The Market Comes to Education in Sweden:
An Evaluation of Sweden’s Surprising School Reforms

by

Anders Björklund†
Melissa A. Clark‡
Per-Anders Edin†
Peter Fredriksson#
Alan B. Krueger*

May 15, 2005

† Swedish Institute for Social Research (SOFI), Stockholm University. E-mail: anders.bjorklund@sofi.su.se
‡ Mathematica Policy Research. E-mail: mclark@mathematica-mpr.com
§ Department of Economics, Uppsala University, and Institute for Labour Market Policy Evaluation (IFAU). E-mail: per-anders.edin@nek.uu.se.
# Department of Economics, Uppsala University, and IFAU. E-mail: peter.fredriksson@nek.uu.se.
* Department of Economics and Woodrow Wilson School, Princeton University. E-mail: akrueger@princeton.edu.
Table of contents

Preface ................................................................................................................................. 5

1 Introduction .................................................................................................................. 7
1.1 Goals of Education in Sweden ............................................................................. 9
1.2 The Swedish schooling system around 1990 ...................................................... 10
1.3 What happened during the 1990s? ................................................................... 11
1.4 Education and growth ....................................................................................... 15
1.5 Questions for the rest of the book .................................................................... 16
References .................................................................................................................. 19

2 Education, equality and efficiency ....................................................................... 21
2.1 Market failures ................................................................................................... 21
2.1.1 Static externalities ....................................................................................... 22
2.1.2 Growth externalities ................................................................................... 23
2.1.3 Credit constraints ....................................................................................... 24
2.1.4 Internality: A paternalistic argument for intervention .................................. 25
2.1.5 Negative externalities from sorting ............................................................ 26
2.2 Pre-existing distortions ..................................................................................... 27
2.3 Targeting ........................................................................................................... 27
2.3.1 Means targeted vs. universal programs .................................................... 28
2.3.2 Targeted for younger vs. older recipients .................................................. 28
2.4 Concluding remarks ......................................................................................... 29
References .................................................................................................................. 31

3 Education, skills and earnings – the Swedish record ........................................... 35
3.1 Inputs .................................................................................................................. 35
3.2 Development of formal schooling ......................................................................
3.3 The skills provided by the Swedish schooling system ..................................... 40
3.3.1 Skills among those still in school ............................................................... 40
3.3.2 Skills in the adult population ................................................................. 41
3.4 Wage inequality and the returns to schooling .................................................. 46
3.5 The dispersion of wages and skills .................................................................. 50
3.6 Summary and conclusions .............................................................................. 51
References .................................................................................................................. 53

4 Resources, decentralization, and student achievement ....................................... 55
4.1 How important are resources for outcomes? .................................................... 56
8 Family background and earnings: What do education and education policy have to do with it? .......................................................... 151
8.1 Intergenerational earnings mobility in Sweden, the United States and other countries ........................................................................ 152
8.2 What do education and education policy have to do with it? .......... 155
8.2.1 Theory ......................................................................................... 155
8.2.2 Empirical results ........................................................................ 156
8.3 Lessons from some specific reforms ......................................... 160
8.4 A backsliding during the 1990s? ................................................. 161
8.5 Conclusions ............................................................................... 163
References ........................................................................................................ 165

9 Conclusions .................................................................................... 167
9.1 Has education policy equalized outcomes and opportunities? .... 168
9.2 Is there an efficiency price for these egalitarian outcomes? ......... 169
9.3 How did the development during the 1990s affect equality and efficiency? ................................................................................ 170
9.4 What is the best use of scarce resources? ......................................... 172
9.5 A call for evaluation ................................................................ 176
References ........................................................................................................ 180
Preface

This book is an expanded, English-language version of a 2003 Report we prepared for the SNS Welfare Policy Group (*Välfärdspolitiska rådet*). This version is more technical than the Swedish version and contains more details on some of the analyses. More background material is also provided to familiarize a non-Swedish audience with the Swedish school system and reforms. The initiative to conduct the study originated from SNS, while the research presented here to a large extent represents ongoing research projects of the individual authors. We are grateful to SNS for their support.

We are all empirical labor economists. Therefore, we are inclined to focus on measurable aspects of what schools produce, in part because making sense of data is our comparative advantage. 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 figure prominently into our analysis. 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 produce. One may, of course, question whether test results accurately measure “knowledge.” Nonetheless, they clearly have some informative value, as evidenced by the fact performance on standardized tests is predictive of adult labor market success.

As economists, we are also interested in studying the production process underlying education. Economists tend to think of schools as human capital factories: they take inputs and convert them into outputs. In a market system, competition ensures that inputs are used effectively. Here we examine how inputs, like the number of teachers per student, translate into outputs, like achievement test scores and grades? We also ask whether achievement test scores and grades relate to outcomes that society ultimately cares about, such as students’ employment prospects. And economics is particularly well suited to consider the intended and unintended consequences of introducing greater competition into the education sector.

---

1 SNS is the Swedish acronym for Studieförbundet Näringsliv och Samhälle, or the Center for Business and Policy Studies, an independent network of decision makers from the private and public sectors who share a commitment to social and economic development in Sweden.
The writing of the book has been funded by The Swedish Council for Working Life and Social Research (FAS), The Swedish National Agency for School Improvement, and the SNS-group. 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

Sweden has long been known for its pursuit of equality – both equality of opportunity and equality of outcomes. This well-earned reputation is reflected in the country’s history of maintaining a large public sector, generous social programs, and high and progressive income taxes. Egalitarian goals have also had a strong impact in other policy areas, especially in education policy. In 1842, for example, Sweden became the first country in the world to introduce compulsory schooling, largely motivated by a desire to pursue egalitarian economic and social goals. Indeed, Sweden has succeeded in producing one of most equal distributions of income in the world, while at the same time affording a high standard of living.

For all these reasons, it will come as a surprise to many readers that Sweden introduced sweeping market-oriented reforms in public education in the 1990s that encouraged private schools to compete with public schools, delegated authority for the public schools from the central government to the local municipalities, and greatly increased variability in the resources available to students in different social classes and geographic areas. This book provides an initial evaluation of the effects of these sweeping reforms on student achievement in Sweden. In addition to considering the effects of the reforms on the level of academic performance, we also consider the effects on the variability in student performance, as well as on Sweden’s rank in international comparisons of educational outcomes.

Before we summarize the plan for the rest of the book and the major findings of our analyses, it is worth asking: What provoked such dramatic changes in public education in Sweden, of all countries? Sweden’s major market-oriented education reforms were put in place when the incumbent Social Democratic party lost office in 1991 and a center right coalition government assumed power for the first time in a decade. However, the movement towards these reforms, and indeed the more general growing public support for decentralization that may have led to the center right government’s success in the 1991 election, began several years earlier, and persisted even after the Social Democrats regained power in 1994.

The groundwork for the reforms may have been laid in the 1980s, when the “spirit of the times” began to change markedly in Sweden, as in many other countries. The concepts of decentralization and “goal steering” – or the practice of governing by defining a broad set of goals for local governments to achieve
rather than by setting strict regulations – became very fashionable. Decentralization was sometimes depicted as a panacea for solving problems in the public sector, which was criticized as being too large, expensive, and inefficient. Several public sector activities were accordingly decentralized. Amidst this growing public support for decentralization, there was also growing dissatisfaction with Swedish schools, which were highly centralized. It was widely argued that Swedish schools were performing poorly, and that Sweden’s once lofty position in international comparisons on reading, math, and science tests had slipped. The chorus of critics espousing a crisis in Swedish public school performance continued to increase in the late 1980s and early 1990s.

The first major steps towards decentralization of public education were taken in 1989, when the Social Democratic government passed legislation that made the local municipalities the main employers for teachers. This reform, passed after a major political battle with the teachers’ unions, allowed for greater variability in teacher pay and working conditions. The government also abolished some of the detailed central regulations of the schools. The major central authority (Skolöverstyrelsen) that oversaw these regulations – and was considered a symbol of Swedish bureaucracy – was closed down and replaced by a new body whose main task was to implement goal steering of the schools.

Around the same time, the Swedish economy suffered a severe setback. Unemployment surged from 1.8 percent of the labor force in 1990 to 9.3 percent in 1993, and crested at 10.1 percent in 1997 – the highest rate since the Great Depression. This economic downturn, coupled with growing public support for decentralization, likely contributed to the 1991 election of the center-right coalition government and may also have provided momentum for the additional school reforms the new government implemented.

In 1993 the center-right government established block grants to the municipalities, which greatly increased the financial responsibility of the local governments. The municipalities could spend the block grant on schools as they saw fit and could even shift the money to other services, such as elderly care. To some extent, the groundwork for these reforms was laid by the Social Democrat’s devolution of responsibility for teacher employment a few years

2 The International Adult Literacy Survey and PISA evaluation of student achievement – which indicated that Swedes continued to do well academically by international standards – came out later, in 1994 and 2000, respectively.
earlier. The economic downturn and resulting government budget crisis also facilitated the reform – since the budget had to be cut, it was politically appealing to delegate authority to the local governments, and to let them take the blame for the cuts. Around the same time the government also introduced the voucher system which led to the expansion of independent schools – schools which are publicly funded but privately run.3

Continuing concern over the performance of the public schools as well as the continuing fiscal crisis may explain why the Social Democrats did not dismantle any of the center right’s reforms when they regained power in 1994. The influential Lindbeck Commission report of 1993 had painted an unflattering picture of the performance of Swedish schools. On the basis of evidence then available, the Commission argued that devoting resources to reducing class size would not improve student achievement, so it instead recommended more homework and larger classes.4 This environment supported the policy reforms that the center-right government advocated and also made it easier for municipalities to implement changes. The economy remained weak, which contributed to the continuing political appeal of the block grants. Finally, the Social Democratic party was also weak, as they did not have a majority in the government and were forced to rely on support from the center right party that had established many of the major reforms a few years earlier.

1.1 Goals of Education in Sweden

Sweden’s radical market-oriented reforms of the 1990s must be evaluated in the context of the longstanding egalitarian goals of Swedish education policy. Ever since the early introduction of the compulsory school in 1842, efforts to promote equality have been central to Swedish education policy. As discussed in a recent policy document (see Regeringens skrivelse 2001/02: 188), these egalitarian goals still feature prominently in Swedish education policy, which attempts to promote both equality of outcomes and equality of opportunities.

The goal of equalizing outcomes, such as cognitive skills, has been evident in many ways. For example, the compulsory schooling system has gradually

---

3 Independent schools are generally non-fee charging, although prior to 1997 they were allowed to charge a small amount of tuition.

4 As with test scores, more evidence would subsequently become available that challenged the conclusion about class size, but the commission did not have access to that information at the time.
become more comprehensive – students are not grouped according to ability, and all follow a similar curriculum. Furthermore, extra resources have been allocated to students with special needs, such as handicapped students and children of immigrants. The goal of equalizing opportunities, on the other hand, has generally been interpreted as an effort to weaken the link between students’ family background and their subsequent educational attainment, which may in turn reduce earnings and income inequality.

Through the 1980s, Sweden appears to have been quite successful in terms of achieving 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. 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 those found in point-in-time income data. However, the contribution of education policy to these egalitarian outcomes remains an unsettled issue.

1.2 The Swedish schooling system in the early 1990s

As a starting point for evaluating the reforms that followed, it is useful to describe the Swedish school and daycare system at the beginning of the 1990s. At the time, Sweden had an extensive public daycare system for children ages 1-6. This system was heavily subsidized, and the fees paid by families covered only 15-20 percent of the average cost per child. The public daycare system had expanded rapidly since the late 1960s, and by 1989 the supply of spaces in the system more or less met demand.

Since the mid 1960s, Sweden has had nine years of tuition-free compulsory education starting at age 7, which is late by American standards. The compulsory schooling is comprehensive, meaning that all children follow essentially the same curriculum, which is determined by the central government. Upper-secondary school is voluntary and offers several programs, ranging from vocational training to programs that prepare students for further studies at the university level. In the early 1990s, about 80 percent of Swedish

---

6 See Aaberge, et al. (2002).
school children continue from compulsory school to any of the study tracks at the upper-secondary level.

For some time, the daycares, primary, and secondary education have been operated by the municipalities. Nonetheless, the system was highly centralized throughout the 1980s. The central government determined the basic goals and curricula and provided earmarked money for the schools. A national body was responsible for evaluation of the schools’ performance. There was little leeway for the individual municipalities to deviate from the national standards, although the rules allowed the municipalities to “top up” their resources with local funding.

Swedish youth could typically apply for university education 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, and they are not allowed to charge tuition. At the beginning of the 1990s, 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 stipend of around $300 per month. An important purpose of the public financial support system was to eliminate any credit constraints that prospective students might face, so that families with more limited resources could still send their children to college. Nonetheless, the financial support system is universal, so even students from wealthy families have been eligible for the loans and the stipend, reflecting the preference for universalism in Swedish education and social policy.

By tradition, prospective university students apply to a specific university and must specify a field of study (such as law, medicine, engineering, business administration, social work, etc.) at the time of application. This may compel young people to wait for a few years to decide which career they would like to pursue before they move on from high school to university, and Swedish university students are relatively old by international standards. The admission rules also tend to favor older applicants.

1.3 What happened during the 1990s?

As mentioned above, Sweden made a number of radical changes to its schooling system during the 1990s. These included changes in governance, changes in the amount of resources allocated to different types of education,

---

7 This corresponds to the 1990 stipend in 2003 Swedish Krona (SEK), which has been converted into US dollars using a conversion rate of 8 SEK/$.
and changes in enrollment at different levels of the system. Some of these changes were motivated by more traditional egalitarian arguments, while others were motivated by efficiency concerns. They all raise important issues about tradeoffs in education policy.

**Governance**

The changes in governance of Swedish schools during the 1990s in many ways represent a radical ideological shift, reflecting a movement towards decentralization, goal steering, accountability, parental choice, and competition. In this respect, Sweden has followed the same route as many other OECD countries. Indeed, Levin (1998) refers to an “epidemic” of decentralizing education policy reforms in the OECD countries. Nonetheless, the quick and radical restructuring of Swedish education during the 1990s seems to have made the school system one of the most decentralized in the entire OECD (Lindblad et al. 2002). Decentralization itself introduces a form of competition into the market for education because people can, and do, vote with their feet in choosing where to reside. The quality of local schools is a major factor in residential choice, with important implications for local property values.

The Swedish central government took the first major step towards its decentralization of public education in 1990, when it transferred the authority for primary and secondary education to the municipalities. As a result of this reform, the municipalities were given full financial responsibility for primary and secondary schools. Although the central government continued to redistribute financial resources from rich to poor municipalities, the fund earmarked specifically for education were gradually reduced and were completely eliminated by 1993. Thus, the scope for differences in expenditures on education across municipalities increased considerably.

Responsibility for teacher employment was shifted from the central government to the municipalities in 1989, although teacher pay negotiations remained centralized. This changed in 1995 when responsibility for teacher pay negotiations was transferred to the school level. Thus, since that time, school-level factors may have affected wages to a greater extent. Many school managers have used the reform as an opportunity to move to more individualized wage setting for teachers. During the 1990s, a long-predicted shortage of trained teachers developed, and schools that sought licensed teachers needed to offer a higher starting salary to fill their vacancies. This
development occurred after an initial wage increase for teachers in 1990. This increase was a visible price that the government paid to persuade teachers' unions to accept the decentralization of primary and secondary education to the municipalities.

School choice was introduced in 1992. The reform allowed parents to choose between all public schools in the municipality, subject to space limitations. However, residing close to a school (the residence principle, närhetsprincipen) is still 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 for their children to attend this school are given little weight. In 2000, however, Stockholm city introduced a major deviation from the residence principle. For upper-secondary education, the city introduced a system where admission is based exclusively on student achievement, as measured by compulsory school grades.

Also beginning in 1992, municipalities were required to fund independent, privately operated schools. In 2002, almost six percent of students at the primary and lower-secondary level attend a private school, a sharp increase from the less than one percent who attended such schools in 1990. Families were given complete freedom to choose between a private and public school, provided that a private alternative was available. Private schools exist in about half of the municipalities, and tend to be located in large urban areas.

In parallel with the move to decentralization and the introduction of school choice, the government placed an increased emphasis on goal steering. Schools were provided with a broad set of goals and were given the responsibility of deciding how to evaluate themselves and determining whether they fulfilled the general goals or not. As an element of their self-evaluation, the schools could use results from the national tests administered to students in grades 5 and 9. The actual use of these tests has varied among schools, however. Some schools and municipalities have published tables showing their average test scores or grade point averages, while others have not.

Resources
The financial crisis in the public sector has also had consequences for the level of resources devoted to daycare and schools. The child-staff ratios in the daycare centers have increased markedly since the late 1980s. During most of the 1990s, daycare fees were also raised considerably, although they were dramatically reduced in 2002 when the central parliament enforced a ceiling on
fees. The magnitude of the reduction varied among municipalities depending on the fee structure previously in place, but proportionate reductions in excess of 50 percent were common. By 2001 the fees had become quite high, so the ceiling implied a large reduction in expenditures for families with small children.

In compulsory schools, the student-teacher ratio increased considerably during the 1990s. After having decreased gradually for many years, the student-teacher ratio increased from close to 11 in 1991 to over 13 in 1997. Expenditures per student in compulsory schools fell markedly from 1990 to 1995, followed by a slight recovery in the second half of the 1990s.

At the same time that resources per student declined in the 1990s, computers were introduced in Swedish schools on a large scale. Teaching techniques changed as well, partly because the computers offered opportunities for new types of instruction. In many schools, the traditional concept of a “school class” lost its meaning due to changes in teaching styles. It is fair to say that there was no overall consensus in the educational community behind these sweeping changes. Thus, the changes made for a turbulent decade in Swedish schools.

**School enrollment**

By 1990, the available number of public daycare slots by and large met the demand, so the expansion of slots tapered off during the ensuing decade. One change to the daycare system, however, was that more emphasis was placed on preparing six-year olds for primary school, so in practice Sweden introduced a Kindergarten-type of program.

School enrollment 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 the vocational programs. A completed curriculum at a vocational program now implies that the student fulfills the so-called general requirement for entering university studies. This in turn means that some of the university programs are available for vocational students. The enrollment rate at upper-secondary school increased during 1990s: by the second part of the decade around 95 percent of each cohort participated in such studies.

During the second half of the 1980s, university enrollment rates had started to increase somewhat, after having fallen sharply in the early 1970s; see Fredriksson (1997). When the youth labor market deteriorated in the early
1990s, enrollment 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 unemployment or participation 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. A motivation for this change was that it would facilitate the recruitment of new students to higher education, especially students from working class families.

Other types of adult education also expanded in the 1990s. Between 1990 and 1993, as employment fell rapidly, labor market training (typically with a classroom training component) became the most common type of labor market program. By the mid-1990s, this measure was largely replaced by work-related programs. Unemployed persons were required to participate in these programs in order to renew their unemployment benefits after their UI eligibility period of 60 weeks had expired.

During the second half of the decade, adult education expanded enormously as a consequence of Sweden’s adult education initiative (kunskapslyftet). This program gave unemployed individuals the opportunity to maintain 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 programs, this new initiative helped low-skilled adults to increase their education levels and, in some cases, qualify for study at the university level.

1.4 Education and growth

So far we have emphasized the egalitarian goals of Swedish education policy. However, growth and efficiency are also prominent concerns of education policy in Sweden. The subject of how Sweden’s economic growth compares to that of other countries, and the role that public policy can play in improving it, has been discussed quite intensively over the past 10-15 years. The discussion has focused on reasons why Sweden’s position in international comparisons of GDP per capita has fallen from a top rank of 3rd or 4th in the world in the early 1970s to a more mediocre rank of only 15 to 18 in 1995.

---

8 See, for example, the interchange between Korpi (2000) and Henrekson (2001). See also Lindbeck (1997) and Freeman, Topel and Swedenborg (1997).
Much of this discussion has centered on the classical question of whether Sweden’s high taxes and high public spending hinder economic growth. In addition to the school reforms, during the 1990s Sweden also implemented a major tax reform, joined the European Union, successfully lowered inflation, and deregulated many markets. These developments were all motivated by a desire to strengthen growth and a drive to enhance efficiency.

Education policy has also been seen as a potential tool for stimulating economic growth and enhancing efficiency. Critics of Swedish education policy have focused on the low estimated wage returns to higher education. These returns were particularly low in the early 1980s when both wage compression and high marginal tax rates contributed to the low private, after-tax return to additional schooling. Some data suggested that Sweden’s labor force was poorly educated compared to that in other countries. Swedish job training programs were also criticized, since most evaluation studies from the 1990s suggested that these programs had poor or modest results.

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

1.5 Questions for the rest of the book

The dramatic changes to Swedish education policy during the 1990s raise a number of interesting research questions. The answers may be of interest to policymakers, school boards, and parents in Sweden as well as in other countries that are concerned about educational performance and are considering market-oriented school reforms. In many ways, Sweden can provide a laboratory for studying the impact of dramatic, market-oriented reforms in education. If such radical reforms to public education were undertaken in Sweden, they could certainly be undertaken in other countries as

---

9 The tax reform, among other things, lowered the highest marginal income tax rate from over 80 percent to 50 percent.
11 However, Edin et al. (1994) have shown this to be incorrect.
well. Furthermore, if the reforms increased inequality of achievement with little gain in efficiency for Sweden, a country with a strong safety net and narrow distribution of income, then other countries could risk an even greater increase in inequality from implementing similar reforms.

We begin in chapter 2 with a discussion of the basic theoretical arguments in favor of a public education policy, and we consider how these arguments relate to both efficiency and equality concerns. In chapter 3 we examine some basic empirical evidence about Swedish education policy. We report how Swedish students have fared in international comparisons and how the skills of the Swedish adult labor force compare to those of adults in other countries. We also report estimates of the private returns to schooling.

Chapters 4 through 6 focus on Sweden’s experience during the 1990s. Chapter 4 shows that the decentralization of education changed the allocation of school resources among Swedish municipalities. We then use this change in resource allocation to study the impact of school resources on student 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 eroded over time. In chapter 6 we investigate whether the competition induced by 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 in guiding parents in their choice of schools. We also examine how well grades and test results for students in the compulsory schooling years predict outcomes in adulthood, such as eventual educational attainment and labor market earnings.

In chapter 8 we examine Sweden’s success in meeting its goal of equalizing educational and labor market performance of individuals from different family backgrounds. In particular, we investigate whether the reforms of the 1990s increased disparities in educational attainment and labor market disparities relative to previous decades. Finally, we summarize our main findings in chapter 9, and also suggest directions for future evaluations of Swedish education policy.

Our overriding conclusion is that the effects of the education reforms have been exaggerated by both sides, by their proponents as well as their opponents. The reforms did increase the efficiency of the school system, but this increase
was modest. Inequality also increased over the decade, but at least so far, the increase in disparities in achievement resulting from the reforms appears to have been modest. Some groups, most notably immigrants, have not benefited from the availability of school choice, but by and large the newly decentralized, choice-driven system has had small positive effects on the academic performance of most Swedish students. The main conclusion from our study is that one should not expect miraculous results even from radical market-oriented education reforms, but neither should one expect a dramatic increase in inequality. We recommend that the best course of action may be to pause to monitor the effects and the implementation of the reforms, as well as to improve academic assessment procedures.
References


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


2 Education, equality and efficiency

In his classic 1975 book, *Equality and Efficiency: The Big Tradeoff*, Arthur Okun (p. 81) suggested that “a vigorous social effort to narrow the educational financing gap can improve both equality and efficiency.” Okun observed that unequal access to credit prevents many talented children from poor families from attending college. By improving educational opportunities for the less advantaged, Okun argued, society could improve both efficiency and equality, with no tradeoff between the two.

Public investment in education may improve equity by narrowing the distribution of income. There are two ways that this may occur. First, a targeted human capital policy can increase access to post-secondary education and to higher quality education for children from lower income families. It may thus improve their economic opportunities and, ultimately, their income relative to their more advantaged peers. A second, and more subtle way in which human capital policy can influence income distribution is by affecting the relative supply of low and high skilled workers. If human capital policy decreases the supply of less educated workers relative to the supply of more educated workers, there will be fewer unskilled workers competing for jobs in the domestic labor market. In addition, an increase in the relative number of skilled workers would be expected to benefit less skilled workers because skilled and unskilled workers are generally found to be complements in production. In the absence of any shifts in employer demand for skilled versus unskilled workers, the economic circumstances of the unskilled workers will improve.

Public investment in education can also improve economic efficiency by ensuring that educational resources are allocated to their most productive uses. In the remainder of this chapter, we review several reasons why subsidizing human capital accumulation, in addition to improving equity, may also improve efficiency, both generally and in the context of Sweden’s particular tax policies and social objectives. This broad discussion of the efficiency-enhancing aspects of educational policy will serve as a framework for the remainder of this book, in which we describe Sweden’s efforts to improve efficiency through its recent education reform efforts in the context of its longstanding egalitarian goals.

2.1 Market failures

Externalities and other market failures may lead individuals to invest less than the optimal amount in education, providing scope for the government to
improve economic efficiency by investing in education. In this section we discuss two types of externalities: those that lead to a one-time increase in social welfare, which we call static externalities, and those that affect society’s growth trajectory, which we call growth externalities. We then discuss capital market failures and other factors that may lead individuals to invest a suboptimal amount in education and provide scope for government intervention. Finally we discuss potential negative externalities from education.

2.1.1 Static externalities

There are several possible static externalities from education. Below we review arguments on, and evidence for, education’s external effects on the democratic process, crime, and unemployment and health care expenditures.

(1) It is widely argued that 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, Imrohoroglu et al., 2001, and Lochner and Moretti, 2004). All else 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) Several studies show that unemployment and education are negatively associated (e.g., Ashenfelter and Ham, 1979), and some experimental evidence from the US suggests that this effect is causal (Eberwein et al., 1997). There are therefore likely to be externalities associated with education that work
through the public budget: if increasing the average level of education in society reduces unemployment, it will increase tax revenues and reduce unemployment expenditures. Because of its generous social welfare benefits for those who are unemployed or employed at low earnings, external effects of education on social spending are therefore likely to be substantially larger in Sweden than in most other countries.

(4) Education may have a positive influence on 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 that increases in education do cause improvements in health (e.g., Lleras-Muney, 2002). If higher levels of education do indeed improve health, education will generate external benefits in the form of reduced health care expenditures in Sweden and other countries with publicly provided health care.

2.1.2 Growth externalities

It is also possible that increasing educational attainment may positively influence economic growth. Several economic models suggest that this may be the case (see, for example, Nelson and Phelps, 1966 and Romer, 1990). However, 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).

Education may also affect economic growth through its effects on the distribution of income. Persson and Tabellini (1994) develop a model 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. 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 rates also show an inverse relationship between GDP growth and inequality.

2.1.3 Credit constraints

In a world with perfect credit markets, all families would have equal access to credit, and all children would invest in educational resources up to the point that their marginal return to education equaled their discount rate. The evidence suggests that education decisions are not made in such a world, however. Children from poor families typically invest less in education than children from wealthier families. One of the most plausible explanations for this phenomenon is that poor families are credit constrained (i.e., cannot borrow at the same rate as everyone else), since students cannot easily use the return on their future human capital as collateral. Poor families may face different borrowing costs than rich ones, leading children from poor families to invest a suboptimal amount in education. If credit constraints do exist, policies to improve access to higher education among students from poor families may improve economic efficiency. (See, however, Carneiro and Heckman, 2002, for an argument that existing credit constraints in the US do not justify policies aimed at subsidizing higher education).

Several findings in the literature are consistent with the view that low-income families face credit constraints when making decisions about education. 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. Shea (2000) looks at the effect of differences in parental income emanating from noncompetitive factors, such as employment in a high-paying union job or industry, on children’s human capital. Income differences due to these factors, he argues, may be based on luck, and 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.”

Ellwood and Kane (2000) find that when the labor market 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,
again consistent with the notion that credit constraints may lead poor families to invest a suboptimal amount in their children’s education. Kane (1999) finds that college enrollment changes more in response to tuition changes, especially at the two-year-college level, than to equivalent, present-value changes in the payoff to education, also consistent with notion that poor families may be credit constrained in their educational decisions.

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 States, so credit constraints are less of an issue for low-income families in Sweden.

2.1.4 Internality: A paternalistic argument for intervention

Economic theory commonly assumes that individuals make educational decisions by implicitly comparing the costs and benefits associated with further investment. If the marginal benefit of an additional year of education exceeds the marginal cost, then a rational individual will invest the time and money in further study. However, 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 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.”

The impulsiveness and irrationality of youth may 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. McClure (2004) provides striking evidence that people’s brains struggle between impulsive response and rational calculation when confronted with problems involving delayed financial rewards. Educational investments require delayed returns. Like externalities, internalities cause the economy to operate at less than peak efficiency. The loss to society due to internalities could be quite large. In the case of cigarette smoking, for example, Gruber and Koszegi
(2001) contend that the societal loss due to internalities greatly exceeds the loss due to externalities (e.g., second-hand smoke). We suspect that the implicit belief that internalities are important is a major reason for public efforts to increase educational attainment. Indeed, because education is expected to improve individuals’ decision making, one can argue that increasing educational attainment and quality 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 may produce negative externalities as well as positive ones. In particular, if education serves only as a sorting mechanism that does not enhance individuals’ abilities, then policies 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 schooling age is increased, higher ability individuals will be forced to increase their educational attainment to distinguish themselves from those with lower ability. This additional schooling will be costly, but, by assumption, will not increase productivity. Wolf (2002) argues that this is the case in England; Card (2002) provides a critique.

In practice, the importance of the sorting externality has been very difficult to assess. Many of the implications of the sorting model are similar to those of a human capital model, in which education does enhance productivity. It has therefore proved difficult to distinguish between the two models. 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, because there is only a relatively small number of countries and many potential influences on GDP, and because it is difficult to measure 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 due to a progressive income tax system also create incentives for suboptimal private investment in education and training. Wage compression and high marginal tax rates reduce the private return to investment in education, so investment decisions are distorted.\footnote{In the special case in which education increases the value of leisure as much as it increases a worker’s productivity, taxes do 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-Ormaetxe (2002) for a discussion of these and related issues.} Such distortions may be particularly important in a socially progressive country such as Sweden. Edin and Holmlund (1995) calculate that the after-tax (pre-subsidy) internal rate of return from completing a university education in Sweden 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.

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

2.3 Targeting

An important issue in education policy is whether subsidies for higher education should be targeted or universal. Targeted subsidies vary with the family background or age of the recipient, while universal subsidies are available to all. In Sweden, subsidies for higher education are universal – most universities are publicly funded and free, and all students are eligible for monthly stipends and subsidized loans.

As we will discuss in Chapter 3, public expenditure on higher education increased considerably in the 1990s even as spending on compulsory and upper secondary education was reduced. Therefore, the question of the efficiency-
enhancing aspects of subsidies for higher education, as well as the question of whether these subsidies should be universal or targeted, is particularly relevant in the Swedish context.

2.3.1 Means tested vs. universal programs

If main the rationale for intervention in the education market is that poor families face credit constraints that lead them to invest a suboptimal amount in their children’s education, then the case for means tested subsidies – subsidies targeted based on income – is stronger. Means testing also increases the redistributive effect of education by increasing the educational attainment of the less advantaged relative to their more advantaged peers.

On the other hand, means tested subsidies provide a disincentive for wealth or income accumulation, since only lower income families are eligible. 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, the disincentive effects of means tested subsidies provide an argument in favor of universal subsidies, such as are common in Sweden. Moreover, because of the compressed after-tax distribution of income in Sweden, credit constraints among the poor may be less severe in Sweden than in most other countries, again alleviating the need for means tested subsidies. Participation in means tested programs could also carry a stigma. And yet another argument for universal programs is that political support for such programs is typically stronger than that for means tested program.

2.3.2 Targeted for younger vs. older recipients

Labor economists often argue 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. However, if there are liquidity constraints that require individuals to work while they are young in order to finance their education, or if the return to education unexpectedly increases, then it may make sense for some workers to make their investment in education at older ages.

Many have argued that human capital programs that invest in disadvantaged young children have a higher rate of return than those that invest in older children, since early skill development begets later skill development (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 returns can also be sizable for educational investments for older workers, at least for a subset of workers who opt for training.

In another study, Jacobson et al. (2003a) find that returns to educational investment 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 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 up 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 allocation of resources under which the marginal return to investment in education is equalized for older and younger workers is a difficult challenge. In the next chapter, we consider evidence on the effectiveness of second chance and adult education programs in Sweden.

### 2.4 Concluding remarks

In this chapter we have discussed reasons why public investment in education may enhance efficiency as well as equity. The arguments for public intervention are based on the positive externalities associated with education, and capital market imperfections and other factors that may lead individuals to make suboptimal decisions about educational investment. In addition, education policy may counteract existing distortions in the market due to economic policies that compress the wage distribution or impose high marginal tax rates.

Of course, the relevant question 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.

Sweden and other 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. In the following chapters we explore whether Sweden’s education reforms efforts of the 1990s managed to successfully enhance efficiency without compromising its longstanding egalitarian goals.
References


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


3 Education, skills and earnings – the Swedish record

What has the Swedish education system accomplished? To what extent has Sweden’s education policy produced more equality in skills? What effects have a compressed distribution of skills had on wage inequality? In this chapter we discuss the effects of the Swedish education system on educational attainment, skills, wages, and wage inequality. We focus on the current situation in Sweden and compare it to that of other (mainly OECD) countries. We also provide information on the historical development of the Swedish education system.

We begin in section 3.1 with a discussion of the resources devoted to education in Sweden. We then discuss various forms of output from the education system: the average level of educational attainment (section 3.2), the average level of skills (section 3.3), and the returns to schooling and wage inequality (section 3.4). In section 3.5 we discuss the relationship between the distribution of skills and the distribution of wages, and in Section 3.6 we conclude.

3.1 Inputs to Education in Sweden

Sweden devotes a large amount of its resources to public education. In 2001, overall expenditures on the education system amounted to 7.7 percent of GDP. This was substantially higher than the education expenditures of most other countries in the OECD area as a share of GDP.

Sweden’s spending on education remained fairly constant over the 1990s – it declined somewhat in the early 1990s but then returned to previous levels in the late 1990s, as shown in Table 3.1, columns 1 and 2. Over this same period, most other OECD countries, including the United States, reduced their education spending relative to GDP. Average OECD education expenditure as a percentage of GDP dropped from 6.1 to 5.6 percent between 1992 and 1999.

Sweden’s educational expenditures per student also remained roughly constant as a share of GDP per capita between 1992 and 1999 (Table 3.1, columns 3 and 4). These numbers are calculated from expenditure data at different levels of education (excluding preprimary education) using a fixed set of weights, so that the differences across countries and over time are not affected by changes in the composition of students across levels. In both 1994 and 1999, Sweden spent about 34 percent of GDP per capita on an “average
student.” Most other OECD countries spent somewhat less per student as a share of GDP per capita, and average per student spending 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

<table>
<thead>
<tr>
<th>Country</th>
<th>Total expenditure</th>
<th>Expenditure/student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>6.8</td>
<td>6.7</td>
</tr>
<tr>
<td>United States</td>
<td>7.0</td>
<td>6.5</td>
</tr>
<tr>
<td>OECD average</td>
<td>6.1</td>
<td>5.6</td>
</tr>
</tbody>
</table>

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.

Sweden’s high overall spending on education does not translate into high spending on teacher salaries, however. The share of total education expenditure devoted to teacher salaries is 48 percent in Sweden, considerably below the OECD average of 65 percent (OECD, 2002, Table B6.3). Expenditures on teacher salaries as a share of GDP in Sweden were 3.2 percent in 1999, also below the OECD average of 3.6 percent.

Table 3.2 provides a breakdown of educational expenditures by level of education. In Sweden, as in most other countries, spending on primary and lower secondary education (grades 1–9 in Sweden) comprises the largest component of total education expenditures. In 1999, Sweden spent 3.0 percent of its GDP on primary and lower secondary education, somewhat higher than the OECD average of 2.3 percent of GDP. Sweden’s expenditures on tertiary education as a share of GDP were also considerably higher than the average for all OECD countries. In contrast, Sweden’s education expenditure per capita on upper secondary education was similar to the average for all OECD countries.
Table 3.2 Total education expenditure as a percentage of GDP by level of education, 1999

<table>
<thead>
<tr>
<th>Country</th>
<th>Primary and lower secondary education</th>
<th>Upper secondary education</th>
<th>Tertiary education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>3.0</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td>United States</td>
<td>NA</td>
<td>NA</td>
<td>2.3</td>
</tr>
<tr>
<td>OECD average</td>
<td>2.3</td>
<td>1.3</td>
<td>1.3</td>
</tr>
</tbody>
</table>

Source: OECD (2002).

Table 3.3 shows a breakdown of education expenditure per student as a share of GDP per capita across countries in 1992 and 1999. For comparability across countries, grade levels are categorized differently here than above.\(^{14}\) In 1999, Sweden’s per student spending on upper secondary education was 25 percent of GDP per capita, the same as the OECD average. However, Sweden spent more per student on primary education as a share of GDP per capita than the OECD average (24 versus 19 percent), and substantially more on tertiary education more than the OECD average (61 versus 44 percent).

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

<table>
<thead>
<tr>
<th>Country</th>
<th>Primary education</th>
<th>Secondary education</th>
<th>Tertiary education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>29</td>
<td>24</td>
<td>33</td>
</tr>
<tr>
<td>United States</td>
<td>24</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>OECD average</td>
<td>21</td>
<td>19</td>
<td>27</td>
</tr>
</tbody>
</table>


There are also striking differences between Sweden and other countries in the change in the distribution of resources across the three levels of education during the 1990s. Although overall educational expenditures per student Sweden remained fairly constant as a share of GDP per capita between 1992

\(^{14}\) The primary education category include Swedish grades 1–6, while grades 7–9 are classified as secondary education.
and 1999, there was a substantial shift in resources across levels of education, as shown in Table 3.3. Per-student resources for tertiary education expanded dramatically, from 43 to 61 percent of GDP per capita, while primary and secondary education both experienced substantial reductions in resources per student as a share of GDP per capita. In contrast, in the United States and the OECD countries on average, the overall reduction in resources for education over this period was fairly symmetrically distributed across the three different levels of education.

Another related feature of the Swedish education system is the large share of resources devoted to relatively older students. This is evident in Table 3.4, which displays enrollment rates by age group in Sweden, the United States, and the OECD countries on average. For instance, enrollment rates among 30–39 year olds in 2000 were about 10 percentage points higher in Sweden than in the United States and the OECD countries on average. Sweden’s relatively high enrollment rates among older age groups are due in part to its high university enrollment age. The median age at university enrollment in Sweden in 2000 was 22.7, compared with an average median age of 19.4 in all OECD countries (last column of Table 3.4). Along with Iceland and New Zealand, Sweden had the highest median age of entry in the OECD area. Sweden’s adult education programs also contribute to its high enrollment rates among older age groups. Over the 1990s adult education programs expanded substantially in Sweden, especially with the introduction of the Adult Education Initiative, or “Kunskapslyftet” (see e.g. Ekström (2003).

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

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

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

The traditionally high age of university entry in Sweden affects calculations of the internal rate of returns to higher education, since older university
students have a shorter time period over which to realize the return to their investment. Based on estimates of the return to education in 1991 from Björklund and Kjellström (2002) and assuming a discount rate of 2 percent, we find that increasing the age of university entry from 19 to 22 years reduces the internal rate of return to a university degree by 7.5 percent. Thus, Sweden pays a non-negligible price for the fact that students are not able (or not willing) to make the transition directly from high school to university.

3.2 Educational attainment in Sweden
At the beginning of the 20th century, higher education was very rare in Sweden. In 1930 less than two percent of the adult population had completed upper secondary education or beyond (Table 3.5). The vast majority of the population had only completed some primary education, of varying length and intensity. Over the next 70 years, however, the situation changed dramatically. By 1970 upper secondary completion rates had increased considerably, but tertiary schooling completion rates were still fairly low. The effects of Sweden’s 1960 reforms of its university system were more evident by 2000, however, at which time over 30 percent of the adult population had completed tertiary education. Even though the expansion of the Swedish university system occurred long after the US expansion, it predated the expansion in many other European countries (Edin et al., 1994).

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

<table>
<thead>
<tr>
<th></th>
<th>1930</th>
<th>1970</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary, lower secondary education and below</td>
<td>92.8</td>
<td>59.0</td>
<td>17.8</td>
</tr>
<tr>
<td>Upper secondary education</td>
<td>0.6</td>
<td>29.6</td>
<td>50.8</td>
</tr>
<tr>
<td>Tertiary education</td>
<td>1.2</td>
<td>7.4</td>
<td>30.5</td>
</tr>
<tr>
<td>No information</td>
<td>5.4</td>
<td>4.0</td>
<td>0.9</td>
</tr>
</tbody>
</table>

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

While educational attainment is still lower in Sweden than in the US, it is higher than in most other OECD countries (Table 3.6). 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 average educational attainment in Sweden will eventually surpass that of the US if current enrollment patterns prevail (see OECD, 2002).

<table>
<thead>
<tr>
<th>Country</th>
<th>Primary and lower secondary education</th>
<th>Upper secondary education</th>
<th>Tertiary education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>19</td>
<td>49</td>
<td>32</td>
</tr>
<tr>
<td>United States</td>
<td>13</td>
<td>50</td>
<td>37</td>
</tr>
<tr>
<td>OECD average</td>
<td>34</td>
<td>41</td>
<td>21</td>
</tr>
</tbody>
</table>

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 only interesting if we have reasons to believe that students actually acquire useful skills in school. Over the past few decades, a number of comparative international studies of various forms of skill 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. Nonetheless, these studies have been carefully executed and provide interesting information. Based on the results of some of these studies, we discuss how the skills of the Swedish population compare to those of other countries, both among students still in school and among the entire labor force.

#### 3.3.1 Skills among those still in school

Since the first International Study of Achievement in Mathematics in 1964 (Husén, 1967), a number of international studies of student achievement have been conducted. These studies have covered students of different ages and in different subjects – mainly reading, mathematics, and science. 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 in mathematics and science tests for students in upper secondary school.

The youngest Swedish children (age 9–10) do very well comparatively on various literacy and science tests. In a recent study of reading skills, PIRLS 2001, Sweden ranked 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 the
ages and grades at which the tests were administered across countries. However, accounting for differences in age and grade across countries does not change the main picture: Swedish primary school students rank at the top in reading skills in international comparison.

The results are more modest for children in lower secondary school (ages 13-15). In most comparisons Swedish children score about average or above average on tests of math and science. Swedish lower secondary students perform quite well on tests of literacy, however. In the PISA 2000 study, Swedish 15-year olds performed significantly better than the average on general literacy tests as well as on tests of mathematical and scientific literacy (OECD, 2002).

Swedish students in their final year of upper secondary school also perform quite well in international comparisons. In the 1995 TIMSS, Swedish students ranked near the top of all upper secondary students in both mathematics and science scores. If we instead focus on “specialists” – students taking advanced courses in these subjects – Swedish students in their final year of upper secondary school rank near the top in science (physics), but mathematics results are closer to the international average.15

A final observation from these international comparisons 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 than 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, Sweden had the lowest between school variance of all the OECD countries, in which the average between-school variance was 35 percent. This may be largely related to the fact that the bottom of the skill distribution in Sweden is fairly compressed, while in the US and other countries there is a long tail of poor performers.

### 3.3.2 Skills in the adult population

The most comprehensive comparative study of adult literacy skills is the International Adult Literacy Survey (IALS) (see OECD and Statistics Canada, 1995). This study provides tests of three measures of literacy (prose, document

---

15 The TIMSS results are available at: http://isc.bc.edu/timss1995i/TIMSSPDF/C_Hilite.pdf
and quantitative) for representative samples of the adult population in several OECD countries.

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 literacy skills of the Swedish population are comparatively very high. The average score in Sweden is the highest among the 16 countries participating in the initial rounds of the survey. The dispersion of literary skills in Sweden is relatively low, but is not among the lowest.

**Table 3.7 Skills in the adult population according to the IALS**

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

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 fall in the US skill distribution. In Table 3.8 we report the quantitative test score for individuals at the 10th, 50th and 90th percentile of the Swedish test score distribution. The final column shows the percentiles of the US test score distribution at which these scores are located. For instance, an individual with a score of 230 would end up at the 10th percentile of the Swedish distribution and at the 30th percentile of the US distribution. The table shows that Swedish adults outperform US adults at all
levels of the skill distribution, and these differences seem to be more pronounced in the lower part of the distribution.

**Table 3.8 A comparison of the Swedish and US skill distribution, quantitative skills IALS**

<table>
<thead>
<tr>
<th>Swedish percentile</th>
<th>Score</th>
<th>US percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>10th</td>
<td>230</td>
<td>30th</td>
</tr>
<tr>
<td>50th</td>
<td>303</td>
<td>69th</td>
</tr>
<tr>
<td>90th</td>
<td>370</td>
<td>95th</td>
</tr>
</tbody>
</table>

Source: Authors’ calculation from the 1994 IALS.

It is also interesting to examine how Swedish schooling system (and other factors) have changed the skills of the population over time. One way to try to gauge this is to look at differences in skills across birth cohorts. The skills of each cohort are affected both by true cohort effects (due to schooling and other factors) and by the age of the cohort. As shown by Nathanaelsson (2003), however, aging effects are relatively unimportant in explaining differences in skills across cohorts in Swedish data (with some qualification for the oldest cohorts). Additionally, when we compare cohort skills across countries, it seems fairly reasonable to assume that the effect of aging on skills is similar across countries. Thus, we can interpret different trends in skills over cohorts as the result of cohort specific effects due to education and other factors.

Figure 3.1 provides an overall picture of skill levels across cohorts of the entire adult population in 1994 in five countries. Across all five countries, skill levels are generally higher in younger cohorts. However, the figure shows that the “growth rate” of skills across cohorts in Sweden seems to be declining in more recent cohorts. In all cohorts, the Swedish rank above their counterparts in the other four countries; however, Finland appears to be closing the gap with its more recent cohorts.
The next set of figures display trends across cohorts by level of education. Figure 3.2.a displays average skills by cohort for individuals with tertiary education. It is evident from this graph that the differences in cohort skill levels across countries are smaller for the highly educated than for the population as a whole. Among the highly educated, Sweden again ranks at the top of the five countries, with Finland closing the gap with its more recent cohorts. Skills are again generally higher among the younger cohorts in all five countries, with the exception of 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 average skill level of all cohorts in this education group in the US is considerably below that of the other four countries. The US is also the only one of the four countries in which average skill level is not generally increasing across cohorts in this education group. Trends across cohorts in this education group in Sweden are quite similar to those in Germany, and the more recent cohorts in Finland again appear to be closing the gap with those in Sweden.
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, and that Sweden continues to fare quite well in international comparison across all cohorts. This is true among both the university educated and those who have only completed compulsory schooling. To the extent that the trends in skills across cohorts reflect changes over time in the quality of education in different countries, these results suggest that the Swedish schooling system has continued to improve over time, and has remained strong in international comparison.

### 3.4 Wage inequality and the returns to schooling

In this section we 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 the assumption that monetary incentives are key determinants of the individual’s decision to invest in formal education. Of course, this may not be the only determinant of the demand for education, but monetary incentives are likely to have an effect at the margin.\(^{16}\) The recurring debate over whether the returns to investing in education are too low in Sweden due to its compressed wage distribution motivates a short review of the evidence.

It is well known that Sweden has a relatively compressed wage distribution compared to other countries. This was particularly evident in the early 1980s, when wage dispersion in Sweden was at its minimum. During the following two decades, wage dispersion increased in Sweden as well as in many other countries. Nonetheless, the wage distribution in Sweden remains more compressed than that of most other countries.

There is concern that this compression of the wage distribution lowers the wage differential between education groups – the wage premium – which is one of the key components of the return to education. Figure 3.5 shows how the wage premium in Sweden has changed over time. The estimates are based on a simple Mincer-specification with hourly earnings regressed on a linear years-of-schooling variable.\(^{17}\) Since the 1930 data do not contain hourly earnings, we

---

16 For evidence of the impact of monetary incentives on education decisions in Sweden, see Fredriksson (1997).

17 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.
also report estimates based on annual income for males for the 1930–1991 period.

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

The estimates for the post-1991 period are based on a different data source than the estimates for earlier years – register data instead of survey data. Therefore, the levels of these estimates may not be completely comparable. This is also likely to be part of the explanation for the break in the trend between 1991 and 1992. Over the 1990s, however, there is a clear pattern of
increasing education wage premiums, in particular at the end of the period.18 By the year 2000, the average wage premium associated with one additional year of schooling had increased to about 6 percent.

The wage premium associated with an additional year of 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 such as tuition fees, student support, and taxes. While tuition fees are nonexistent in Sweden, the effects of the tax system and student support (grants and subsidized loans) are non-negligible. Calculations of the internal rate of return to higher education that incorporate these two factors suggest 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. Replicating their calculations for the year 2000 suggests an internal rate of return of about 10 percent. The slight decrease in the return is due to a decrease in the generosity of student subsidies (loans and grants) between 1999 and 2000. Note that these calculations are based on survey data (the LNU), in which the increase in the wage premium over this period is less pronounced than in the register data.

An international comparison of the return to schooling is provided by OECD (2002). Their computations of the internal rates of return are not directly comparable with those presented above, as they are based on annual earnings instead of hourly wages and they also include differences in unemployment rates across education groups. Nonetheless, they provide a useful measure for comparison across countries.

Table 3.8 reports the OECD computations of the internal rates of return for both upper secondary and tertiary education, for males and females separately. These computations of the returns to tertiary education in Sweden are fairly close to those reported above – around 11 percent. The Swedish returns to tertiary education are substantially below those in some countries, like the US and UK, but are comparable to or above those in other countries.

Table 3.8 also demonstrates how much internal rates of returns vary across levels of education within countries. The Swedish return to upper secondary education is strikingly low compared both to the returns in other countries and

18 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. Reweighting the estimates to account for this change in coverage does not affect the overall trend in the estimates, however.
to the return to tertiary education. At least part of this low return is likely due to the increased supply of workers with an upper secondary education, due to the massive expansion of upper secondary education that has taken place in Sweden over the past several decades.

Table 3.9 Internal private rates of returns to upper secondary and tertiary education, 1997-2000

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

Source: OECD (2002).

* 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 comparatively 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 of selection into adult education which are difficult to account for in empirical work. For instance, participants may be very different from nonparticipants in terms of both skill and motivation, so simply comparing the earnings of the two groups will not provide a reliable estimate of the programs’ effects. Studies that do try to address these selection issues suggest that the average returns to adult secondary education are even lower than those presented above. Ekström (2003) estimates the earnings premium for various subgroups of adult education recipients about 10 years after they began 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 to Sweden – estimates for female immigrants are positive although not statistically
significant at conventional levels. Similar results are obtained by Stenberg (2003) in his study of the Swedish Adult Education Initiative. He finds that participants in this initiative have lower incomes than comparable participants in labor market training programs.19

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 skill dispersion? 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 at each level will affect the price of skills at that level. We summarize the evidence on these two mechanisms in turn.

Evidence suggests that the direct effect of skill dispersion on wage dispersion may be fairly small. 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 skill distribution in Sweden would explain only 12.5 percent of the difference in the standard deviation of earnings across the two 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), who use a decomposition method that allows for different effects of skill at different part of the wage distribution. An interesting aspect of their results is that the contribution of measured characteristics (including skills) to wages is higher at the lower end of the wage distribution. Their findings suggest that differences in age, schooling, and skills account for 26 percent of the 50-10 log wage differential between the US and Sweden.

19 Available evaluations of labor market training programs in Sweden suggest that these programs do not have positive returns (see e.g. Calmfors et al., 2001).
Both of these studies measure only the direct effect of differences in the distribution of skills across countries. Leuven et al. (2002) attempt to include the indirect effect as well – the effect of the net supply of skills on the price of skills. Applying the methodology of Katz and Murphy (1992), they find much stronger effects of cross-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 differences in relative wages in the lower parts of the wage distribution, accounting for about 60 percent of the variation. 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 affect differences in wage inequality across countries.

3.6 Summary and conclusions

In general, Sweden appears to devote considerable resources to education relative to other OECD countries. Both total education spending as a share of GDP and education expenditures per student as a share of GDP per capita are considerably higher in Sweden than the OECD average. However, Swedish spending on teacher salaries as a share of GDP is actually considerably lower than the OECD average.

During the 1990s, total education spending in Sweden as a share of GDP remained roughly constant, but the structure of the expenditure changed dramatically. There were large reductions in resources per student relative to GDP for primary and secondary education, while resources per student in tertiary education increased substantially. How the dramatic reduction in resources allocated to primary and secondary education, in conjunction with the other reforms to these levels of education, affected student achievement is a question we will explore empirically in the following chapters.

The Swedish population appears to be highly educated and highly skilled in an international comparison. The average level of formal schooling is comparatively high, as are available direct measures of skills. In addition, the dispersion of skills in Sweden is fairly small. This low dispersion of skills contributes to the compressed wage structure in Sweden, although wage setting institutions such as unions probably also play an important role.
The return to higher education in Sweden is comparable to that in many other OECD countries. The return to upper secondary school in Sweden, however, is very low by international standards. Whether these returns are too low is an open question. Currently a very large share of each cohort enters upper secondary school in Sweden, and the large supply of workers with an upper secondary education may partially explain the low returns to this level of education.

The return to adult education in Sweden appears to be even lower than the return to upper secondary education. It is therefore hard to 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 students are relatively old by international comparison, and the public return to education is unambiguously decreasing for education received at older ages. We will return to a discussion of the allocation of resources to different levels of the education system in the final chapter.
References


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


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


4 Resources, decentralization, and student achievement

The public education system underwent dramatic changes in Sweden in the 1990s. One of the most important was the substantial decline in resources devoted to primary and secondary education, as discussed in the previous chapter. Although early in the decade spending on primary and secondary education in Sweden was quite high by international comparison, by the end of the decade the share of GDP devoted to these levels of education was only slightly above the OECD average.

Part of the decrease in resources can be attributed to the severe economic slump that hit Sweden in the beginning of the 1990s. Another important factor, however, was the decentralization of authority for the schools from the central government to the localities. Formally this change took place in 1991. Between 1991 and 1993, however, the central government continued to provide the localities money ear-marked for primary and secondary education, although with fewer guidelines and restrictions than in the past. It was not until 1993 that the system was fully decentralized, and localities were given full authority to decide on the allocation of spending among education and their various other responsibilities.

In this chapter we first show that the decentralization of the 1990s induced a sharp change in the distribution of resources to Swedish primary and secondary schools across the localities. We then use this sharp change to estimate the effects of resources on student achievement. This analysis provides specific evidence on the effects of the Swedish reforms over this period as well as more general evidence on the effects of spending on student achievement. We focus our analysis on the compulsory school system (grades 1-9).20

We begin in section 4.1 with a review of the literature on the effect of resources on student performance. We discuss general evidence from the

---

20 The main reason for this focus is that students self-select into upper secondary school, but not into compulsory school, so we do not have to be concerned with selection bias when we analyze the effects of resources on student achievement at the compulsory school level. Furthermore, compulsory school spans nine grade levels, while upper secondary only spans three. As discussed in Chapter 3, the resource changes during the 1990s were similar at the compulsory and upper secondary levels, although student-teacher ratios did not increase as much in upper secondary schools as in the compulsory schools.
In Section 4.2 we focus in greater detail on the change in resources to the Swedish compulsory schools following the decentralization, examining both the overall changes in resources as well as changes in the distribution of resources across municipalities. Finally, in Section 4.3 we estimate the effect of these changes in resources on student achievement, both on average and among students from more disadvantaged backgrounds. In Section 4.4 we summarize our findings and offer conclusions.

4.1 How important are resources for outcomes?

In this section we provide a brief review the economics literature on the effects of resources on student performance, devoting particular attention to experimental and quasi-experimental studies and studies that relate specifically to Sweden.21 We discuss the available evidence on three types of resources: class size, teacher qualifications, and computer aided instruction.

4.1.1 Class size

Class size and the student-teacher ratio are major determinants of the cost of schooling, as teacher salaries are a major component of school budgets. There is a very broad literature on the effect of class size on student achievement, but the findings from this literature are quite mixed. Indeed, even the quantitative reviews of the literature themselves are mixed, with some (e.g., Hedges, et al. 1994; Krueger, 2003) concluding that there are 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 class sizes can be increased without consequences. Many of the existing studies have various limitations – for instance, it is often difficult to distinguish the effect of class size from the effect of background factors and other school inputs that are likely to be correlated with student achievement. It is also difficult to account for potential

21 There is also a broad literature on the effects of resources on student achievement from the education field. In the interest of space, we do not review this literature here; however, the findings from the education literature are generally mixed, as are those from the economics literature. The studies from the education literature also face many of the same potential limitations as those from the economics literature.
selection bias due to the fact that school authorities may deliberately assign weaker (or stronger) students to smaller classes. Comparisons of the performance of students in small and large classes may reflect this selection rather than the true causal effect of class size on student achievement.

Despite these limitations, we believe a reasonable interpretation of the available research suggests that smaller classes have a beneficial effect on student achievement, especially if the population under consideration includes young children or disadvantaged children. This interpretation is based on comprehensive reviews of the literature, with particular weight placed on the findings from experimental and quasi-experimental studies that attempt to overcome some of the limitations of the broader literature, as well as on the available evidence from Sweden that we describe below. Despite the potential benefits, however, reducing class sizes may be quite costly. The benefits may or may not outweigh the costs, particularly for older students or students from more advantaged families. Nonetheless, in deciding whether to increase class sizes, policymakers should take into account the likely costs in terms of reduced educational achievement.

Several literature reviews, such as Hanushek, Hedges et al., Krueger, and Gustafsson (2003), provide a broad account of the literature and the debate over methodological issues and interpretation. Below we provide a brief summary of the literature, the surrounding issues, and the ongoing debate.

In an influential review of the class size literature, Hanushek (1986, p. 1162) argues that the studies which he reviews “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.” His conclusion is challenged by Hedges, et al. (1994), who point out that Hanushek does not take into account the magnitude of the estimated effects from the studies he reviews, but only whether the effects are positive or negative. When Hedges, et al. take the magnitude and precision of the estimates into account using formal meta-analysis techniques, they find a large impact of small classes and other resources on student achievement.

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 others. If each study in the literature is accorded equal weight, which he argues is more defensible than attributing more weight to studies from which Hanushek extracted a larger number of estimates, Krueger finds that the literature as a whole suggests a beneficial effect of
smaller classes: studies that find a positive effect are 57 percent more prevalent than those that 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 of smaller classes 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 student-teacher ratio while holding the amount of expenditures per student constant, thereby confounding the effects of smaller classes with the effects of reductions in other inputs. Hanushek (2003) argues 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 kindergarten through third grade. Within each participating school, students and teachers were randomly assigned to one of three types of classes: small classes (15 students on average); regular size classes (22 students on average); and regular size classes with a teacher’s aide. After four years, all students returned to regular size classes. Because students and teachers were randomly assigned, biases due to selection and omitted variables are eliminated, although some have expressed concern 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). On average, the estimated effect on test scores was on the order of 0.2 standard deviations. The effects were also larger for African American students and low-income students than for other students.

---

22 The assignment to a class with a teacher’s aid, however, was found to have little impact on achievement.
There is other convincing quasi-experimental evidence of the effects of class size on student achievement. 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 sizes below 40. 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 enrollment approaches the cap, class size rises, and then it drops discretely once the cap is exceeded and students are split into smaller classes. Variability from this source is close to random, as there is little rhyme or reason for the jumps in class size other than variation in the number of students enrolled in a grade level. They use these discontinuous changes in class size 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 are mirrored in test scores: reductions in class size are found to induce a significant 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 to the Swedish case is an open question. The socioeconomic characteristics of Swedish students and their families are clearly different from those of students in the other two locations. In considering the effect of resources on student achievement in Sweden, however, it is therefore important to also consider studies of class size effects that focus on Sweden. We know of only one study in the past 20 years that has made a serious attempt to examine the impact of class size on student achievement in Sweden, a study by Lindahl (2004).

In this study, Lindahl collected and analyzed longitudinal data on 556 students 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, with which he can look at the gains in scores over the school year. The typical value added model in the previous literature has received some criticism because it ignores the fact that family background may also affect the trajectory of student
achievement, independently of what goes on in school. Lindahl is able to 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 the student achievement trajectory, in order to disentangle the causal effect of class size on achievement.

Class size in his study was measured by the number of students taught in each math class. In sixth grade, the average math class size was 19.9 students and ranged from 6 to 25. In fifth grade, the average class size was 22.9 students and ranged from 3 to 32. When Lindahl estimates level regressions (e.g., 6th grade test scores as the outcome) or fall-to-fall changes in test scores (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 net of changes over the summer months as the outcome variable (or by including 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 Project STAR. A reduction of class size by seven students increases average test scores by 2.8 to 7.0 percentile ranks according to his preferred set of estimates; this corresponds to a gain of 0.10 to 0.24 standard deviations. In accord with the STAR findings, Lindahl finds a larger gain for immigrant children, but he does not report separate estimates by socioeconomic status.

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 Swedish Krona (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 rate of return Krueger (2003) finds for the United States based on the STAR experiment.
4.1.2 Teacher qualifications

There can be little doubt that teachers matter. A large literature finds “teacher effects” – that is, that 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 effect 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 based on observable characteristics and credentials, is a much more difficult matter. For instance, 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 on teacher competencies.

In general, the literature on the importance of teachers suffers from a similar problem as the observational literature on class size. Teachers having certain characteristics might be systematically allocated to either high or low achieving groups of students. A few studies manage to circumvent this problem, however. For instance, since teachers were randomly assigned to classes in STAR, these data can be used to examine the effect of observed teacher characteristics on student performance. 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). A study by Rockoff (2003), which also attempts to circumvent the problems caused by the systematic allocation of teachers to students, reaches a similar conclusion, although his study finds teacher experience to be more important than the STAR results suggested.
A related question concerns the effect of continuing training for teachers on student achievement. 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 examining which modes of teacher training are most successful. This topic offers many interesting research possibilities, especially because randomized experiments could be designed to assess the impact of various teacher training programs on student achievement.

### 4.1.3 Computers, information technology, and new teaching techniques

Despite broad enthusiasm for 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. However, it is difficult to conclude that his findings reflect the causal effect of these different modes of computer instruction. For instance, it seems plausible that teachers in classes with more advanced students may be more likely to use computers for applications and simulations, and teachers with lower achieving students may be more likely to use traditional drill and practice exercises. This assignment mechanism, rather than a causal effect of computer-aided instruction on achievement, could explain his findings.

To avoid such problems, Angrist and Lavy (2002) analyze a natural experiment in Israel. The program Tomorrow-98 distributed computers to several elementary and middle schools in Israel. By 1996, 10 percent of elementary schools and 40 percent of middle schools had applied for and received computers under the program. Yet Angrist and Lavy find no evidence that the computers actually raised students’ test scores. However, it is unclear what instructional packages were being used or whether the teachers had the expertise to use the packages and computers correctly 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* was developed by neuroscientists to aid reading. It is considered to be
at the leading edge of scientifically-based computer technology in schools and is one of the more expensive programs available, so it affords a test of a state-of-the-art computer application. 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 be in a control group that did not use the software. 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 conflicts of interest when for-profit educational companies evaluate the effectiveness of their own products.

So what does one make of this literature? Computer-aided instruction and information technology 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 information technology innovations in schools, particularly in light of the conflicts between the business interests of the firms that develop and market the equipment and the interests of the students who use them.

### 4.2 Inputs to Swedish compulsory schools during the 1990s and early 2000s

In this section we describe the changes in resources to Swedish compulsory schools over the 1990s and early 2000s. We focus on the teacher/student ratio as a measure of resources, since our reading of the literature suggests reducing class size (increasing the teacher/student ratio) may be one of the most effective uses of school resources.\(^{23}\) We begin by discussing the aggregate changes in the teacher/student ratio in compulsory schools and we then discuss changes in the distribution of this ratio across the municipalities over this period.

\(^{23}\) Unfortunately, direct measures of class size (as opposed to teacher/student ratios) are not readily available.
4.2.1 Aggregate changes in resources

Figure 4.1 shows changes in the median number of teachers per 100 students (the median teacher/student ratio times 100) in compulsory schools over the 1990s. These numbers are calculated from the compulsory school teacher/student ratio in each municipality weighted by the number of compulsory school students in each municipality. They therefore reflect the change in resources experienced by the median compulsory school student in Sweden over this period.

![Graph showing median teacher/student ratios from 1990 to 2003.](image)

**Figure 4.1** Median teacher/student ratios, times 100, 1990-91 – 2002-03.
Source: Calculations using data from the Swedish teacher register.
Notes: Certified teachers are those with a degree in teacher education. For the 1990-91—1997-98 period, the number of teachers has been converted to full time equivalents using 25.3 teaching hours per week as the measure of full-time teaching load. For later years there is a (reliable) measure of teaching hours directly available in the data. In doing these calculations, we have imposed the restriction that a teacher holding more than one position can work no more than 120% of the full time 25.3 hours.

During the 1990s there was a continuous decrease in the median teacher/student ratio, as shown by the solid line. From school year 1990-91 to 1999-2000, the ratio decreased from 9.1 to 7.4 teachers per 100 students,
although it rebounded somewhat in the subsequent two years. The decline over the 1990s was driven by the fact that the number of teachers did not keep pace with the increase in the student population during those years. Between school years 1990-91 and 1999-2000, the median number of teachers per 100 students would have declined by only 0.4 had the size of the student population remained fixed at its 1990 level.

The literature on the effect of resources on outcomes has primarily focused on class size. Unfortunately, the most recent information available on average class size in Sweden is from the beginning of the 1990s. In the 1991-92 school year, the average class consisted of 21.8 students. We can use this information along with the information in Figure 4.1 to roughly estimate the average class size later in the decade. The information presented in Figure 4.1 indicates that the number of students per teacher increased by 18 percent between 1991-92 and 1999-2000. Assuming the average class size of 21.8 in 1991-92 also increased by 18 percent over this period provides an estimated average class size of 25.8 students in 1999-2000, an increase of four students per class. The estimated class size for 2002-03, calculated in the same way, is 25 students.

In our survey of the evidence on the effects of teacher quality in section 4.1, we documented that teachers appear to have an effect on student achievement, but that this effect appears to be uncorrelated with the teacher’s educational attainment. Nevertheless, it is interesting to examine the change in the number of certified teachers per student over the 1990s, as shown by the lower dotted line. The increase in the teacher/student ratio that began in 1998-99 was accomplished by an increase in number of non-certified teachers – teachers without formal pedagogical training. The ratio of certified teachers to students declined rather dramatically over this period – from 8.6 percent in 1990-91 to 6.2 percent in 2002-03. By 2002-03 almost 19 percent of teachers were uncertified.

In contrast to the decrease in teacher inputs, investment in computers and information technology in the Swedish schools 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 a major computer investment initiative which began in 1994 and continued through 1999, the student/computer ratio decreased from 38 in 1993 to 8 in 2001. Among the OECD countries, Sweden ranked second only to Canada in terms of the average number of computers per student in primary school (Statistics Sweden, 2002).
### Table 4.1 The number of students per computer in compulsory schools

<table>
<thead>
<tr>
<th>Year</th>
<th>1993</th>
<th>1995</th>
<th>1997</th>
<th>1999</th>
<th>2001</th>
</tr>
</thead>
<tbody>
<tr>
<td># students per computer</td>
<td>38</td>
<td>19</td>
<td>13</td>
<td>10</td>
<td>8</td>
</tr>
</tbody>
</table>


In light of the existing evidence on the effects of resources on student achievement reviewed in Section 4.1, the changes in resources in the Swedish compulsory schools over the 1990s may well have had a negative effect on student achievement. The number of teachers per student decreased considerably, which, according to our reading of the literature, is detrimental to student achievement. Although there is less evidence that teacher certification is an important determinant of student achievement, the decline in the number of certified teachers per student may also be worrisome. On the other hand, the number of computers per student increased, but there is little evidence in the literature to suggest that this will benefit academic performance.

#### 4.2.2 The distribution of resources

It is also important to examine how the distribution of resources for education may have changed during the 1990s, since after the reform, each municipality was given authority to decide how to allocate resources between compulsory education and its other responsibilities. In this section we describe how the compulsory school finance system changed in Sweden after the decentralization. We then examine how this change affected the distribution of resources across the municipalities, focusing on the teacher/student ratio as a measure of resources. In the following section, we use the decentralization-induced change in the distribution of the teacher/student ratio to estimate the effects of class size on student achievement.

Prior to the 1991-92 school year, the allocation of school funding to the municipalities was determined through a strict system, and the funds were earmarked for specific purposes. The central government – via the regional schooling authorities (länsskolnämnderna) – essentially determined the amount of resources allocated to each school within the municipality. 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 expenditures within the compulsory schooling system.

The decentralization of the finance system began in the 1991-92 school year. The central government continued to provide a grant to the municipalities
to fund their compulsory schooling systems. However, the municipalities were given broad authority to allocate the funds across schools and across various expenditures within the compulsory schooling system as they saw fit. The real (inflation-adjusted) total amount of resources for schools provided to each municipality remained the same as under the previous system.

The system was further decentralized in January 1993, when the money previously earmarked for education was incorporated into an overall equalization grant to each municipality. From 1993 onwards, the municipalities were given responsibility to freely allocate resources over their various responsibilities, with no minimum requirements for education expenditures.

We begin by examining the overall spread of the distribution of the teacher/student ratio over municipalities. Figure 4.2 shows percentile ratios of the distribution. The solid line shows the 90/10 ratio (i.e., the ratio of the 90th percentile of the teacher/student ratio to the 10th percentile of the ratio). The dotted lines “decompose” this ratio into 90/50 ratio and the 50/10 ratio. The spread of the distribution was lowest in the 1993-94 school year, and increased fairly steadily throughout the rest of the decade. 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.

It is also interesting to look at whether the relationship between compulsory school resources and mean income in the municipalities has changed over the years. For instance, when school funding was determined by the central government, we might expect there to have been no relationship between average income in the municipality and the teacher/student ratio, or even a negative relationship if the central government provided more money for teachers to poorer municipalities. We might expect this relationship to change after the decentralization if wealthier municipalities had different spending priorities than poorer municipalities.

Figure 4.3 reports the coefficients (solid line) from year-by-year regressions of the log teacher/student ratio on the log average income in each municipality. The dotted lines indicate the 95 percent confidence bands on the estimates from each year. The regressions are run separately for each year from 1988 to 2000 and are weighted by the number of students in each municipality. Because of this weighting procedure, the regression coefficients capture the relationship between resources and income facing the typical student in each year.
The allocation of resources appears to have been redistributive throughout the decade. On average across the years, a one percent increase in mean income was associated with a decrease in the teacher/student ratio of 0.1 percent. However, the relationship changed sharply in 1993, the year the decentralization was fully implemented, and, perhaps surprisingly, became considerably more compensatory. Over the remainder of the decade, however, the negative association between resources and income became weaker, and by the end of the decade the relationship between income and teacher/student ratio had returned to its pre-1993 level. Therefore, it seems that the decentralization had little long-term effect on the association between the resources devoted to teachers and average income across the municipalities.
Figure 4.3 The relationship between teacher/student ratio 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 log teacher/student ratio on log income in each municipality. 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 was supplied by Inge Göransson at Statistics Sweden.

Another way of examining whether decentralization induced a change in the distribution of resources is to rank the municipalities according to their teacher/student ratio in each year and then examining the correlation of each municipality’s rank across years. Figure 4.4 displays the rank correlation of teacher/student ratios over time. The solid line shows the correlation with the rank in the preceding year (for example, the entry for 1991 shows the correlation between the 1991-92 and 1990-91 school years). The dotted line shows the correlation with the rank three years prior (for example, the entry for 1991 shows the correlation of the 1991-92 and 1988-89 rankings).
Figure 4.4 The rank correlation in the current and lagged teacher/student ratio, 1986-2001.
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 also affected the ranking of municipalities according to their teacher/student ratios. The solid line shows that the correlation is markedly lower between school years 1992-93 and 1993-94 than between previous years. The dotted line indicates that the correlations are substantially lower between years that span the 1993 reform than between consecutive years that are either before or after the reform. After the reform, the correlation between adjacent years increases almost to its pre-reform level, although the relationship seems somewhat less stable than in the pre-reform years. This is perhaps not surprising given that the state of the municipal budget became a more important determinant of school resources following the reform.

Of course, some of the observed changes in the distribution of resources to the municipalities may be due to changes in the characteristics of students. The characteristics of students typically change slowly over time, so it is unlikely that such changes would account for the sudden shifts in the distribution shown.
in Figures 4.3 and 4.4. Nonetheless, in order to generate a more accurate picture of the reform-induced change in the distribution of resources, we can take changes in the characteristics of students and municipalities into account. We do this by first selecting a pre-reform year (1991-92) and regressing the teacher/student ratio on a set of municipality and student characteristics that are known to have been determinants of resource allocation prior to the reform. These characteristics include the fraction of 9th grade students who are female, the fraction who are foreign-born, the fraction with university educated parents, average income in the municipality, average size of schools, and the number of students in the municipality. The regression results provide an estimate of the resource allocation formula prior to the reform.

We then use the estimated coefficients from this regression to predict the teacher/student ratio in each municipality in a post-reform year (2000-01), based on the characteristics of the municipality in that year. This provides an estimate of the teacher/student ratio in each municipality in 2000-01 if the 1993 reform had not occurred and the resource allocation formula had remained unchanged. The difference between the predicted and actual values of the teacher/student ratio in 2000-01 (the residual) provides a measure of the reform-induced change in municipality resources, holding constant changes in student and municipality characteristics.

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 Section 4.1). 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 believe the characteristics we include in the regression capture these various aspects of the resource allocation system well.

Ideally we would have used information on the actual resource allocation formula rather than estimating it. However, it is impossible to construct this formula, in part because resource allocation was subject to the discretion of the regional schooling authorities. See Du Rietz et al. (1987) for an excellent description of the system prior to the reform.

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

$$R_{m(2000/01)} = \beta_{1991/92} x_{m(2000/01)} + (\beta_{2000/01} - \beta_{1991/92}) x_{m(1991/92)} + \epsilon_{m(2000/01)}.$$
Figure 4.5 plots the densities of the residuals generated by the above procedure. Since we are interested in the distribution rather than the level of resources, we have normalized the data to have same mean in both 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 depicts the same for 2000-01. As we would expect, the residuals from the 1991-92 regression are evenly distributed around zero. The plot of the density of the residuals from the 2000-01 regression shows how the distribution of resources has shifted relative to what would have been expected had the 1993 reform not occurred. The results shown in this figure indicate that the reform induced a shift of the distribution of teacher inputs over the municipalities. Thus, the distribution of resources clearly changed after the reform, even when changes in observed characteristics of the students and municipalities are taken into account.

The first component reflects the predicted level of expenditure had the pre-reform formula remained in place. The sum of the second and third components is the deviation of resources from this predicted level.
Notes: The figure shows the densities of the deviations of log teacher/student ratio from the predicted level. Predicted teacher/student ratio 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.

4.3 Resources and achievement: Swedish evidence

As discussed in section 4.1, it is difficult to estimate the “causal” effect of resources on student performance. The difficulties stem from potential selection bias due to the fact that educational authorities may assign more (or fewer) resources to lower achieving students. Simply comparing the average achievement of students in low- versus high-resource schools or municipalities will confound the true causal effect of resources on achievement with differences due to the compensatory behavior of school officials and other factors. Identifying the true causal effect therefore requires some sort of exogenous variation – a variation that is unrelated to unobserved characteristics of the students or municipalities that are correlated with test scores. In this section we make use of a potentially exogenous source of variation – the change in schooling expenditure induced by the decentralization. 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 shifted the distribution of resources across the municipalities, it is not evident that the reform-induced change in resources is exogenous to student achievement. For instance, the exogeneity assumption would be violated if local governments targeted resources to weaker students more effectively than did the central government. With this caveat in mind, let us proceed to the evidence.

4.3.1 The effect of resources on achievement

Our basic strategy is to calculate difference-in-differences estimates of the effect of the reform. The difference-in-differences approach amounts to comparing the average performance of students graduating from compulsory schools in a particular municipality in 2001 to those graduating in 1992 in the same municipalities and relating this change in performance to the change in
resources in the municipality over this period. We also control for a set of student and municipality characteristics. Therefore, the variation we are utilizing to estimate the effect of resources is the reform-induced shift in the distribution of resources shown in Figure 4.5. We use percentile ranked grades of 9th graders as our measure of student performance.

Our main data source is pooled data on 9th graders leaving compulsory school in the spring of 1992 and 2001 in each municipality. These data contain the grade point averages of these students along with information on their individual and family background characteristics. We match information about the municipality onto these data. The key municipality-level variables are measures of schooling inputs such as the teacher/student ratio and teaching expenditures per student.

Table 4.2 presents the regression results. The dependent variable is the percentile ranked grade point average. The regression standardizes for student gender, age, whether the individual is a first or second generation immigrant, whether he or she 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 inclusion of municipality fixed effects allows us to use the variation within a municipality over time to estimate the effects of resources on achievement. The time effects control for resource changes that are common across municipalities (for instance, the resource change induced by the national economic downturn).

We begin with a specification that relates 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 better teachers which, in turn, may 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 the remainder, which is mostly wages. The results suggest that only the number of teachers per student affects achievement; the variation in “wages” (although crudely measured) appears to have no effects. The estimated coefficient 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.
Table 4.2 The relationship between student achievement and resources.
Dependent variable: percentile ranked grade point average.

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

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 uncertified 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 is based on information from 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 students in the municipality.

We next attempt to address the question of whether teacher certification affects student performance. The model presented in column (2) of Table 4.2 controls for the log number of certified teachers and the share of uncertified teachers. The estimates of the coefficients are not statistically different from

To obtain this specification we decompose the log teacher student ratio as follows:

\[
\ln(t / s) = \ln((t' / s)(1 + t'' / t')) = \ln(t' / s) + \ln(1 + t'' / t'),
\]

where \(s\) denotes the number of students, \(t\) the number of teachers, and superindex \(c\) \((nc)\) indicates whether the teachers are certified (or not). Notice that if \(t''\) is small relative to \(t'\), then the second component of the decomposition is approximately equal to the share of uncertified teachers. By estimating separate coefficients for the two components of the decomposition, we can examine whether teacher certification affects student achievement or not, by testing whether the two coefficients are...
each other, suggesting that teacher certification does not affect student performance, consistent with findings from the previous literature. The results in column (1) and (2) thus lead to the parsimonious specification in column (3), which relates student outcomes only to the log teacher/student ratio and student and municipality characteristics. Multiplying the estimate reported in column (3) with the actual decline in the teacher/student ratio between 1990-91 and 2002-03 (about 17 log points) implies that the actual resource change lowered student performance by 1.2 percentile points.

Alternatively, we can compare these estimates to the ones in Krueger (1999) and Lindahl (2004). The average class size in Sweden in the 1991-92 school year was about 22 students, and the median teacher/student ratio was 8.8 percent. If class size were reduced by seven students – roughly the magnitude of the reduction in the Project STAR experiment – this would correspond to an increase in the teacher/student ratio of about .59 log points. Thus, the estimates in column (3) suggest that an increase in the teacher/student ratio equivalent to the seven-student class size reduction in STAR would produce a gain in achievement of about 4.2 percentile ranks. This gain corresponds to an increase of 0.14 standard deviations, which is within the range of the estimates reported in Lindahl (2004).

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/student ratio, we have interacted the ratio with indicators reflecting socioeconomic disadvantage (immigrants and low-educated parents). The regression includes an interaction of the ratio with low

significantly different from one another. If they are not, then (ostensibly) teacher certification does not affect student achievement. It is important to note, however, that, although the reform appears to have generated exogenous variation in the total amount of resources, it is not clear that it generated exogenous variation in the allocation of teachers by certification status. If the exogeneity assumption is violated, the results may be biased.

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 teacher/student ratio to be 7.01, which is virtually identical to the estimate presented in column (3) with a larger sample.

A limitation of this computation is that it is nearly outside the range of the data. There were only two municipalities in 1991-92, and none in 2000-01, that had a teacher/student ratio of this implied magnitude.
parental education and an interaction with recent immigrant status. These estimates indicate that the academic achievement of students from disadvantaged families is more susceptible to variations in resources than that of students from more advantaged families. For instance, the results suggest that a 10 percent reduction of the teacher/student ratio lowers the performance of recent immigrants by 3.5 percentile points. This is similar to the findings of Krueger (1999). Note also that Lindahl (2004) found that class size had more of an effect for immigrant than for native children, although he did not report size effects by socioeconomic status more generally.

### 4.3.2 Resources and the between school variance in achievement

The results shown above suggest that resource changes do affect student achievement. They also suggest that students from disadvantaged background are particularly susceptible to variations in resources. This suggests that the variance of student achievement may have increased as a consequence of the decline of the teacher/student ratio during the 1990s.

One way to examine whether the decline in resources has affected the variance in achievement is to look at the change in the between-school variance in student achievement over time, as shown 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 comparable estimate standardized for student characteristics. By comparing these two lines we can get a sense of whether changes in the distribution of different types of student across schools are important in explaining changes in the between-school variance.

---

30 We also ran specifications that included interaction terms of teacher/student ratio with immigrant status and having foreign born parents, but we did not find any statistically significant effects.
Figure 4.6 The evolution of the between school variance in outcomes, 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 sharp increase in the between-school variance between 1997 and 1998. This jump coincides with the introduction of a new grading system in Sweden. However, since we are converting grades to percentile ranks, 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 mechanically related to the changes in the grading system. Nonetheless, it is difficult to argue that the sudden change is related to an increase in the variance of resources – these processes are presumably smooth. One possible interpretation of the change is that the new grading system may have increased the scope for school policies to determine what constitutes a “pass” or a “pass with distinction.” Under the old system, students’ grades were “anchored” using a national standardized test. Under the new system, national standardized tests are still conducted, but they are used simply to help teachers in grading the students. Therefore, the sharp increase in
the between-school variance may reflect the fact that the schools were given more discretion in determining grades under the new system.

There is some evidence of a general upward drift in the between-school variance over the decade. The variance (standardized for student characteristics) increased by a third from 1992 to 1997 (from 0.152 to 0.201). There is also some evidence suggesting that sorting on observed 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 the median teacher/student ratio in compulsory schools, from 9.1 per 100 students in 1990-91 to 7.4 per 100 students in 1999-2000; the early 2000s have seen a slight rebound in the number of teachers per student. The number of certified teachers, however, declined precipitously throughout the entire period of study.

The reduction in inputs has of course been driven to a large extent by the economic downturn in Sweden. It is impossible to discern whether decentralization also contributed to the resource decline. However, decentralization and the subsequent abolition of earmarked central government funds for schools do appear to have shifted the distribution of resources across municipalities.

Our literature review and the new evidence we have presented also suggest that the resource decline had a negative effect on student achievement. Increases in the teacher/student ratio appear to improve student outcomes on average. The average effect we estimate is similar in magnitude to the estimates reported in Lindahl (2004). The effect is larger for students from disadvantaged family backgrounds – recent immigrants in particular – which also accords with the results from the previous literature. The fact that the effect varies by family background characteristics implies that the resource decline contributed to the upward trend in the between school variance in student outcomes that we observe in the data.

Should Sweden increase the number of teachers per student? 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 would be considered high among the possible alternative public investment projects. The important qualification is that their calculation involves several assumptions
regarding the unknown components of the cost-benefit calculation. In addition, it may be difficult to argue in favor of reducing class size in light of the predicted future teacher shortages in Sweden. Lastly, we would emphasize that the benefit of smaller classes appears to be greater for children in immigrant families and in families in which the parents have a relatively low level of education.
References


Nye, B., J. Zacharias, 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.


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.


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


5 Teacher supply

In chapter 4 we reviewed the literature on effects of teacher qualifications on student achievement. The quality of teachers is clearly one of the most important inputs for student achievement, although the literature has been unable to determine which characteristics contribute to teacher quality. The purpose of this chapter is to present evidence relating to the change in the quality of the teacher supply in Sweden during the 1990s.

As shown by le Grand et al (2001), the public sector wage premium in Sweden has decreased quite substantially since the late 1960s. Therefore, one might expect that the monetary incentives for becoming a teacher have also declined over time. Additionally, working conditions have probably worsened in the Swedish schools in recent years, as the ratio of teachers to students has declined. Both of these factors are likely to contribute to the predicted teacher shortages in Sweden in coming years (see National Agency for Education, 2003, on the predicted demand for teachers through 2020).

![Figure 5.1](image-url)

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

Figure 5.1 presents information on the change in the age distribution of certified teachers in Sweden over the past 25 years. The figure shows that the
supply of teachers has aged considerably over the past 25 years; the median age of certified teachers increased from 32 years in 1975 to 46 years in 2000. This suggests that the supply of new entrants to the teaching profession has fallen over this time period. Moreover, the supply of certified teachers seems to have fallen relative to the number of individuals with a similar educational attainment. The median university graduate was 33 years of age in 1975; in 2000 the median age had increased to 40. In this chapter we will explore the factors underlying these developments.

5.1 The incentives to become a teacher

We begin by looking at changes in the incentives for becoming a teacher in Sweden over the past several decades. 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.31 Since 1945, the relative wage has declined by almost 50 percent. The decline in the relative wage has slowed somewhat over the past thirty years, but the general decline has continued.

The decline of teacher relative wages is not a unique Swedish phenomenon. For almost all countries for which there are data, teacher relative wages have declined since the mid-1960s. However, the international evidence suggests that the decline in Sweden was particularly dramatic; 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 earn only 82 percent of the OECD average for teachers at this level. The relative wage gap is lower in primary schools and also in terms of starting wages.

Although Figure 5.2 is informative about the long-run changes in teacher salaries, there are a number of qualifications to keep in mind. For instance, the share of females in teaching has increased over time, and it is well known that women are generally paid less than men. 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 over this period.

Of course there are potential comparison occupations other than production workers. Nursing provides a useful comparison since nurses and teachers have

31 We thank Anders Nilsson for supplying the pre 1975 data on teacher wages.
similar gender composition and education level. Over the past thirty years, teacher wages have declined relative to nurses as well: the relative wage decreased by 20 percent between 1970 and 2001.\footnote{See Statistics Sweden’s wage statistics (\textit{Kommunal personal} 1970 and \textit{Statistical Yearbook of Salaries and Wages} 2001) for information on the wages of nurses.}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.2.png}
\caption{Index of teacher wages relative to production worker wages, 1920-2001, 1920=100.}
\end{figure}

\textbf{Figure 5.2} Index of teacher wages relative to production worker wages, 1920-2001, 1920=100.
\textbf{Notes:} During the period 1920-89, teacher wages refer to the wage in the highest wage category for teachers in primary schools (\textit{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 (\textit{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.

Nonetheless, comparisons with single occupations do not answer the most relevant question: How have the incentives to pursue a teaching career once one has decided to pursue a university education evolved over time? Figure 5.3 presents the standardized (annual) earnings premium for male and female teachers relative to those with the same educational attainment; see Figure \footnote{See Statistics Sweden’s wage statistics (\textit{Kommunal personal} 1970 and \textit{Statistical Yearbook of Salaries and Wages} 2001) for information on the wages of nurses.}
5.3. The calculations are based on the LINDA data base (see Edin and Fredriksson, 2000). The figure shows that female teachers (the vast majority of teachers) saw no earnings decline relative to other women with similar educational attainment during the 1970s. There is a downward jump for female teachers around the mid-1980s, however, when their relative earnings fell by five percent.

![Figure 5.3 Teacher earnings relative to those with the same educational attainment, 1968-2000.](image)

Notes: When making this calculation we have standardized with respect to age and educational attainment (short and long university education).

---

33 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 the evolution shown in Figure 5.3.

34 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 women in other professions for most of the period. The opposite is true for male teachers. In the late 1960s male teachers earned ten percent less than men with a similar educational attainment, and there has been a downward trend in the relative earnings of male teachers. By the year 2000 their relative earnings had fallen to less than 80 percent of the earnings of men with comparable education levels in other professions.


Of course, there are also non-monetary reasons why one might chose a career in teaching, and these may also have changed over time. In Figure 5.4, 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 workers as well. During the first half of the 1990s there was a small difference between teachers and the average employee. However, from

35 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.
1995 onward the gap between teachers and other employees widened rather dramatically. In 2002, 21 percent of teachers report suffering from mental stress, compared with only 10 percent of the workforce as a whole.

It is of course relevant to ask whether the differential change in reported stress levels is related to the changing gender and age composition of the teaching profession. Unfortunately, we do not have the information necessary to adjust these figures for potential differences in gender and age. It seems highly unlikely, however, that the differences shown in Figure 5.4 would disappear when standardized with respect to age and gender.

Table 5.1 reports the shares of teachers, white collar employees, and all workers 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 general increase in mental stress and health problems in all categories. In the final two rows we compare the changes among teachers to those faced by workers in other fields. These computations show that there has been a significant increase of the mental stress in teaching relative to that faced by all workers and to white collar workers. 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 there has been no increase in the work-related health problems of teachers relative to white collar workers as a whole over this period.

In summary, the relative wages of teachers have declined fairly steadily over the past several decades, and this trend continued into the 1990s. During the 1990s, the earnings of teachers declined by 2-3 percent relative to those of individuals with a similar educational attainment. It also seems that working conditions have become substantially worse in Swedish schools during the 1990s. We think that one of the most plausible explanations for this development is the reduction in the teacher/student ratio 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.

36 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.

37 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.

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

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 the teacher supply is sensitive 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 in Sweden. However, after reviewing the literature on the effects of 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 the supply of teachers is sensitive to variations in working conditions.

In this section we examine whether the changes in wages and working conditions faced by teachers, as described in the previous section, had any consequences for the supply of skills to the teaching profession. We first examine changes in the number of applicants per slot in teacher education programs and the average grades among those admitted to these programs. We next examine whether there have been changes over time in the standardized test scores of those who later on opted for teacher education.

**Figure 5.5** The number of qualified 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 at its 1981 or 1987 level, as indicated. Teacher education was reformed in 1988. 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 in teacher education programs. At this stage, we restrict our attention to programs

38 Admission to teacher education and other programs are determined at a centralized level.
preparing teachers for teaching at 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 2000, so we have not extended the graph beyond 2000. The figure shows a dramatic reduction in the number of applicants to teacher education programs in the 1980s. This decline is not due to an expansion in the number of slots in teacher education programs – the dashed line illustrates the number of applicants per slot when the number of slots are held constant at their 1981 level (in the years prior to the 1988 teacher education reform) and 1987 level (in the years following the reform). In fact, the number of slots also fell throughout most of the 1980s. During the 1990s, there appears to have been a slight rebound in the number of applicants to primary school-level teacher education programs. However, it is possible that this increase was driven by the economic downturn; as unemployment rose in the beginning of the 1990s, the overall demand for university education presumably increased.

To eliminate the variation due to changes in the overall popularity of university education it is natural to examine the demand for teacher education relative to the demand for other university programs. Figure 5.6 shows the number of qualified university applicants who indicated pedagogical training as their preferred choice relative to the overall number of qualified 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 2001-2003, relative demand was even lower than in 1990.

Thus, there has been a decline in the popularity of teacher education programs during the past two decades. It is also interesting to examine the average skill level of applicants to these programs. Figure 5.7 presents the minimum grade level of those admitted to teacher education programs from 1980-2000. Teacher education was reformed in 1988. Prior to 1988 there was only one track for university students aspiring to teach at the primary level, but from 1988 and onwards there were two separate tracks – students could choose either Swedish/Social Science or Math/Science as their major subject. From

---

39 Pedagogical training programs include primary-level teaching programs analyzed in Figure 5.4 and other education programs. We examine this broader measure of applicants to teaching programs in Figure 5.5 in order to be better able to compare the periods before and after the reforms to teacher education in 1988 and 2001.
1988 and onwards we show the marginal admittance grade for applicants to the Swedish/Social Science major. We do not show the evolution of entry grades for those majoring in Math/Sciences for a very simple reason – everyone who applied was admitted! Figure 5.7 illustrates that the formal qualifications of the marginal entrant to Swedish/Social Science teaching programs declined along with the demand for teacher education.

**Figure 5.6** The relative number of qualified university applicants who indicated pedagogical education as their preferred choice of program, 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.

We can also look at changes in the average ability levels of the teacher supply over a longer time period. The Departments of Education in Göteborg and Stockholm have conducted tests of random samples 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 include scores on verbal, inductive, and spatial ability tests, which are comparable across all four cohorts. We have matched information on educational attainment, and the type of education each student pursued later on in life, to their 6th grade test scores.40

---

40 We are very grateful to Jan-Eric Gustafsson and Åsa Berndtsson for letting us access and assisting us with the data.
Using the matched data sets we can examine the relative performance on tests at age 13 for those who later decided to pursue a program in teacher education. These four cohorts made their career choice at very different time points. The majority of those born in 1948 presumably made their career choices around 1970 while those born in 1972 made their choice in the mid 1990s. The question is whether the overall decline in the relative wages of teachers has affected the relative ability level of those who choose to enter teacher education programs. As a measure of ability, we use the sum of the scores from the three tests mentioned above, which we refer to as the “ability score.”

Table 5.2 presents the average scores for students in each cohort who later entered a teacher education program. (See Nickell and Quintini, 2002 and Evans et al., 2002 for similar analyses for the UK and US respectively.) For comparative purposes we also report the scores of the average university graduate. For both teachers and other university-educated individuals, we focus on those who completed at least three years of tertiary education. As shown in the second column of the table, in the cohort born in 1948, the average college educated individual had a percentile ranked ability score of 72. There is a decrease in the average rank in the later cohorts. 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 not surprising, given that the share of the population proceeding to tertiary education increased by almost 7 percentage points between the two cohorts. If there is a positive association between performance on the test and the probability of pursuing a university education, average relative performance must decrease as the percentage of the population obtaining a university education increases. However, this cannot be the only explanation for the decline in the relative rank of university graduates over time, as the share going on to attend university remained roughly constant between the 1948 and 1967 cohorts, but the relative rank of university graduates was significantly lower in the 1967 cohort.

41 Results are very similar when the analysis is conducted for each test separately.
### Table 5.2 Average percentile ranked ability scores for individuals with teacher and university education.

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

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 t-statistics from (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: 10,560; 9,372; 8,098; and 7,938 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.
Between 1948 and 1972, there was also a significant reduction in the relative performance of teachers – the average percentile rank score of those completing teacher education programs declined by 6.6 percentile points. To put this decline into perspective we have calculated the average difference between individuals of varying levels of educational attainment for the cohort born 1972. According to this measure, the decline in the relative performance of teachers roughly corresponds to the average difference between individuals having a 3-year and a 2-year upper secondary education. This is a substantial decline both in its own right and when compared to the average individual with a university degree. Thus those opting to pursue teacher education programs in the early 1990s appear to be of lower relative ability than those entering teaching in the late 1960s.

The evidence presented in this section tells a consistent story: The decline in teacher quality, as measured by scores on standardized tests of those entering teacher education programs, generally mirrors the decline in the relative wage of teachers 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, Evans et al. (2002) note that, over time, it has become less likely that females in the top decile of the ability distribution opt to enter the teaching profession in the US. The relative attractiveness of teaching in the Swedish primary schools appears to have declined substantially during the 1980s. There appears to have been a slight increase in the number of applicants to teacher education programs in the early 1990s, but this rebound subsided during the second half of the decade.

5.3 Teacher mobility

In the previous section we focused on the size and characteristics of those entering teacher education programs. However, changes in the quality of teaching jobs may also affect outflows from the profession. Figure 5.8 displays

---

42 The gap between these two groups was 8.2 percentile points. In passing, we note also that the gap between individuals having a 2-year tertiary and a 3-year upper secondary education was the same size.

43 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 teachers, 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.
the outflow from the teaching profession for individuals under the age of 60.\textsuperscript{44} We present separate calculations for teachers initially employed at the compulsory and upper secondary school levels. The figure indicates that there has been a general increase in the probability of leaving the teaching profession. This increase has occurred among teachers at both the compulsory and upper secondary school levels, although the increase is more pronounced for compulsory school teachers. The probability of leaving teaching within the next year for those originally employed in compulsory schools was 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.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{probability_of_leaving.png}
\caption{The probability of leaving the teacher profession, percent.}
\textbf{Source}: Calculations based on the Teacher register.
\textbf{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.
\end{figure}

\textsuperscript{44} 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.
Among teachers teaching at particular levels, there has been a similar increase in total outflow (both from the profession entirely and to other levels of teaching). However, there was an upward jump in the total outflow around 1998, which is when enrollment in adult education programs peaked. Therefore we take a closer look at the destination of those that left compulsory and upper secondary teaching for other forms of teaching. We are particularly interested in the outflow from upper secondary schools to adult teaching. Is the introduction of the Adult Education Initiative (AEI) visible in the data? As mentioned earlier, the AEI was introduced in 1997. Total enrollment in adult education peaked in 1998 and remained high during 1999 and 2000. As shown in Figure 5.9, there was a corresponding peak in the outflow rate of teachers to adult education from the upper secondary level. Thus, the development of the adult education market appears to have drawn teachers away from youth education. This may reflect preferable working conditions for teachers in adult education programs.45

There is some evidence that the competition for teachers has increased. The establishment of privately run schools has implied that teachers in public school have more alternatives available than were previously available, and the data show an upward trend in the probability of leaving public schools for private schools. The increased probability of leaving teaching in the public schools for private schools mirrors the expansion of private schools over the 1990s. Also, a system of individual wage setting was introduced during the second half of the 1990s, which enabled schools to attract teachers by offering higher wages. This may have increased mobility both across schools within municipalities and across municipalities. Unfortunately, we cannot examine whether mobility across schools has increased because we are only able to identify the school at which the teachers work in the later years of our data. However, we can examine teacher mobility between municipalities, which has increased. The probability of a teacher changing municipalities was twice as large in 2000 as in the mid-1990s.

45 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, however. In 1996, the earnings of adult education teachers were higher than the earnings of teachers in youth education. In 1998, this relationship had been reversed, 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.
We can also examine the characteristics of those who leave the teaching profession. Table 5.3 presents the results of (logit) regressions in which the probability of leaving the profession is related to a set of characteristics of the individual and the teaching 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—1992 (column 1) and 1999 (column 2).

The regressions for both points in time look as one might expect. For instance, certified teachers and teachers on permanent contracts are less likely to leave the profession than those without these characteristics. The more interesting question is whether there has been a change in the importance of any characteristic between the two time points. Comparing columns (1) and (2) suggests there have been some changes. However, the major changes relate 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. There appear to have been no major changes in the importance of characteristics relating to teacher quality, such as teacher certification and...
teacher experience (as approximated by age). However, we note that this does not necessarily imply that there has been no change in the relationship between teacher quality and the outflow rate. As our review of the literature suggested teacher quality (as measured by a teacher’s contribution to student performance) appears to be uncorrelated or only weakly correlated with observable indicators of teacher quality.

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

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

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.
In summary, there appears to have been an increase in outflows from the teaching profession in recent decades. There is also evidence suggesting that the growth of adult education programs in recent years has led teachers to leave primary and secondary schools for adult education programs. Given that the relative wages of teachers have not changed very much since the mid 1980s, it seems likely that the changes documented here are more closely related to changes in working conditions for teachers in primary and secondary schools than to changes in monetary incentives.

5.4 Summary and conclusions

The relative wages of primary and secondary-level teachers in Sweden have declined considerably over the past several decades. To some extent this decline is due to a decrease in the wage differentials between workers with different levels of educational attainment. But there has also been a decline relative to those with similar education levels: for female teachers, relative earnings decreased by 6 percent between 1968 and 2000; for male teachers, relative earnings declined by almost 13 percent over this same time period. However, the changes in relative wages were fairly minor during the 1990s. Working conditions faced by teachers at the primary and secondary levels appear to have become substantially worse during the 1990s, at least as evidenced by the relative increase in health problems caused by mental stress. This increase is probably related to cut-backs in resources and a consequent increase in class size.

In sum, it seems that teaching at the primary and secondary levels in Sweden has become less attractive over time. This is likely to affect the number of teachers as well as the composition of the pool of teachers. We have shown that the relative demand for slots in teacher education programs has declined. In addition, it seems that high ability individuals have become less inclined to enter the teaching profession over time. The approaching retirement of the baby boomers of the 1940s implies future shortages that will hit Swedish schools particularly hard. The increasing shortage of teachers is not a problem unique to Sweden, but is faced by many advanced countries.

The national and local governments in Sweden, at least to some extent, control several instruments that could potentially improve the attractiveness of teaching and, hence, the supply of teachers. The availability of spaces in teacher education, teacher wages, and school resources in general are a few, but important, examples. The number of slots in teacher education programs seems
to be the main instrument the Swedish government has used to influence teacher supply. This may be an ineffective policy without due consideration to changes in incentives to enter the teaching profession – pecuniary as well as non-pecuniary. This is clearly illustrated by the fact that the pool of qualified applicants for teacher education programs in math and science is not sufficient to fill the available slots. Prospective teachers in math and sciences apparently have better options outside teaching.

Are there any developments in the teacher labor market in Sweden that imply that future teacher 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. The greater flexibility provided to schools will probably lead to greater pay differentials across both subject areas and municipalities. Such differentials may be necessary to attract appropriately qualified teachers to the profession in certain areas.

The analysis in this chapter indicates that basic economic concepts apply to the teacher labor market in Sweden. If school authorities decide to reduce the amount of resources going to schools by, for instance, increasing class size, this may lead to a decline in teacher quality, as working conditions in the profession become less attractive. The quality of teachers is, in turn, important for student performance, although it is interesting to note that the far the performance of Swedish students. The findings in the chapter suggest Swedish policy makers must take into account potential effects of changes in resources on the supply of qualified teachers when deciding how to allocate scarce resources.
References


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


6 The consequences of school choice

Following in the wake of the decentralization of the Swedish education system, another market-oriented reform was implemented in July of 1992. This reform required municipalities to provide funding for privately run independent schools. Additionally, parents were given the right to choose among all schools – public as well as private. As a general rule, the parents’ choice is made subject to the availability of slots. In cases of excess demand, slots in public schools are allocated on the basis of residence while slots in private schools are allocated on a first-come–first-served basis. An important exception to this general rule is that Stockholm city, in 2000, began to admit students to upper secondary schools solely on the basis of performance.

Arguably, Sweden’s school choice reform has increased 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 recent attention. A report by Bergström and Sandström (2001) suggested that competition from independent schools increased the performance of public schools. Their conclusion was questioned by Wibe (2002) who argued that the results were 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. They have accused each other of being “scientific humbugs” and producing research of “deficient standards.” To 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 achievement.

The remainder of this chapter proceeds as follows. In Section 6.1 we describe the main features of the Swedish reform. In Section 6.2 we summarize the available international evidence on the effects of school competition. In Section 6.3, we present descriptive evidence on the prevalence of school choice in Sweden. In sections 6.4 and 6.5 we examine the consequences of school choice; Section 6.4 examines the effects of school choice on segregation and

---

46 We will use “independent schools” and “private schools” interchangeably. The independent schools, although privately run, are publicly funded and generally free to students (prior to 1997 they were allowed to charge a small amount of tuition). Thus, they are more similar to US charter schools than to US private schools.
costs, while Section 6.5 examines the effects on student achievement. In Section 6.6 we conclude.

6.1 Main features of the Swedish reform

As described above, the Swedish reform introduced universal school choice by allowing parents to choose between public schools. It also allowed for the establishment of privately run independent schools. The schools are funded by the municipal governments on a per-student basis in an amount determined in negotiation with the municipal government. The determination is based on the general principal the funds to the independent schools are to be allocated on the same basis as the funds for the public schools in the municipality. The rules thus imply that there will be variation in the grant per student across municipalities. Within a municipality, however, the local government may not discriminate between public and independent schools – the grant amount per student must be the same for these two types of schools. The independent schools are not allowed to select students based on particular admissions criteria, although this rule may be difficult to enforce in practice. The Swedish system has some similarity to the charter school system in the US as well as to systems operating in Chile and New Zealand.

The establishment of independent schools is determined in an application procedure administered by the National Agency for Education (NAE). The municipality in which the school will be located is allowed to express its opinion on the application, but it has no veto power over the application.

6.2 International evidence on school competition

The effect of private school attendance, and school competition more generally, on student achievement is a contentious subject in many countries. The research community has not reached a consensus as to the effects 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

---

47 For example, some municipalities include compensation for pupils in need of extra resources and rental expenses when defining per pupil cost while others do not. The rules for determining this grant have been changed on several occasions. Initially, there was a nationally determined floor of 85 percent of the cost per student in local public schools. Required compensation was less than 100 percent since private schools were exempt from certain administrative duties and the value added tax (VAT).
and state – are likely to condition the extent and direction of any differences in outcomes produced by public and private schools. Here we primarily address a fairly narrow question: What has previous research found regarding the effect of private school attendance on student achievement? We also review evidence on the more difficult question of the effect of competition among schools on student achievement.

We first consider the evidence on 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, and they are frequently administered by private contractors. Charter schools receive public funding, and do not charge tuition; in this respect, they are public schools. However, they are exempt from many of the rules and regulations that govern other public schools. For example, charter schools are often exempt from teacher certification as well as curriculum and testing requirements that apply to public schools. An important way that charter schools differ from private schools is that they are not allowed to charge tuition, but instead receive public support. In this respect, independent schools in Sweden are more similar to American charter schools than to American private schools.

The literature on charter schools is relatively 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 in which charter and traditional public schools in a five mile radius are matched to each other (based on initial test scores) and the performance of their students compared. He also implements an instrumental variables strategy in which the exogenous variation in the existence of charter schools is derived from proximity to a state university, as the Michigan charter school law allows state universities to approve charter schools. Moreover, 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 that “when charter schools are compared to public schools with similar pre-charter characteristics, students in charter schools score no higher, on average, and may even be doing worse.” He also concludes that “charters have had little effect on student achievement in neighboring public schools.”
Bettinger’s work is limited to Michigan. The US Department of Education regularly arranges for contractors to collect nationwide achievement test performance of 4th, 8th and 12th grade students on the National Assessment of Educational Progress (NAEP). In the Spring of 2003, the Department included a sample of fourth and eight graders in charter schools in the regular NAEP assessment of reading and mathematics skills. The availability of these data set off a controversy. The Department, which has strongly backed charter schools under the Bush administration, had analyzed the data, but twice delayed the release of a public report on charter school performance. The NAEP charter school data were available via the Internet tool that gives public access to NAEP data by the summer of 2004, however. Researchers at the American Federation of Teachers promptly analyzed the data, releasing their findings in Nelson, et al. 2004. The AFT report concluded, “Compared to students in regular public schools, charter school students had lower achievement both in grade 4 (six scale points lower in math, seven scale points lower in reading) and grade 8 (five points lower in math, two points lower in reading). These differences were all statistically significant, except for grade 8 reading, and translate into about a half year of schooling.” The lower performance of charter school students held up when the AFT looked separately at students on free and reduced price lunch, those in central cities and black students; the NAEP tool does not allow one to simultaneous control for these characteristics. The AFT’s findings were prominently displayed in the *New York Times*, which generated some controversy.

The effectiveness of charter schools in the US remains a controversial and politically charged issue. Shortly after the AFT’s study was released, Hoxby (2004) released a report that claimed charter school students were more likely to be proficient at math and reading than their (geographically) closest public school counterpart. Hoxby argued that her analysis had advantages over the AFT’s study of the NAEP data because she matched schools to their closest geographic competitor, and had a sample that included almost all of the charter schools in the country. Her results must be viewed with caution, however, as Scott (2005) finds qualitatively different results when she reanalyzes Hoxby’s data and matches schools geographically herself. Roy and Mishel (2005) cast further doubt on Hoxby’s main conclusions by showing that her findings are sensitive to including controls for student backgrounds. They find that Hoxby’s findings of a positive effect of attending a charter school on math and reading
scores disappears once the racial composition and family income status of
students are controlled for in the equation.

The literature on private schools in the US is much more extensive but also
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.48 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 (for example, Neal,
1997 and Evans and Schwab, 1995) used religious affiliation or distance from a
Catholic school as a potentially exogenous source of variation in private school
attendance. Recent work by Altonji et al. (2000) casts 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 closer 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 random lotteries to select students from the pool of applicants. These
lotteries were not intended to facilitate research, but instead to allocate scarce
slots in the program. As a result, complete information on the composition of
the applicant pools is typically unavailable in these studies. In the first studies
in this genre, Rouse (1998a) finds that students selected to receive vouchers to
attend private schools outperformed those in the public schools in terms of
gains in their math test scores, but not in reading scores. 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, had similar gains in
math test scores to those who received vouchers to attend private schools.
Furthermore, students in the P-5 schools outperformed the students who
received vouchers in reading. She concludes that “one potential explanation for

48 For a critique, see Goldberger and Cain (1982).
these results is that 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) to other researchers. In these experiments, low-income applicants for a privately funded voucher program were randomly assigned to a treatment group that was offered a private school 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 (2004) reanalyze the data from New York, and their findings raise doubts that the offer of vouchers improved scores of African American students in the experiment. For example, they show that the results for African Americans become statistically insignificant if students with missing baseline scores (almost a third of students) are included in the sample. 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.

The evidence from outside the US is mixed as well. Angrist et al. (2002) study a program in Colombia that provided vouchers to 125,000 children from poor neighborhoods. The vouchers covered roughly half of the cost of private secondary school attendance. Many of the vouchers were allocated by random 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 from 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. Moreover, when they look across 300 municipalities, they find evidence of an increase in sorting by ability, as the higher achieving students were more likely to switch to private schools. Hsieh and Urquiola make the important point that it is difficult to estimate the effects of school competition \textit{per se} on student achievement if there is a change in sorting: higher (or lower) achieving students could disproportionately leave the public schools for private schools, for example, causing the average achievement level of the students remaining in the public schools to fall (or rise) just because of a change in the composition of students.

Even if private schools do not improve the achievement of students who switch from public to private schools, \textit{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. The opportunity to vote with one’s feet introduces competition even in the absence of private schools. In an influential study, Hoxby (2000) examines the effect of the level of “Tiebout choice” in 316 metropolitan areas on student achievement in public schools and on private school enrollment. Tiebout choice is measured by (one minus) the Herfindahl index of school districts’ share of total enrollment within the metropolitan area. The logic of this measure, which is widely used in studies of product market concentration, is that metropolitan areas with more school districts for a given number of enrolled students should provide parents with greater choice of education providers.

Hoxby’s Ordinary Least Squares regression results indicate no effect of district concentration on student achievement. School district concentration may be related to unobserved characteristics of the metropolitan area that are also correlated with student achievement, however, which would bias the OLS results. In an effort to address this possible problem, Hoxby identifies her model by using two variables as instrumental variables: the number of streams and large rivers in the area. She argues that these provide valid instruments because streams form natural boundaries that historically influenced the development of school district boundaries. (Of course, to serve as valid instruments, it is also necessary that the presence of large and small rivers is uncorrelated with other factors that might be related to student achievement and that are uncontrolled in the model.) Her estimates suggest that less choice among school districts is associated with lower achievement of public school
students and higher private school enrollment, as well as higher expenditures per student. She concludes that, “Tiebout choice raises productivity by simultaneously raising achievement and lower spending” (p. 1236-7).

Hoxby’s conclusion, however, has been the subject of considerable criticism. Most importantly, in a reanalysis of her data, Rothstein (forthcoming) finds: (1) Hoxby’s published results cannot be replicated; (2) her computer program mistakenly assigned some schools to the wrong metropolitan area; (3) her results stemmed entirely from a rather idiosyncratic and subjective categorization of large rivers that could not be replicated; (4) qualitatively different results emerge if a more standard definition of large rivers (e.g., inter-county rivers) is used; and (5) the first-stage estimates underlying Hoxby’s specification are misreported in the published version of the paper; the actual first-stage results underlying her estimates reveal an inverse relationship between the number of large streams in an area and her index of school district competition, contrary to the logic underlying her estimation strategy. Hoxby (2005) disputes these findings in a reply, but in our judgment her initial results should be viewed with skepticism.

Even despite these criticisms, it is still unclear how Hoxby’s results should be interpreted in the context of Sweden. First, Hoxby finds that Tiebout competition already exists among many public schools in the US. It is possible that competition among public schools was already sufficient in Sweden prior to the school choice reform and that the increased competition due to the reform will have only a minor effect. Second, Hoxby’s results relate to competition between public schools, and do not address the question of how private school competition affects public school performance. It is possible that Sweden’s school choice reform affected the sorting of students between schools in a way that could either offset any gains from increased competition or further enhance those gains. Only a direct examination of school choice in the Swedish context can shed light on these issues.

New Zealand recently introduced an education reform which is fairly similar to the Swedish reform, and Fiske and Ladd (2000) have analyzed this reform. They find that school attendance patterns became more polarized within five years after New Zealand adopted parental school choice and promoted school competition. In particular, they find that there was a dramatic increase in stratification across schools by students’ socioeconomic status.

Based on this review of the literature, we tentatively conclude that the injection of more private schools into the Swedish education system is unlikely
to be a panacea. Private schools could perform better or worse than public schools, but there is little reason to expect large differences from the previous empirical literature, or to generalize the results from countries with very different educational institutions or demographics. The academic benefits of school choice and private school attendance 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 in Sweden. First, there are no data on mobility between municipalities due to variations in school quality, although 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 determine the fraction of families that choose where to reside based on their choice of school. Furthermore, until recently there was little information on families that exercised their option to attend a public school outside their immediate area of residence. However, the fraction of students attending independent schools is readily available.

Prior to Sweden’s school choice reform in 1992, there were around 90 independent compulsory schools. By the 2002-03 school year, the number of independent compulsory schools had increased to 539, and 5.7 percent of children attended these schools. Independent schools tend to be located in urban areas. This is illustrated in Figure 6.1, which shows the evolution of the private school share of total enrollment across different municipalities grouped by population density. In the most urbanized areas (4th quartile of population density) 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 only 0.1 to 1.2 percent during the same time period.

In 1992 the majority of independent schools had a special pedagogical profile, such as Montessori. The growth in private schools, since the introduction of school choice, has to a large extent been concentrated among schools with more general pedagogical approaches. In the 2002-03 school year, 44 percent of private schools had a general approach or focused on a particular
subject area, 17 percent had a religious or ethnic profile, 32 percent had a specific pedagogical approach, and 7 percent had other approaches.\(^{49}\)

**Figure 6.1** The share of students attending an independent compulsory school by population density of municipality, percent, 1992-2001.

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 4\(^{th}\) quartile refers to the most densely populated areas and the 1\(^{st}\) quartile to the least densely populated areas.

As mentioned above, until recently there was little information available on parental choice between public schools. However, a recent study by the National Agency for Education (2003) asked parents questions about their choice of school. The study found that it was twice as common for parents to choose a public school outside the one in their immediate area of residence than to choose a private alternative. Around a quarter of parents reported having chosen the closest public school.

\(^{49}\) These numbers are taken from the National Agency for Education website.
6.4 The effects of school choice: segregation and costs

In this section we examine the consequences of school choice in Sweden. We begin by looking at the effects of school choice on two intermediate outcomes: segregation and costs. In section 6.5, we turn to an examine of the effects of school choice on student achievement.

6.4.1 The effect of school choice on segregation

Does school choice lead to less mixing of children from different backgrounds? This question may be a particularly important one in the presence of what are known as peer effects. There is fairly convincing evidence suggesting that the behavior and performance of one’s peers influence one’s own behavior; see, e.g., Sacerdote (2001). However, the nature and quantitative importance of peer effects in the classroom have yet to be determined, and the measurement of peer effects can be difficult, as discussed by Manski (1993). Nonetheless, if peers do have an important effect on student achievement, and if increased school choice leads to more segregation across schools and less opportunity for students from disadvantaged backgrounds to interact with their more advantaged peers, there may be cause for concern.

Therefore, a primary question is whether the existence of independent schools in Sweden has increased sorting of students from different backgrounds across schools. Independent schools tend to be located in urban areas, and the characteristics of the population in these areas tend to differ from those of the Swedish population as a whole (for instance, the urban population is more highly educated and contains a higher fraction of immigrants). A simple comparison of the characteristics of students in public versus independent schools across all municipalities would be confounded by these differences in the urban and non-urban populations. Therefore, we instead examine whether the probability of attending an independent school varies with observed family characteristics within each municipality.

To answer this question we run individual level (logit) regressions relating private school attendance to individual characteristics. We run these regressions for students who completed compulsory schools in the spring of 2001. Consistent with previous research (e.g. Hsieh and Lindahl, 2003), our results suggest that parental education and immigrant status are important predictors of private school attendance. Students with university-educated parents are 4.5 percentage points more likely to attend a private school than students with parents who have completed only a compulsory school 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 when considering only enrollment in private schools with a general pedagogical profile (the majority of independent schools). There is no statistically significant association between attendance at a private school with a general profile and immigrant background or parental education. This suggests that much of the relationship between private school enrollment and student background is driven by attendance at private schools with a particular profile, be it a focus in a particular subject area, a specific pedagogical approach, or a specific religious or ethnic profile.

We can also analyze whether the increase in private schools appears to have increased segregation across schools, using information from Stockholm upper secondary schools. In 2000, a new admissions system was introduced in the Stockholm municipality which admitted students on the basis of their grade point average in compulsory school. We examine the effect of private schools on segregation by comparing the situation in (private and public) schools prior to the reform to the situation after the reform. Of course, one must be cautious in generalizing from the Stockholm experience to the system operating in the rest of the country, since the effects of private school on segregation may be substantially different under the Stockholm system than in a system without explicit sorting on performance.

The statistical office in Stockholm has evaluated the effects of the reform by comparing sorting on observed characteristics in Stockholm upper secondary schools in 2001 and 1999; see USK (2002). Since this evidence is based on a before-and-after calculation, the estimates may be biased if there is an underlying trend in sorting that is incorrectly attributed to the reform. Nevertheless, the magnitudes of the changes found by the study seem too large to be attributed solely to such a trend. One pattern that emerges from the evaluation is that mobility increased over this period, in that the probability of attending a school in the neighborhood of residence declined. Moreover, segregation across schools increased along several dimensions, including immigrant status, parental income, and parental education. 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). This means 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
across schools will affect the outcomes for the least well-off students depends on the nature of peer effects and is a question that cannot yet be answered with the available data.

6.4.2 The effect of school choice on costs

A key argument for introducing school choice is that 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 student increases along with the independent school share. A percentage point increase in the independent school share raised total costs per student in private schools by SEK 25,000 according to their analysis. However, these conclusions were based on a cross-section regression for 2001, which examines whether costs are higher in municipalities with a greater share of students in independent schools. We would argue that the relevant question is instead whether increases in the private school share raise total cost. 50

Table 6.1 illustrates the limitations of the cross-section regression approach used by Fransson and Wennemo. 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 are related to costs. The coefficient on the private school share of enrollment is statistically significant; total costs in public schools are 0.6 percent higher when the private school share is one percentage point higher. 51 In Column (2) we present the results from an analogous regression for 1992. At that time there were only a few private schools, as the choice reform had not yet been implemented. The coefficient on independent school share of enrollment is again statistically significant (at the 10 percent level of significance): a one percentage point increase in the private school share is related to an increase in costs of 1.1 percent. Finally, in Column (3) we illustrate the problems associated with the

50 In addition, one might want to hold student performance constant, since proponents of school competition argue that it reduces the cost of producing a given level of student achievement.
51 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 the only available measure for 1992. Using total costs in all schools in 2001 gives a lower and somewhat less precise coefficient: the estimate is 0.49 (t-statistic: 2.78).
cross-section analysis. This column presents results of a regression of total costs in 1992 on the private school share in 2001. The results of this regression clearly do not reflect the causal effect of private school share on school expenditures, as the private school share is measured at a considerably later point in time than expenditures. Nonetheless, the coefficient on private school share in 2001 is positive and statistically significant. This indicates that unobserved characteristics of the municipalities, rather than causal effects of private schools themselves, are responsible for the cross-section results.

Table 6.1 The relationship between total costs and the independent school share. Dependent variable: logarithm of total cost per student

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of students in independent schools</td>
<td>0.57</td>
<td>1.09</td>
<td>0.42</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(3.12)</td>
<td>(1.84)</td>
<td>(2.41)</td>
<td>(0.72)</td>
</tr>
<tr>
<td>Adjusted R-square</td>
<td>0.304</td>
<td>0.486</td>
<td>0.490</td>
<td>0.794</td>
</tr>
<tr>
<td># Municipalities</td>
<td>263</td>
<td>263</td>
<td>263</td>
<td>526</td>
</tr>
</tbody>
</table>

Notes: *-statistics 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 students 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.

We can instead examine whether changes in school expenditures between 1992 and 2001 are related to changes in the private school share of enrollment, using the fixed effects regression as shown in Column (4).52 The results of this regression suggest 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 statistically significant. Of course, causality may run in both directions – arguably, the incentives are such that it is potentially more attractive to start an independent school in regions in which the costs in public school are high and rising.

52 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.
In summary, 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. We explore this issue in the following section.

6.5 The effects of school choice: student achievement

The primary difficulty in estimating the causal effect of school choice on student achievement is that the amount of choice available to parents is not randomly determined. In particular, it is reasonable to suspect that independent schools locate in areas in which parents demand them. The demand for alternative schools is likely to be higher when parents are unhappy with the performance of existing schools. The number of independent schools in a municipality may thus be systematically related to the average achievement level in the public schools, and simply comparing the average test scores of students in municipalities with higher and lower shares of independent schools will fail to capture the true causal effect of school choice on student achievement. To avoid this problem requires an exogenous source of variation (an instrumental variable) that is correlated with the level of school choice in a municipality but otherwise uncorrelated with student achievement.

6.5.1 Previous Swedish studies

The analysis of school competition in Sweden began with Bergström and Sandström (2001). Sandström and Bergström’s (2005) article in the *Journal of Public Economics* uses data on a single cross section of 9th graders who finished public compulsory schools in 1998. The key independent variable is the share of students in the municipality who attended an independent compulsory school. The outcome was the performance on tests (*Nationella prov*) in Math, English, and Swedish. The data consist of a sample of students from about 30 municipalities – this small number of municipalities is one potential limitation of the study.

Sandström and Bergström adopt a complicated approach to solve for two difficult econometric problems: possible simultaneous determination between student achievement and the fraction of students attending independent schools and the possibility that independent schools attracted a nonrandom sample of students from public schools, thereby affecting the composition of those who
remained in public schools. They apply an instrumental variables approach for the independent school share, and they estimate a parametric sample-selection correction to address composition changes. Both of these steps require, we would argue, rather arbitrary assumptions, and it is crucial that all characteristics relevant for student performance are included in the regression. For example, the key identifying assumption for their simultaneous equation model is that the propensity of municipalities to contract out non-school services is unrelated to factors that affect student achievement. This may or may not be true. For example, municipalities that contract out childcare services may have a higher proportion of women who work or may have greater academic ambitions for their children; these characteristics could affect student achievement.

Ahlin (2003) uses data similar to that used by Bergström and Sandström. However, Ahlin has two observations on test scores for each individual in the dataset, from both 9th grade and 6th grade tests. Therefore, the fact that the private school share may be a function of prior achievement is less of a concern, since she can control for prior achievement in the regressions with 6th grade test scores. In order to avoid bias due to the selection of stronger (or weaker) students into private schools, she includes students in both public and private schools in the regressions. This appears to be a sensible specification, since the prospective benefits of school competition should accrue in all schools. Therefore, Ahlin avoids some of the methodological limitations of Bergström and Sandström’s studies.

Nonetheless, the results of her study are very similar to those of Bergström and Sandström. An increase in the independent school share by ten percentage points is associated with a 0.19 standard deviation improvement in math scores in Sandström and Bergström’s study and a 0.17 standard deviation improvement in Ahlin’s study. Ahlin’s results translate into a six point increase in percentile rank scores. None of the studies find an effect of school choice on student performance in Swedish or English. Moreover, in both studies there appears to be only a slight downward bias in the cross-sectional OLS estimate.

53 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 first-stage regression. If not, the proposed instrument may be invalid because it is correlated with the omitted characteristics. In practice, Sandström and Bergström (2005) use the share of young children attending independent preschools in the municipality to identify the influence of the independent school share.
of the coefficient on the independent school share, suggesting that endogeneity of the independent school share is not a major issue in their analyses. There is no evidence from these studies suggesting that disadvantaged students gain less from competition than their more advantaged peers. In the following section we reanalyze the data from these studies. Our reanalysis is also inspired by the work of Runeson (2003) and Hsieh and Lindahl (2003).

6.5.2 A reanalysis

We have access to repeated cross sections of data where the outcome measure is either students’ test scores or grades for the time period 1998-2001. Between 1998 and 2000, the test score data were constructed in a similar fashion – the data includes observations from about 30 municipalities. In 2001, the data were obtained by random sampling at the school level. About 30 municipalities have at least two schools in the test score data during 1998-2001. The grade data, on the other hand, are for the entire population of students. A nice feature of the repeated cross-sections is that we can control for unchanging unobserved differences across municipalities that potentially bias the results. Thus, our estimates are robust to the fact that the level of student performance in the municipality may influence private school entry.54

Given a choice between test score data and grade data of equivalent quality, we would prefer the test score 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, as we have noted, is not standardized across the country.55 This is not to say that grades are uninformative about student performance, but the metric may vary across municipalities.

However, there are several limitations of the test score data. First, only 30 municipalities are observed more than once over the period, so there is only limited variation in the key explanatory variable, the independent school share of enrollment. Second, test score data are not available for all students due to absenteeism on the day of the test, which may be highest among lower performing students. If absenteeism is systematically related to the degree of

54 However, if private school entry is influenced by trends in student achievement, our results may be biased.

55 Chapter 4 shows, for instance, that the variation in grading standards seems to have increased with the introduction of the new grading system in 1998.
school choice in a municipality, then the estimate on the independent school share will be biased. Third, independent schools are not required to report test results. For our current purposes this is partially mitigated because we are examining the performance of both public and private school students combined, but it seriously compromises any attempt to estimate the test score gains from private school attendance. The grade data do not suffer from these problems since they are available for the full population of students graduating from compulsory school in the municipalities in the sample.

The data contain scores from two tests in Math (tests A and B) and two tests in English (tests B1 and B2). Math test A measures 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 a reading comprehension test in Swedish (test A), but unfortunately the actual score on this test is unavailable in the data, which provides only the grade implied by the test score. The five sets of test scores are comparable over the period of study.

Our regressions include municipality fixed effects, which take into account observed and unobserved characteristics of the municipality that are fixed over time. This implies that the regressions relate the change in outcomes to the change in private school enrollment. Of course, we also hold constant observable characteristics associated with student achievement. The covariates in the regression include 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). The dependent variables (grades and test scores) are measured for both private and public school students in the municipality.

Consistent with the previous studies, we measure competition from independent schools as the share of students enrolled in these schools. Since this measure varies only at the municipality level, it is thus vital to have an estimator that is robust to unobserved municipality characteristics. There is also an issue of when to measure this key independent variable. Using independent school share contemporaneous to test scores assumes that the feed-back from competition to student performance is instantaneous. This may not be entirely plausible – one would also think that competitive pressure from preceding years is (perhaps more) relevant. We attempt to deal with this problem by also including the private school enrollment share averaged over three years – the
year in which test scores are measured (year \( t \)) as well as in the two preceding years (years \( t - 1 \) and \( t - 2 \)). For ease of interpretation, we report results that have constrained the coefficients to be the same on each of the three independent share variables.\(^5\)

Tables 6.2.a-c present the regression results. Table 6.2.a presents results for Math, 6.2.b for English, and 6.2.c for 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 includes only a main effect of independent school enrollment, and an extended specification which also includes interactions between private school enrollment and indicators for characteristics related to socioeconomic disadvantage. The estimates of the coefficients on these interaction terms reflect the extent to which foreign born students and students whose parents have low levels of education are differentially affected by competition from independent schools.\(^7\) The tables also report estimates of the coefficient 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 causal 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 of higher (or lower) achieving students into private schools is driving the estimates.

Before we review the results, it is useful to outline what we believe would constitute robust evidence of positive effects of competition on achievement. If competition from private schools is truly beneficial for students, we would expect the estimated effects to be roughly similar across subject areas.

\(^5\) Throughout this analysis \( t \) refers to the spring of the school year. Thus grades awarded in the spring of 2001 (pertaining to the 2000-01 school year) are related to private school enrollment share in October 2000. When we experimented with controlling for the independent school share in years \( t, t - 1, \) and \( t - 2 \) separately, generally enrollment in year \( t \) was the most significant determinant of test scores, but in a few cases one of the lagged values was more important than the contemporaneous value. Since it is difficult to interpret such variation 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 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.

\(^7\) The indicator variable for low educated parents is defined to equal one if both parents have only a comprehensive education, zero otherwise.
Additionally, the estimated effects should be similar whether the outcome variable is test scores or grades. The estimated effect on grades may be weaker since different standards are used across the country, but we would expect the estimated effects to be in the same direction regardless of whether test scores or grades are used as the outcome variable. Finally, we would expect test scores and grades to have the greatest correspondence in Math. There may be a greater deviation in Swedish since, for instance, the ability to write essays (which is not explicitly tested) is a vital component of the subject.

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

<table>
<thead>
<tr>
<th>Math</th>
<th>Test score</th>
<th>Problem solving</th>
<th>Final grade</th>
<th>Entire pop.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basic</td>
<td></td>
<td>Sampled municip.</td>
<td>Entire pop.</td>
</tr>
<tr>
<td>Independent school share</td>
<td>1.01</td>
<td>1.02</td>
<td>1.11</td>
<td>1.13</td>
</tr>
<tr>
<td>…interacted with foreign born</td>
<td>-1.18</td>
<td>-1.19</td>
<td>-.14</td>
<td>-.19</td>
</tr>
<tr>
<td>…interacted with low-ed. parents</td>
<td>-1.40</td>
<td>-1.58</td>
<td>-.36</td>
<td>-1.45</td>
</tr>
<tr>
<td>Attending independent school</td>
<td>2.46</td>
<td>1.93</td>
<td>6.62</td>
<td>5.61</td>
</tr>
<tr>
<td>…interacted with foreign born</td>
<td>3.20</td>
<td>4.50</td>
<td>6.08</td>
<td>3.32</td>
</tr>
<tr>
<td>…interacted with low-ed. parents</td>
<td>1.23</td>
<td>5.94</td>
<td>2.50</td>
<td>-2.73</td>
</tr>
<tr>
<td># individuals</td>
<td>49,298</td>
<td>49,298</td>
<td>49,298</td>
<td>49,298</td>
</tr>
</tbody>
</table>

Notes: t-statistics 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), foreign-born parents, at least one parent with an upper secondary degree, at least one parent university-educated, the share of the population with low education levels, immigrant density, and mean income in the municipality. Results are weighted by the inverse of the probability of selection into the sample.

Overall, the evidence based on the limited sample of 30 municipalities (first three columns of each table) does not meet these expectations. For example, the
estimated effects of competition on math achievement are significantly positive when test scores are used as the outcome variable, but significantly negative when final grades within the same population of students are used as the outcome variable. Therefore, the results from this subset of municipalities do not seem particularly credible. This is most likely due to the small number of municipalities in the sample.

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

<table>
<thead>
<tr>
<th>Test score</th>
<th>Final grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading comprehension</td>
<td>Listening comprehension</td>
</tr>
<tr>
<td>English</td>
<td></td>
</tr>
<tr>
<td>Independent school share</td>
<td>2.06</td>
</tr>
<tr>
<td>(4.19)</td>
<td>(4.59)</td>
</tr>
<tr>
<td>…interacted with foreign born</td>
<td>-.37</td>
</tr>
<tr>
<td>(1.95)</td>
<td>(1.64)</td>
</tr>
<tr>
<td>…interacted with low-ed. parents</td>
<td>-.77</td>
</tr>
<tr>
<td>(4.20)</td>
<td>(3.78)</td>
</tr>
<tr>
<td>Attending independent school</td>
<td>9.36</td>
</tr>
<tr>
<td>(6.13)</td>
<td>(4.37)</td>
</tr>
<tr>
<td>…interacted with foreign born</td>
<td>4.31</td>
</tr>
<tr>
<td>(3.20)</td>
<td>(0.69)</td>
</tr>
<tr>
<td>…interacted with low-ed. parents</td>
<td>4.57</td>
</tr>
<tr>
<td>(0.80)</td>
<td>(1.00)</td>
</tr>
<tr>
<td># individuals</td>
<td>47725</td>
</tr>
</tbody>
</table>

Notes: See Table 6.2.a

In contrast, the results based on the entire population of students (final column of each table) seem to be much more robust across the three subjects areas, with estimated coefficients on independent school share ranging from .36 to .46. To illustrate the implications of these estimates, we can evaluate them at a point corresponding to the typical variation in the data. The independent school share has increased by one percentage point between 1998 and 2001, and the standard deviation of the share is 1.2 percentage points. The estimates therefore imply that student test scores improve by approximately half a percentile rank in response to a one standard deviation increase in the
independent school share.\textsuperscript{58} The estimated effect in Math is slightly smaller than the one estimated by Ahlin (2003).\textsuperscript{59} Disadvantaged students also appear to benefit less from increases in competition, as illustrated by the estimates on the interaction terms. Across the board, students attending private schools score higher than their public school counterparts, although, as emphasized earlier, this may reflect self-selection of higher achieving students into private schools rather than causal effects of private school attendance.

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

<table>
<thead>
<tr>
<th>Swedish</th>
<th>Test grade</th>
<th>Final grade</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading</td>
<td>Test score pop.</td>
</tr>
<tr>
<td><strong>Independent school share</strong></td>
<td>1.96</td>
<td>.91</td>
</tr>
<tr>
<td></td>
<td>(2.57)</td>
<td>(1.22)</td>
</tr>
<tr>
<td>...interacted with foreign born</td>
<td>-34</td>
<td>-32</td>
</tr>
<tr>
<td></td>
<td>(2.06)</td>
<td>(1.85)</td>
</tr>
<tr>
<td>...interacted with low-ed. parents</td>
<td>-42</td>
<td>-31</td>
</tr>
<tr>
<td></td>
<td>(3.08)</td>
<td>(1.54)</td>
</tr>
<tr>
<td><strong>Attending independent school</strong></td>
<td>7.73</td>
<td>4.83</td>
</tr>
<tr>
<td></td>
<td>(5.33)</td>
<td>(3.97)</td>
</tr>
<tr>
<td>...interacted with foreign born</td>
<td>1.96</td>
<td>-2.81</td>
</tr>
<tr>
<td></td>
<td>(0.40)</td>
<td>(0.75)</td>
</tr>
<tr>
<td>...interacted with low-ed. parents</td>
<td>-3.55</td>
<td>-2.71</td>
</tr>
<tr>
<td></td>
<td>(1.69)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>#individuals</td>
<td>51828</td>
<td>51828</td>
</tr>
<tr>
<td></td>
<td>385054</td>
<td>385054</td>
</tr>
</tbody>
</table>

Notes: See Table 6.2.a

\textsuperscript{58} This is the population-weighted average of the growth within municipalities.  
\textsuperscript{59} 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.
In summary, there appears to be no evidence suggesting that students are harmed by competition from private schools. However, competition from independent schools is no panacea either. The gains we estimate for native-born students whose parents are relatively highly educated are fairly small. Perhaps this is not too surprising, as parents have always had the option of moving to a different location in order to choose a different school for their children. Additionally, the school choice reform also provided parents with the option to choose between public schools; therefore, the added competitive pressure introduced by the availability of independent schools in the municipality is probably relatively minor.

6.6 Summary and conclusions

As discussed above, the international evidence on the effects of private school attendance and competition between schools is mixed. There are some studies that find positive effects, but according to other studies neither private school attendance nor school competition has an effect on student performance.

We have presented evidence from Sweden suggesting that independent schools increase segregation across schools: immigrants and children with highly-educated parents are more likely to attend independent schools. The independent school share of enrollment appears to be unrelated to total costs per student. In particular, independent schools do not appear to increase costs as was argued in a recent report. Ninth grade performance of the average native-born student is positively related to the independent school share in the municipality, but this association appears to be relatively modest. There is a much weaker and often insignificant association between achievement and the independent school share for foreign born students and children with parents with low levels of educational attainment.

It is important to keep in mind that we have examined only one component of the school choice reform in Sweden: the expansion of independent schools. The reform also introduced the possibility of choosing a public school other than that in the student’s immediate area of residence. Little is known about the effects of this facet of expanded school choice in Sweden. This is unfortunate, since this aspect of the reform may be at least as important as the expansion of the independent schools.
References


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


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

7 Quantitative tests as an evaluation device

In order for school choice to improve school productivity via competitive pressure, parents must be reasonably well informed about the quality of the schools from which they are choosing. This requires that parents are able to obtain and evaluate information on various facets of school quality. Parents may collect this information in a number of ways – through exchanging information with other parents, visiting the schools, and also by considering quantitative information on the school’s effectiveness, such as test scores and grade point averages.

In this chapter we discuss how quantitative information such as grade point averages (GPAs) and scores from standardized tests such as the Swedish national tests can be used to help parents choose among schools, as well as to help policymakers evaluate school effectiveness. The availability of such information at the school level has increased dramatically in recent years, in Sweden as in many other countries. The National Agency for Education in Sweden has developed a special data base (called SIRIS), which is available on its website and contains information on average test scores and grade point averages at the municipality level and, to a lesser extent, for individual schools. Many schools have also started to report such information on their own web pages. Some schools even report tables comparing their own average GPAs and test results to those of competing schools.

It is thus important to know the benefits and limitations of such information as an indicator of school quality, particularly in light of growing interest in accountability policies which reward or penalize schools based on student performance on tests. For instance, in some Swedish public policy discussion, it has been suggested that resources be allocated to schools based on their students’ performance national tests, in order to provide incentives for schools to improve performance.\(^{60}\) Others have suggested that teacher salaries include a bonus component based on changes in their students’ test scores over the school year.\(^{61}\) Similar use of test score results to improve incentives has been considered and implemented in many countries. Such systems require testing

---

\(^{60}\) See Storesletten and Zilibotti (1999) for a discussion of such proposals.

\(^{61}\) See Lazear (2003) for a discussion.
instruments that can reliably measure changes in performance for small units such as an individual class.

Tests can also be used for more general evaluations of schools. When Sweden’s National Agency for Education was given greater responsibility for evaluation in 2003, 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 measures of the results and their quality.” (Dagens Nyheter, March 3, 2003). It is difficult to see how this goal could be achieved if national tests and other quantitative tests are not used more frequently than previously.

Quantitative tests as an evaluation device are, however, frequently controversial, both in Sweden and in other countries. Teachers often emphasize that most quantitative tests measure only a limited set of skills and do not contain enough information to effectively assess the school’s contribution to student achievement. Quantitative tests are also controversial 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 unfortunate, as qualitative research based on in-depth interviews and observations can serve as a complement to – rather than a substitute for – statistical analysis based on quantitative tests.

It is important, however, to know what information is (and is not) contained in data from grades and tests, and the potential benefits and limitations of systems which reward or penalize schools based on test performance. The purpose of this chapter is to contribute to the discussion about the appropriate role of test score data in Swedish education policy. We begin in section 7.1 with a general discussion the potential limitations of systems which use test scores and grades to evaluate school quality, drawing on recent US studies. In Section 7.2 we describe Sweden’s national tests, and in Section 7.3 we assess whether these tests provide useful information about student achievement, by analyzing the relationship between grades and test results in compulsory school and subsequent educational attainment and earnings in adulthood. Finally, in Section 7.4, we summarize our results and discuss how tests results might be used more effectively in future evaluations of Swedish schools.
7.1 Problems and limitations of standardized tests

There is a growing international literature on the benefits and limitations of using standardized tests for accountability purposes. Koretz (2002) discusses long-run trends in political interest in using standardized tests for such purposes. In recent years such interest seems to have increased dramatically, particularly in the United States. At the same time, recent research has shed new light on the potential problems associated with using test scores for accountability purposes. While some research suggests there may be significant benefits from such test-based evaluation schemes, the literature also clearly demonstrates a number of associated problems and limitations that must be taken into account. We review these problems and limitations below.

All skills and knowledge cannot be easily measured

All skills and knowledge cannot be easily measured – this problem may be illustrated by the fact that the Swedish national tests over the years have only been conducted in a limited number of subjects, presumably those in which skills and knowledge are most easily measured. Additionally, in Sweden, as in many other countries, many of the goals of public education policy are rather broad and extend beyond measurable skills and knowledge. In a recent policy document, for example, the Swedish government wrote that “the school has an important up-bringing role to play, not least to teach and consolidate society’s basic values.” (Translation from Regeringens skrivelse 2001/2002: 188, page 4.). It is not easy to measure the extent to which such goals are assessed successfully with a standardized test.

Teaching to the test

If schools and teachers are evaluated based on the test scores of their students, and if the outcomes of the evaluations have meaningful consequences, teachers will have incentives to “teach to the test” – devoting extra class time to test preparation while deemphasizing material not covered by the test. Because the domains of the tests are typically limited to specific skills and knowledge, other important goals for education policy that are not measured by the test may be neglected by teachers in their efforts to improve their students’ test scores.

Recent research suggests that teaching to the test is a problem in accountability systems that evaluate schools based on students’ test performance. For instance, Jacob (2002) studies the effects of an accountability policy implemented in the Chicago public schools in 1996-1997. He finds
statistically significant and substantial improvements in subject areas covered by the tests: reading and mathematics. But the gains were concentrated in the specific skills which were tested, and did not extend to other dimensions of reading and math. He observes that teaching time in fields that were not tested (like physics and social science) was reduced. He also finds evidence of other strategic responses of teachers, such as increased placement of students in special education classes that are exempt from the tests.

Cheating by teachers and students
Accountability systems that reward or penalize teachers based on student test performance also provide teachers and students with incentives to cheat on tests. Cheating, broadly interpreted, may include making extra efforts to insure that the best students show up on the day of the test, as well as more direct efforts to inappropriately influence scores, such as providing students extra time to complete the exams or changing students’ incorrect answers. Recent research has suggested that teacher cheating may be an important problem. Jacob and Levitt (2003) develop a statistical technique to infer cheating from suspicious patterns of answers on tests used in Chicago’s public schools. They estimate that cheating occurred in at least 4-5 percent of elementary school classrooms. More importantly, they show that cheating became more prevalent when the school district began to use the test results for evaluation purposes. They also find that poor test performance in a classroom in one year increased the likelihood of cheating in that classroom the following 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 must be taken into account if tests are to be used for evaluation purposes.

Low precision at the classroom level
An additional limitation of standardized tests for evaluating school or teacher quality is that the statistical precision of average student performance on a specific standardized test can be very low when the average is computed for a single class or a small school. Many temporary factors may affect the performance of a single student at a specific test occasion – the student might be temporarily ill, or something might disturb the student’s concentration before or during the test. When calculating the mean for a class or school, such temporary factors specific to a single student will average out. But some
temporary disturbing factors (a loud lawnmower, a coughing classmate) could also affect all students in a class simultaneously, and then the problem will remain.

Even more important in this respect is the fact that changes in test scores are often considered theoretically more appealing than test score levels for evaluating the contribution of teachers and schools to student achievement. However, the statistical precision in estimates of changes is typically lower than the precision of levels. Recent work by Kane and Staiger (2001, 2002) show quite convincingly that measured changes in test scores for units as large as schools can be due to temporary and irrelevant factors rather than to real ones. Thus the allocation of resources and bonus payments based on measured test score 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.

A final concern with changes in scores, regardless of the number of students tested, is that tests are often not designed in a way to permit longitudinal inferences. In other words, a gain in scores from 20 to 25, say, may not indicate the same amount of learning as a gain from 80 to 85. Unless considerable thought goes into the vertical scaling of tests, there is no guarantee that changes in scores will yield better information than the levels of scores.

### 7.2 National tests in Sweden

National tests have a long history in Sweden, see Ljung (2000). Prior to 1944, admission to secondary school (realskolan) was based on scores on a special admissions test of Swedish and Math, which students could take only once. This test was regarded as mentally quite demanding for the children and not very reliable, and critics suggested that admission instead be based on GPA, which might reflect student performance more reliably over a longer period of time. But GPA can only be used to compare student performance if comparable across classes and schools, so national tests were introduced in 1944 in order to establish a standardized grading system for the whole country.
Ever since 1944 such national tests have been used in Swedish compulsory schools, and their main purpose has been to make grades comparable throughout the country. In general, they have been conducted in Swedish, English, and Math, but sometimes also in other subjects. Since 1997 national tests have been voluntary in fifth grade and compulsory in ninth grade. In both grades the tests are conducted in Swedish, English and Math. National tests in upper secondary school are a more recent development. They were introduced in 1966 after a major reform of upper secondary school. They have been conducted 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 overuse – of quantitative tests can diminish their informative value, it is useful to know how much information the Swedish national tests provide about school quality and student achievement. Therefore, in this section we examine how well scores from the Swedish national tests predict two important future outcomes – annual earnings and educational attainment.

It is also interesting to know whether test scores provide useful predictive information about future outcomes over and beyond the information provided by student grades. In Sweden, teachers have the test results at their disposal when they grade their students. However, they are also free to use additional information on the students’ daily achievement in the classroom, which may be quantitative or qualitative in nature. Thus it is interesting to know whether grades have stronger predictive value than test scores.

We therefore run several regression models which examine how well grades and test scores predict the two outcomes of interest – annual earnings and educational attainment. The models are run separately for each subject – Math, Swedish, and English – and include dummy variables for both student grades and level of performance on the national tests. We divide the test results into five levels, in a way that corresponds to the Swedish grading system’s recommendations for the fraction of students at each level (7 percent get 1 and 5 respectively, 24 percent get 2 and 4 respectively, and 38 percent get 3).

We conduct this analysis using the UGU-data for the 1948 and 1953 birth cohorts. The detailed results are reported in a set of tables in an appendix. The main results are described below:
Both test scores and grades have strong predictive power, as indicated by the fact that the coefficients on these variables are highly statistically significant. Tests scores and grades are particularly strong predictors of educational attainment, but also predict earnings even conditional on educational attainment.

Test results predict educational attainment and earnings even conditional on grades. This indicates that the tests have additional predictive information that is not contained in grades.

The predictive power of test scores in the three subject areas is about the same. Thus the results do not support the common claim that math skills are particularly important predictors of future outcomes.

In the 1953 data we could do separate analyses for one group of students who were graded according to the old Swedish 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

In this chapter we have discussed the potential value and limitations of national standardized tests for evaluating Swedish schools and Swedish education policy. Our analysis suggests that the Swedish national tests measure skills and knowledge that are highly correlated with future educational attainment and labor market earnings in adulthood. Even controlling for grade received in the subject area, there is additional explanatory value in the scores from the national tests. Thus one cannot simply dismiss the value of tests by claiming that they do not contain useful information. On the contrary, test scores may be useful for evaluating school quality – both for parents choosing between different schools for their children as well as for policymakers assessing the school performance.

Nonetheless, there are many risks involved in using standardized tests to evaluate school quality. Estimates of average test score performance from small units such as classes, small schools, and single teachers may provide a very imprecise measure of true student achievement. Additionally, policies which evaluate teachers and schools based on students’ test performance may affect teachers and schools in non-productive ways – providing incentives for teaching to the test and teacher cheating.
Although we have stressed the potential negative behavioral effects of evaluation policies based on test scores, we do not dismiss the possibility that such policies may also have productive effects – namely, that they may provide incentives for teachers to improve teaching quality. Indeed, we believe that there is much potential for such positive effects, and two recent studies using Israeli data (Lavy 2002, 2003) suggest that these effects can be substantial. Our main conclusion from this chapter is that any accountability system based on standardized test scores must be implemented with much care to be successful and to avoid the potential pitfalls we have discussed.
References


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


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

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.377</td>
<td>0.370</td>
<td>0.387</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Grade B</td>
<td>0.129</td>
<td>0.080</td>
<td>-0.205</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.045)</td>
<td>(0.218)</td>
</tr>
<tr>
<td>Grade Ba</td>
<td>0.205</td>
<td>0.126</td>
<td>-0.140</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.048)</td>
<td>(0.218)</td>
</tr>
<tr>
<td>Grade AB</td>
<td>0.282</td>
<td>0.184</td>
<td>-0.056</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.052)</td>
<td>(0.218)</td>
</tr>
<tr>
<td>Grade a</td>
<td>0.339</td>
<td>0.223</td>
<td>-0.076</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.061)</td>
<td>(0.219)</td>
</tr>
<tr>
<td>Grade A</td>
<td>0.469</td>
<td>0.344</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.177)</td>
<td>(0.184)</td>
<td>(0.294)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-0.106</td>
<td>-</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.037)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Test 3</td>
<td>-0.129</td>
<td>-</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.040)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Test 4</td>
<td>-0.157</td>
<td>-</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.046)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Test 5</td>
<td>-0.177</td>
<td>-</td>
<td>0.070</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.058)</td>
<td>(0.059)</td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.151</td>
<td>0.152</td>
<td>0.145</td>
</tr>
<tr>
<td># observations</td>
<td>7657</td>
<td>7657</td>
<td>7657</td>
</tr>
</tbody>
</table>

Notes: Constants and coefficients for six educational levels not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
Table A.2 Regression estimates, linear probability models. 1948 cohort of UGU-data. Dependent variable: University degree.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.019</td>
<td>0.004</td>
<td>0.056</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Grade B</td>
<td>0.000</td>
<td>-0.010</td>
<td>-0.069</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.023)</td>
<td>(0.114)</td>
</tr>
<tr>
<td>Grade Ba</td>
<td>0.072</td>
<td>0.028</td>
<td>-0.017</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.025)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>Grade AB</td>
<td>0.255</td>
<td>0.143</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.027)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>Grade a</td>
<td>0.481</td>
<td>0.274</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.031)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>Grade A</td>
<td>0.768</td>
<td>0.491</td>
<td>0.655</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.094)</td>
<td>(0.153)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>0.005</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.019)</td>
<td>(0.019)</td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>0.050</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>0.156</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.023)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>0.285</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.029)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.162</td>
<td>0.177</td>
<td>0.159</td>
</tr>
<tr>
<td># observations</td>
<td>7657</td>
<td>7657</td>
<td>7657</td>
</tr>
</tbody>
</table>

Notes: Constants not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
Table A.3 Regression estimates, linear probability models. 1948 cohort of UGU-data. Dependent variable: less than or equal to compulsory school.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.041</td>
<td>0.047</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.006)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Grade B</td>
<td>-0.074</td>
<td>-0.039</td>
<td>-0.154</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.020)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Grade Ba</td>
<td>-0.160</td>
<td>-0.101</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.021)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Grade AB</td>
<td>-0.225</td>
<td>-0.014</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.023)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Grade a</td>
<td>-0.266</td>
<td>-0.163</td>
<td>-0.383</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.026)</td>
<td>(0.096)</td>
</tr>
<tr>
<td>Grade A</td>
<td>-0.280</td>
<td>0.170</td>
<td>-0.392</td>
</tr>
<tr>
<td></td>
<td>(0.077)</td>
<td>(0.080)</td>
<td>(0.129)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>-0.074</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>-0.092</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.018)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>-0.131</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.120)</td>
<td>(0.120)</td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>-0.142</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.025)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.058</td>
<td>0.063</td>
<td>0.058</td>
</tr>
<tr>
<td># observations</td>
<td>7657</td>
<td>7657</td>
<td>7657</td>
</tr>
</tbody>
</table>

Notes: Constants not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.401</td>
<td>0.392</td>
<td>0.410</td>
</tr>
<tr>
<td></td>
<td>(0.041)</td>
<td>(0.042)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Grade B</td>
<td>0.103</td>
<td>-0.073</td>
<td>-0.454</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.132)</td>
<td>(0.739)</td>
</tr>
<tr>
<td>Grade Ba</td>
<td>0.015</td>
<td>-0.081</td>
<td>-0.467</td>
</tr>
<tr>
<td></td>
<td>(0.120)</td>
<td>(0.139)</td>
<td>(0.737)</td>
</tr>
<tr>
<td>Grade AB</td>
<td>0.053</td>
<td>-0.088</td>
<td>-0.410</td>
</tr>
<tr>
<td></td>
<td>(0.123)</td>
<td>(0.150)</td>
<td>(0.738)</td>
</tr>
<tr>
<td>Grade a</td>
<td>0.180</td>
<td>0.012</td>
<td>-0.399</td>
</tr>
<tr>
<td></td>
<td>(0.132)</td>
<td>(0.165)</td>
<td>(0.740)</td>
</tr>
<tr>
<td>Grade A</td>
<td>0.540</td>
<td>0.343</td>
<td>-0.044</td>
</tr>
<tr>
<td></td>
<td>(0.388)</td>
<td>(0.410)</td>
<td>(0.904)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>0.053</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.095)</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>0.155</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.105)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>0.174</td>
<td>0.145</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.121)</td>
<td>(0.1369)</td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>0.221</td>
<td>0.178</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.151)</td>
<td>(0.162)</td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.098</td>
<td>0.098</td>
<td>0.093</td>
</tr>
<tr>
<td># observations</td>
<td>1302</td>
<td>1302</td>
<td>1302</td>
</tr>
</tbody>
</table>

Notes: Constants and coefficients for six educational levels not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
Table A.5 Regression estimates, linear probability models. 1953 cohort of UGU-data. Dependent variable: University degree. Old system, letter grades.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>-0.007</td>
<td>0.023</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.018)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Grade B</td>
<td>0.040</td>
<td>0.009</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.318)</td>
<td>(0.316)</td>
</tr>
<tr>
<td>Grade Ba</td>
<td>0.059</td>
<td>0.038</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.317)</td>
<td>(0.317)</td>
</tr>
<tr>
<td>Grade AB</td>
<td>0.162</td>
<td>0.144</td>
<td>0.064</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.317)</td>
<td>(0.318)</td>
</tr>
<tr>
<td>Grade a</td>
<td>0.353</td>
<td>0.278</td>
<td>0.143</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.318)</td>
<td>(0.320)</td>
</tr>
<tr>
<td>Grade A</td>
<td>0.253</td>
<td>0.073</td>
<td>0.712</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.174)</td>
<td>(0.387)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>0.017</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.050)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>0.044</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.053)</td>
<td>(0.046)</td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>0.124</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.050)</td>
<td>(0.058)</td>
<td>(0.050)</td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>0.287</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td>(0.068)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.096</td>
<td>0.105</td>
<td>0.082</td>
</tr>
<tr>
<td># observations</td>
<td>1302</td>
<td>1302</td>
<td>1302</td>
</tr>
</tbody>
</table>

Notes: Constants not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
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.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.100</td>
<td>0.061</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.024)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>Grade B</td>
<td>-0.122</td>
<td>-0.496</td>
<td>-0.404</td>
</tr>
<tr>
<td></td>
<td>(0.0070)</td>
<td>(0.421)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Grade Ba</td>
<td>-0.223</td>
<td>-0.633</td>
<td>-0.488</td>
</tr>
<tr>
<td></td>
<td>(0.067)</td>
<td>(0.422)</td>
<td>(0.079)</td>
</tr>
<tr>
<td>Grade AB</td>
<td>-0.312</td>
<td>-0.776</td>
<td>-0.580</td>
</tr>
<tr>
<td></td>
<td>(0.068)</td>
<td>(0.423)</td>
<td>(0.080)</td>
</tr>
<tr>
<td>Grade a</td>
<td>-0.464</td>
<td>-0.0896</td>
<td>-0.653</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.425)</td>
<td>(0.086)</td>
</tr>
<tr>
<td>Grade A</td>
<td>-0.545</td>
<td>-0.997</td>
<td>-0.724</td>
</tr>
<tr>
<td></td>
<td>(0.218)</td>
<td>(0.511)</td>
<td>(0.518)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-0.068</td>
<td>-0.126</td>
<td>-0.028</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.067)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>Test 3</td>
<td>0.049</td>
<td>-0.152</td>
<td>-0.111</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.070)</td>
<td>(0.061)</td>
</tr>
<tr>
<td>Test 4</td>
<td>-0.52</td>
<td>-0.219</td>
<td>-0.153</td>
</tr>
<tr>
<td></td>
<td>(0.069)</td>
<td>(0.077)</td>
<td>(0.067)</td>
</tr>
<tr>
<td>Test 5</td>
<td>-0.016</td>
<td>-0.273</td>
<td>-0.181</td>
</tr>
<tr>
<td></td>
<td>(0.086)</td>
<td>(0.090)</td>
<td>(0.084)</td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.079</td>
<td>0.078</td>
<td>0.083</td>
</tr>
<tr>
<td># observations</td>
<td>1302</td>
<td>1302</td>
<td>1302</td>
</tr>
</tbody>
</table>

Notes: Constants not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
Table A.7 Regression estimates, 1953 cohort of UGU-data. Dependent variable: log annual earnings 1993. New system, figure (relative) grades.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.434</td>
<td>0.430</td>
<td>0.443</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.020)</td>
<td>(0.020)</td>
</tr>
<tr>
<td>Grade 2</td>
<td>0.019</td>
<td>0.039</td>
<td>0.113</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.060)</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>0.066</td>
<td>0.067</td>
<td>0.124</td>
</tr>
<tr>
<td></td>
<td>(0.052)</td>
<td>(0.065)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>0.109</td>
<td>0.077</td>
<td>0.151</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.070)</td>
<td>(0.064)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>0.176</td>
<td>0.122</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>(0.060)</td>
<td>(0.082)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>-0.044</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.051)</td>
<td></td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>-0.002</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.056)</td>
<td></td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>0.038</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.078)</td>
<td></td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>0.058</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.078)</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.120</td>
<td>0.120</td>
<td>0.118</td>
</tr>
<tr>
<td># observations</td>
<td>5776</td>
<td>5776</td>
<td>5776</td>
</tr>
</tbody>
</table>

Notes: Constants and coefficients for six educational levels not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
Table A.8 Regression estimates, linear probability models. 1953 cohort of UGU-data. Dependent variable: University degree. New system, figure (relative) grades.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.009</td>
<td>0.000</td>
<td>0.073</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Grade 2</td>
<td>0.031</td>
<td>0.014</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.028)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>0.078</td>
<td>0.042</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.030)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>0.236</td>
<td>0.141</td>
<td>0.277</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.032)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>0.498</td>
<td>0.329</td>
<td>0.545</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.037)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>0.024</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.023)</td>
<td></td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>0.043</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>0.130</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.029)</td>
<td></td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>0.218</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.035)</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.147</td>
<td>0.155</td>
<td>0.143</td>
</tr>
<tr>
<td># observations</td>
<td>5776</td>
<td>5776</td>
<td>5776</td>
</tr>
</tbody>
</table>

Notes: Constants not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
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.

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Swedish</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.079</td>
<td>0.087</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Grade 2</td>
<td>-0.100</td>
<td>-0.032</td>
<td>-0.083</td>
</tr>
<tr>
<td></td>
<td>(0.026)</td>
<td>(0.029)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>-0.230</td>
<td>-0.099</td>
<td>-0.247</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.031)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Grade 4</td>
<td>-0.343</td>
<td>-0.172</td>
<td>-0.346</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.034)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Grade 5</td>
<td>-0.388</td>
<td>-0.195</td>
<td>-0.384</td>
</tr>
<tr>
<td></td>
<td>(0.027)</td>
<td>(0.039)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Test 2</td>
<td>-</td>
<td>-0.110</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.024)</td>
<td></td>
</tr>
<tr>
<td>Test 3</td>
<td>-</td>
<td>-0.184</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.027)</td>
<td></td>
</tr>
<tr>
<td>Test 4</td>
<td>-</td>
<td>-0.022</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.030)</td>
<td></td>
</tr>
<tr>
<td>Test 5</td>
<td>-</td>
<td>-0.239</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.037)</td>
<td></td>
</tr>
<tr>
<td>Adjusted R-sq.</td>
<td>0.089</td>
<td>0.098</td>
<td>0.086</td>
</tr>
<tr>
<td># observations</td>
<td>5776</td>
<td>5776</td>
<td>5776</td>
</tr>
</tbody>
</table>

Notes: Constants not reported. Standard errors in parentheses. Test scores range from 1 (lowest) to 5 (highest); possible grades (from lowest to highest) are b, B, Ba, AB, a, and A.
8 Family background and earnings: What do education and education policy have 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 addressed the issue of equality of opportunity has also used such a framework. Sociologists have 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 and reduce inequality in opportunities along ethnic lines. Yet another example is regional policy that has aimed at equalizing outcomes among people born in different parts of the country. And, of course, racial gaps in the United States pose a special dilemma.

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 students, 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 explore whether the association between parent and childrens’ socioeconomic status is weaker in Sweden than in other countries. 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 budgets, 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) association as indicating 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. 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.

---

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 0.40. A recent study by Mazumder (2002) suggests that the elasticity in the US might be as high as 0.60. 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 correlations between brothers 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 elasticity reported in Table 1. The Nordic and US estimates were significantly different from one another.

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

<table>
<thead>
<tr>
<th>Country</th>
<th>Study</th>
<th>Elasticity</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sweden</td>
<td>Gustafsson (1994)</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Björklund and Jäntti (1997)</td>
<td>.28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Österberg (2000)</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Björklund and Chadwick (2003)</td>
<td>.25</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>Jäntti and Österbacka (1996)</td>
<td>.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Österbacka (2001)</td>
<td>.13</td>
<td></td>
</tr>
<tr>
<td>Norway</td>
<td>Bratberg et al. (2002)</td>
<td>.12</td>
<td>.12 refers to the 1960-cohort and .17 to the 1950-cohort</td>
</tr>
<tr>
<td>Canada</td>
<td>Corak and Heisz (1999)</td>
<td>.23</td>
<td></td>
</tr>
<tr>
<td>UK</td>
<td>Atkinson et al. (1983)</td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dearden et al. (1997)</td>
<td>.57</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Couch and Dunn (1997)</td>
<td>.11</td>
<td>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.</td>
</tr>
<tr>
<td></td>
<td>Wiegand (1997)</td>
<td>.34</td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>Solon (1992)</td>
<td>.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zimmerman (1992)</td>
<td>.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mazumder (2002)</td>
<td>.60</td>
<td></td>
</tr>
</tbody>
</table>
8.2 What do education and education policy have to do with it?

8.2.1 Theory

The natural question that follows from this short survey is what education and education policy have to do with intergenerational earnings mobility. 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 (2004) 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 parents’ 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 parents’ 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
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.\(^{63}\)

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:

\[
Y_f = \alpha_f + \beta_f X_f + e_f
\]

\(^{63}\) This exercise is inspired by Österbacka (2001).
\[ Y_c = \alpha_c + \beta_c X_c + e_c \]  
(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 \( \beta_f \) and \( \beta_c \), \( \alpha_f \) and \( \alpha_c \) are intercepts, and \( e_f \) and \( e_c \) are errors terms.

It follows that the intergenerational elasticity equals:

\[
\frac{\text{Cov}(Y_f, Y_c)}{\sigma^2} = \left[ \beta_f \beta_c \text{Cov}(X_f, X_c) + \beta_f \text{Cov}(X_f, e_c) \right] / \sigma^2 \\
+ \left[ \beta_c \text{Cov}(e_f, X_c) + \text{Cov}(e_f, e_c) \right] / \sigma^2
\]  
(8.3)

where \( \sigma^2 \) 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 intergenerational covariances in years of schooling? 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 \( \beta \)'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. 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 had returns equal to those in the US? And what would the US intergenerational elasticity be if the US had returns equal to those 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 be 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. When we apply the US covariance to the Swedish data (but keep the Swedish returns) the Swedish estimate only increases from 0.211 to 0.220. When the Swedish covariance is applied to the US data, the US estimate only falls from 0.343 to 0.323.

Although this exercise is mechanical (and is not based on a sophisticated behavioral model), it clearly suggests the importance of the differences in the

---

64 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.
earnings returns to education as a crucial factor behind the cross-national differences in intergenerational earnings elasticities.
Table 8.2 Actual and counterfactual components of the intergenerational income elasticity.

<table>
<thead>
<tr>
<th></th>
<th>$\frac{\text{Cov}(Y, Y)}{\sigma^2}$</th>
<th>$\frac{\beta , \text{Cov}(X, X)}{\sigma^2}$</th>
<th>$\frac{\beta , \text{Cov}(e, e)}{\sigma^2}$</th>
<th>$\frac{\beta , \text{Cov}(e, X)}{\sigma^2}$</th>
<th>$\frac{\text{Cov}(e, e)}{\sigma^2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Swedish estimates</td>
<td>0.211</td>
<td>.089 $\times$ .075 $\times$ 2.733</td>
<td>.089 $\times$ .037</td>
<td>.075 $\times$ .119</td>
<td>.022</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.249</td>
<td>.249</td>
<td>.249</td>
<td>.249</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$= .073$</td>
<td>$= .013$</td>
<td>$= .037$</td>
<td>$.088</td>
</tr>
<tr>
<td>US estimates</td>
<td>0.343</td>
<td>.107 $\times$ 136 $\times$ 3.315</td>
<td>.107 $\times$ 253</td>
<td>.136 $\times$ .155</td>
<td>.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.423</td>
<td>.423</td>
<td>.423</td>
<td>.423</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$= .114$</td>
<td>$= .064$</td>
<td>$= .050$</td>
<td>$.115</td>
</tr>
<tr>
<td>US with Swedish returns</td>
<td>.247</td>
<td>.052</td>
<td>.053</td>
<td>.027</td>
<td>.115</td>
</tr>
<tr>
<td>Sweden with US returns</td>
<td>.329</td>
<td>.160</td>
<td>.016</td>
<td>.065</td>
<td>.088</td>
</tr>
</tbody>
</table>

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 might 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 aspects of Swedish educational policy in this respect include 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 available (or has not yet been discovered).
There is, however, one interesting exception. Meghir and Palme (2003) have recently exploited the variation in schooling 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 was deliberately 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 students’ 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 the 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. Recall 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 was no backsliding in intergenerational mobility during the 1990s as long as we focus on performance in compulsory school. The same result applies to the big-city areas where privatization went further and parents had more options for their children’s school choice. It may be the case that the genetic and family-based environmental factors that are responsible for the intergenerational and sibling correlations render everything else mere background noise when it comes to these correlations. 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](image-url) 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. We return to this issue in our concluding chapter.
References


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


Solon G (2004), A Model of Intergenerational Mobility over Time and Place, in M Corak (ed.), *Generational Income Mobility in North America and Europe*, Cambridge University Press.

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


9 Conclusions

The goal of this book has been to provide an initial evaluation of the effects of Sweden’s sweeping market-oriented reforms of the 1990s. Our analyses have been able to cover considerable ground. In this concluding chapter we return to the general issues that we raised in the introduction – equality and efficiency – and reflect on the policy implications that follow from our analyses.

Equality of opportunities and outcomes has been a longstanding goal of Swedish education policy, and the effects of the reforms must be evaluated in this context. We therefore begin in Section 9.1 with a discussion of whether the “traditional” Swedish educational policy that was in place prior to the market-oriented reforms of the 1990s was successful in meeting its egalitarian goals. In Section 9.2 we then discuss whether the equalization achieved by the system prior to the 1990s was obtained at a cost to efficiency. In Section 9.3 we turn to a discussion of the effects of Sweden’s market-oriented reforms of the 1990s on both equality and efficiency. In Section 9.4, based on the evidence presented throughout this book, we address the question of how Sweden can make the best use of scarce resources as it forges an education policy for the future. The predicted teacher shortage is an important part of this discussion.

Many of our conclusions are quite cautious, as we do not consider the available empirical evidence sufficiently convincing to render a strong verdict. One reason for this is that it is intrinsically difficult to conduct convincing causal analyses of many important educational policy questions. Another reason is that relevant data for such analyses are often unavailable. Based on our examination of Swedish and international research, we are convinced that evaluations of Swedish educational policies can be improved in several respects. We therefore 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 additional quantitative information on student performance. 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?

Did Sweden’s education policy prior to the 1990s successfully equalize earnings? It is difficult to establish a direct causal link from education policy to the skill distribution, and from the skill distribution to the earnings distribution. However, we have presented suggestive evidence that Sweden’s education system prior to the 1990s was effective in promoting equality of both skills and earnings.

International comparisons have shown that Swedish students have done very well on standardized tests. The variation among students is not particularly low, but the variation between schools is low. Similar patterns are evident among adults. Data from the International Adult Literacy Survey, a literacy test that was conducted in 1994, suggest that Swedish adults also perform well in international comparison. This is true for all birth cohorts (cohorts born from 1929 until 1970) and all education levels. In general, Swedish adults in the lower part of the skill distribution perform very well compared to their counterparts in other countries, in particular relative to those in the United States. These patterns do not directly establish the effectiveness of the Swedish education system in narrowing the skill distribution prior to the 1990s, but they are suggestive.

We have also discussed the results of recent attempts to estimate how much of the differences in the earnings distributions between Sweden (and other European countries) and the United States is due to differences in the skill distributions. We conclude from these estimates that a significant portion of the lower earnings variation in Sweden can be attributed to the more equal skill distribution of the Swedish population. Educational policy is likely to have been an important factor in equalizing the skill distribution, but other factors, such as union wage compression, may also have contributed.

A related goal of Swedish education policy is to reduce the impact of family background on educational achievement and thus increase equality of opportunity. A consistent finding 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. Family background therefore appears to be a less important determinant of opportunity in Sweden than in the US.
We also examined 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 educational 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 returns to schooling in Sweden. This outcome is intuitively reasonable: if children’s educational attainment is influenced by their family background, a lower return to education produces a lower impact of family background on earnings. The fact that the difference in the return to education is the main explanation for the difference in intergenerational earnings mobility between Sweden and the US does not necessarily mean that educational policies are unimportant. On the contrary, the return to education itself is presumably a function of education policy.

Nonetheless, this analysis suggests a potential trade-off for education policy. If it is necessary to increase the returns to schooling to compel students to pursue higher education, this may lead to 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 union wage compression and progressive income taxes. The fact that Swedish adults rank very high on literacy and numeracy tests does not rule out the possibility that higher private incentives would have enhanced productivity and economic growth.

The private rate of return to university studies in Sweden was quite low in the 1980s, a period during which gross returns to education were low and extremely high marginal taxes further reduced the after-tax return to education. Even when the subsidy component of the student loans is taken into account, the estimated rates of returns for this period are low. As in many other countries, however, market forces have increased the gross return to education
since then. Further, the Swedish tax reform in 1990-91 reduced the marginal tax rates to a maximum of around 50 percent. When the net value of student loans is taken into account, the return to education appears to be even 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 about average at this time.

The Swedish returns are, however, substantially bolstered by the subsidies imparted by the student loans (and the tax free stipend of about $300 a month). Some have argued that this system does not create optimal incentives, since the returns are contingent on being a student, rather than on 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 group of students was offered financial rewards ranging from 200–600 Euros if they completed their studies 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, suggesting that monetary incentives may play less of a role in students’ decisions than critics might suggest.

However, there are other potential efficiency problems in the Swedish education system. One possible problem is that Swedish university students are relatively old by international standards. This is the result of restricted university admissions and generous admittance rules for older students. In addition, the possibility of enrolling in adult education to improve one’s grades prior to applying to university was introduced in 1997, further delaying the transition to university education for many students. The costs of this delay are non-negligible – our crude calculation suggesting delaying the age of university entry from 19 to 22 reduces the present value of earnings associated with a university degree by 7.5 percent.

### 9.3 How did the reforms of the 1990s affect equality and efficiency?

The decentralization and school choice reforms of the 1990s were far reaching both from a Swedish historical perspective and from an international perspective. Since the students who were affected by these reforms have not
yet entered the labor market, our analyses of these issues have focused on achievement measures such as grades and test scores.

One important question we examined is whether the increased competition among schools due to the school choice reforms has enhanced school productivity. Previous (mainly US-based) research suggests that private schools *per se* are no more efficient than public ones. There is also little evidence from the international literature that the competitive pressure from private schools raises productivity in nearby public schools.

Evidence on the effects of Sweden’s school choice reform is mixed. A study published by Bergström and Sandström (2001) attracted considerable attention by arguing that the presence of independent schools in a municipality is associated with better school achievement in the public schools in the same municipality, at least in math. However, there is concern that independent schools may locate systematically in districts with higher (or lower) achieving students, and that this, rather than a causal effect of independent schools on student achievement, may explain Bergström and Sandström’s result. In an effort to address this problem, we conducted our own analysis, which examined whether *increases* in the share of students enrolled in private schools are associated with *increases* in student achievement. We found that final grades in math, English and Swedish did improve more in municipalities in which independent schools were expanding than elsewhere. We found similar results for test scores in a smaller sample of municipalities.

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

We also explored whether students from wealthy and highly educated families appear to have benefited more from the reforms than students from other family backgrounds. We showed that there was an increase in segregation of students by family background across schools and in resource differentials
between municipalities. We also presented evidence suggesting that increasing class size is particularly detrimental for immigrant students. Thus the decline in resources over the decade might have been more harmful for students from less advantaged families than for those from more advantaged families.

Despite evidence of increased segregation and increased dispersion of student performance, the relationship between family background and school performance (as measured as grade point average at age 16) in Sweden has changed very little from 1988 to 2000. The association between grade point average and parental income has been very stable, as has the correlation of grades between siblings. Thus, the composite importance of families, neighborhoods, and schools appears not to have increased during the 1990s.

What then is our final verdict on the effects of Sweden’s education reforms? The evidence suggests that the positive effects on school productivity are not as apparent or as large as many advocates of school reforms argued they would be. But neither have the reforms greatly reduced equality of outcomes with respect to family background, as many skeptics had feared. At present, it seems unrealistic to argue that Sweden should go back to the centralized system 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 for Sweden to introduce a new series of major school reforms without solid evidence about the consequences of such reforms. Our final recommendations in Section 9.5 provide some examples of how Sweden and other countries might better evaluate teaching techniques, schools, and the quality of the current school system.

9.4 What is the best use of scarce resources?

Resource allocation at the school level

The available research on the best allocation of resources to education has many limitations, and findings from the literature are often mixed. However, we believe that recent research on the importance of school resources offers some useful lessons for politicians. Our survey of recent research suggests that class size can significantly affect student performance. Ten years ago, a reading of the international literature might have suggest that class size had little effect on student achievement. This view was also widespread in Sweden despite the lack of evidence from Swedish data. Today the consensus has changed, mainly
due to a number of recent studies that have been able to circumvent some of the problems involved in studying the causal impact of class size (see, for example, Finn and Achilles, 1990; Nye et al., 1994; Krueger, 1999; Krueger and Whitmore, 2001; and Angrist and Lavy, 1999).

Moreover, a recent Swedish study (Lindahl, 2004) estimates class-size effects of similar magnitude to those found in the recent studies from the international literature. These results are also consistent with our own estimates. We used the sharp change in the teacher-student ratios across Swedish municipalities following the decentralization of the school system to estimate the effects of class size on student achievement. Our analysis indicates that changes in the teacher-student ratio are positively associated with changes in student achievement. Our findings also suggest that the impact of small classes on school achievement is greater for students from more disadvantaged family backgrounds, consistent with the findings from recent international studies.

A cost-benefit analysis of class size reduction using Swedish data (Krueger and Lindahl, 2002) suggests that the internal rate of return on investments in smaller classes is around five percent. Although there are many caveats to this calculation, it is clear that the cost savings of increasing class size may not be able to justify the potential negative impact on student achievement.

So what are reasonable guidelines for politicians and school leaders trying to decide how to allocate scarce resources for schools? In Sweden, 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 indicated that teachers’ relative wages have been declining for many years, 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. In Sweden, the main policy instrument used to affect teacher supply to date has been the number study slots at teachers colleges. We believe that this policy is not likely to be successful. To increase the supply of teachers, and the quality of the teacher pool, it is important to improve incentives to enter the teaching profession.

In surveys teachers often respond that smaller classes would make their job more attractive. So one possible policy recommendation is to reduce class sizes. Such a policy might fulfill the dual objective of increasing the attractiveness of the teaching profession and improving student performance. Given the present situation in Sweden, however, such a policy could be
counterproductive if it is phased in too quickly, as reducing class sizes would increase the demand for teachers at the same time that the supply of teachers is falling. In turn, this might have severe consequences for schools that are less attractive to teachers. In a situation with excess demand for teachers, it is possible that good teachers would leave troubled neighborhoods in order to take positions in more pleasant work environments.

So what then is the best approach? Since the decentralization of schooling authority it has become more difficult for the central government to influence the amount of resources going to the Swedish schools. At the same time we think that decentralization and individual wage setting may have introduced a self-correcting mechanism that did not exist in the old system. Market forces imply that teachers in unattractive areas and unattractive subjects will be offered a more favorable “compensation package” than those in more attractive areas. 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 levels 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 such as 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 that should guide decisions on how to allocate funds, including the considerations about the stage of life in which individual learning capacity is highest, 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 book 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 on which to base credible estimates.\(^{65}\)

We do think that too much money was allocated to adult education at the expense of youth education during the second half of the 1990s. Recent studies such as Ekström (2003) suggest markedly lower returns to Swedish adult education than to other levels of education. Additionally, adult education in the late 1990s was subsidized at an unprecedented rate. Individuals may have been attracted to adult education by the fact that they received student pay at the level of unemployment benefits, which in Sweden replace 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 in Sweden. Prospective university students can improve 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; this policy seems particularly undesirable given that Swedish university students are relatively old in international comparison, there is a shortage of qualified teachers in youth education, and one can enter the university via the Swedish equivalent to the SAT (Högskoleprovet). We think that the incentives for improving grades via adult education programs should be drastically reduced.

Having raised these critical points about adult education, we wish to emphasize two points. First, we do not believe that individual learning capacity declines so quickly that you cannot “teach old dogs new tricks.” Indeed, some recent US evidence that we have referred to suggests 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 simply think that the scale and the rate of subsidization of adult education in Sweden have exceeded the optimal levels during recent years.

\(^{65}\) 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. Selection issues are likely to be a big concern.
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 tests such as those conducted by the Department of Education in Gothenburg. Our empirical analysis in Chapter 7 demonstrated that the national tests 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 because of sampling variability. Allocating resources based on test results may also compromise the informational content of the tests by providing incentives for teaching to the test or teacher cheating.

The Swedish school system does not currently use high-stakes testing. But Swedish schools increasingly use results from national tests for marketing purposes in the wake of the school choice reform. 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 that have arisen in the US under No Child Left Behind – which requires annual testing in grades 3-8 but decentralizes selection of the tests and the standards used to judge proficiency -- 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.

Test score 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. We have presented several analyses in which changes in municipalities’ school policies are related to changes in student performance. Such evaluations are, however, limited by the fact that Sweden’s national tests are currently only conducted in fifth and ninth grades. Further, until recently, data made available for evaluation purposes have only been collected for the 9th grade test, and only for a sample of municipalities. These facts have limited the opportunity to gain valuable information about the effects of Swedish education policies in general and the reforms of the 1990s in particular.
Swedish policymakers have recently decided that 9th grade test results will be collected for all test takers. All schools must participate in the 9th grade test, but participation in the test in the fifth grade is voluntary. Even if participation in the fifth grade test were compulsory, test score data collected at an even earlier age could be beneficial. Presumably, it is easier to correct problems in schools the sooner they are detected. We therefore recommend that students be tested at an earlier age, possibly as early as first grade. Testing students at regular intervals would enable school authorities to monitor individual and school performance more closely.

The regular assessment of Swedish schools and education policy would also benefit from information on student achievement that is comparable over time. The national tests do not have this property. Currently, the only comparable information on compulsory school student achievement over time is provided by the UGU-data, administered by the Department of Education in Gothenburg. A natural task for the Swedish National Board of Education would be to ensure that such studies are conducted regularly for random samples of the student population. It makes sense to have such time-series information not only for 6th graders, but for students at other grade levels as well.

We should also mention a potentially useful lesson for the US based on Sweden’s experience: giving teachers the ability to combine class performance and standardized test results helps to humanize assessment, and increases the teacher’s authority. As things currently stand in the US, testing and teacher assessment are two entirely independent endeavors. The evidence from Sweden shows that teacher assessments add value over and above standardized test scores when it comes to predicting labor market success, so it may make sense to combine the two to generate an overall assessment.

More randomized experiments
Swedish educational policy would benefit if more randomized experiments were conducted to evaluate educational initiatives. By randomly selecting classrooms, or schools, to participate in new initiatives, and selecting others to

66 Researchers at this department have administered one verbal, one spatial and one inductive test to random samples of 6th grade students for a quite long period of time. The tests have been done in 1961, 1966, 1980, 1985, 1990 and 1995; see Härqvist (2000). The next test will not be done until 2005.
serve as a control group that does not participate, the effects of such initiatives can be reliably and definitively measured. Random assignment insures that there are no differences in either observable or unobservable factors between the two groups that may be correlated with student achievement. Thus, any difference in outcomes between the “treated” and “non-treated” groups can be attributed to the education initiative itself. In contrast, in observational studies that simply compare outcomes of those who self-select to participate in an intervention and those who do not, it is unclear whether observed differences in outcomes can be attributed to causal effects of the intervention or simply to unobserved differences between participants and nonparticipants that are not measured by researchers.

The United States has a long tradition of running such experiments and utilizing the results in public policy deliberations. In particular, US job training programs have been evaluated in this way since the 1970s. The US welfare reform of the 1990s was also preceded by several randomized experiments. As we have shown at several places throughout this book, such experiments have also been used to evaluate education initiatives; examples are privately-funded voucher programs, computer-aided instruction, and the Project STAR study of smaller class sizes. Sweden (as well as most European countries) is lagging behind in using this evaluation approach.

Our recommendation is thus that randomized experiments be used as a regular evaluation device within Swedish education policy. One solution would be for the National Board of Education to be provided with resources to do the required field work. It is important 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 to be the most sensible approach to obtaining objective and reliable results.

Nonetheless, 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 addition, individuals who are aware that they are taking part in an experiment may modify their

---

67 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.
normal behavior, particularly if there is something at stake for the subjects involved in the experiment. Because of its lack of experience with randomized evaluations, we believe that Sweden should not start with a large-scale experiment that randomly assigns thousands of students to participate in a particular initiative. It would be better to start with small-scale experiments to gain methodological experience and to introduce the education community to the approach. Alternative teaching methods, such as alternative use of computers in the classroom, or vouchers for summer courses may be reasonable candidates for the first wave of experimental evaluations in Sweden.

_Evaluation and the public policy discussion_

We are convinced that a well-designed strategy for evaluation will be beneficial for Swedish education policy. For instance, a strategy involving more frequent tests will enlighten public policy discussions and will prevent undue criticism. In all countries, critics frequently claim that the school system is in crisis. Reliable information on student achievement is an effective way of monitoring school quality in order to prevent unnecessary school reforms in response to such alarmist claims. Cross-country comparisons may fulfill the same objective.

In a similar vein, we think that reliable experimental information may enlighten the policy discussion about, for example, the effectiveness of different teaching methods. Changes in school policies – be they minor or major – should always be based on credible evidence.

---

68 This phenomenon is referred to as a “Hawthorne effect”; see, e.g., Heckman and Smith (1995).
References


Nye, B., J. Zacharias, 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.