IFN Working Paper No. 779, 2008

One Size Fits All? The Effects of Teacher Cognitive and Non-cognitive Abilities on Student Achievement

Jonas Vlachos

Research Institute of Industrial Economics
P.O. Box 55665
SE-102 15 Stockholm, Sweden
info@ifn.se
www.ifn.se

One size fits all?

The effects of teacher cognitive and non-cognitive abilities on student achievement[a]

by

Erik Grönqvist[b] and Jonas Vlachos[c]

December 2, 2008

Abstract

Teachers are increasingly being drawn from the lower parts of the general ability distribution, but it is not clear how this affects student achievement. We track the position of entering teachers in population-wide cognitive and non-cognitive ability distributions using school grades and draft records from Swedish registers. The impact on student achievement caused by the position of teachers in these ability distributions is estimated using matched student-teacher data. On average, teachers' cognitive and non-cognitive social interactive abilities do not have a positive effect on student performance. However, social interactive ability turns out to be important for low aptitude students, whilst the reverse holds for cognitive abilities. In fact, while high performing students benefit from high cognitive teachers, being matched to such a teacher can even be detrimental to their lower performing peers. Hence, the lower abilities among teachers may hurt some students, whereas others may even benefit. High cognitive and non-cognitive abilities thus need not necessarily translate into teacher quality. Instead, these heterogeneities highlight the importance of the student-teacher matching process.

Keywords: Cognitive and non-cognitive ability, Teacher quality, Student achievement
JEL-codes: I21, H4, J4

---

[a] We have benefited from discussions with and comments from David Figlio, Peter Fredriksson, Per Johansson, Victor Lavy, Andrew Leigh, Mikael Lindahl, Assar Lindbeck, Erik Lindqvist, Oskar Nordström Skans, Torsten Persson, Per Pettersson-Lidbom, Jonah Rockoff, David Strömberg, Justin Wolfers, Björn Öckert, and from seminar participants at the Institute for International Economic Studies (IIES), the Institute for Labour Market Policy Evaluation (IFAU), the Institute for Social Research (SOFI), and the Research Institute for Industrial Economics (IFN), the RATIO institute, the Tinbergen institute, the a23 COST-meeting in Paris 2008, Nordic Summer Institute in Labour Economics in Aarhus 2008, ESPE 2008, EEA 2008, EALE 2008, and the 4th EEEPE network conference 2008 in Amsterdam. We are grateful to various people in the municipal administration in Linköping, Göteborg, Stockholm, Malmö, Uppsala, Örebro, Halmstad, Västerås and Jönköping for making their school records available. We are also grateful the Swedish National Service Administration and the Swedish War Archive for providing us with military draft data. Staffan Brantingson at Statistics Sweden has provided services beyond the call of duty.

[b] IFAU, P.O. Box 513, SE-75120 Uppsala, email: erik.gronqvist@ifau.uu.se

[c] Department of Economics, Stockholm university, IFN and CEPR, SE-106 91 Stockholm, email: jonas.vlachos@ne.su.se

Table of contents

1 Introduction

No one involved in education would deny the importance of teacher quality for student performance. Indeed, there is a large body of research showing that "teacher fixed effects" are systematically related to student outcomes. With the exception of teacher experience however, it has proven remarkably difficult to pinpoint observable teacher characteristics that raise student achievement (e.g. Rockoff, 2004 and Rivkin et al., 2005). The conjecture has been that the position of teachers in some general, but hard to observe, ability distribution is what matters for student outcomes. The worry about teacher quality has therefore been fuelled by studies from several countries showing that the ability ranking, gauged by aptitude tests or standardised subject tests, of new teachers and individuals entering teacher education has declined substantially over time.[1]

Despite widespread beliefs to the opposite (McKinsey, 2007 and *Economist*, 2007), a causal link between the position of teachers in the population-wide ability distribution and student achievement has been assumed rather than shown.[2] In this study, we document the position of entering teachers in three population-wide cognitive and non-cognitive ability distributions. Our main contribution is that we use the same ability measures to estimate the causal effects of teacher ability rankings on student achievement. Our findings suggest that the position of teachers in the overall ability distributions has no statistically significant effect on average student achievement.

[1] See Nickell and Quintini (2002) for the UK; Corcoran et al. (2004) and Bacalod (2007) for the US; Leigh and Ryan (2006) for Australia; Fredriksson and Öckert (2008) for Sweden. Hoxby and Leigh (2004) and Lakdawalla (2006) are other studies documenting the decline of teacher aptitude and ability in the US. These studies are all based on ability measures that are (more or less) comparable across cohorts. Importantly, ability is measured prior to the start of teacher education so they do not reflect the impact of the educational and professional choices themselves. While the mentioned studies attempt to explain the decline in teacher ability, this issue is beyond the scope of our paper.

[2] See Wayne and Youngs (2003) and Hanushek and Rivkin (2006) for surveys of this extensive literature. Several papers use the selectivity of teachers' undergraduate institution as a proxy for the position in the ability distribution. This is at best a crude measure of individual ability that may also reflect the quality of the education that the teacher has received. Other studies find that the scores on teacher licensure tests affect student outcomes, but this again has little to say about the teacher's position in the general ability distribution. Ehrenberg and Brewer (1995) find a "verbal ability test" to be positively related to student outcomes, but the measure is aggregated to the school level and its relation to the general ability distribution is unclear. Ferguson and Ladd (1996) find a positive relation between college entrance ACT scores and student achievement gains among 3 and 4 graders. The ACT is, however, taken by an already selected group of individuals. Close to our study is also Hanushek (1992) who finds that gains in reading performance among 2-6 graders are greater if the teacher has scored high on the "Quick word test", sometimes seen as a substitute intelligence test.

However, we do find important asymmetries both between students and across male and female teachers. Equating teacher quality with IQ-like measures of human capital thus seems questionable.

We track the position of teachers in the distribution of abilities using measures of cognitive ability and non-cognitive social interactive ability from the military draft; the latter being aimed at capturing leadership capacity under war-time stress. The draft data are available for men only so we also rank teachers according to their upper-secondary grade-point average (GPA). The cognitive draft evaluation is close to a standard IQ-test. The non-cognitive social abilities being gauged by a standardized psychological evaluation are emotional stability, psychological endurance, the ability to take initiatives, social outgoingness, and sense of responsibility—all personality traits that should be important to teachers. GPA-scores capture a mix of cognitive and non-cognitive abilities such as ambition, conscientiousness and self-discipline. As the ability data cover up to 30 consecutive cohorts, it is important to note that all abilities are evaluated using consistent procedures prior to post-secondary education. Thus the position of teachers in these population-wide ability distributions does not reflect changes in the quality of teacher training, or changes in the ability evaluations.

Using a database matching a large number of individual teachers to individual students we directly relate the position of each teacher in the respective ability distribution to standardized student test-scores. The decline in social ability is found to have had a negative impact on low-aptitude students and students of foreign background. On the other hand, the teachers' social interactive ability appears to be close to irrelevant for the highest performing students. Similar asymmetric effects are also found for teachers' GPA-rank. The reverse pattern holds for the position of teachers in the cognitive ability distribution: the average insignificant effect hides that high-aptitude students benefit from high cognitive teachers, while being matched to such teachers can even be detrimental to the achievements of low-aptitude students. These asymmetries are in line with Clotfelter et al. (2006) who document that the impact of teacher's mathematical ability differs substantially across students from different backgrounds.

Regarding teacher asymmetries, our results show that male teachers with a high GPA are highly positive for student achievement while female teachers are not. This may indicate that grades capture different capacities for men and women.[3] Alternatively, it could reflect that the selection into teaching differs substantially between men and women. Other important capacities, such as the motivation to teach, may therefore differ between male and female teachers with high GPA-scores. Yet another possibility is that the school environment itself hampers the performance of high-GPA female teachers.

Our data set is very rich in the sense that we match individual students to the responsible teacher for a large sample of students during their last year of middle school. We observe standardized test scores in several core subjects. By using student fixed effects, we control for average student ability across subjects and use the within-student variation to identify the effects of teacher abilities. This identification strategy deals with all sensible selection patterns in the student-teacher matching process.

In what follows, we start by describing the different ability measures and document the decline in teacher abilities along these dimensions. We then discuss our identification strategy in the light of the institutional features of the Swedish school system and thereafter present our results. In the final section, we conclude and discuss policy implications of our findings.

2 The evolution of teacher abilities

As mentioned in the introduction, evidence from several countries shows that teachers over time have become increasingly likely to be drawn from the lower parts of the ability distribution, as measured by aptitude or standardised subject tests.[4] In this section, we start by presenting three different ability measures and then proceed to describe how the position of teachers in these ability distributions has changed over time.

2.1 Ability measures

In order to derive the position of teachers in some overall ability distribution it is necessary to use ability data based on large representative samples of the population. We

[3] For example, Lindahl (2007) finds that girls' grades deviate more from their test scores than boys'. This may indicate that boys' grades reflect actual subject understanding to a greater extent than girls'.

[4] See for example Nickell and Quintini (2002), Corcoran et al. (2004), Leigh and Ryan (2006), Bacolod (2007), and Fredriksson and Öckert (2008).

have access to three such measures. The first is a measure of cognitive abilities from the military draft, available for essentially all Swedish men. The second, also from the military draft, is an evaluation of non-cognitive social interactive ability. Both these ability measures have been found to be strongly related to future earnings.[5] Finally, we use information on upper-secondary school GPA. It is not fully understood which capacities that are captured by GPA-scores, but both Björklund et al (2005) and Lindahl (2001) have shown that school performance is a good predictor of future earnings, even when controlling for cognitive ability. The main benefit of the draft data is that the tests are designed for capturing particular cognitive and non-cognitive capacities. The main drawback, though, is that these data are only available for men.

All Swedish men are by law obliged to go through the military draft, if called upon.[6] In most cases, the draft occurs the year the man turns 18. Up until the late 1990s, more than 90 percent of all men in each cohort went through the whole draft procedure, with only the physically and mentally handicapped being exempted.[7] Since then, the need for conscripts has declined dramatically, and as a consequence the draft procedure underwent a major change in 2000.

The draft consists of a series of physical, psychological and intellectual tests and evaluations. For the purpose of this study, we have acquired data on the draft tests of cognitive ability and on the standardised psychological evaluation of social interactive ability under war-time stress. Comparable data are available from 1969 to 1999, which means that our data will contain information for draftees born approximately between 1951 and 1981.

The evaluation of cognitive ability consists of several subtests of logical, verbal, and spatial abilities, as well as a test of the draftees' technical understanding. The results on these subtests are combined to produce a general cognitive ability ranking on a 1-9 scale.

[5] Lindqvist and Vestman (2008) find that a one standard deviation increase in cognitive ability is associated with 8.4 percent higher income. For social interactive ability, an increase of the same magnitude is associated with 5.7 percent higher income (these estimates are unadjusted to several selection issues discussed at length by Lindqvist and Vestman).

[6] This discussion of the draft data draws heavily on an interview with Johan Lothigius, chief psychologist at the National Service Administration, carried out by Erik Lindqvist (August 25, 2004). We are grateful to Erik for sharing his notes with us.

[7] The consequences of refusing the draft include fines and being round up by the police, and ultimately prison in up to one year (1994:1809 Lag om totalförsvarsplikt, kap 10).

This test has been subject to evaluation by psychologists and appears be a good measure of general intelligence (Carlstedt, 2000). In order to account for general trends in test-taking capacity and for minor changes in the draft tests, we percentile rank the cognitive ability separately for each draft cohort.

The other main measure from the draft is based on a standardized psychological evaluation aimed at determining social interactive abilities under war-time stress. The evaluation is performed by a certified psychologist who conducts a structured interview with the draftee. As a basis for the interview, the psychologist has information about the draftee's results on the tests of cognitive ability, physical endurance, muscular strength, grades from school and the answers on questions about friends, family and hobbies, etc. The interview follows a specific, and secret, manual that states topics to discuss and also how to grade different answers. As in the case for cognitive ability, the social interactive ability is recorded on a 1-9 scale which we use to construct a year-by-year percentile ranking. The personality traits evaluated in the draft procedure are *psychological endurance*, *emotional stability*, *the ability to take initiative*, *social outgoingness*, *sense of responsibility*, and *ease to adjust to a military environment*. Motivation for doing the military service is, however, explicitly *not* a factor which is to be evaluated. The evaluation instrument is based on the experiences from the Korean War, adapted to Swedish circumstances. The experiences of Swedish UN peacekeeping troops have also been important.

One concern with using the draft data is that some subjects may not aim at receiving a maximum score at the cognitive tests—potentially in hope to avoid the military service altogether. As measurement error of this type is likely to be more pronounced among individuals scoring low at the evaluations, we drop the five percent lowest performing from the analysis when estimating the impact on student achievement.

The final measure of teachers' ability is their upper-secondary school GPA, generally set at age 19. This is a very general ability measure capturing not only cognitive ability, but also personality traits like adaptability, ambition, motivation, maturity and conscientiousness. Grading data from the upper-secondary school is available from the cohort graduating in 1985 and onwards; that is, those born approximately 1966 and later. Since we are interested in the position of teachers in the overall ability distribution, we percentile rank the GPAs for each cohort of graduates. This way we also take account of any potential grade inflation.[8]

[Table 1. Correlation between cognitive abilities and social ability]

To sum up, we use three different measures of abilities—upper-secondary school GPA, cognitive ability, and non-cognitive social interactive ability—all measured at about same age. Since all abilities are measured prior to entering tertiary education, they are not affected by any changes in the teacher education that may have occurred over time. The measures are all related but still capture different aspects of the personality: *Table 1* shows the correlation between the ability measures for the full population. Between social and cognitive ability, the correlation is 0.36. It is worth noting that this correlation is close to the correlation between cognitive and non-cognitive personality factors reported by Cunha and Heckman (2008). The correlation between GPA and cognitive ability is 0.47, and 0.23 between GPA and social interactive ability.

Both cognitive and non-cognitive abilities are highly significant when running an OLS regression with GPA as the dependent variable.[9] Combined, the social and cognitive abilities only pick up 25 percent of the total variation in GPA, which indicates that a substantial part of the variation in GPA may capture other personality traits and capacities; such as, adaptability, ambition, motivation, maturity or conscientiousness.

2.2 The evolution of the teacher pool

In Sweden, all teachers are registered in the *Teacher register* from 1979 onwards. By matching our ability measures to this register, we can track the evolution of cognitive and non-cognitive abilities recorded at the draft for entering teachers from 1980 and onwards. For teacher GPA, we can track abilities from 1993 and onwards. As mentioned, the

[8] In upper-secondary school there are different programs, and grading standards may differ between programs. However, since most teachers have graduated from three year theoretical programs we believe any differences in grading standards to be a negligible problem. Still, in all analyses using the GPA measure we control for upper-secondary school program. Further, in 1992 there was a minor change in the grading system as it was no longer possible to exclude the two lowest grades from the GPA when applying to higher education, and in 1996 there was a major change as the system of relative grades was replaced with goal related grades. The year-wise percentile rank takes care of the change in grading system, to the extent that these changes did not affect the rank of grades in the distribution.

positions in the respective ability distribution are measured prior to entering tertiary education and are therefore unaffected by any changes in the quality of teacher education that may have occurred over time.

The ideal way to measure the evolution of abilities in the teacher pool would be to track the average ability scores of the whole teacher stock over time. However, with the teacher register being available from 1979 and the draft data only being available between 1969 and 1999 this is not possible. As draftees are around 18 years old, abilities are only observed for teachers aged 29 and younger in 1980. The available draft data would therefore not allow us to paint a comparable picture of the teacher stock over time.[10]

For this reason we instead track the average annual values of cognitive and non-cognitive abilities for teachers *entering* the teacher register. This gives us a flow measure of the evolution of teacher abilities between 1980 and 2006. Due to the age restrictions that our data imposes, however, we do not capture teachers entering the profession at a relatively high age in the beginning of the period. We therefore make the series comparable by only analyzing entering teachers between 25 and 30 years of age.[11] Similarly, ability rankings based on the GPA is available for entering teachers aged 25 to 30 between 1993 and 2006.

[Figure 1. Ability ranks of new subject teachers (ages 25-30), 1980-2006]

We restrict our attention to teachers in theoretical subjects, so called subject teachers.[12] The evolution of cognitive ability, non-cognitive social interactive ability, and teacher GPA among new subject teachers in the middle school system is depicted in *Figure 1*. There has been a marked decline in all ability measures, most pronounced in cognitive

---

[9] The point estimate is 0.43 for cognitive and 0.14 for social ability (the number of observations is 596,143).

[10] In particular, if there has been a gradual decline in teacher abilities, the data limitations mean that we will understate the degree of this decline. If teacher quality has improved, the opposite naturally applies.

[11] It is indeed the case that the decline in teacher abilities is more severe when analyzing all entering teachers without imposing any age restriction.

[12] This means that we do not include teachers in athletics, aesthetics, music, home economics, shop, and similar subjects. The main reason for this exclusion is that we estimate student outcomes only on theoretical subjects. Further, the turnover of non-theoretical subjects in the curriculum is much higher that in the core subjects. By excluding the

ability. According to the cognitive draft test, the average ability has declined by close to 20 percentile ranks since the peak in the early 1990's. The decline in social ability and GPA is between 10 and 15 percentile ranks over the same period. The decline in GPA is of similar magnitude for both men and women, although female teachers on average tend to have a higher GPA.

[Table 2. The evolution of abilities of new subject teachers (ages 25-30)]

To get at the rate of the decline, we regress the average ability rank on a time trend. As can be seen in *Table 2*, the average cognitive ability and social ability has declined at an average annual rate (measured over the whole time period) of -0.57 and -0.34 percentile rank points, respectively. The average GPA, in turn, has declined with an annual rate of -0.66 since 1993. For male teachers, the rate is -0.60 and for female ones -0.76, but the trends are not statistically different from each other.

The fact that the rate of decline in GPA is similar for men and women suggest that there are no important gender differences in ability trends among teachers. When it comes to the abilities recorded at the draft a direct comparison across genders is of course impossible. However, we can still get a picture of the evolution of abilities for female teachers by comparing the draft records for the full brothers of female and male teachers. Under the assumption that ability correlations between siblings have not changed over time, this approach should yield informative answers.

[Table 3. The evolution of abilities of new male and female subject teachers (ages 25-30)]

In *Table 3* we compare the evolution of new male and female middle school subject teachers, as measured by the average abilities of their brothers. For cognitive ability there is a clear declining trend for both male and female teachers. The trend coefficient is larger (in absolute values) for women, albeit not statistically different from the male trend. For social interactive ability, we find a statistically significant negative trend for

---

practical subjects we thus increase the comparability of the teacher pool over time. For details concerning the female teachers and an insignificant trend for males. Again, the difference is not statistically significant.

All in all, the results show that the decline in teacher abilities has, if anything, been even more dramatic among female teachers than among male ones. This is in line with the findings in Bacolod (2007) who shows that the decline in teacher abilities in the US has been much more pronounced among women than among men. Corroborating evidence for Sweden can be found in Fredriksson and Öckert (2008) who, using an alternative measure of cognitive ability, find that the decline among those graduating from teacher education has been slightly larger for women. We now turn to the question whether or not this decline actually matters for student achievement.

3 The school system and empirical strategy

To estimate the causal effect of teacher characteristics on student performance, teachers with different abilities would ideally be randomly assigned to students. In our setting, this is not the case. Rather, students and teachers are sorted into schools and classes in non-random ways that would bias the results unless the selection process is properly handled.[13] In this section, we provide a brief introduction to the Swedish school system and then describe our identification strategy in light of these institutional features. This strategy deals with all sensible selection problems that could arise.

3.1 The Swedish school system

Compulsory schooling in Sweden usually starts at age seven and lasts for nine years. Five years of primary/elementary school are followed by four years of middle school (grades 6-9). Thereafter, a non-compulsory three year upper-secondary program follows. The municipalities are responsible for all tiers of schooling. The 1985 *Education Act* (Ministry of Education and Research, 2000) sets the national educational standards which are overseen by the Swedish National Agency of Education. The middle school system is organized around municipal schools that all students within a municipality formally are

classification of teachers, see the appendix.

[13] Clotfelter et al. (2006) documents this type of sorting between and within schools in the North Carolina elementary school system, and discuss the biases that arise when not taking sorting into account.

free to apply to. Actual admittance is in practice highly regulated as priority has to be given to the students residing within the school's catchment area.[14] The *Education Act* provides detailed requirements that all schools have to fulfil, leaving schools with limited discretion regarding the curriculum.[15]

In the last year of middle school all students take nation-wide exams in Swedish, English and Mathematics, for which the scores are filed in central registers. These standardized test scores—graded on the scale *No-Pass*, *Pass*, *High-Pass*, and *Pass-with-Distinction*—are the outcome variable in this study. The tests scores shall be used by the teacher when setting students' final grades (Skolverket 2004). These grades, in turn, should reflect how well the students live up to national pre-defined standards. The middle school grades are used to sort students when applying for upper-secondary school.

3.2 Identification

Among Swedish middle schools there is substantial sorting of students between schools, reflecting the socio-economic situation in different residential areas. Within schools there may also be sorting in the sense that students from different locations are not randomly assigned to different classes. Schools have varying policies in this regard, but it is common that students living close to each other are grouped together. In addition, ability tracking is not allowed.[16] Thus, while students are definitely sorted in the Swedish school system, sorting mainly occurs along the lines of general ability and motivation and not due to subject-specific student proficiency. As teachers are likely to be matched to students in non-random ways based on these general characteristics, we need to control for average student ability.

Each middle school student is observed across several subjects, but only once for each of these. This allows us to hold general student ability constant by controlling for student fixed effects. As middle school students are primarily sorted on general ability, this approach accounts for most serious selection problems. Further, there may be a

---

[14] Since 1992, Sweden also has a comprehensive voucher school system described in Björklund et al. (2005). As we are only dealing with municipal schools in this study, the voucher schools will not be discussed further. It should be noted that the *Education Act* regulates private as well as the public schools.

[15] Out of 6,665 compulsory school hours, the schools are free to decide on less than 10 percent; 600 hours (Skolverket, 2007).

correlation between the relative difficulty of a subject and teacher ability. If, for example, teachers in mathematics on average have a high ability ranking while it is difficult to achieve a high test result in this subject, our ability estimates will be downward biased. We control for this by also including subject fixed effects. Hence, we estimate the following relationship:

$$Test\ score\ rank_{its} = a\ Ability\ rank_t + X_t'b + \mu_i + \mu_s + \varepsilon_{its}.$$

The outcome is the ranked test score for student *i*, in subject *s*, taught by teacher *t*. We are primarily interested in estimating the parameter *a*, the impact of teacher ability rank on student achievement. Other teacher characteristics—birth cohort indicators and, where applicable, a gender indicator—are captured by the vector X_t,[17] μ_i are student fixed effects, and μ_s are subject fixed effects. The birth cohort indicators deal with any trends in test taking capacity, such as the Flynn (1984) effect,[18] changes in teacher education that may have occurred over time, as well as potential changes in the motivation to become a teacher based on unobservable characteristics. Standard errors are clustered on teachers, and we include time effects to account for general trends in test results.

We are interested in estimating the full impact of the position of teachers in the ability distribution on student achievement. As both the educational attainment of teachers and their experience level are likely to be endogenous to ability rankings, we only include controls for birth cohort and gender indicators in the final regressions. The approach to exclude variables such as educational attainment is standard when estimating the full effect of personality factors such as IQ on earnings (e.g. Neal and Johnson, 1996).

Under the plausible assumption that students are assigned teachers based on the same mechanism across all subjects, this *within-student estimator* captures the causal effect of teacher characteristics. The strategy is related to a value-added approach (Hanushek and Rivkin, 2006) in that we control for average student performance across subjects. In order to appreciate the within-student estimator, it is useful to consider the situations in which it

[16] In Sweden, ability tracking was gradually abandoned with the introduction of the new middle school curriculum, Lpo94, in 1995 (Skolverket, 2006). As of 1998 tracking was completely abolished.
[17] We use biennial cohort indicators since there are very few teachers in some of the cells when using annual indicators.

would not yield unbiased estimates on teacher characteristics. For this to occur it needs to be the case that students, within a school, are assigned to teachers whose characteristics systematically differ between subjects. This would be the case if, for example, high-ability English teachers were systematically assigned to highly motivated students at the same time as high-ability Swedish teachers were systematically assigned to poorly motivated ones. Similarly, we would not get unbiased estimates if past educational experiences were asymmetric across subjects in the sense that students with a good background in English and a poor background in Swedish were systematically assigned both highly-skilled English teachers and Swedish teachers. While this can certainly be the case in individual schools, it is unlikely to be a general scenario.

4 Data

To estimate the effect of teacher abilities on student achievement, we use detailed data matching individual students to individual teachers. These data are linked to teacher ability based on their upper-secondary school GPA and on the cognitive and non-cognitive ability ranks from the military draft, as described in section 2. Such a linkage is possible since all Swedish residents have a unique personal identifier that follows them throughout life and is used in all contact with the authorities. In this section we describe the data coming from different sources and how the data set is compiled.

4.1 Schooling data

In Sweden, there is no central authority keeping records that allows the individual grade setting teacher to be matched with the individual student. Some municipalities, however, have computerized student records allowing such a link to be created. We have been able to acquire such data from nine of the largest municipalities in Sweden, covering roughly 20 percent of all Swedish compulsory school students in each cohort. These schooling data are available for the years 2003-2007 and the coverage varies between municipalities.[19]

[18] Flynn (1984) observed substantial increases in population wide IQ-scores over time.
[19] We contacted the 20 largest municipalities—in terms of compulsory school students—with a request for data matching students with grade setting teachers. Of these, nine had computer systems that made it possible to fully meet

The data files for the compulsory schools contain information on test scores from national standardized tests in Swedish, English, and Mathematics, taken during the last year of compulsory schooling (grade 9). The final teacher grades in these and other subjects are also recorded in the files. Since these grade records are used for allocating individuals to upper-secondary school programs they are of high quality.

Both test scores and teacher grades can take one of four ordinal values: No Pass/Fail, Pass, High Pass and Pass with Distinction. In the analysis, these are percentile ranked in the full sample of students in order to facilitate comparison with other studies.

One caveat has to be mentioned. Usually, the same subject teacher is responsible for a subject throughout middle school. However, due to parental leave, teacher and student mobility, retirement and so on, there is some turnover in the student-teacher match. As no records are kept prior to the final year, we have no way of determining how many years students and teachers have actually been matched.

4.2 Individual level data

In order to undertake an analysis of asymmetric effects across different student groups we add student background information to our dataset. These data are from Statistics Sweden's population wide register datasets, based on tax records and population censes, and they include high quality information on student gender, parental educational attainment and immigration status. From these records we also collect information on teacher age and gender.

Using this information, we classify a student as coming from an *academic home* if both parents have some level of university education. A student has a *foreign background* either if he or she is born abroad or if both parents are born abroad.

4.3 The matched data set

The base for the analysis is schooling data between 2003 and 2007 from nine municipalities containing information on individual grades and test scores for each student and the identity of the subject teacher.

---

this request. These are Stockholm, Göteborg, Malmö, Uppsala, Jönköping, Örebro, Västerås, Linköping and Halmstad. The reason for contacting the largest municipalities was that the data request was both time-consuming and expensive.

Since teachers are recorded using their unique personal identifier, they can be matched both to their upper-secondary school GPA and their draft records. The GPA data are available for teachers graduating from upper-secondary school in 1985 and later. Draft data, in turn, are available for the draft cohorts 1969 to 1999 and made available by the National Service Administration and the Swedish War Archive. This means that only male teachers born from 1951 to 1981 who were Swedish citizens at the time of the draft can be matched to the school records. Even if we have draft data for a longer period we have more observations for the GPA data; by using the draft data where only men are available we loss of approximately two thirds of all grade setting teachers compared to the original data.[20] This also reduces the number of observations per students, relative the GPA data.

In total, we have 1,589 (704) teachers for whom we observe their GPA (draft record), administering 70,305 (29,749) test scores to 45,428 (24,847) students. Summary statistics of the data is shown in *Table 1*.[21]

[Table 4. Summary statistics]

In the GPA-sample of teachers, 69 percent are female and the average age is 33. The average GPA rank is 63.5 (66.6 for women, 56.7 for men). In the draft sample, the average age is close to 39 years and the cognitive rank is 64.5. The mean rank in social interactive ability is 54. Student characteristics vary little between samples: 12 percent are from an academic home, 22 percent have a foreign background, and 49 percent are female.

[20] This also means that the group of teachers we analyze are relatively homogenous, which is an advantage since we want to isolate the effects of the ability rankings.
[21] As can be seen, the mean values for the percentile ranked outcomes is not exactly 50. The reason is that we are using two different, only partly overlapping, samples. We have therefore percentile ranked using the whole population of outcomes, prior to dropping observations for which we lack teacher data.

5 Results

In section 2 we showed that over time teachers are increasingly drawn from lower parts of the ability distributions, a pattern also found for several other countries. Whether this development is a matter for concern crucially depends on whether the position of teachers in these ability distributions actually matters for student achievement. In this section we first present regression results of the causal link between the rank of teachers' abilities and student outcomes. We then interpret these effect estimates in the light of the declining position of teachers in the overall ability distributions. In sum, we find little evidence that teachers from higher parts of the distribution of cognitive or social interactive abilities would improve the achievement for the average student; though teachers possessing both high cognitive and high non-cognitive skills may be more capable of raising student achievement. However, there are important asymmetries between different student types, and between male and female teachers.

5.1 Baseline effects of teacher abilities

We begin by analysing the average impact of teacher abilities on student achievement using the identification strategy presented in section 3.[22] In all specifications, we control for student, subject and year fixed effects, as well as teacher birth cohort. Student fixed effects deal with the sorting of students to teachers and subject fixed effects take care of the selection of teachers to different subjects. Birth cohort dummies control for changes in teacher education as well as potential changes in the ability evaluations. In addition, birth cohort is a close proxy for teacher experience. In the teacher GPA regressions we also take teacher gender and upper-secondary program fixed effects into account. The upper-secondary program effects allow for potential differences in grading standards across these programs and the gender indicator for systematic differences between male and female teachers.

[22] An implicit assumption in our identification strategy of comparing students' performance in different subjects across different teacher abilities is that there are no spill-over effects of performance between subjects; such an effect would bias our results downwards. As a consistency test we have therefore included the abilities of the teachers in the students other subjects in the analysis, and find no evidence that the abilities of teachers in one subject have an influence on the results in other subjects.

The baseline results of how teachers' ability ranks effect student outcome are shown in *Table 5*. The first column shows that the estimated effect of cognitive ability rank on student outcomes is close to zero. The effect of social ability rank (column 2) is positive, but not statistically different from zero. In the third column, both the cognitive and social ability ranks are included but the estimates are the same as in columns one and two. In other words, there is no clear indication that a higher cognitive or social ability rank among teachers will lead to better (or worse) student performance on standardized tests.

The benefit of using cognitive and social abilities from the draft is that these measures are designed and validated to capture specific personality traits, but the drawback is that they are only available for male teachers. We therefore turn to teachers' ranked upper-secondary school GPA, which capture a mix of cognitive ability and characters traits like adaptability, ambition, motivation, maturity and conscientiousness. As a high upper-secondary school GPA gives access to selective tertiary education programs, the GPA measure is also interesting since it provides a measure of alternative career opportunities. In column 4 of *Table 5* we find that, on average, teachers with higher GPAs will not result in better student performance. The estimated coefficient is in fact even negative, but not statistically significant.

The next question is whether the effects of teacher GPA differ across teacher gender. We test this in column 5 and 6 by analysing male and female teachers separately. In column 5 we find a large positive and statistically significant effect for male teachers (the point estimate is 0.093), indicating that male teachers with higher GPA are more productive. In order to appreciate the magnitude of this effect we can think of a student switching to a male teacher with a one standard deviation higher GPA. This would increase the average student's performance by almost 10 percent of a standard deviation, indeed a substantial improvement. For female teachers (column 6) the estimate is substantially smaller in absolute size—about 25 percent of the male coefficient—and surprisingly indicates a *negative* effect from having a female teacher with higher GPA.

In essence, the general ability captured by teachers' upper-secondary GPA has different implications for male and female teachers. While male teachers with a higher ability rank are more productive, female teachers with higher upper-secondary school GPA may actually be slightly worse teachers. One reason for this difference may be that grades capture different capacities for men and women. In fact, Lindahl (2007) find evidence that girls' school grades to a larger extent capture other competences than what is measurable in objective test scores. Another potential explanation can be gender differences in the selection into the teacher profession; for example males who—despite having all the career opportunities a high GPA-score entails—chose to become teachers may be highly motivated, while a different selection process, for some reason, may be present among women with high GPA-ranks. Yet another possibility is that the school environment itself for some reason hampers the performance of high-GPA female teachers.

[Table 5. Baseline within student estimates]

It is worth noting that the point estimates will change slightly for cognitive and social rank—although remain statistically insignificant—if we depart from our preferred specification. In *Table A 1* in the Appendix we only control for school fixed effects and observable student characteristics (gender, parental education, and foreign background). This way, we do not capture student selection on unobservable characteristics. The estimated effect for teacher GPA becomes larger, even positive and significant, when we do not control for student fixed effects, suggesting that more able students are systematically matched to high GPA teachers. However, the results also show that the matching of high GPA teachers to more able students differs between male and female teachers. This stresses the importance of using the within-student variation when estimating the effect of teacher characteristics, in particular in the current setting where student sorting into classes is reasonably based on general motivation.

In addition to the baseline results in *Table 5*, we consider a number of extensions. In order to test the importance of functional forms, we add squared ability terms to the analysis in *Table 6* (columns 1-2 and 4-6). We do not find any effect of teachers' abilities on student outcomes in these regressions. In particular, we do not find any clear evidence of non-linear effects.

[Table 6. Baseline extensions: Functional form]

There is ample evidence suggesting that cognitive and non-cognitive personality traits can reinforce each other in test taking situations (see Borghans et al, 2008 for a survey). It is therefore possible that different teacher abilities influence each other in the actual teaching situation. For these reasons, we in column 3 introduce an interaction term between the cognitive and social ability rankings. We find a positive cross-term indicating that teachers with a high ability to interact socially are particularly productive if also equipped with a high cognitive ability. In fact, the estimates indicate that high cognitive teachers with low social skills are detrimental to student achievement, the breakpoint being at about the median level of social ability.

For all ability measures we use, teachers are ranked according to their abilities at about the age of 18. This begs the question regarding the stability of ability rankings over time. Regarding cognitive ability, there is evidence (Hopkins and Bracht, 1975; Schreuger and Witt, 1989) that the rank-order correlation over time is high and plateaus long before age 18. At the same time, the mean levels of cognitive skills decline substantially with age (Schaie, 1994). The rank-stability of non-cognitive abilities is lower but still substantial, at least when these abilities are evaluated at age 18 (Roberts and DelVecchio, 2000).[23] As opposed to cognitive skills, ability traits such as emotional stability and conscientiousness are increasing rather than declining over time (Roberts et al, 2006).

In *Table 7* we therefore test if the importance of ability rank changes by interacting ranked ability with teacher age. Columns 1-3 do not give any indication that the effect of cognitive and social ability rank would change with age. For female teachers on the other hand, the effect of the personality traits that are captured by the GPA-rank is not stable with age. In effect, early in their career female teachers with higher GPA appear to be more productive than are their low GPA sisters. This productive advantage declines with age, and at around 30 years of age, female teachers with high GPA have become less productive. For male teachers the effect of having a high GPA does not change with age.

[23] The rank-correlation between cognitive tests taken today compared to tests taken ten years ago is about 0.78. The rank-correlation between non-cognitive abilities evaluated with an average time-interval of seven years is about 0.5 at age 18 (see Borghans et al, 2008, figures 5a and 5b).

[Table 7. Baseline extensions: Teacher age]

These results reinforce the earlier complex picture for female teachers with high GPA. With our data, only linking teachers and 9th grade students between the years 2003-2007, we cannot determine whether these female teachers become less productive with age, or if highly productive female teachers with high GPA are more inclined to leave the teachers profession as they grow older. Men with high GPA who choose to become teachers (or not leave the profession), on the other hand, appear to be committed and motivated. Turning to the ability ranks assessed at the draft, our results indicate that a high cognitive rank or social ability rank need not cause better performance for the average student, unless the teacher is ranked high in the joint distribution of these abilities.

5.2 Heterogeneous effects of teacher abilities

As our general evidence is mixed concerning the impact of teachers' abilities, we turn to analyzing if different types of students respond differently to the same teacher abilities. That such heterogeneities may be of importance has previously been suggested by Clotfelter et al. (2006) who document that teachers with stronger math credentials generate larger achievement gains among relatively advantaged students. We find important heterogeneities across student aptitude, foreign background and student gender.

We first examine if the average effects hide heterogeneities along the dimension of students' aptitude, by analysing if the effects vary across students with different (adjusted) middle school GPA. As the observed GPA is endogenous to abilities of the teachers in Swedish, English, and Mathematics—the subjects for which we have test score results—we therefore adjust the GPA measure by dropping all subjects taught by each student's Swedish, English, and Mathematics teachers when calculating students' adjusted GPA-score.[24] It is also important to bear in mind that the main effect of student aptitude is captured by the student fixed effects. Still, there may be spill-over effects across teachers in different subjects, but as long as any potential spill-over has the same

[24] As we drop different subjects for different students, one minor problem with this approach is that the GPA-scores are not fully comparable across students.

effect for students with different aptitude this is not a problem. If, on the other hand, any spill-over effects were larger for high aptitude students our estimates would be lower bounds.[25]

With this caveat in mind, in column 1 of *Table 8* we interact teachers' ranked cognitive and social abilities with students' percentile ranked adjusted GPA. According to this estimate, high-aptitude students will gain from teachers with a high cognitive rank, whereas low-aptitude students will in fact suffer. The point estimates of the direct effect (-0.112) and the interaction effect (0.002) suggests a breakpoint at the 56th student GPA percentile. For social interactive ability the pattern is reversed: the lowest performing students are those who benefit particularly from being matched to teachers with a high social ability rank. For high aptitude students the effect of teachers' social ability is all but negligible.

As a robustness check of these heterogeneities in columns 2 and 3 we allow for more flexibility by splitting data at the student with median GPA. By estimating the effects of cognitive and social ability for high and low GPA students separately we impose little structure on the heterogeneities, at the expense of discarding a substantial amount of the variation in the data. In these flexible specification we, reassuringly, find a positive estimate for the effect of teacher cognitive rank for students with above median GPA (0.029) and a negative estimate for those below the median (-0.028), where the effect estimates are significantly different from each other (p=0.047). Similarly, we find the estimated effect of social rank to be larger for the low GPA sample (0.048) than for the high GPA sample (0.019). Again the effects are significantly different (p=0.023).

[Table 8. Heterogeneous effects for student aptitude]

Moving on to ability captured by teachers' upper-secondary school GPA, we find the lowest performing students to be benefiting particularly from being matched to teachers with a high GPA, whereas high performing students may actually suffer from such a match. In our specification with a linear interaction (column 4) the breakpoint is at the

[25] See footnote 22 for a discussion on potential spill-over effects.

41st aptitude percentile. This pattern is corroborated in the more flexible specification where we split data. We find a significant difference (p=0.055) between the estimated effects in the low aptitude sample (0.003) and the high aptitude sample (-0.021).

We next turn to the question if the effects of teachers' ability ranks vary between students from different academic backgrounds. In *Table 9* the effects of the ability ranks are estimated separately for students coming from a home where both parents hold a university degree (Edu high=1) and from a non-academic home (Edu high=0). Columns 1 and 2 do not indicate that effects of teachers' cognitive rank or social ability rank differ across students' academic background. Similarly, there is no indication that the effect of teachers GPA is asymmetric across students' educational background (columns 3 and 4).

[Table 9. Heterogeneous effects for educational background]

When it comes to students with a foreign background—students born abroad or students whose parents are both born abroad—*Table 10* shows that there are no asymmetries related to the effects of teachers' GPA or cognitive rank. However, students with a foreign background will benefit from being matched to a teacher ranked high on social interactive ability. The point estimate (0.081) is statistically significant, and quite substantial. It indicates that foreign students who are matched to a teacher at a one standard deviation higher position in the social ability distribution on average would have improved their test score performance by 8 percent of a standard deviation. For students without a foreign background, on the other hand, there is no gain from being matched to a teacher with high social rank.

[Table 10. Heterogeneous effects for foreign background]

In *Table 11* we estimate separate effects for girls and boys to see if their test score performance respond differently to having teachers' drawn from different parts of the ability distributions. For the abilities assessed at the draft (columns 1 and 2) we find that boys benefit relative to girls from having teachers with high cognitive rank. The point estimates for girls (-0.030) and boys (0.023) are statistically different (p=0.022). Girls, on the other hand, benefit from having teachers with high social rank, both in absolute terms—the point estimate (0.034) is statistically significant—and also relative to boys (p=0.043). Boys do not appear to benefit from teachers with high social interactive ability.

[Table 11. Heterogeneous effects for female student]

Turning to teacher GPA, boys' school performance seems to suffer from being matched to a high GPA teacher. The effect for boys is negative and statistically significant, whereas the effect for girls is close to zero and insignificant (columns 3 and 4). The difference between these estimates is statistically significant (p=0.056). Interestingly, these overall effects hide important differences between male and female teachers. Female teachers with high GPA-ranks are not good for any students (columns 7 and 8), and particularly bad for boys: the point estimate (-0.037) is statistically significant. In columns 5 and 6 we see that male teachers with high upper-secondary school GPA are good both for girls (0.113) and for boys (0.079). These effects are both statistically significant and not statistically different from each other (p=0.384).

When we also split boys and girls into high and low aptitude groups, based on their adjusted GPA (discussed above), it turns out that male teachers with high GPA's are uniformly good for all four subgroups. Female teachers with high GPA:s are detrimental to the performance of both high and low aptitude boys. They are also worse for high performing girls compared to low performing ones, but the absolute effects are not statistically significant for girls. (See *Table A 2* in the Appendix for these results).

All in all, teachers with higher social interactive ability are particularly good for low-aptitude students and for students with a foreign background, while it may be detrimental for weak students to be matched to a high cognitive teacher. Low aptitude students also benefit from teachers with a high GPA-rank. Boys—despite having a lower average performance than girls—appear to benefit from high cognitive teachers relative to girls. The negative effect from having a female teacher with high GPA—as indicated by the baseline estimates in the previous section—appears to emanate from these teachers being particularly bad for boys and high aptitude girls. Male teachers with high a GPA are, on the other hand, equally productive for students of both genders. This suggests both that usual indicators of human capital need not be indicative of higher productivity in teaching, and that the selection to the teacher profession may differ substantially across gender.

5.3 Consequences of the evolution of the teacher pool

Entering subject teachers in the Swedish middle school are increasingly drawn from lower parts of the distribution of cognitive ability, social interactive ability, and the abilities captured by the upper-secondary school GPA. As seen in section 2, there has been a drastic decline in the position of new teachers for all these ability measures since the early 1990's. The crucial question is whether such a decline in teacher abilities has implications for student achievements.

At first glance this development does not appear to have had any major consequences for the average student, but there turns out to be heterogeneities that are particularly important for certain groups. The gradual decline in social interactive abilities with around 10 percentile ranks has made it relatively more difficult for weak and low achieving students to reach high educational standards. To appreciate the size if this impairment in educational attainment, we can think of the lowest aptitude students being matched to a new subject teacher with 10 rank points lower social ability. This will reduce their expected school performance with 0.7 percentile ranks (2.5 percent of a standard deviation).

The decline in cognitive ability among new teachers with about 20 percentile ranks does not appear to have such harmful consequences for weak students; in fact, for some student groups, teachers with high cognitive ability may have a negative effect on achievement. This said, the decline in teacher cognitive ability has had a detrimental effect on the highest performing students and also appears to have widened the gender gap in student achievement. The highest aptitude students will loose on average 1.8 percentile ranks (7 percent of a standard deviation) in performance if being matched to a new subject teacher with 20 rank point lower cognitive ability, while the lowest performing student will gain 2.2 rank points (9 percent of a standard deviation). Similarly, such a change in teacher cognitive ability will widen the relative difference in performance between male and female students with about 1 percentile rank, in favour of girls.

The marked decline in GPA rank for male and female teachers has very different implications. While the drop in ability rank among male subject teachers entering the teacher profession has been detrimental for student performance—and equally bad for boys and girls—the similar drop in the ability rank among female teachers has been positive for educational attainment, particularly for boys and high aptitude girls. Being matched to a male teacher with 10 rank points lower GPA will decrease average test scores, for both boys and girls, with 0.9 percentile ranks (3 percent of a standard deviation). A similar reduction in GPA for female teachers will instead increase test scores for boys with 0.4 percentile ranks.

The picture is complex, and suggests that it is difficult to draw a general conclusion about malign consequences of the successive decline in ability rank among teachers. For some student groups this development may have been harmful, while others may have benefited.

6 Conclusions

In this paper we document a marked decline in teacher abilities, regardless of how these abilities are measured. Over a 15 year period, the average cognitive ability among new teachers has declined by about 20 percentile ranks in the Swedish middle school. For non-cognitive social interactive ability and upper-secondary GPA the decline is between 10-15 percentile ranks. Even if we lack data for women in some ability dimensions, our results indicate that the decline is—if anything—even more dramatic among female teachers than among males.

The main findings of this paper are that this decline has had small effects for the average student, but that this hides important asymmetries. The decline in teacher non-cognitive social interactive ability has had a negative impact for low achieving students, as have the decline in general abilities captured by teachers' GPA. Our results also indicate that while high-ability students benefit from being matched to a high cognitive teacher, such a match is even detrimental for lower achieving students. It further appears as if the gender gap in school results can in part be explained by boys being harmed relative to girls by the decline in teacher cognitive ability. On the other hand, the decline in social ability has had the opposite effect on gender differences. In a broad sense, these findings support the conclusion reached by Clotfelter et al. (2006) that it is not just the average teacher ability that matters for student outcomes, but also how students and teachers are matched.

Our results further indicate important gender differences among the teachers. Changing to a male subject teacher whose GPA rank is 10 percentiles lower will on average reduce test scores by 0.9 percentile ranks. For female teachers, the same reduction in GPA-rank would *increase* test scores by 0.4 percentile ranks. Female teachers with high GPA-scores appear to be particularly detrimental to the performance of boys.

These findings suggest that school grades may capture different capacities for men and women, as suggested by Lindahl (2007), or that there are other important differences between men and women when selecting into the teacher profession. For example, as a high GPA implies a wide variety of career opportunities, one interpretation is that men who actively forego these opportunities are highly motivated teachers. For some reason, the same mechanisms do not appear to be present among men and women. Yet another possibility is that schools are particularly bad at motivating and retaining female teachers with high GPA-scores. An important venue for future research is to understand these gender differences. Perhaps such an understanding can help finding the right ways to attract, and screen, the teacher candidates, as well as keeping teachers motivated throughout their career.

Our results indicate that a general increase in teacher cognitive abilities would increase the achievement gap between high and low-performing students, both by raising the achievement of high-performing students and lowering it for low-performing ones. It is therefore difficult to draw any general conclusions regarding the desirability of policies aimed at attracting high-cognitive individuals to the teaching profession. Policies aimed at raising teacher cognitive abilities can be put into question from a different perspective as well. After all, cognitive skills can be put to good use elsewhere in the economy, since high-cognitive individuals have higher earnings (see for example Heckman et al, 2006, Lindqvist and Vestman, 2008). Thus, any policy aimed at attracting high-cognitive individuals to the teaching profession must consider the alternative costs such a policy involves. As we find zero average effects on student outcomes by teacher cognitive ability, the objective function of the policy maker must be skewed towards the highest performing students to make such a policy welfare improving. Of course, it is important to keep in mind that our study is silent concerning the potential effects of the decline in teacher abilities on long-term outcomes such as educational choices and earnings.

In sum, the picture on what abilities are productive for teachers is complex, and it is difficult to draw conclusions on the desirability of having teachers from the upper part of the overall ability distribution; that is, equating teacher quality with measures of human capital like cognitive and non-cognitive abilities seems questionable. What our results clearly show, however, is that the process matching students to teachers is important, and that, given a suitable teacher, there may be positive aspects to segregating students along the ability dimension. The teacher who is good for the best is not necessarily good for the rest.

References

Bacolod, M. (2007), "Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality", *Review of Economics and Statistics*, 89(4), 737-751.

Björklund, A., M. Clark, P.-A. Edin, P. Fredriksson, and A. Krueger (2005), *The market comes to education in Sweden: An evaluation of Sweden's surprising school reforms*, Russel Sage Foundation, New York.

Borghans, L., A.L. Duckworth, J.J. Heckman and B. ter Weel (2008), "The economics and psychology of personality traits", NBER Working Paper #13810.

Carlstedt, B. (2000), *Cognitive abilities: aspects of structure, process and measurement*, Ph.D. dissertation at the Department of Psychology, Gothenburg University. Göteborg: Acta Universitatis Gothoburgensis.

Clotfelter, C., H. Ladd and J. Vigdor (2006), "Teacher-student matching and the assessment of teacher effectiveness", *Journal of Human Resources*, 41(4), 778-820.

Corcoran, S., W. Evans, and R. Schwab (2004), "Changing labor-market opportunities for women and the quality of teachers 1957-2000", *American Economic Review*, 94, 230-235.

Cunha, F. and J. Heckman (2008), "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation", *Journal of Human Resources*, 43(4), 738-782.

Economist, The (2007), "The race is not always to the richest", December -6.

Ehrenberg R.G. and D.J. Brewer (1995), "Did teachers' verbal ability and race matter in the 1960s? *Coleman* revisited", *Economics of Education Review*, 14(1), 1-21.

Ferguson, R.F. and H.F. Ladd (1996), "How and why money matters: An analysis of Alabama schools". In H.F. Ladd (ed.), *Holding Schools Accountable: Performance-based Reform in Education*. Brookings Institution: Washington DC.

Flynn, J. (1984), "The mean IQ of Americans: Massive gains 1932-1978", *Psychological Bulletin*, 95, 29-51.

Fredriksson, P. and B. Öckert (2008), "The supply of skills to the teacher profession", mimeo Uppsala.

Hanushek, E. (1992), "The trade-off between child quantity and quality", *Journal of Political Economy*, 100(1), 84-117.

Hanushek, E. and S. Rivkin (2006), "Teacher quality" in *Handbook of the Economics of Education, Volume 2*, Elsevier.

Heckman, J.J., J. Stixrud and S. Urzua (2006), "The effects of cognitive and noncognitive abilities on labor market outcomes and social behaviour", *Journal of Labor Economics* 24(3), 411-482.

Hopkins, K.D. and G.H. Bracht (1975), "Ten-year stability of verbal and nonverbal IQ scores", *Amercan Educational Research Journal* 12(4), 469-477.

Hoxby, C. and A. Leigh (2004), "Pulled away or pushed out? Explaining the decline of teacher aptitude in the United States", *American Economic Review Papers and Proceedings*, 94(2), 236-240.

Lakdawalla, D. (2006), "The economics of teacher quality", *Journal of Law and Economics*, XLIX, 285-329.

Leigh, A. and C. Ryan (2006), "How and why has teacher quality changed in Australia?", *Australian Economic Review*, 41(2), 141-159.

Lindahl, E. (2007), "Comparing teachers' assessments and national test results - evidence from Sweden", IFAU working paper 2007:24

Lindahl, M. (2001), "Summer learning and the effect of schooling: Evidence from Sweden", IZA discussion paper 262.

Lindqvist, E. and R. Vestman (2008), "The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment battery", mimeo IFN.

Neal, D. and W. Johnson (1996), "The role of pre-market factors in black-white wage differences", *Journal of Political Economy*, 104(5), 869-895.

Nickell, S. and G. Quintini (2002), "The consequences of the decline in public sector pay in Britain: A little bit of evidence", *The Economic Journal*, 112, F107-F118.

McKinsey & Company (2007), *How the Word's Best-Performing School Systems Come Out On Top*, McKinsey Corporation.

Ministry of Education and Research (2000), Education Act 1985: 1100, http://www.sweden.gov.se/content/1/c6/02/15/38/1532b277.pdf (2007-10-15).

Rivkin, S., E. Hanushek, and J. Kain (2005), "Teachers, schools, and academic achievement", *Econometrica*, 73, 417-458.

Roberts, B.W. and W.F. DelVeccho (2000), "The rank-order consistency of personality traits from childhood to old age", *Psychological Bulletin* 126(1), 3-25.

Roberts, B.W., K.E. Walton, W. Viechtbauer (2006), "Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal data", *Psychological Bulletin* 132(1), 1-25.

Rockoff, J. (2004), "The impact of individual teachers on student achievement: Evidence from panel data", *American Economic Review*, 94, 247-252.

Schaie, K.W. (1994), "The course of adult intellectual development", *American Psychologist* 49(4), 304-313.

Schreuger, J.; and A.C. Witt (1989), "The temporal stability of individually tested intelligence", *Journal of Clinical Psychology*, 45(2), 294-302.

Skolverket (2004), "Allmänna råd och kommentarer: Likvärdig bedömning och betygsättning", http://www.skolverket.se/publikationer?id=1368 (2008-10-22)

Skolverket (2006), *Curriculum for the compulsory school system, the pre-school class and the leisure-time center*, Lpo 94, Fritzes, Stockholm. http://www.skolverket.se/publikationer?id=1070 (2007-12-03).

Skolverket (2007), "Fact sheet: Compulsory school", http://www.skolverket.se/content/1/c4/12/15/en_grundskola.pdf (2007-10-15).

Wayne, A.J. and P. Youngs (2003), "Teacher characteristics and student achievement gains: A review", *Review of Educational Research*, 73(1), 89-122.

Appendix

Table A 1. Baseline OLS estimates with school fixed effects and student controls

	(1)	(2)	(3)	(4)	(5)	(6)
Teacher sample	Draft	Draft	Draft	All GPA	Male GPA	Fem GPA
	Test score	Test score	Test score	Test score	Test score	Test score
Cognitive	0.0068		0.0071			
	(0.0174)		(0.0175)			
Social		-0.0053	-0.0055			
		(0.0121)	(0.0121)			
Teacher GPA				0.0230**	0.0530***	0.0146
				(0.0090)	(0.0189)	(0.0109)
					[0.08]	
Observations	28378	28378	28378	67266	19552	47714
# students	23692	23692	23692	43322	16871	33863
# teachers	703	703	703	1587	498	1089
# schools	202	202	202	224	189	219
R-squared	0.17	0.17	0.17	0.16	0.18	0.16

Note: School and time period fixed effects and controls for *academic* home, student *gender* and *foreign background* always included, as well as biennial teacher birth cohort dummies. In (4) a teacher gender dummy is included. Robust standard errors, clustered by teacher, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A 2. Heterogeneous effects for student and teacher gender by student aptitude for teacher GPA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: Percentile ranked test scores							
Teacher sample	Fem GPA	Fem GPA	Fem GPA	Fem GPA	Male GPA	Male GPA	Male GPA	Male GPA
Student sample	Girls	Girls	Boys	Boys	Girls	Girls	Boys	Boys
	High GPA	Low GPA	High GPA	Low GPA	High GPA	Low GPA	High GPA	Low GPA
Teacher GPA	-0.0198	0.0215	-0.0272*	-0.0458***	0.1427***	0.0791**	0.1080*	0.1016***
	(0.0162)	(0.0158)	(0.0156)	(0.0142)	(0.0468)	(0.0351)	(0.0555)	(0.0330)
	[0.041]		[0.299]		[0.149]		[0.900]	
Observations	14233	10131	10578	14733	6110	3896	4262	6031
# students	10044	7323	7339	10638	5267	3433	3666	5230
# teachers	1022	1024	1018	1069	449	458	438	486
R-squared	0.85	0.87	0.84	0.86	0.93	0.95	0.92	0.93

Note: The dependents variable is student test scores in Swedish, English, and Mathematics. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, time period, and teacher upper-secondary school program. [a] Is the p-value from testing for equality of coefficients between samples. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Tables

Table 1. Correlation between cognitive abilities and social ability

	GPA	Cognitive ability
Cognitive ability	0.47	
Social ability	0.27	0.36

Note: The number of observations is 633149 for the GPA-cognitive correlation, 597307 for the GPA-social correlation, 423743 for the GPA-leader correlation, 1450084 for the cognitive-social correlation, 938021 for the cognitive-leader correlation, and 938364 for the social-leader correlation.

Table 2. The evolution of abilities of new subject teachers (ages 25-30)

	(1)	(2)	(3)	(4)	(5)
	Average teacher ability				
	Cognitive	Social	GPA	GPA men	GPA wom
Trend	-0.577***	-0.342***	-0.656***	-0.591***	-0.757***
	(0.094)	(0.067)	(0.074)	(0.183)	(0.094)
Diff-trends				-0.166	
				(0.206)	
Observations	28	28	14	14	14
R-squared	0.64	0.58	0.84	0.60	0.86

Note: Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 3. The evolution of abilities of new male and female subject teachers (ages 25-30)

	(1)	(2)	(3)	(4)
	Teacher sibling abilities			
	Cognitive		Social	
	Men	Women	Men	Women
Trend	-0.289***	-0.361***	-0.050	-0.170***
	(0.065)	(0.081)	(0.063)	(0.054)
Diff trends	-0.072		-0.120	
	(0.104)		(0.083)	
Observations	28	28	28	28
R-squared	0.41	0.54	0.02	0.33

Note: Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Diff trends is the estimated differences in the trend coefficients for male and female teachers. Robust standard errors in parentheses. Regressions in columns (3)-(6) are based on the brothers of male and female teachers.

Table 4. Summary statistics

Sample	(1) GPA	(2) Draft
Student level		
Test score	48.0	45.7
	(25.6)	(26.9)
GPA9	53.2	53.6
	(27.6)	(27.6)
Academic home	0.12	0.12
	(0.32)	(0.33)
Foreign background	0.21	0.21
	(0.41)	(0.41)
Girl	0.49	0.49
	(0.50)	(0.50)
No of students	45428	24847
Teacher level		
GPA teacher	63.5	
	(23.1)	
Social ability		54.1
		(27.9)
Cognitive ability		64.5
		(22.1)
Age	32.8	38.7
	(3.9)	(8.1)
Female teacher	0.69	
	(0.46)	
No of teachers	1589	704

Note: Mean values of all variables and standard deviations in parentheses

Table 5. Baseline within student estimates

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome: Percentile ranked test scores			
Teacher sample	Draft	Draft	Draft	All GPA	Male GPA	Female GPA
Cognitive	-0.0036		-0.0034			
	(0.0204)		(0.0202)			
Social		0.0199	0.0199			
		(0.0158)	(0.0158)			
Teacher GPA				-0.0104	0.0928***	-0.0227**
				(0.0073)	(0.0278)	(0.0106)
					[0.0001][a]	
Observations	29749	29749	29749	70305	20505	49800
# students	24847	24847	24847	45428	17710	35422
# teachers	704	704	704	1589	498	1091
R-squared	0.93	0.93	0.93	0.85	0.94	0.88

Note: The dependents variable is student test scores in Swedish, English, and Mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (4)-(6) also include teacher upper-secondary school program fixed effects, and column (4) a teacher gender indicator. [a] Is the p-value from testing for equality of coefficients between samples. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Table 6. Baseline extensions: Functional form

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: Percentile ranked test scores					
Teacher sample	Draft	Draft	Draft	All GPA	Male GPA	Fem GPA
Student sample	All	All	All	All	All	All
Cognitive	-0.0649		-0.0855*			
	(0.1120)		(0.0439)			
Cognitive^2	0.0005					
	(0.0009)					
Social		-0.0064	-0.0864			
		(0.0645)	(0.0552)			
Social^2		0.0003				
		(0.0006)				
Cognitive×Social			0.0016**			
			(0.0008)			
Teacher GPA				0.0339	0.2570*	0.0686
				(0.0400)	(0.1539)	(0.0549)
Teacher GPA^2				-0.0004	-0.0014	-0.0007*
				(0.0003)	(0.0013)	(0.0004)
Observations	29749	29749	29749	70305	20505	49800
# students	24847	24847	24847	45428	17710	35422
# teachers	704	704	704	1589	498	1091
R-squared	0.93	0.93	0.93	0.85	0.94	0.88

Note: The dependents variable is student test scores in Swedish, English, and Mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (5)-(7) also include teacher upper-secondary school program fixed effects, and column (5) a teacher gender indicator. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Table 7. Baseline extensions: Teacher age

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: Percentile ranked test scores					
Teacher sample	Draft	Draft	Draft	All GPA	Male GPA	Fem GPA
Student sample	All	All	All	All	All	All
Cognitive	-0.0324		-0.0199			
	(0.0795)		(0.0801)			
Cognitive×Age	0.0007		0.0004			
	(0.0020)		(0.0020)			
Social		0.0229	0.0221			
		(0.0818)	(0.0815)			
Social× age		-0.0001	-0.0001			
		(0.0020)	(0.0020)			
Teacher GPA				0.1713**	0.1625	0.1512*
				(0.0684)	(0.2041)	(0.0907)
Teacher GPA× Age				-0.0054***	-0.0020	-0.0052*
				(0.0020)	(0.0058)	(0.0027)
Observations	29749	29749	29749	70305	20505	49800
# students	24847	24847	24847	45428	17710	35422
# teachers	704	704	704	1589	498	1091
R-squared	0.93	0.93	0.93	0.85	0.94	0.88

Note: The dependents variable is student test scores in Swedish, English, and Mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (4)-(6) also include teacher upper-secondary school program fixed effects, and column (4) a teacher gender indicator. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Table 8. Heterogeneous effects for student aptitude

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome: Percentile ranked test scores					
Teacher sample	Draft	Draft	Draft	All GPA	All GPA	All GPA
Student sample	All	High GPA	Low GPA	All	High GPA	Low GPA
Cognitive	-0.1116***	0.0285	-0.0275			
	(0.0401)	(0.0242)	(0.0248)			
Cognitive× student GPA	0.0020***	[0.047] [a]				
	(0.0006)					
Social	0.0671**	0.0021	0.0475***			
	(0.0268)	(0.0188)	(0.0177)			
Social× student GPA	-0.0008**	[0.023] [a]				
	(0.0004)					
Teacher GPA				0.0328**	-0.0206**	0.0030
				(0.0130)	(0.0096)	(0.0082)
Teacher GPA× student GPA				-0.0008***	[0.055] [a]	
				(0.0002)		
Observations	29465	14711	14754	69974	35183	34791
# students	24662	12315	12347	45258	22214	23044
# teachers	698	647	692	1583	1516	1572
R-squared	0.93	0.91	0.91	0.85	0.79	0.81

Note: The dependents variable is student test scores in Swedish, English, and Mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (4)-(6) also include teacher upper-secondary school program fixed effects, and column (4) a teacher gender indicator. In columns (1) and (4), teacher abilities are interacted with student GPA calculated using the subjects not taught by the Swedish, English, or mathematics teachers. The student sample is split according to the median value of this GPA. [a] Is the p-value from a t-test of equality of coefficients between samples. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Table 9. Heterogeneous effects for educational background

	(1)		(2)	(3)		(4)
			Outcome: Percentile ranked test scores			
Teacher sample	Draft		Draft	All GPA		All GPA
Student sample	Edu high=1		Edu high=0	Edu high=1		Edu high=0
Cognitive	-0.0185		0.0015			
	(0.0381)		(0.0219)			
		[0.629] [a]				
Social	0.0040		0.0192			
	(0.0260)		(0.0165)			
		[0.565] [a]				
Teacher GPA				0.0062		-0.0137*
				(0.0134)		(0.0078)
					[0.153] [a]	
Observations	3427		24951	8029		59237
# students	2908		20784	5094		38228
# teachers	546		703	1306		1587
R-squared	0.93		0.93	0.81		0.84

Note: The dependents variable is student test scores in Swedish, English, and mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (3) and (4) also include teacher upper-secondary school program fixed effects and a teacher gender indicator. The student sample is split according to their parents' educational attainment. *Edu high* equals one if both parents have some level of post-secondary education, zero otherwise. [a] Is the p-value from a t-test of equality of coefficients between samples. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Table 10. Heterogeneous effects for foreign background

	(1)		(2)	(3)		(4)
			Outcome: Percentile ranked test scores			
Teacher sample	Draft		Draft	All GPA		All GPA
Student sample	Foreign=1		Foreign=0	Foreign=1		Foreign=0
Cognitive	-0.0171		0.0013			
	(0.0339)		(0.0210)			
		[0.606] [a]				
Social	0.0809***		0.0064			
	(0.0243)		(0.0160)			
		[0.006] [a]				
Teacher GPA				-0.0067		-0.0097
				(0.0138)		(0.0077)
					[0.835] [a]	
Observations	6284		23465	13632		56673
# students	5242		19605	9435		35993
# teachers	633		694	1456		1570
R-squared	0.93		0.93	0.87		0.84

Note: The dependents variable is student test scores in Swedish, English, and mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (3) and (4) also include teacher upper-secondary school program fixed effects and a teacher gender indicator. The student sample is split according to their background. *Foreign* equals one if either the student or both parents are born abroad. [a] Is the p-value from a t-test of equality of coefficients between samples. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

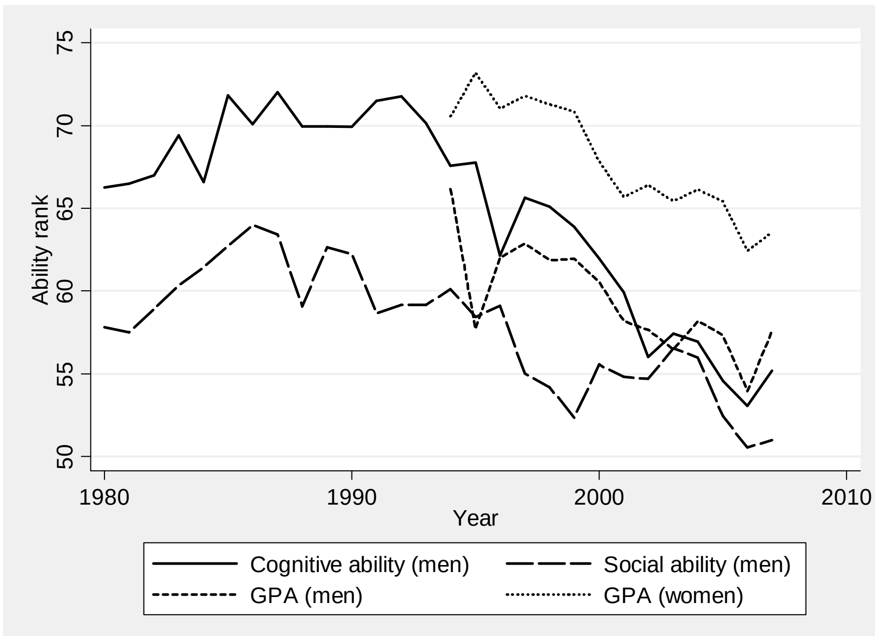
Table 11. Heterogeneous effects for female student

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Outcome: Percentile ranked test scores					
Teacher sample	Draft	Draft	All GPA	All GPA	Male GPA	Male GPA	Fem GPA	Fem GPA
Student sample	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys
Cognitive	-0.0296	0.0233						
	(0.0224)	(0.0245)						
	[0.022] [a]							
Social	0.0345**	-0.0003						
	(0.0175)	(0.0183)						
	[0.043] [a]							
Teacher GPA			-0.0006	-0.0190**	0.1134***	0.0786**	-0.0037	-0.037***
			(0.0085)	(0.0089)	(0.0326)	(0.0351)	(0.0128)	(0.0120)
			[0.056] [a]		[0.384] [a]		[0.010] [a]	
Observations	14558	15191	34523	35782	10101	10404	24422	25378
# students	12151	12696	22340	23008	8757	8953	17403	18019
# teachers	679	692	1553	1575	480	493	1073	1082
R-squared	0.94	0.93	0.86	0.85	0.95	0.94	0.89	0.88

Note: The dependents variable is student test scores in Swedish, English, and Mathematics. *Cognitive* is the teacher's percentile ranked cognitive ability from the military draft. *Social* is the teacher's percentile ranked social interactive ability from the military draft. *Teacher GPA* is the teacher's percentile ranked upper-secondary GPA. Control variables include fixed effects for teacher biennial birth cohort, subject, student, and time period. Columns (3)-(8) also include teacher upper-secondary school program fixed effects, and columns (3)-(4) a teacher gender indicator. The student sample is split according to gender. [a] Is the p-value from a t-test of equality of coefficients between samples. Robust standard errors in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%, Standard errors are clustered by teacher.

Figures

Figure 1. Ability ranks of new subject teachers (ages 25-30), 1980-2006



Note: The graph plots the average cognitive and social interactive abilities, as well as the average GPA rank of all new middle school subject teachers ages 25-30 in the Swedish teacher register.