grade. Testing students at regular intervals enables the school authorities to monitor individual and school performance more closely.

The regular assessment of Swedish schools and education policy would also benefit from more time-series information on pupil attainment. Such information requires tests that are comparable over time. The national tests do not have this property. They serve the important purpose of assisting teachers in their grading and, therefore, these tests cannot be identical over time.

Presently the only consistent time-series information on pupil achievement in compulsory school is provided by so-called UGU-data administered by the

Department of Education in Gothenburg.[63] A natural task for the National Board of Education would be to ensure that such studies are done regularly for random samples of the student population. It makes sense to have such time-series information not only for 6th graders, but also for, say, 3rd and 9th graders.

More randomized experiments

Swedish educational policy would benefit if more randomized experiments were conducted to evaluate educational initiatives. By randomly selecting some classrooms, or some schools, to implement a new teaching method (like using new equipment as a teaching device) and selecting others to serve as a control group that does not use the new method, the effects of the method could be much more accurately and confidently measured. Random assignment ensures a balance in terms of variables like student ability, motivation, parental involvement – and any other variable that might differentiate the groups. Thus, any difference in outcomes between the "treated" and "non-treated" groups is most likely to be due to the education initiative itself, and not to other factors. In observational studies, it is unclear whether classrooms that use different techniques have different outcomes (or the same outcomes!) because of factors that are not identified and measured by researchers.

The United States has a long tradition of running such experiments and utilizing the results in public policy deliberations. In particular, job and training programs have been evaluated this way since the 1970s. The US welfare reform in the 1990s was also preceded by several randomized experiments. As we have shown at several places in this report, such experiments have also been used to evaluate education initiatives; examples are privately-funded voucher programs, computer-aided instruction, and a large-scale study of smaller class sizes.[64] Sweden (as well as most European countries) is lagging behind in using this evaluation approach.

---

[63] Researchers at this department have administered one verbal, one spatial and one inductive test to random samples of 6th grade pupils for a quite long period of time. The tests have been done in 1961, 1966, 1980, 1985, 1990 and 1995; see Härnqvist (2000). The next test will not be done until 2005.

[64] Interestingly, in contrast to training and welfare, support for randomized experiments in the education field has not come mainly from the federal government. The recently launched No Child Left Behind Act, however, requires "scientifically based" research for education practices. Although this term is vaguely defined in the Act, we would predict more support for randomized experiments in the education field from the federal Department of Education in the future.

It is interesting to note that seemingly good education initiatives have not always worked well when evaluated in this way. The same conclusion applies to some of the US training programs. Two US economists, who summarized the US job and training programs, found that some programs had been successful and some had not. So they concluded by emphasizing that "good intentions are not enough". This is a general lesson that also applies to education initiatives.

Our recommendation is thus that randomized experiments are introduced as a regular evaluation device within Swedish education policy. One solution would be that the National Board of Education be provided with resources to develop the competence to do the required field work. We want to stress, however, that such evaluations should be done by autonomous bodies that are not subject to the influence of politicians. In this respect, contracting out of the field work and analysis seems a more sensible approach to us.

But it is important to realize that experimental studies can be difficult to conduct successfully. In particular, this is the case if the policy to be evaluated is not accepted by the teachers, school leaders, or school administrators who are directly affected by the policy. In general, individuals who are aware that they take part in an experiment may depart their normal behavior, in particular if there is something at stake for the subjects involved in the experiment.[65] Because of lack of experience, we believe that Sweden should not start with large-scale experiments like one that randomly assigns, say, 15 thousand first graders to small and regular-sized classes. It would be better to start with small-scale experiments to get methodological experience and make the education community used to the approach. Therefore, we believe that alternative teaching methods, like alternative use of computers in the class room or vouchers for summer courses, are more suitable candidates for the first wave of experimental evaluations.

Evaluation and the public policy discussion

We are convinced that a well-designed strategy for evaluation will be beneficial for education policy. For instance, a strategy involving more frequent tests will enlighten public policy discussions. In particular, we think that such information can prevent unjustified criticism. It seems to us that all school systems across the world are in crises more or less all of the time. Reliable

[65] This phenomenon is referred to as a "Hawthorne effect"; see, e.g., Heckman and Smith (1995).

information on the evolution of student achievement is an effective way of disciplining a destructive policy discussion that in the end may result in unjustified school reforms. Cross-country comparisons may fulfill the same objective.

In a similar vein, we think that reliable experimental information may discipline the discussion about, e.g., the effects of different teaching modes. Changes in school policies – be they minor or major – should be based on credible evidence and not on fads that have gained foothold among various consultants and in the media.

References

Heckman, J. and J. Smith (1995), Assessing the Case for Social Experiments, Journal of Economic Perspectives, 9, 85-110.

Härnqvist, K. (2000), Evaluation through Follow-up, in C-G. Jansson (ed.), Seven Swedish Longitudinal Studies in the Behavioral Sciences, Forskningsrådsnämnden, Stockholm,

Leuven, E, H. Oosterbeek, and B. van der Klaauw (2003), The Effect of Financial Rewards on Students' Achievement: Evidence from a Randomized Experiment, SCHOLAR Working Paper 2003:38, University of Amsterdam.

Söderström, L, A. Björklund, P-G. Edebalk, and A. Kruse (1999), Från dagis till servicehus – välfärdspolitik i livets olika skeden, Välfärdspolitiska rådets rapport 1999, SNS Förlag, Stockholm.