An Evaluation of Selected Reforms to Education and Labor Market Policy in Sweden*

Alan B. Krueger
Princeton University and NBER
and
Mikael Lindahl
Uppsala University and IFAU

Finanspolitiska rådet är en myndighet som har till uppgift att göra en oberoende granskning av regeringens finanspolitik. Rådets uppgifter fullföljs framför allt genom publiceringen av rapporten **Svensk finanspolitik** som lämnas till regeringen en gång per år. Rapporten ska kunna användas som ett underlag bland annat för riksdagens granskning av regeringens politik. Som ett led i uppdraget anordnar rådet även konferenser och utger skrifter om olika aspekter på finanspolitiken. I serien **Studier i finanspolitik** publiceras fördjupade studier eller rapporter som härrör från externa uppdrag.

1 Introduction

In this report we discuss the expected effects of reforms in education and labor market policies already implemented or suggested by the current liberal-conservative government in Sweden. According to our guidelines, the main focus is on whether the reforms are suited for the purpose of increasing employment, hours worked and productivity.

The reforms that we analyze have wide ranging objectives, approaches and timeframes. Some are designed to directly affect employment and productivity in the short run (including active labor market programs (ALMPs) and reforms to high school with the purpose of facilitating the transition to work). Others are likely to have little immediate effect, but possibly long-term effects on productivity, employment and living standards (such as reforms designed to put more emphasis on scholastic curriculum at young ages). And others can have effects on output and employment over varying time horizons (such as an increase in resources for university research). These differences should be borne in mind when we discuss expected effects of various possible reforms.

Our comparative advantage is to analyze likely effects of general policy changes, with relatively little attention to the details. Of course, actual results will depend very much on the details of the way policy changes are implemented. Still, we believe that some general tendencies can be inferred from the relevant evaluation literature. We have divided our report into six separate sections. For education, we consider the economic effects of preschool education, efforts to raise teacher skills, school accountability policies, vocational high school and increased funding of university research. We divide our analysis of the effects labor market policies into reforms of active labor market programs and of the unemployment insurance system.

2 What is the payoff to preschool education?

2.1 Description of the reforms

There are two reforms proposed by the government that will have a direct impact on the amount and focus of pre-schooling in Sweden. First, the availability of pre-schooling will expand to include all children from the age of three. This is an extension of an earlier reform in 2003 which introduced general pre-schooling for four and five year olds. These changes have had the effect that although most children previously experienced pre-schooling at a day care center, it is now available for all children. Since having a place at a day-care center is costly for parents, the reforms made pre-schooling free-of-charge. Second, pre-schooling will become more "school-like" as there will be an increased focus on learning (this is the so called "Förskolelyftet"). This will be accomplished in two ways: 1) Additional resources will be used for continued education and training of pre-school teachers and day-care personnel. The goals are that pre-school teachers should be present in all

preschools and that teachers should be given a more specified pedagogical role as well as more overall responsibility. 2) The objectives of the pre-school curriculum should be more explicit, especially regarding the language and math skills taught to children.

2.2 A survey on the effects of pre-school education

We survey the best available evidence on the effects of pre-compulsory school education on short- and long-term outcomes. Much evidence will be drawn from U.S., since the U.S. traditionally has put much effort into evaluating preschool programs. Due to the lack of Swedish studies of pre-school effects, we are forced to rely on evidence for other countries. This is unfortunate, but is likely a result of the absence of registry information on child care and pre-schooling attendance for Sweden, which makes studies of pre-school effects in Sweden difficult.

We will mainly focus on universally available pre-schooling (i.e., made available for a large and representative part of the population). An example of such a program is Oklahoma's universal pre-K program. We will then compare the results found in these universal programs to small-scale experiments of preschool programs targeting disadvantaged children, such as the Perry Preschool (which also included home visits) and the Carolina Abecedarian (which included a number of interventions from ages 0-5) programs. These evaluations have the advantage that they are based on a randomized design, and hence estimates are particularly credible. The main drawback is that these programs were targeted to children from (very) disadvantaged backgrounds and included other interventions in addition to pre-school education. The external validity to the Swedish setting can be questioned. There are also relevant results to be drawn from a recent study for Netherlands, where compulsory schooling starts at the age of four.

Oklahoma established a universally available pre-kindergarten program for 4 year old children in 1998. It is funded by the state and is free-of-charge. School districts are free to participate and an overwhelming majority of them do so. The program can be run either as a full-day or half–day program. All teachers are required to have a Bachelor's degree and an early childhood certificate, and they are paid at the same rate as elementary and secondary school teachers. The class size maximum is set to 20 and most classrooms have an assistant teacher. The participation rate is quite high; over 60 percent of the 4 year olds in Oklahoma start pre-school. And those children who start preschool are from a more diverse backgrounds than participants in Head Start (discussed below).

In a series of papers, Gormley Jr. and coauthors (Gormley Jr. and Gayer, 2005, Gormley Jr. et al, 2005, Gormley Jr., Philips and Gayer, 2008) have studied the effects of this program in Tulsa, the largest school district of Oklahoma, on early achievement score measures directly after participation in the program. An identical test was administrated to children just starting pre-school (k-1) and those just starting kindergarten (k). Over 1500 pre-k students and over 3000 k-students were tested in 2001. In 2003 and 2006 follow up studies have

been performed for more recent cohorts. The test participation rate was high (above 80%). The children were tested on their reading (identify letters and pronounce words), spelling (tracing letters and spell simple words) and math skills (simple calculations).

The challenge with isolating the causal effect of such a program is: 1) how to separate out the effect of maturity or age from the effect of spending time in the program, and 2) how to control for non-random program participation (selection bias).[1] The research design in these studies utilized the existence of a cutoff date (1 September) for admittance to preschool. Children turning age 4 by this date were allowed to attend preschool whereas children not yet 4 were not. It is therefore possible to control for age, either parametrically or by using sub-samples born close to around this cutoff. The authors compare test scores for children born after the cutoff (with 1 year of program treatment at date t), with test scores for those children born before the cutoff (which has no program treatment at date t). The former is the treatment group and the latter controls, and both are tested at date t. If the control for age is done properly and if the adjacent cohorts have similar average characteristics, one can argue that the effect of the program on the treated is estimated.

The authors find (Gormley et al, 2005, Gormley Jr., Philips and Gayer, 2008) that one year of preschool increase reading by 0.8-1.0 standard deviation (SD), spelling by 0.6-0.7 SD and math by 0.4 SD. The effects are similar across racial-ethnic and socioeconomic groups. Separate effects were not estimated by sex. There is no evidence that effects of the half-day program are smaller than for the full-day program. If these gains are converted to age-equivalent scores, children gained about 8 months in reading, 7 months for spelling and 4 months for math. Given that the preschool lasted 1 year, these gains are very large. The effect sizes in Gormley Jr. and Gayer, 2005, are somewhat smaller, but still impressive.

The research design in these studies is good, but some caveats should be kept in mind. Given their small sample size, they are not able to compare children born very close to one side or the other of the cutoff date (they use samples of ±3 months or a quadratic age fit), so their effect estimates might still be confounded by age/maturity. They are also unable to control for relative age effects. If the youngest individual in preschool gains disproportionally because of their older peers make them mature for their age, effects will be overestimated. Since the tests are taken without any further schooling, we do not know if the very large effects are long lasting. If effects are non-linear and decreasing there will be smaller effects found when these children are older. Also, ideally all children whether or not they participated in the pre-school program) should have been tested on both occasions. Finally, they do not look at catch-up effects: those who did not participate may catch up very quickly when they start school.

---

[1] An unresolved issue in the literature on education effects for young children is how to separate the effects of time in school, school start age and age at test (general maturity) from each other. Since age at test is exactly equal to school start age and time in school, these are not separately identified. However, for the purpose of our question we are interested in the effect of adding time in school for young children, which is the combined effect of a lower school start age and longer time in school. Hence, the important things in the studies we review are how they control for age, as well as how they control for selection bias.

Berlinski et al., 2008, estimate effects of pre-school participation in Argentina during the 1990s, when pre-schooling increased rapidly. They exploit variation in expanding pre-school facilities across regions and cohorts, and find that pre-school participation leads to higher test scores in primary school: one year of pre-schooling increase third-grade test scores by 0.2 S.D.. The effects estimates are similar for boys and girls, and somewhat higher for children from impoverished households. They also found that pre-schooling positively impacted psycho-social behavior as measured by attention, class participation and discipline.

In most countries children born within a year typically starts school once in a year, usually at a date sometime in late August or early September. In the Netherlands, children are instead allowed to start school upon attaining a certain age, the day when they turn four years of age. This is also what most students do. Leuven et al (2007) estimates the impact of time in school, by using the fact that the summer break (6 weeks) introduces exogenous variation in the age at which children are allowed to start school. The authors evaluate effects on scores from tests in reading and math which were given at the same time for all children during second grade. Consider the date of the test, t. At t, those who had their 4th birthday just after the summer holiday are six weeks younger than children who had their 4th birthday at the beginning of the summer holiday. However, they have spent the same amount of time in school. At t, those children who had their 4th birthday six weeks after the end of the summer holiday are six weeks younger and have six weeks more schooling than children who had their 4th birthday just after the summer holiday. Since, the summer holiday creates a discontinuous relationship between time in school and age at test date (perfectly correlated for those children with birthdays falling at a time when schools are in session, and uncorrelated for those with birthdays falling within the time period of the summer holiday) the authors can estimate the effect of (potential) time in school and at the same time control parametrically for age effects. They find that one-month of additional (allowed) schooling generates higher scores for those children from disadvantaged backgrounds, but statistically insignificant effects for other pupils. However, the estimates are fairly imprecise. For non-disadvantaged children they can reject effect sizes of 0.03 SD for an additional month of school. Given that not all children start school as soon as they are allowed to, the estimates for disadvantaged students are probably an underestimate of effect of making enrollment compulsory at age four.

A drawback with these studies is that they only investigate short term effects of the programs. It very well might be the case that the positive effects fade away over time. Hence, we would like to know of any evidence of long-term effects on educational attainment, employment and earnings.
Head Start is the largest preschool program in the U.S. It is a program for disadvantaged 3-5 year old children. It contains not only early childhood education but also other child care services (encouraging parents to get more involved, various social services and improved nutritional intake for instance). Recently, studies have been able to evaluate the long-term effects of this program. Garces, Thomas and Currie (2002) use variation in Head Start participation across siblings, i.e., they compare outcomes for head start

children and their siblings who did not attend. They find long-term effects on educational attainment and reductions in criminal behavior for some groups. One drawback of this study, however, is that Headstart is intended to involve entire families, so there may be selection issues if a sibling does not participate. Ludwig and Miller (2007) avoid this concern by exploiting a discontinuity in the funding of head start across counties. A policy of providing assistance to develop Head Start funding proposals for the 300 poorest counties had a large impact of head start participation in those counties. Hence, they can compare counties with poverty rates just above the cutoff with those just below. They find a large drop in mortality rates of children five to nine years of age, and suggestive evidence for a positive effect on schooling. However, they find no statistically significant effects on math and reading scores in 8th grade, or for absence from school or time spent doing home work.

We can also draw on evidence from a few small-scale experiments, where children have been randomized into whether or not they attain preschool. Such randomization controls for selection bias, since the treated and control children on average will have the same characteristics. The Carolina Abecedarian Program was an intensive program which lasted from age 0-5. The study group consisted of 111 infants, of which 57 were randomly selected to receive child care services that very intensively focused on language development. The infants were healthy but from very disadvantaged backgrounds. They were followed throughout their childhood and youth up till the age of 21. Those from the treated group were at age 21 much more likely to attending college than the controls. Similarly very impressive long-term effects are found in another small-scale experiment, the Perry preschool. The treated group consisted of 123 African-American children born in poverty. There, surprisingly, few beneficial effects are found in the early grades, but at older ages very large positive effects were found for a number of outcomes. At age 40 treated children had higher earnings, more years of schooling, were more likely to be employed and had lower arrest rates (Schweinhart et al, 2005). Note that these experiments are for programs which also included treatments other than just preschooling: The Perry Preschool also included visits to the children' home and the Carolina Abecedarian included a number of interventions. Hence, the effect estimates are likely too high if inference is to be drawn for the impact of pre-schoolings.

A more recent program is "Early Head Start," which extended the Head Start concept to low-income pregnant women and families with infants and toddlers. This program has been evaluated by a randomized research design by Mathematica Policy Research (Love et al., 2002). They found that participating children outperformed their non-participating peers in cognitive, language, and social-emotional development. The score gains on the cognitive and language tests amounts to about of around 0.1-0.2 SD for one year of program attendance. However, the children in the control group receive some treatment as well (some type of childcare). Hence, these estimates are very likely an underestimate of attending an Early Head Start program.

Andersson (2008) analyzed several of the existing small-scale pre-school experiments in an integrated framework. He finds that although girls gained

significantly both in the short- and long-term, there were no statistically significant long-term effects found for boys.

Lastly, we will mention results from some studies which have estimated effects of school-starting age, which have found mixed effects of lowering the school-starting age. On the one hand, Angrist and Krueger, 1992, found positive effects on education and earnings using U.S. data, and Black, Devereux and Salvanes, 2008, similarly found positive effects using Norwegian data. On the other hand, Fredriksson and Öckert, 2006, found negative effects on education and no effects for earnings, for Sweden. So how can we reconcile this evidence of large effects of pre-schooling with small effects of lowering the school starting age (at least in Sweden)? First, the school-starting age literature estimates effects of school-starting age (attempting to control for effects of schooling and age) whereas the pre-school literature attempts to estimate the combined effects of school-starting age and schooling. Second, the school-starting age literature estimate effects of attending school using the same curriculum as for older pupils, whereas the preschool literature estimates effect of programs designed to enhance learning for young children. Third, the school start age literature looks at earnings and other long-term outcomes, whereas most of the pre-school literature looks at short-term effects on test scores.

2.3 Conclusions and policy suggestions

In a survey of pre-school programs, Currie (2001) concludes that the benefits of Head Start would cover the costs if Head Start generates one-fourth of the long-term benefits found for the small-scale experimental programs. Our survey shows that the short-run returns to universal pre-school education are positive and can be very high (as for Tulsa, Oklahoma) and that these returns are positive across socioeconomic and ethnic groups, although it is not clear that they are as large for boys as for girls. It therefore seems very likely that universal pre-school education in Sweden would be good policy and that pre-school investments will have at least as high returns as other human capital investments. For instance, Hansen et al. (2004) find that one more year of high school generates about 0.3 S.D. high scores on military (AFQT) tests. This effect is lower than most estimates of one-year of pre-schooling surveyed above.

We further believe that mandatory schooling for children is preferable to having universally available pre-schooling. One reason is that we worry that children from disadvantaged background are more likely to not participate in pre-school education. Given its increased focus on learning and the high returns found in the literature, a voluntary pre-school system will likely increase inequality in skills at the time when compulsory school starts. Such initial skill differences might then be very difficult to overcome. Another reason is that a first grade with a mixture of children, with and without pre-school experience, will make it very difficult for teachers to adjust their teaching to the right level of difficulty. Either, the children without pre-schooling will find it hard to keep up with the pace or those with pre-school experience will have to wait for the

other children to catch up. For these reasons we recommend mandatory preschool already from the year children turn 5.

3 What are the effects of policies to raise teacher skills?

3.1 Policy reforms

With the so called "lärarlyftet," the government will put forward extra resources to increase the opportunity for teachers to obtain further training and education, including graduate education. In addition, there will be additional education available for teachers that need to complement their education so as to become certified teachers.

3.2 Empirical evidence

How important are teachers? On the one hand, the identity of teachers matters: some are better than others. On the other hand, it has proved difficult to find strong associations between student performance and teacher characteristics such as education and experience. So predicting which teachers will succeed *ex ante* is very difficult.

One reason for the lack of correlation between teacher characteristics and student performance could be that experienced and well educated teachers are assigned to classes with weaker students, leading to underestimated effects.

One approach does not look at teacher characteristics, and instead focuses on the effect on performance of being assigned to one teacher as opposed to another. Specifically, researchers have attempted to estimate the overall effects of teacher quality, utilizing longitudinal data on students, teachers and schools. This approach estimates teacher specific effects, controlling for student, school and year fixed effects, hence capturing also unobservable teacher skills that are important for student achievement. Findings from this literature (Rivkin et al, 2005, and Rockoff, 2004) are that teacher assignments explain a relatively large part of the variation in overall tests core gains: a one standard deviation (S.D.) increase in teacher quality raises tests scores in math and reading by about 0.1 S.D. Note that because of the importance of pupil and family background characteristics in explaining student achievement, the importance of teacher quality as a fraction of overall "school factors" are much higher. Neither study find effects of teacher education being associated with teacher quality. Teacher experience does have some impact, but in Rivkin et al (2005) it fades away after a few years.

The importance of teacher characteristics has also been estimated using experimental data from the Tennessee STAR experiment. In STAR, kindergarten students were randomly assigned to small or large classes, where the large classes consist of three types of classes: a small class, a large class

without a teacher aide and a large class with a teacher aide. Importantly, teachers were also randomly assigned to different class types. In an analysis of the STAR data, Krueger (1999) finds little effects of having a teacher aide and of observable teacher characteristics such as teacher experience and having a master degree. Nye et al (2004) use the STAR data to estimate overall teacher effects, utilizing across-grades within- school and class-type variation, and find effects of similar magnitude as in the non-experimental studies referred to above.

That teaching methods matter and that teaching effectiveness can be improved is illustrated in Machin and McNally (2008), who find very large benefits of the introduction of the literacy hour in U.K. primary schools. The literacy hour, a daily hour of reading and writing, placed a much more rigid and structured form on the English curriculum than what previously was the case. The program also meant instructing teachers how to change their teaching practises with regard to content and pedagogy. Machin and McNally (2008) evaluate the introduction of the literacy hour using a difference-in-differences methodology, where changes in pupil test scores in areas with teachers participating in the program was compared to score changes in areas not affected by the program. They find positive effects of students' performance in reading and understanding of English. The policy was not expensive, and highly cost effective. It is possible that the large estimated effects not only are due to changed teaching methods, but also to some extent reflect more hours of English teaching.

3.3 Policy recommendations

Unsurprisingly, teachers matter a lot for student achievement. However, it has been difficult to pinpoint just what characteristics that make a successful teacher. Teacher education and experience have consistently been found to have no or a small impact on student achievement. An exception is that the first years of teaching appear to be important. These findings make it difficult to draw policy conclusions.

4 What are the benefits of school accountability?

4.1 Policy reforms

The new government will implement reforms with the purpose of putting more emphasis on monitoring pupils' learning development in central subjects in compulsory school. There will be clearly stated knowledge goals in grades 3, 6 and 9 in key subjects, with national tests in these subjects distributed to all pupils in these grades. The purpose is to monitor pupils' learning progress, and to compare it across schools and groups of pupils. Pupils will also be given marks in earlier grades.

We will discuss the advantages and disadvantages of a school system that have clearly stated standards for learning, assessment of learning progress through

national tests and that emphasize accountability for students, teachers and schools. Such a policy of test-based accountability has not been popular among policy makers and educators in Sweden in the past. Note that the proposed changes increase standards for learning and assessment of learning progress, but place little emphasis on accountability of pupils, teachers and schools. The proposed changes discussed above are still significant: before, national tests were used for fewer subjects in later grades, and a decade ago national tests did not exist.

However, before we discuss accountability policies, we will show some evidence on how satisfied Swedes are with their schools.

4.2 Evidence of the satisfaction with the Swedish education system

Before we discuss the pros and cons of introducing standards, assessments and accountability in Swedish schools, we will present some new evidence on how satisfied Swedish citizens are with their schools. Utilizing newly available data from Gallup, we can compare the public's satisfaction with schools in Sweden relative to that in other countries, and how satisfaction relates to performance in tests from the most recent PISA and TIMSS evaluations.

Our tabulations of the percentage of people in OECD countries who express satisfaction with the schools in their area are reported in Table 1. Regarding satisfaction with the educational system, Sweden ranks in the middle of OECD countries. Some 79.5 percent of Swedes say they are satisfied with the education system. Interestingly, this figure is down from 84 percent in 2004, and this decrease is statistically significant (standard error is 1.7).

We correlated the percentage of the population in a country satisfied with the schools in their area with their countries' average test scores. The correlations were not statistically significant, although the correlation was positive and significant for the PISA scores if South Korea was dropped from the sample. We conclude from this analysis that the percentage of people in a country who express satisfaction with their school system is fairly disconnected with objective test-related measures of student performance. Thus, there is a risk of complacency if data on school performance are not provided. Casual impressions and anecdotal evidence that affect the public's satisfaction with the schools in their area are not highly correlated with the measures of performance that are available.

Table 1 Percentage of individuals who are satisfied with the educational system in their area, and average test scores from PISA and TIMSS

Country	Satisfaction with Educational System	Mean Test Scores			
		Reading		Math	
		PISA	PISA	TIMSS (4th grade)	TIMSS (8th)
Singapore	94.7%			599	593
Belgium	86.0	501	520		
Netherlands	85.1	507	531	535	
Finland	84.3	547	548		
Denmark	82.6	494	513	523	
Ireland	82.3	517	501		
New Zealand	81.4	521	522	492	
Austria	80.8	490	505	505	
Norway	80.5	484	490	473	469
Sweden	*79.6*	*507*	*502*	*503*	*491*
France	79.6	488	496		
United Kingdom	79.4	495	495	541	513
Canada	78.3	527	527		
Australia	76.5	513	520	516	496
Mexico	75.2	410	406		
Spain	71.9	461	480		
Italy	64.2	469	462	507	480
Latvia	63.9			537	
Japan	59.9	498	523	568	570
Turkey	57.4	447	424		432
South Korea	51.2	556	547		597

Note: Educational satisfaction percentages are the authors' calculations from the Gallup World Poll for 2007. PISA is from 2006 and TIMSS is from 2007. The average test scores have mean 500 for all participating countries (the mean is slightly higher for OECD countries) and a standard deviation of 100.

4.3 Theoretical arguments for school accountability

The arguments for using a system of test-based accountability are that it will create incentives for students and school personnel to put more effort into learning and to help authorities to identify students and schools that are performing poorly. A potential danger of this system is that teachers and school administrators will adapt and focus on only those outputs that are measured. These are imperfect proxies for achievement and only moderately correlated with the long-term outcomes that are of ultimate interest, such as earnings. For instance, teachers can "teach-to-the-test" (so as to improving the

skills specific to a test, but perhaps ignoring more general skills that are difficult to test for but that are important for a deeper understanding of the subject) and spend more time on the students who are on the margin of obtaining passing grades (who matter most in the evaluations). Schools can put weak students into special classes, perhaps not included in the evaluations. That teachers and school administrators would respond strategically to the introduction of a test-based accountability system also has some support by economic theory. Holmstrom and Milgrom (1991) show that managers (i.e., teachers), performing under an incentive scheme constructed so as to rewarding certain measurable outputs, will respond strategically by focus more effort on easily observable tasks, and less on other factors relevant for overall output.

4.4 Survey of relevant empirical evidence

Jacob (2005) estimates the effects of test-based accountability (i.e., high-stakes testing) by focusing on the introduction of an accountability policy in Chicago in 1996-1997. The policy was designed to raise achievement and involved holding students accountable (by not promoting them to the next grade without demonstrating sufficient skills, which could mean summer-school attendance before retaking exam) as well as holding teachers and schools accountable by putting them on academic probation (with a risk of teachers and principals of eventually losing their jobs) if performance was low. Jacob performs two types of analyses. He uses a panel of student-level data, and compares actual scores to predicted trends from prior scores across groups of students that were or were not affected by the accountability policy, controlling for student, school and neighborhood characteristics. He also uses district-level panel data, and compares achievement changes in districts in Chicago that implemented the policy with achievement changes in similar districts in other cities that did not implement the policy.

Jacob finds both math and reading scores to increase following the introduction of the accountability policy. However, performance did not increase for younger students (where tests did not have consequences for pupils, teaches and schools). There are also results suggesting that achievement gains were driven by test-specific skills. Also, there is evidence that teachers respond strategically to incentives by focusing on high stake subjects, and increase placement of weak students in special classes. In some cases, teachers also responded by cheating, filling in some students' answers (Levitt and Jacob). Jacob concludes that test-based accountability can improve achievement but that undesired responses by teachers might also follow. This means that the design of such accountability systems is of great importance

Hanushek and Raymond (2005) analyze achievement changes across U.S. states that introduced accountability systems at different times during the 1990s. They found that accountability accompanied increases in student achievement, but that it did not close the achievement gap between black and white students. They did not find any unintended consequence on special education placement rates in response to the introduction of accountability systems. Wong (2008), estimates effects on long-term outcomes (also using across state

implementation of accountability policies) and finds positive effects on schooling and earnings for some groups.

In 2002, the "No child left behind" (NCLB) was put in effect in the U.S. This landmark legislation involved a number of federal requirements with the aim of improving the performance of schools by, among other things, increasing the standards of accountability for schools. NCLB requires all public schools to administer a standardized test each year to all students. If a school fails to make enough yearly progress, it is "stigmatized" as a failing school, and (if the school does not improve after a period of time) parents can transfer their children to another school within the district. Initially funding is increased to help failing schools improve and special tutoring for its students must be provided.

Neal and Whitmore-Schanzenbach (2008) evaluate the introduction of NCLB in Chicago in 2002. In the design of this program, rewards and punishments depended on the number of students in certain demographic groups that scored above certain levels. They compare individual changes in test scores, where one group of students took the second test after the implementation of NCLB and the other group took the second test at the same time but before the implementation of NCLB, but both groups took their first test just before the implementation of NCLB. Hence, they can compare test score gains for those going from a system of low- to high-stakes testing (i.e., when test scores did not and did matter for accountability purposes) with gains for those always experiencing a system of low stakes testing. This makes it possible to control for unobserved pupil heterogeneity that are important for achievement levels.

Neal and Whitmore-Schanzenbach find that math and reading scores did increase substantially, but that this was entirely due to scores at the middle of the test score distribution, which was where the tests were high-stakes (i.e., most important for satisfying the accountability policy). Students at the bottom of the distribution, who were a long way from performing satisfactorily on the test, actually exhibited a decline in performance. Their findings are consistent with teachers reallocating their effort levels to those particular groups of pupils which matters for sanctions and rewards, and possibly neglecting others.

One possible way to make teachers and schools perform better is to put forward monetary incentives that reward teachers when their students do well on some measurable output. In Israeli secondary schools, several performance-based pay programs have been implemented and recently evaluated in credible ways. One such program, which was implemented in the mid 1990s, rewarded schools that performed well additional money to be spent or the school's teacher and staff. The empirical results from evaluating this program suggest that such monetary incentives caused significant test score gains and lowering of drop-out rates (Lavy, 2002). In a recent paper Lavy (2008) evaluates individual monetary incentives based on an experiment among teachers in Israel. Teachers were rewarded with bonuses for improving their students' performance in high-school exit exams. Significant improvements in students' performance were found. There is also evidence that the test score outcomes improved through changes in teaching methods and increased teacher

involvement for students' performance. Interestingly, no evidence of manipulation of test scores by teachers was found.

Björklund, et al. (2005; Chapter 7) provide a discussion of the potential value and limitations of using national standardized tests to evaluate Swedish schools. They conclude that Swedish national tests do measure skills and knowledge that are highly correlated with future earnings and educational attainment, even after controlling for grades. Thus, they conclude, "test scores may be useful for evaluating school quality – both for parents choosing between schools for their children as well as for policymakers assessing school performance." They also highlight the risks involved in using tests to draw inferences about teacher performance (e.g., small sample sizes and potential for teaching toward the test). They conclude, "Any accountability system based on standardized test scores, however, must be implemented with much care to achieve success and avoid the potential pitfalls we have discussed."

4.5 Summary and policy recommendations

We conclude that there are benefits to clearly formulated standards and assessment of pupils' learning progress. However, it is clearly both up- and down-sides to making teachers and schools accountable for pupils' performance on tests. There are benefits since teachers respond to incentives (as evidences by Lavy's work for Israel) but also drawbacks since teachers are likely to respond strategically as to maximize results on those outcomes that matter for evaluations. An optimally structured accountability system will maximize students learning progress and minimize strategic responses by teachers. How this should be done is not clear, but what is needed is likely a system that takes into account performance on many different types of tasks and performance at different parts of the distribution. We believe that teacher involvement is key for this to work, and therefore recommend that teacher representatives should be involved in the design of an accountability system, regarding how to test and evaluate learning development.

We also believe that in the Swedish school system, with decentralized financing and both publicly and privately run schools, it is especially important to have a system where learning progress is monitored and compared across schools and groups of pupils.

5 What are the effects of vocational high school education?

5.1 Policy reforms

In the government's latest budget, resources are proposed to be used for more apprentice slots in high school, to create a vocational college and for more vocational slots in the adult education system. There are also plans to shift to a high school system that would contain three tracks: academic (university

preparing), vocational and apprentice high-school programs. All of these are still three-year programs and it is stated that all these high school programs should be organized in a way that it is still possible to take enough academic courses to qualify for university studies. However, they are not compulsory, and in reality, the division will likely be decisive. Currently, all high school programs have a uniform construction. They all last three years, have a common set of core-subjects and can all be used to qualify for university studies. Even the vocational programs are quite academically oriented.[2]

Some important questions are: Will vocational high school tracks facilitate school-to-work transitions for students less likely to continue to university? What will be the effect on youth unemployment? What are the effects on the high school drop-out rates?

5.2 Theoretical arguments

An argument in favor of vocational programs is that they will facilitate acquiring occupational-specific skills and therefore probably increase school-to-work transitions. An opposing argument is that acquiring general academic skills make people "learn how to learn." This is deemed especially important in a fast-changing economy, where existing occupational skills rapidly become obsolete. There is also a related literature on student tracking, i.e., the sorting of students into different levels based on students' ability. Tracking at earlier levels than high school has been found to increase inequality in skills, without increasing average skills (Hanushek and Woessmann, 2006). It is unclear, though, how related these results are for choices made just before highs school.

5.3 A short survey of the relevant literature

Several papers have estimated the effects of curricula with different degree of vocational courses on labor market outcomes in the U.S. Typically, outcome variables (wages, employment) measured a few years after high school are regressed on type of high-school courses (academic or vocational course indicators) and a large set of control variables (i.e., family background and high school achievement measures) for a sample of individuals who choose not to attend post high-school education.

Kang and Bishop (1989) use data for the U.S. (the High School and Beyond data), where students were surveyed and tested in the last year of high school in 1980. They were followed up two years later collecting information on labor market outcomes: employment (number of months worked), wages and earnings in the short-run. They restrict the data to those who never attended college. They find that taking more vocational (relative to academic) courses is associated with higher hourly wages, yearly earnings and months of work. Effects, if they are causal, are very large: if a typical individual take four

---

[2] However, note that almost one-fourth actually choose programs outside the regular university qualifying programs. These are the so called "individual program."

vocational courses, instead of four academic courses, this increases hourly wage rate with 5% for women and 7% for men, and months employed by 22% for women and 8% for men.

Mane (1999) uses data sets for three cohorts of U.S. high-school graduates, those leaving high school in 1972, 1980 and 1992. The samples are restricted to those individuals with at most 6 months of college within the two years following graduation. Labor-market outcomes are employment, wages and earnings in the short-run (within 2 years of high school graduation) for all cohorts and in the medium run (up to 7 years) for the first two cohorts. Those who took more vocational courses in high school are found to have higher labor market payoff than those that attended more academic courses (true for all except for males in 1972 where payoff is similar). These vocational educated high school students are more likely to be employed and earn a higher wage both immediate after high school as well as in the medium term (up to 7 year later). When the relative returns to academic and vocational course work are compared over time, there is evidence of earnings returns to vocational courses to having increased between 1972 and 1980 and remained high in 1992. These results contradict the hypothesis that firms in later years rewarded general education more highly, and instead point toward occupation specific skills necessary to do the job being more desirable from the employers standpoint of employers. Hence, for high school graduates, vocational high school tracks are still (and even to a higher degree) more beneficial than basic academic courses in terms of earnings and employment.

Light (1999) uses data from the NLSY on male high-school graduates in the early 1980s. She finds that vocational courses have a positive effect on wages measured 6 years after high school. The estimated effects are smaller than in Kane and Bishop (1989) and Mane (1999). This might be because she controls for post high-school work experience (which likely is positively correlated with the number of vocational courses taken for this sample).

These three studies focus on high school students with either no or very limited college experience. So they limit the sample used in the estimations based on an outcome variable. Hence, there are potential sample-selection issues that can potentially bias the estimates (to be too high) since those academic high school students not continuing to college will be those that have the least talent for college, but (perhaps) mistakenly chooses academic high school courses. They also control for test scores in the last year of high school, which probably removes part of the academic-course effect. Also, labor market outcomes measured just a few years after high school graduation are clearly imperfect measures of long-run labor-market outcomes. Lastly, even though the authors control for a rich list of explanatory variables, there are still concerns that relevant factors are omitted from the estimations that confound the vocational effect. Whether these issues lead to small or large biases, or even what is the net sign of these biases, is very difficult to evaluate.

The Career academies have existed for long in the US, but have not been very widely implemented. Their curriculum combines academic and vocational courses. The academies also establish partnership with local employers. The

aim is to prepare students for both work and college, and it is aimed at both low- and high performing students.

The Career Academies have been evaluated with a random assignment research design by MDRC, a non-profit research organization in New York City. About 1700 students who applied for enrollment were randomly assigned to participate or not participate in Career Academies (Kemple and Scott-Clayton, 2004), where non-participation means attending a regular high-school education (including mainly academic courses). The evaluators found a decreased probability of dropping out of school but no impact on achievement test scores for the average student. For those with the highest drop-out risks, treatment in the program decreased drop-out rates, increased attendance and positively affected the probability of graduating on time. In a later follow-up for 1400 of these students (Kemple and Willner, 2008), positive effects on earnings were found, but no effect on post-secondary education was found.

Important evidence regarding likely effects of the proposed high school reform is available from a recent study by Caroline Hall at IFAU, Uppsala, for Sweden (Hall, 2008). Since this study is still in progress, results are somewhat preliminary. She evaluates a reform in 1991 that increased the academic content in all high school tracks (through a higher share of theoretical subjects and through more theoretically advanced vocational relevant subjects), lengthening of vocational tracks from 2 to 3 years, and that made students in all high school programs eligible for university studies. This system is currently in effect in Sweden. Since, the change proposed by the incumbent government partly is a return to the old more vocationally oriented pre-1991 high school system (an important difference is that the vocational tracks then were 2 years) results from this study will be very relevant for the proposed reform. In the old system, some vocational programs were also closely connected to workplaces, so as to be more apprentice-like.

Hall exploits that the three-year more academically oriented high school system that were introduced nationally in Sweden in 1991, were preceded by a six-year pilot period in parts of the country. Hence, she can exploit variation in exposure to the new high school system across regions and cohorts. This makes it possible to compare changes in average outcome variables over time in municipalities with a high fraction of students and a low fraction of students in the new high-school system. She studies effects on high-school drop-out rates, educational attainment and earnings.

Findings are that the introduction of the three-year more academically oriented high school increased the probability of dropping out of upper secondary school and decreased the probability of beginning university studies. It had no statistically significant effect on the probability of completing a university degree, having positive wage earnings or on wage earnings up to 16 years after high school attendance. The findings of no or negative effects of a more academic oriented high school are in line with results for US discussed above.

There is also some work on apprenticeship training for Germany and Austria, where such systems have been effect for long. Krueger and Pischke (1995) use data for East and West Germany from 1988. They estimate Mincer-equations,

except that the variable "years of schooling" is replaced by four types of post-compulsory schooling variables, including completed apprenticeship training. They find that the return to apprenticeship training is 19% in West Germany, and slightly lower in East Germany. This amounts to about 5-6 percent per year, which is fairly close to the payoff for a year of schooling in West Germany during this time.

Ferstener, Piscke and Winter-Ebner (2008) estimate returns to apprenticeship training in Austria. They argue that selection problems might bias estimates from simple OLS estimations, since firms select apprenticeship participants. They therefore focus on an estimation sample consisting only of small firms that close down. This means that variation in the length of apprentice periods among participants arguably will be created due to factors uncorrelated with the characteristics of the individual participants. They find that returns to apprenticeship training that is about 2.5-4 percent per year, only slightly lower than the OLS estimate for the same sub-sample. If those participants whose training was interrupted early acquire other labor-market training or experience during the time where they otherwise would have had more apprentice training, this is an underestimate of the effect for this subsample. An important issue is whether results for this subsample can be generalized to all apprentice trainees. The OLS estimates for this group seems in line with earlier apprentice estimates for Austria, suggesting that the estimates can be generalized. However, working in failed firms means that the participant's training is suddenly interrupted, which probably is associated with atypical problem which might make estimates difficult to generalize and likely understate the benefit of training.

Parey (2008) estimate effects of apprenticeship training on unemployment and wages (at age 23-26) for German individuals. An advantage in using data for Germany is that the early tracking system produces a clear counterfactual in full-time vocational high schooling, whereas college is a very unlikely option. He uses variation in the availability of apprenticeship places (across time and space), stemming from fluctuations in firm demand for apprenticeships, as an instrument for apprenticeship training. He finds no effect on wages from participating in apprenticeship training (relative to attending a vocational high school). However, those in apprenticeship training are less likely to becoming unemployed.

5.4 Conclusions and policy recommendations

Overall, there is much evidence that at least some vocational tracks in high school education have positive effects on short- and medium-run outcomes. Preliminary findings in a study for Sweden point to similar conclusions. The most convincing evidence of beneficial effects of a mixture of vocational and academic high-school courses comes from evaluations of Career Academies in the US, which has been evaluated with a random assignment research design. Hence, we conclude that there are at least short-run benefits from offering at least partly vocational high school tracks.

6 What is the return to public investments in university research?

6.1 Proposed reforms

A lot of additional resources will be used for university research funding: about 15 billion SEK, the next four years. This is a large increase, which amounts to about 8% of the total educational expenditures by the government over the same period. The resources will be distributed in the following ways. First, there will be a general resource increase. Second, there will be more resources distributed through scientific councils and for research infrastructure, etc. Third, there will be a particular focus on some "strategic areas", as defined by the government. These are medicine, technology and climate. Recently, the government has specified which subfields (a total of 24, where Energy, Molecular Bioscience and Transport research will receive the most) within these areas that will receive funding. There will also be increased emphasis on the research quality (through citations/publications and the ability to receive external research grants) of universities as a means of distributing resources. There are also plans ("Självständiga lärosäten," SOU 2008:104) to make universities more autonomous. Universities will no longer be government agencies. They will still be overwhelmingly financed by public resources (although they can receive additional funding through donations for instance) but they will be more freely governed.

We will first briefly review some theoretical arguments for why research is important for economic growth, with a special emphasize on the arguments for why research should be publicly funded and carried out at universities. We will further survey evidence from empirical studies attempting to estimate the effects of university resources on innovations. Much of the empirical evidence is for the U.S., a country with a very different system for higher education than Sweden (in particular, Sweden has less private funding, less autonomous universities, and no student fees). The question is how a Swedish system for higher education can best attain high quality research with more resources. Is it through more incentive-based distribution of resources based on quality indicators, and through more autonomous universities? Also, note that Sweden already, in an international perspective, spends a lot of resources (as a fraction of GDP) on research. So, what is the payoff to additional investments?[3]

6.2 Theoretical background

Innovations are central to economic growth in new endogenous growth theories, the so called "innovation-based" growth models. The key argument in these types of models is that overall productivity increases because of innovations that lead to development of new (and improved) products (Romer, 1990, Grossman and Helpman, 1991, and Aghion and Howitt, 1992). In these

---

[3] The reader should bear in mind that both authors are employed by universities, so we acknowledge the obvious potential for a conflict of interest.

theories it is not the addition of physical and human capital that leads to sustainable growth, but instead the level of human capital that, through innovations, raises productivity. Since the number of innovations is central to growth, the factor of production of most importance is naturally R&D.

In recent work by Aghion and others (summarized in Aghion and Howitt, 2006), theory has been modified to allow for heterogenous effects of education. Different levels of human capital play different roles at different stages of development. For countries close to the technological frontier, resources to research and higher education have the highest payoff, whereas for countries far away from the technological frontier, investments in primary and secondary education will be wiser. The reasons are that innovations are growth driving for high developed countries, whereas imitation and adoption of existing technologies are more important for low developed countries. Research is a key factor for the former, but not for the latter, types of countries. Hence, there is strong theoretical support for (university) research to impact economic growth, at least for more developed countries.

Researchers at universities have less incentive to conceal their research findings as opposed to researchers working in private companies. But this does not mean that universities necessarily should be publicly funded. However, in addition to the importance of published research (which is available to everyone regardless of how far away they are from the university), universities are perhaps also important sources of knowledge spillovers through informal interactions. And there are obvious complementarities between having researchers who are on the knowledge frontier teaching the top graduate students and next generation of researchers. Whether or not spillover effects are large enough so as to motivate public funding of research is mainly an empirical question.

6.3 A short survey on the effects of university research on innovations

We next review the literature that has estimated effects of university research on innovations. What are the returns to university research spending, and are they large enough to motivate public investments in research carried out at universities? What is typically estimated is a so called "knowledge production function," where some measure of the number of innovations in a region is regressed on the amount spent on university research in the same region, controlling for industry R&D and some exogenous variables.

Jaffe (1989) looked at spillovers from university research by estimating the impact of university resources on corporate patents, controlling for industry R&D and population size, across U.S. States. Looking separately for different research areas, positive effects are found, especially for medicine and electronics. The estimated elasticities are between 0.04 and 0.28, the latter for drugs & medical research and the former for mechanical arts. The average elasticity is about 0.10, which means that a 10 percent increase in university research resources is associated with 1 percent more patents, on average. Note that these estimates are conditional on industry R&D. If university research

leads to more industry R&D spending (and he finds a positive correlation for most "technical" areas), the elasticity capturing the total effect of university research is much larger: about 0.60 on average. These effects are found to be even stronger in Acs, Audretsch and Feldman (1992), which, instead of the number of patents, used a direct measure of innovative activity (the number of innovations recorded by the U.S. small business administration from leading industry journals).

Jaffe, Trajtenberg and Henderson (1993), investigate if knowledge spillovers to a high degree stay within the region where new knowledge was created. If this is the case, it means that a region cannot rely on investments in research in other regions, but instead needs to invest resources locally. This conclusion is supported by their analysis. Looking at U.S. data on patent citation across SMSAs and states, they do find that patent citations are more likely to come from regionally developed patents. Hence, new knowledge is more likely to stay in the region where that knowledge was originally developed. Geographic concentration seems to matter, which probably applies to universities as well.

Anselin, Varga and Acs (1997) extend the data set employed in Jaffe (1989) to more states and also use a lower level of aggregation (MSAs or counties). They also use the same proxy for innovations as in Acs, Audretsch and Feldman (1992), and estimate a similar model to the one in Jaffe (1989). Using state level data, they find an effect of university research that is at least as large as in Jaffe (1989). Using county-level data, they find that university research partly spills over outside the county where the university is located.

In a recent paper for Sweden, Andersson and Ejermo (2005) use data on the number of patents in 130 Swedish corporations in 1993-1994. The number of patents per corporation is estimated as a function of the number of own researchers and of proximity to other firms research and to university research. They find that access to university research does have a positive impact on the number of patents, and that proximity to a university is important. Hence, these results are in line with the results for the U.S., although it is difficult to compare the size of the estimates in this study to the ones found for U.S..

We note that there are many estimation challenges in this literature that are difficult to overcome. These are discussed extensively in Griliches (1979), and relate to issues of measurement of innovations, which level of aggregation to use (to adequately allow for spillover effects), and endogeneity problems such as omission of relevant unobserved variables related to compositional changes due to mobility of firms and researchers. Overall, our impression is that research in this area has put considerably effort into exploring the effect of various units of measurement and what proxy for inventive activity to use, but has made much less progress on solving possible endogeneity issues. In many studies some form of instrumental variable analysis is used to isolate the causal effect from university research on innovations. However, the instruments are usually of questionable validity.

6.4 Conclusions and policy suggestions

We conclude that investments in university research can have very high payoffs. This is based on theoretical arguments as well as evidence of large spillovers in several papers for the U.S. However, there remain concerns as to whether the estimates should be interpreted as representing the causal effect of university research given the econometric challenges. Also, even if we could make a causal interpretation, it is uncertain whether results for U.S. data can be generalized to Sweden. The theoretical literature, however, gives strong reasons to presume that R&D, spillovers from university research, and technological advances are critical for growth.

7 Active labor market policies

According to the government, the overall goals for labor market policies are to contribute to a well functioning labor market through improved matching of job seekers and vacant positions, permanently increasing employment in the long run, and for the labor market insurance programs to facilitate workers transitions between jobs (Arbetsmarknad and arbetsliv in "Budget 2009 - för jobben, framtiden och välfärden," Chapter 14, p.24). A priority is for resources to be focused on job matching and on those with weakest labor market attachment.

The active labor market programs (ALMPs) can be subdivided into three main categories: training, subsidized employment and work practice programs. During the economic crisis in the early 1990s, first training programs and then subsidized employment programs and work practice programs expanded rapidly. Subsidized employment programs reached very high levels, whereas practice and training programs decreased in volume. When the labor market improved in the late 1990s, participation in all three programs declined, and in the early 2000s, the three types of programs were approximately of equal size.[4] The last few years has seen an increase in subsidized employment and a decline in training and work practice programs.

7.1 Description of the reforms

In the latest budget, there is evidently an increased focus on ALMPs for the long-term unemployed and youths at risk of becoming long-term unemployed. Few activities are proposed for the initial 300 days of unemployment (90 days for youths of 16-24 years of age and 450 days if the individual is not receiving unemployment benefit) of unemployment After this initial period, the ALMPs commence, and consist of three sequential steps: 1) job coaching, providing intensive assistance and advice for searching for a job, 2) work-related training and work practice, and 3) as a last resort, after 450 days, job in some work that is deemed important for society.

---

[4] For a fuller description of the history of the ALMPs in Sweden see Forslund and Krueger (2008).

In the old system, the unemployed could participate in the ALMP programs longer than 450 days (by re-qualifying for participation). Whether or not an unemployed was to enter a program depended on assessments made by labor market employment officers. In the new system, program qualification will only depend on the length of the unemployment spell. The target and design of employment subsidies will change, as well as increase in relative importance among ALMPs. They will be given to a broader group of unemployed (also including those not in labor force such as those on long-term sick leave) that have not worked in one year. The subsidy is somewhat smaller than before but can be provided for a longer period. Longer unemployment spells mean that the subsidy lasts for a longer time.

In light of these policy changes there are a number of questions that are of interest. Is it better to put all resources into programs for the long-term unemployed or should services be provided earlier? What are the most efficient ALMPs, and does this change depending on the economic climate? What is the effect of employment subsidies for increasing the transition to regular employment? Should the unemployed be profiled (e.g., based on their education, training, work experience and other background characteristics), and the services be provided to individuals based on their predicted needs early on in a spell of unemployment.

7.2 Relevant evidence

Effect of employment subsidies

The purpose of employment subsidies is to facilitate the transition to regular jobs for the long-term unemployed. Employment subsidies have been found to have a positive effect on employment probabilities in a number of studies (see Kluve, 2006; Katz,1998).

The efficiency of employment subsidies depends, in part, on the reasons why the job-finding rate among the long-term unemployed is so low. If it is because the productivity of the unemployed temporarily decreases (e.g., because of skill obsolescence), employment subsidies are a reasonable policy measure. During the period of the subsidy, the workers can regain their working capacity lost during the unemployment spell. If, on the other hand, their productivity is permanently lower, an employment subsidy is unlikely to permanently solve the problem. Forslund, Johansson and Lindqvist (2004) find that a large part of the positive effect of employment subsidies derives from a large flow to (non-subsidized) employment at the point in time at which the subsidy expires. This is consistent with the hypotheses that employers do not hire long-term unemployed because of a temporary decrease in their productivity, which speaks in favor of employment subsidies, or because employers are uncertain of the productivity of the long-term unemployed. On the other hand, Calmfors, Forslund and Hemström (2004), conclude that employment subsidies cause more displacement of regular employment, compared to labor market training.

An important change regarding employment subsidies is that, instead of targeting only the long-term unemployed, the "new-start jobs" program targets a larger group of long-term non-employed. It is therefore unclear how much one can generalize results referred to above to the new system.

Job Search Assistance and Job Coaching

One of the more robust findings in the literature is that job search assistance, a short training course to help the unemployed search for work, is cost effective (see for example Meyer, 1995, and Blundell, et al. 2004). Job search assistance in these programs is typically provided when workers first become unemployed or reenter the labor force from welfare. Job search assistance has been found to be effective for diverse populations (Meyer, 1995). We suspect that this conclusion carries over to job coaching in Sweden given the universality of the evidence regarding job search assistance. Moreover, job search assistance is generally found to pass benefit-cost tests, although the cost of job search assistance is relatively low so the program should not be expected to yield huge benefits. Nevertheless, we would recommend that job coaching be provided early on during a spell of unemployment in Sweden.

Profiling

If job coaching is to be provided earlier, without costs increasing, there is a need to find a cheap way of identifying those most helped by job search assistance. One such way is to use statistical evaluation rules, profiling, to assign coaching to unemployed individuals that benefit the most from such policies.

Black et al. (2003) estimate effects of the "Worker Profiling and Reemployment Services" in Kentucky. This program assign mandatory reemployment services such as job training and job search workshops early in the unemployment spell based on predicted lengths of UI spells. The predictions are based on past employment and UI benefit receipt for the individual and on local labor market conditions. Note that variables such as sex, age, race and ethnicity are prohibited from being used. UI claimants are assigned integer scores based on their predicted UI-spell. Because of program capacity constraints, random assignment is used to allocate the services for those with the same integer score. The treated group had to participate in reemployment services as a condition for conditioned UI receipt. The control group was exempt from this requirement. The program is found to increase mean earnings, as well as reducing the mean duration and amount of UI benefits received. The effects of the program are mainly due to a sharp increase in number of unemployment exits early in the unemployment spell. The program is very cheap, and hence very cost effective. Black et al. (2007)

use the discontinuities in the assignment of this program to analyze effects for a broader population during the same time period. Their analysis gave similar results as found with the experimental design. Since the Kentucky economy was strong during the period analyzed (mid 1990s), it is unclear how this study generalizes to a period with low demand for labor.

Lechner and Smith (2006) investigate caseworker allocation of unemployed individuals to subprograms within Swiss active labor market policy in 1998. They compare the caseworker allocation to alternatives including random assignment to services and allocation via statistical treatment rules based on observable participant characteristics. They find that Swiss caseworkers obtain roughly the same employment rate one year after the program initiation as would result from random allocation o unemployed to services. However, statistical evaluation rules do substantially better.

Thus, evaluating workers early in a spell of unemployment to determine the services that would help them return to work most rapidly has proved effective.

7.3 Conclusions and policy recommendations

Is it better to focus ALMPs on the long-term unemployed or should more be done for those at the start of an unemployment spell? The policies discussed above should be evaluated in light of the high number of layoffs recently, which, by some, are predicted to result in an almost 11% unemployment rate in Sweden in 2010 ("The Swedish Economy, March 2009," National Institute of Economic Research). Calmfors, Forslund and Hemström (2004) conclude that using ALMPs as such a large scale in the 1990s was inefficient employment policy. Forslund and Krueger (2008) conclude that the ALMPs of the 1990s were not well adapted to combat the unemployment that arose in the first years of the 1990s. The question is then how to avoid the same mistakes now.

Those becoming unemployed have different re-employment probabilities and different needs. Starting job coaching early in the unemployment spell will not help everyone. However, it will decrease the duration of unemployment for those unemployed who in a better economic climate would find a new job quickly, but now have trouble doing so. Job coaching would prevent them from losing job skills and from perhaps eventually leaving the labor force.

To make extended job coaching and other services more cost effective, we also recommend considering worker profiling to help identify those groups of individuals with characteristics that are correlated with the highest re-employment rate and most likely to benefit from job coaching, and those with the lowest rates who are most in need of employment subsidies or retraining.

8 Unemployment insurance

Unemployment is rising in Sweden and most countries. The worldwide recession makes it critical to ensure that Unemployment Insurance (UI) is achieving its goals with minimal undesirable side effects.

Sweden has made important adjustments to its UI system. Most importantly, benefit generosity now varies over the course of a spell of unemployment. Starting in last year, during the first 200 days of unemployment the benefit level is 80% of prior earnings, during days 201-300 it falls to 70%. After 300 days, workers can qualify for the program "job and development guarantee" at a rate of 65%.[5] Previously, the benefit equaled 80% of prior earnings during the first 300 days of unemployment, and after that point workers could qualify for an additional period of benefits at this rate by participating in a labor market program. The benefit ceiling decreased somewhat for the first 100 days of unemployment: from 730 to 680 SEK per day. For longer unemployed periods than 100 days, the benefit ceiling has not changed: it is still 680 SEK per day.

Another important change is that after 300 days the unemployed are provided with job coaching, which provides intensive assistance and advice for searching for a job. If they still have not found a job after job coaching, they are provided work-related training and work practice programs. For youths age 16-24, the job coaching starts after 90 days of unemployment.

8.1 Goals and tradeoffs

Unemployment Insurance has three main goals: 1) to lessen the hardship experienced by workers during a spell of unemployment; 2) to enable the unemployed to search for an appropriate job instead of accept a job that is a poor match for their skills out of desperation; 3) to help stabilize the macro economy by automatically causing spending to rise during a downturn. Economists and policymakers have long recognized that the achievement of these goals may be accompanied by undesirable consequences. As with all insurance programs, tradeoffs are involved in UI.

Providing UI benefits increases the opportunity cost of work, leading some unemployed workers to delay their return to work.[6] This effect is commonly known as a "moral hazard" effect. A second unintended consequence of UI is that some people may fail to adequately save resources on their own to cushion periods of unemployment if generous UI benefits are provided. Finally, if employers in an industry do not bear the full cost of paying UI benefits for their workers (i.e., if there are cross-industry subsidies in the financing of the program), then employment across industries is expected to be distorted, resulting in high-layoff sectors being larger than they otherwise would be absent UI. The moral hazard, savings disincentive and sectoral distortion

---

5 For those with children under age 18, UI benefits are available for another 150 days (day 301-450) at a level of 70% of past earnings.

6 See Atkinson and Micklewright (1991) and Krueger and Meyer (2002) for surveys.

caused by UI, however, are not necessarily signs of failure. They simply mean that the unintended consequences of UI must be weighed against the desired effects of the program, and an appropriate balance struck. It should also be noted that social norms in many communities can encourage the unemployed to search for work more intensively than they otherwise would, and mitigate the moral hazard effect.[7]

Economists have developed theoretical models to guide the tradeoffs involved in setting optimal UI benefits. Raj Chetty, for example, argues that it is desirable from society's perspective to provide job seekers who have inadequate savings sufficiently generous UI benefits to enable them to stay out of work longer and search for an appropriate job, and he balances this objective against the moral hazard effect.[8] In his model, longer spells of unemployment resulting from UI benefits are not necessarily undesirable if they enable workers to find jobs that fully utilize their skills. Thus, longer unemployment spells are not always an unintended consequence of UI. Recognizing the competing effects of UI benefits, Chetty provides evidence for the U.S. suggesting that the optimal benefit level exceeds 50 percent of a worker's pre-layoff wage. Benefits in the U.S. typically replace around 50 to 60 percent of previous earnings, although the existence of maximum benefit levels in every state means that replacement rates in practice are lower, closer to 35 percent on average.

A particularly relevant strand of the theoretical literature for Sweden's recent reforms considers the optimal timing of UI benefit generosity in relation to unemployment duration (i.e., should benefits fall or rise over a spell of unemployment?) under alternative assumptions. Several studies, which make the assumption that an unemployed worker's consumption is exactly equal to his UI benefit because of borrowing and savings constraints, conclude that UI benefits should fall over the course of a spell of unemployment, consistent with the schedule recently put in place in Sweden.[9] The intuition for this result is that a declining sequence of benefits encourages workers to find a job sooner rather than later, and therefore offsets the moral hazard effect (while at the same time reducing the insurance benefit of UI). Werning (2002), however, extends the model to allow for borrowing and saving and concludes that the optimal profile of UI benefits should rise very slightly over time, in contrast to the previous work.[10] Werning nonetheless concludes that with the optimal benefit profile, consumption should fall the longer that someone remains unemployed, similar to the earlier models. By allowing saving and borrowing, Werner breaks the automatic relationship between UI benefits and consumption. Because people who can borrow or draw down on their assets face a decline in wealth the longer they stay unemployed, their consumption falls even if their benefit rises slightly over time in his model.

---

7 For a model of social norms in income-maintenance insurance programs, see Lindbeck and Persson (2006). Although they model work absences, their insights can be applied to job search intensity an UI as well.

8 See Chetty (2008).

9 Examples are Shavell and Weiss (1979) and Hopenhayn and Nicolini (1997).

10 http://econ-www.mit.edu/files/1267

Despite the sensitivity of the optimal benefit profile over time to alternative modeling assumptions, we think that Sweden's decision to institute a declining benefit profile over the course of an unemployment spell is sensible.[11] The reason for this conclusion is that savings are low in Sweden compared with the EU, and UI benefits probably represent the lion share of consumption for most unemployed individuals.

Finally, it is worth noting that before the adjustments to the UI schedule were implemented, Sweden was tied for the third highest after-tax benefit replacement rate for unemployed workers during the initial period of unemployment out of 26 OECD countries.[12] And Sweden was in 6th place out of 26 OECD countries for the net replacement rate over the first 5 years of a spell of unemployment. Thus, reducing benefits after a spell of unemployment lasted for a while is one way of bringing UI benefit generosity closer in line to that in other countries.

8.2 Relevant evidence: a brief summary

Benefit generosity

A larger international literature finds that more generous benefits are associated with longer spells of unemployment. The evidence for Sweden is also consistent with this literature. Carling, Holmlund and Vejsiu (2001), for example, studied a short-lived change in the replacement rate from 80 % to 75 % that was introduced on January 1, 1996. This reform only affected a sub-sample of the unemployed (those unaffected by the benefit ceiling). The authors estimated the effect of the reform by comparing job-finding rates before and after the reform of those affected and unaffected. Those unaffected were high earners who were constrained by the maximum benefit, so their replacement rate was unaffected. The estimates suggest that the cut in the replacement rate resulted in an increase in the job-finding rate by about 10 %.

Benefit duration

One of the more robust findings in the literature is that the job finding rate rises when benefits are close to expiring (Katz and Meyer, 1990a, 1990b, and Card, Chetty and Weber, 2007). Recent work also finds that job search effort rises in the U.S. when benefits are close to expiring for those who are eligible for UI (Krueger and Mueller, 2008).

A recent study for Sweden is also relevant. Bennmarker, Carling and Holmlund (2007) investigated the effects of several changes in the

---

11 Fredriksson and Holmlund (2003) similarly conclude, "the case for a declining time profile is reasonably well developed."

12 http://www.oecd.org/dataoecd/28/9/36965805.pdf

unemployment insurance system in 2001 and 2002. In 2001, a two-tiered benefit system was introduced that paid higher compensation during the first 20 weeks of an unemployment spell. In 2002, benefits were raised both for spells exceeding 20 weeks and for the first 20 weeks of spells. The changes applied only to sub-samples of the unemployed, namely those who previously had high earnings. This reform was used to identify the effects of changes in benefit generosity of time on job-finding rates. Somewhat surprisingly, there was a striking difference between the results for males and females. The female job-finding rate increased whereas the male job-finding rate (in accordance with theoretical expectations) fell. There was no significant effect on the aggregate job-finding rate. The authors speculate that the difference between males and females were due to a reform in child-care that took place simultaneously. The expected effect of the latter reform was to increase female labor supply.

An accumulation of evidence suggests that the opportunity to renew UI benefit eligibility through participation in labor market programs in Sweden in the 1990s was a disincentive to work, where individuals cycled between unemployment and program participation, with adverse results for program participation.[13] For example, Ackum Agell, Björklund and Harkman (1995) showed that long spells in the registers of the National Labour Market Board (including both UI and program periods) were common in the early 1990s and Carling, Edin, Harkman and Holmlund (1996) show that UI compensated unemployment spells close to benefit exhaustion were significantly more likely to lead to program participation than were uncompensated unemployment spells. Sianesi (2004) found that program participation increased the probability of future benefit-compensated unemployment as well as subsequent program participation. These effects were both of a non-trivial size and long-lasting.

The psychological and material well-being of the unemployed

Numerous studies have found that unemployment has a lasting effect on people's psychological well-being, leading to lower life satisfaction and even depression (Björklund, 1994, and Clark and Oswald, 1994). A longitudinal study of German workers, for example, found that individuals' reported life satisfaction dropped in the year they became unemployed and stayed low for at least three years afterwards, even after they found new employment (Clark, Georgellis and Sanfey, 2001).

In addition to psychological scars, unemployment has serious economic consequences for the unemployed. Gruber (1997), for example, has found that consumption of food, a basic necessity, falls for the unemployed. He further finds that the provision of UI benefits reduces the drop in food consumption of the unemployed.

---

13 Provided that the programme lasted long enough, participation in any program counted as "employment and, hence, gave renewed benefit eligibility. Before 1986, only participation in "job-like" programs, such as relief jobs, qualified. See Forslund and Krueger (1997).

Experience rating

In the U.S., the financing of UI is experience rated, meaning that employers who have a worse record of laying off workers are charged higher tax rates. Although experience rating has lapsed in recent years, this feature of the program still servers as a disincentive for layoffs and causes employers to internalize the cost of laying off workers. Much evidence suggests that experience rating reduces unemployment. For example, Card and Levine (1994) estimate that the unemployment rate would decline by six-tenths of a percentage point if industries were fully experience rated – that is, if employers in an industry were required to pay the full additional costs of unemployment benefits for layoffs in that industry.

8.3 Job search in Sweden in international comparison

Of particular concern for Sweden is whether the UI system, and extensive labor market programs more generally, depress job search on the part of the unemployed. Despite relatively generous benefits, the evidence that job search in Sweden is unusually low is mixed. Using time diary data, Krueger and Mueller (2008) found that 11% of unemployed individuals searched for a job on any given day in Sweden in 2000-01, exactly matching the average of the 14 countries that they examined, but about half the proportion in the U.S. However, the amount of time that job seekers devoted to searching for a job was low in Sweden compared with the other countries.

8.4 Policy considerations

Confronted with rising unemployment, Sweden and other countries are likely to shore up their UI systems to help the jobless in the current downturn.

One limitation of the extant literature is that studies have not examined the effect of the business cycle on the unintended consequences of UI benefits. Governments often increase the generosity or duration of benefits in economic downturns. The presumption is that when fewer jobs are available, there is less of a moral hazard effect of UI and the consumption smoothing benefits of UI are greater. While this is very plausible, the argument rests on theoretical reasoning rather than evidence. We also note that there are some advantages to designing UI programs that automatically extend benefits in a region when economic conditions weaken there, and then automatically turn them off when conditions improve. One advantage of such a system is that benefits are extended rapidly on a regional basis where they are needed most, which strengthens the stabilization effect of UI. In addition, having benefits return to normal when economic conditions improve insulates politicians who vote to end extended benefits and wards against a tendency for benefits to be ratcheted upward.

We would also emphasize that UI could play an even greater role in discouraging layoffs if benefits were experienced rated.

9 Conclusion

Rather than repeat the conclusions and substantive recommendations that we provided at the end of each section, we conclude by noting that we have summarized evaluation results from a range of countries with different contexts and different programs. Even for the same program, such as job search assistance or early education, there is variability within a country in the design, content and implementation of the program. The details of the programs, which we have not emphasized, are relevant for evaluating their effectiveness. A preschool program can be well designed and implemented or poorly designed and implemented. Moreover, the most effective design and implementation will likely vary with the target population.

Despite glossing over the details of the programs, we believe that the broad brush approach that we have taken is useful in that it highlights the most robust findings. The fact that job search assistance programs do, in fact, vary, yet have generally been found to be cost-effective suggests that job search assistance can be effective for a wide range of populations under diverse circumstances. In addition, we have tried to emphasize the studies that we think provide the most credible estimation strategies in the literature, often involving randomized assignment to "treatment" and "control" groups, so that the effect of the programs themselves can be separated from extraneous factors. This has naturally led us to draw on evidence from a wide range of settings.

References

Ackum Agell, S., Björklund, A. and Harkman, A. (1995), Unemployment insurance, labour market programmes and repeated unemployment in Sweden, Swedish Economic Policy Review 2.

Acs, Z. J, Audretsch, D. B. and Feldman, M. P. (1992), Real Effects of Academic Research: Comment, American Economic Review 82.

Aghion P. and Howitt P. (1992), A Model of Growth through Creative Destruction, Econometrica 60.

Aghion P. and Howitt P. (2006), Appropriate Growth Policy, Schumpeter Lecture, Journal of the European Economic Association 4.

Andersson, M. and Ejermo, O. (2005), How does accessibility to knowledge sources affect the innovativeness of corporations? — evidence from Sweden, The Annals of Regional Science 39.

Anderson, M. (2008), Uncovering gender differences in the effects of early intervention: A reevaluation of abecedarian, Perry preschool and early training projects, forthcoming in Journal of the American Statistical Association.

Angrist, J, and Krueger, A.B. (1992), The Effect of Age at School Entry on Educational Attainment: an Application of Instrumental Variables with Moments from Two Samples, Journal of the American Statistical Association 57.

Anselin, L., Varga, A. and Acs, Z. (1997), Local geographic spillovers between university research and high technology innovations, Journal of Urban Economics 42.

Atkinson, A. and Micklewright, J. (1991), Unemployment compensation and labor market transitions: A critical review, Journal of Economic Literature 29.

Bennmarker, H., Carling, K. and Holmlund, B. (2007), Do Benefit Hikes Damage Job Finding? Evidence from Swedish Unemployment Insurance Reforms, Labour 21.

Berlinski, S., Galiani, S., and Gertler, P. (2008), The effect of pre-primary education on primary school performance, forthcoming Journal of Public Economics.

Björklund, A. (1985), Unemployment and mental health: Some evidence from panel data, Journal of Human Resources 20.

Björklund, A., Clark, M., Edin, P-A, Fredriksson, P and Krueger, A.B. (2005), The Market Comes to Education in Sweden: An Evaluation of Sweden's Surprising School Reforms, Russell Sage Foundation, New York.

Black, D.A., Galdo, J. and Smith, J.A. (2007), Evaluating the Worker Profiling and Reemployment Services System Using a Regression Discontinuity Approach, American Economic Review Papers and Proceedings 97.

Black, D.A., Smith, J.A., Berger, M.C. and Noel, B.J. Brett J Noel (2003), Is the Threat of Reemployment Services More effective Than the Services Themselves? Evidence from Random Assignment in the UI System, American Economic Review 93.

Black, S., Devereux, P. and Salvanes, K. (2008), Too Young to Leave the Nest? The Effects of School Starting Age, IZA WP 3452.

Blundell, R., Costa Dias, M., Meghir, C. and Van Reenen, J. (2004), Evaluating the Employment Impact of a Mandatory Job Search Program, Journal of the European Economic Association 4.

Calmfors, L., Forslund, A. and Hemström, M. (2004), The Effects of Active Labor-Market Policies in Sweden: What is the evidence?, in Agell, J., Keen, M. and Weichenrieder, J. (ed), Labor Market Institutions and Public Regulation, MIT Press.

Card, D., Chetty, R. and Weber, A. (2007), The Spike at Benefit Exhaustion: Leaving the Unemployment System or Starting a New Job? American Economic Review: Papers and Proceedings 97.

Card, D. and Levine, P. (1994), Unemployment Insurance Taxes and the Cyclical and Seasonal Properties of Unemployment, Journal of Public Economics 53.

Carling, K., Edin, P-A., Harkman, A. and Holmlund, B. (1996), Unemployment duration, unemployment benefits, and labor market programs in Sweden, Journal of Public Economics 59.

Carling, K., Holmlund, B. and Vejsiu, A. (2001), Do Benefit Cuts Boost Job Finding? Swedish Evidence from the 1990s, Economic Journal 111.

Chetty, R. (2008), Moral Hazard versus Liquidity and Optimal Unemployment Insurance, Journal of Political Economy 116.

Clark, A., Georgellis, Y. and Sanfey, P. (2001) Scarring: The Psychological Impact of Past Unemployment, Economica 68.

Clark, A. and Oswald, A. (1994), Unhappiness and Unemployment, The Economic Journal 104.

Currie, J. (2001), Early Childhood Education Programs, Journal of Economic Perspectives 15.

Fersterer, P., Pischke, J-S. and Winter-Ebner, R. (2008), Returns to Apprenticeship Training in Austria: Evidence from Failed Firms, Scandinavian Journal of economics 110.

Forslund, A., Johansson, P. and Lindqvist, L. (2004), Employment subsidies - A fast lane from unemployment to work?, IFAU Working Paper 2004:18.

Forslund, A. and Krueger, A. (1997), An evaluation of the active Swedish labor market policy: New and received wisdom, in: Freeman, R., Swedenborg, B. and Topel, R. (eds), The welfare state in transition, Chicago University Press, 1997.

Forslund, A. and Krueger, A. (2008) Did Active Labour Market Policies Help Sweden Rebound from the Depression of the Early 1990s? forthcoming in Freeman, R., Swedenborg, B. and Topel, R. (eds), Reforming the Welfare State: Recovery and Beyond in Sweden.

Fredriksson, P. and Holmlund, B. (2003), Improving incentives in unemployment insurance: A review of recent research, IFAU Woring Paper 2003:5.

Fredriksson, P. and Öckert, B. (2006), Is Early Learning Really More Productive? The Effect of School Starting Age on School and Labor Market Performance, IZA WP 1659.

Garces, E., Thomas, D. and Currie, J. (2002) Longer-Term Effects of Head Start, American Economic Review 92.

Gormley, Jr., W. T. and Gayer, T. (2005), Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program, Journal of Human Resources 60.

Gormley, Jr., W. T., Gayer, T., Phillips, D. and Dawson, B. (2005), The effects of universal pre-K on cognitive development, Developmental Psychology 41.

Gormley Jr., W.T., Gayer, T. and Phillips, D. (2008) Preschool Programs Can Boost School readiness. Science 320.

Griliches, Z. (1979), Issues in Assessing the Contribution of Research and Development to Productivity Growth, Bell Journal of Economics 10.

Grossman, G.M., and Helpman, E. (1991), Innovation and Growth in the Global Economy, MIT Press, Cambridge, MA.

Gruber, J. (1997), The Consumption Smoothing Benefits of Unemployment Insurance, The American Economic Review 87.

Hall, C. (2008), The introduction of a more comprehensive upper secondary school system in Sweden: the effects on drop-out rates, educational attainment and earnings, work in progress.

Hansen, K. T., Heckman, J. J. and Mullen, K. J. (2004), The Effect of Schooling and Ability on Achievement Test Scores, Journal of Econometrics 121.

Hanushek, E.A. and Raymond, M.E. (2005), Does School Accountability Lead to Improved Student Performance? Journal of Policy Analysis and Management 24.

Hanushek, E.A. and Woessmann, L. (2006), Does Educational Tracking Affect Performance and Inequality? Differences-in-Differences Evidence across Countries, Economic Journal 116.

Holmstrom, B. and Milgrom, P. (1991) Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design, Journal of Law, Economics and Organization 7.

Hopenhayn, H.A. and Nicolini, J.P. (1997), Optimal Unemployment Insurance, Journal of Political Economy 105.

Jacob, B. (2005), Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools, Journal of Public Economics 89.

Jaffe, A.B. (1989), Real Effects of Academic Research, American Economic Review 79.

Jaffe, A.B., Trajtenberg, M and Henderson, R. (1993), Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, Quarterly Journal of Economics 108.

Kang, S. and Bishop, J. (1989), Vocational and Academic Education in High School: Complements or Substitutes? Economics of Education Review 8.

Katz, L. (1998), Wage Subsidies for the Disadvantaged, in Freeman, R. and Gottschalk, P. (eds), Generating Jobs: How to Increase Demand for Less-Skilled Workers, Russell Sage.

Katz, L. and Meyer, B. (1990a), The impact of the potential duration of unemployment benefits on the duration of unemployment. Journal of Public Economics 41.

Katz, L. and Meyer, B. (1990b), Unemployment Insurance, Recall Expectations, and Unemployment Outcomes, Quarterly Journal of Economics 105.

Kemple, J.J. and Scott-Clayton, J. (2004), Career Academies Impacts on Labor Market Outcomes and Educational Attainment, MDRC, New York.

Kemple, J.J. and Willner, C.J. (2008), Career Academies Long-Term Impacts on Labor Market Outcomes, Educational Attainment, and Transitions to Adulthood, MDRC, New York.

Kluve, J. (2006), The Effectiveness of European Active Labor Market Policy, IZA WP 2018.

Krueger, A. (1999), Experimental Estimates of education Production Functions, Quarterly Journal of economics 114.

Krueger, A. and Meyer, B. (2002), Labor supply effects of social insurance, in Auerbach, A. J. and Feldstein, M. (eds), Handbook of Public Economics.

Krueger, A. and Mueller, A. (2008), Job Search and Unemployment Insurance: New Evidence from Time Use Data, IZA working paper 3667.

Krueger, A. and Pischke, J-S. (1995), A Comparison of East and West German Labor Markets Before and After Unification, in Freeman, R. and Katz, L. (eds.), Differences and Changes in Wage Structures, University of Chicago Press, Chicago.

Lavy, V. (2002), Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement, Journal of Political Economy 110.

Lavy, V. (2008) Performance Pay and Teachers' Effort, Productivity and Grading Ethics, forthcoming in American Economic Review.

Lechner, M. and Smith, J.A. (2006), What is the value added by caseworkers? Labour Economics 14.

Leuven, E., Lindahl, M., Oosterbeek, H. and Webbink, D. (2007) Expaning Schooling Opportunities for 4-year Olds, Mimeo, University of Amsterdam.

Light, A.L. (1999), High School Employment, High School Curriculum, and Post-School Wages. Economics of Education Review 18.

Lindbeck, A. and Persson, M. (2006), A Model of Income Insurance and Social Norms, CESifo Working Paper No. 1675.

Love J.M., Eliason-Kisker, E., Ross, C.M., Schochet, P.Z., Brooks-Gunn, J., Paulsell, D., Boller, K., Constantine, J., Vogel, C., Sidle Fuligni, A. and Brady-Smith, C. (2002), Making a difference in the lives of infants and toddlers and their families: The impacts of early head start, Technical Report PR02-30a, Mathematica Policy Research Report.

Ludwig, J. and Miller, D.L. (2007) Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design, Quarterly Journal of Economics 122.

Mane, F. (1999), Trends in the payoff to academic and occupation-specific skills: the short and medium run returns to academic and vocational high school courses for non-college-bound students, Economics of Education Review 18.

Machin, S. and McNally, S. (2008), The Literacy Hour, Journal of Public Economics 92.

Meyer, B. (1995), Lessons from the U.S. Unemployment Insurance Experiments, Journal of Economic Literature 33.

Neal, D. and Whitmore-Schanzenbach, D. (2008) Left Behind By Design: Proficiency Counts and Test-Based Accountability, forthcoming in Review of Economics and Statistics.

Nye, B., Konstantopoulos, S. and Hedges, L.V. (2004), How Large are Teacher Effects? Educational Evaluation and Policy Analysis 26.

Parey, M. (2008) Vocational Schooling versus Apprenticeship Training: Evidence from Vacancy Data, Mimeo, University College London.

Rivkin, S., Hanushek, E. and Kain, J. (2005) Teachers, Schools, and Academic Achievement, Econometrica 73.

Romer, P.M. (1990), Endogenous Technological Change, Journal of Political Economy 98.

Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., and Nores, M. (2005), Lifetime Effects: The High/Scope Perry Preschool Study through Age 40 . The High/Scope Press, Ypsilanti, MI.

Shavell, S. and Weiss, L. (1979), The Optimal Payment of Unemployment Insurance Benefits over Time, Journal of Political Economy 87.

Sianesi, B. (2004), An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s, The Review of Economics and Statistics 86.

Werning, I. (2002) Optimal Unemployment Insurance with Unobservable Savings, Mimeo, MIT.

Wong, K. (2008), Looking Beyond Test Score Gains: State Accountability's Effect on Educational Attainment and Labor Market Outcomes, Mimeo, University of California, Irvine.