I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17815

Long-Term Effects of Phonics Approaches to the Development of Reading Skills

Carsten Elbro
Nicolai Kristensen
Peter Rohde Skov

MARCH 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17815

Long-Term Effects of Phonics Approaches to the Development of Reading Skills

Carsten Elbro
University of Copenhagen

Nicolai Kristensen
VIVE - The Danish Centre of Applied Social Science and IZA

Peter Rohde Skov
The Rockwool Foundation Intervention Unit

MARCH 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Long-Term Effects of Phonics Approaches to the Development of Reading Skills

Many studies have documented that a phonics approach is an essential component in the teaching of reading for beginning readers, especially for students at risk of reading disabilities. We study whether phonics approaches, as indicated by the choice of basal readers (materials for the initial teaching of reading) in Grade 1, have long-term effects on adult education levels and labour market outcomes 23 years later. The study shows evidence of a causal link between the choice of basal reader in primary school and school results at the end of lower secondary school as well as labour market outcomes at age 30. Choosing a basal reader that is easily compatible with a phonics approach is a highly cost-effective policy choice.

JEL Classification: H52, I28, I38, J13
Keywords: basal reader, long-term outcomes, reading

Corresponding author:
Nicolai Kristensen
VIVE - The Danish Centre of Applied Social Science
Herluf Trolles Gade 11
1052 Copenhagen K
Denmark
E-mail: nikr@vive.dk

1. Introduction

In this paper, we investigate whether the choice of basal reader, i.e., materials for the initial teaching of reading, have long-term impacts on the students' level of education and labour market outcomes measured at age 30. In particular, we investigate whether basal readers that are highly compatible with a phonics approach have a long-term impact on the students. Following a phonics approach, beginning readers are explicitly taught to identify speech sounds and to associate letters with these sounds and to blend the identified speech sounds into words. Many studies, e.g., Machin et al. (2018), have reported positive short- and medium-term effects of phonics approaches. To the best of our knowledge, we provide the first (very) long-term study of the importance of phonics approaches.

Literacy is arguably the most important skill learned in school, but although some progress has been made over the past decades, the proportion of poor readers remains above 10 percent at the end of compulsory schooling, (OECD, 2012). Poor reading ability negatively affects children's experience of time in school, limit their level of further education, and can lead to a poor working life with limited prospects and early retirement. Hence, poor literacy is often associated with low social mobility, low-skilled jobs, and unemployment (Vignoles et al. 2011, Cherry and Vignoles, 2020). Poor literacy is, therefore, immensely costly for the individuals as well as for societies.

A vast literature has been devoted to studies of how to improve the teaching of reading of students in primary school (e.g. Rose 2006). A key finding from this research is that teaching methods matter. In particular, there is evidence in favour of phonics approaches, i.e., teaching methods that focus explicitly on letter-sound connections and how to employ them to decode written words (Machin et al. 2018, Slavin et al. 2009). Especially for struggling readers, it is important that initial teaching contains an element of explicit instruction in the alphabetic code

in addition to a number of other elements (Brady, 2011; National Reading Panel, 2000; National Early Literacy Panel, 2008; Rose, 2006). This means that teaching should comprise examples and ample practice with the basic letter-sound correspondences and with blending sounds into words. In short, it is helpful for beginning readers to be explicitly taught how to make use of the basic alphabetic principle of the orthography to identify written words.

Phonics approaches are closely linked to the choice of basal reader. Typically, a basal reader will not only provide texts with a vocabulary that is suitable for reading practice, but will also be tailored for a sequence of instructions and activities for students. 'Basal reader' is here short for a whole set or package of published materials for teaching initial reading, usually comprising a reading book, one or more work books, a teacher's guide and additional easy materials for independent reading. The basal reader is thus likely to have a significant influence on both teaching contents and methods (cf. Machin et al. 2018).

Several countries have made a phonics approach mandatory. For instance, phonics was implemented in the US through the federal Reading First program following evidence-based recommendations by the US National Reading Panel (National Reading Panel 2000, Machin et al. 2018). In the UK, the National Literacy Strategy from 2006 followed the recommendations of a "synthetic phonics"[2] approach given by the Rose report (2006). In several other countries where the orthography is much more regular than the English (or Danish), phonics have been the dominant approach for generations. Examples are Finland, Greece, Italy, Germany and the Netherlands. The variation in and debate about initial teaching methods is almost exclusively found in countries where the orthography is less regular, that is, with many deviations from simple letter-to-sound relationships. Denmark is such an example (Seymour et al., 2003; Elbro,

[2] Rose (2006) describes how the National Literacy Strategy, first implemented in 1998, resulted in major changes to the teaching of literacy. "The Strategy engaged schools in developing a structured teaching programme of literacy that included not only what phonic content should be taught but also how to teach it, with a subsequent rise in standards", Rose (2006:3).

2005). Denmark continues to differ from other countries with irregular orthographies in that phonics approaches are *not* mandatory. In Denmark, methods for the initial teaching of reading are entirely the teachers' choice.

An important limitation of the literature is that the existing studies of phonics approaches have been relatively short in time, with only a few follow-up studies at the completion of lower secondary (ninth grade) school (Johnston and Watson 2005). Machin et al. (2018) evaluate the short-medium-term impact of a change in UKs reading strategy towards more emphasis on "synthetic phonics". They find positive initial effects on average test scores at age five. At age 11, the average effect has faded away while it remains significant for disadvantaged children (low SES and/or non-native speakers).

To the best of our knowledge, no published study has looked at very long term, adulthood outcomes of phonics approaches, e.g., in terms of education and labour market outcomes. This study aims to fill this gap.

In a comprehensive assessment of the importance of a range of basal readers in Denmark, Borstrøm, Petersen & Elbro (BPE) (1999) followed and repeatedly tested all pupils from the 1st to the 3rd grade across 124 schools, roughly 3,500 pupils - from 1995 to 1998. They also interviewed the pupils' teachers and collected detailed information about the basal readers. Focusing on the five most popular readers at the time, BPE found that one basal reader in particular was characterized by a focus on the direct instruction in letter-sound connections and sound blending. Use of this basal reader was found to be associated with significantly fewer participants being categorized as "poor readers", i.e., pupils with reading scores below the 10th percentile across all participants. The BPE (1999) report established that, for students with poor prerequisite skills, it is important that the teaching of reading comprises direct instruction in the alphabetic code, covering the basic links between letters and their sounds and how to

blend sounds into words. The most supportive basal reader was also found to comprise predominantly regularly spelled words that allowed students to practice the basic letter-sound relationships.

The present paper reports from a follow-up study of BPE (1999). We estimate the impact of basal readers for students who started in first grade in one of the 124 schools in 1995. These children were born in 1988 and we were able to follow them in the Danish administrative registries up until 2018, i.e. the year they turned 30.

Even though BPE (1999) found the expected short-term effects associated with basal readers such effects may cease to exist in the longer run. Some well-known and thoroughly studied US initiatives may spur some optimism though. For instance, the Student Teacher Achievement Ratio (STAR) initiative, which reduced the class size for randomly selected classes, resulted in positive long-run outcomes measured at age 27, (Chetty et al. 2011)[3]. Similarly, the Head Start program, which promotes school readiness, has also resulted in positive long-run effects in adulthood (see Deming 2009). The Chetty et al. (2011) and Deming (2009) studies establish that improvements in short-run test scores translate into improvements in adult labour market outcomes.[4] We investigate whether a similar long-run impact can be found for the choice of basal reader in primary school. In the absence of a

[3] Early studies of STAR are reviewed in Schanzenbach (2006). Generally, the effects remained up to third grade and were especially high for black students.

[4] There are several other studies of programs and interventions that have shown long-term effects, see fx. Fredriksson et al. (2013) on long-term effects of class size, and a series of papers written by James J. Heckman and co-authors on the Perry Pre School program (e.g. Baulos et al. 2024), an experimental study of an early childhood education program. Heckman & Garcia (2023) on the Carolina Abecedarian Project (see also Ramey & Campbell 1984). Carneiro & Ginja (2014) investigate the medium to long-term impact of HEAD start on health and behavioral problems and find long-term effects when the children have become of age 12-13. Hoynes et al. (2016) also investigate long-term impact of program changes in childhood. They analyze the introduction of the Food Stamp Program, rolled out across counties between 1961 and 1975, and find that access to food stamps in childhood leads to a significant reduction in the incidence of metabolic syndrome (related to obesity) and, for women, an increase in economic self-sufficiency.

randomized experiment or a natural experiment, we lay out in detail reasons why the links between basal reader and outcomes 25 years later may nonetheless be interpreted as causal.

We find that students who in grade one was taught how to read using the most "phonics-compatible" basal reader obtain better outcomes both in terms of test scores in the short-term (BPE's original study), medium-term test scores in Danish measured at the end of the ninth grade (upper secondary exam), and at age 30 in terms of length of education. Applying an unconditional quantile regression approach, we establish that the average results are driven entirely by improvements in the lower echelons of the distribution of the various long-term outcomes, and, when adapting quantile regression, we also find statistically significant effects of the phonics-compatible basal reader in the lower end of the distribution of income and fulltime employment at age 30.

Choosing the best basal reader has very low (or no) costs so any positive long-run impact will inevitably imply that choosing a phonics-compatible basal reader is highly cost effective. This is in sharp contrast to much more costly initiatives usually analysed in social science, such as reduced class size (Krueger 1999, Angrist & Lavy 1999, Heinesen 2010), increased instruction time (Lavy 2015, Rivkin & Schiman 2015, Andersen et al. 2016), two-teacher classes (Andersen et al. 2020, Andersen et al. 2025), and special provisions for students with learning disabilities (Schwartz et al. 2021).

2. Some background on basal readers and phonics instruction

Typically, the basal reader will not only provide the texts and words for reading practice, but will also be tailored for a sequence of instructions and activities for students. The basal reader will thus have a potentially significant influence on both teaching contents and methods (in line with Machin et al. 2018).

As mentioned, research on reading has established that, especially for struggling readers, it is important that the initial teaching has an element of direct instruction in the alphabetic code ('phonics' instruction) in addition to a number of other elements of teaching (Brady, 2011; National Reading Panel, 2000; National Early Literacy Panel, 2008; Rose, 2006). Some basal readers are more suitable for such instruction and practice than others. A basal reader that lends itself well to phonics instruction is one with which students can apply the most common letter-sound correspondences to short, common, and regularly spelled words that are repeated many times through the text. With such a text, the beginner can practice letter-sound correspondences while being able to read independently from an early point.

A phonics-compatible basal reader would thus be expected to be beneficial for students of all abilities, because all students need to acquire proficiency with letter-sound correspondences in order to attain a robust foundation for development of further abilities in reading and writing. Indeed, a phonics-compatible basal reader is expected to be particularly supportive for students at risk for reading disabilities who need more explicit instruction and practice with this foundation.

In the first half of the 1990s, the predominant view on basal readers in Denmark, as in most other countries, was the same as it had been at least since the 1960's – that children should be taught multiple strategies for decoding written words. The view was described by Mogens Jansen, the long-time president of the Danish reading teachers' association, as "the teaching of reading without really any method" in a report on reading instruction in Denmark (Jansen, Jacobsen, and Jensen 1978), and elsewhere as an eclectic or mixed cue method (Dalby et al., 1983; 1989). Others have described it as a "searchlight" method (e.g. Rose 2006) by which children are taught to decode words by means of multiple cues, such as context (including illustrations), prior knowledge of content, whole word shapes, orthographic knowledge, and

letter-sound correspondences. The idea was that the beginning reader would (and should) use whatever cue to try to identify printed words, similarly to the many strategies one may apply to solve a rebus.

This dominant searchlight or mixed cue method was *not* criticized in Denmark in the early 1990s for making all kinds of cues equal and thus failing to recognise the primacy of the letter-sound correspondences. [5]

Following the searchlight method, all the available basal readers in Denmark in the 1990s would contain richly illustrated narratives about children in more or less familiar situations and surroundings. This would provide students with context cues for guessing words rather than recognising them via their letter-sounds. All of the basal readers would also support teaching of at least some "sight words", i.e., words to be recognised as wholes without any encouragement to "spell out" the words letter by letter.

While basal readers that were fully compatible with a phonics approach were still unavailable in Denmark in 1995, some of the basal readers available were more compatible with a phonics approach than other basal readers (BPE, 1999). The more phonics-compatible basic reader would be one with a majority of short, easy-to-decode words with high-frequent letter-sound correspondences, and with many repetitions of these words in rather short sentences. A few words would always be presented as wholes, e.g. the names of the main characters, and a few irregularly spelled, very high-frequency words, corresponding to words like English *was* ["wæz"] and *are* ["ar"]. Examples of such basal readers from English are *Peter and Jane*, and *The Alison books* in the US.

[5] Though it was noted with some concern that a high proportion (more than one in four) of Danish students took at very long time (more than three years) to learn to decode written words fluently. This fact was widely referred to in Denmark following the first international comparison of students' reading (Elley 1992).

At the beginning of the 1990's, such phonics-compatible materials were simply viewed as very old-fashioned in Denmark. There was no knowledge available to teachers about which basal readers were better compatible with any particular, well-researched method of instruction. This has important implications for the econometric identification of estimated parameters, cf. section 4.

3. Data

3.1 The basal readers of 1995

In the estimations we distinguish between four types of basal reader that were all common in 1995.

First, we have chosen a basal reader based on the *searchlight method* as the reference. By 1995, 'The Danish Books' (*Danskbøgerne*) (Lau 1987) was the most typical basal reader of the searchlight tradition from the 1960's and suitable as a reference in the estimations.[6]

Second, as mentioned above, basal readers that were fully compatible with a phonics approach were unavailable in Denmark in 1995. However, the basal reader that was *the most compatible with a phonics* approach was *Soren and Mette* (*Søren and Mette*, Hermansen and Jensen 1954). It was also the oldest available basal reader at the time. With *Soren and Mette*, students would first practice letter names and letter-sounds under the teacher's instruction and guidance *before* reading any words or texts in the main textbook. Students would know all the relevant letters, their shapes and sounds enabling them to read the first pages of their basal reader on their own or with minimal support. In comparison with the other basal readers of the

[6] It replaced the earlier basal reader 'The house Højbo' (*Huset Højbo*, Bjerg and Jansen, 1961) which had dominated the market in the 1970's and early 1980's. Jansen, who was the main author of *The house Højbo*, put together the group of authors and continued as the editor of *The Danish books*.

time, *Soren and Mette* employed words that were very short (often just one or two letters, *i*, *ø* and *å* are words in Danish), of very high frequency, and each with many repetitions (Borstrøm et al., 1999). Consequently, *Soren and Mette* was singled out as a *phonics-compatible* basal reader in the present study

Third, a special case was presented by the basal reader *Helhedslæsning* (Frost, 1983), literally 'Holistic Reading' but in effect '*Integrated phonemic awareness*'.[7] The basic idea was to help beginning readers become 'phonemically aware', that is, aware of the sound segments of speech (the phonemes) that correspond to the individual letters. This awareness is important for the beginning reader as a basis for acquiring letter-sound correspondences and learning to decode written words (National Early Literacy Panel, 2008; Elbro and Petersen 2004; Kjeldsen, Saarento-Zaprudin, and Niemi 2019). With '*Integrated phonemic awareness*', this awareness was kindled via a number of steps beginning at the level of 'meanings', i.e., whole sentences. The students would then proceed to literally cut up whole printed sentences into words, and words into syllables, and ultimately syllables into single letters. At each step, the students would attend to both the spelling and the pronunciation of the segments of the sentence. Since this teaching method could be beneficial for students at risk for reading difficulties, *Integrated phonemic awareness* was also singled out for analysis in the present study.

Fourth, a further two modified searchlight readers were entered into the empirical analysis. They were Brudholm (1988) and Kokborg & Rosenberg (1984). Both were similar to the reference reader (Lau 1987) in that they were meant to be used to support the teaching of a searchlight method. All three 'searchlight'-compatible readers differed from the phonics

[7] This basal reader was written by a special education teacher and reading consultant, Jørgen Frost, who was inspired by the work of the research group around Isabelle and Alvin Liberman (e.g., Liberman, 1973; Liberman, Shankweiler, Fisher & Carter, 1974) at the Haskins Laboratories affiliated to Yale University. In collaboration with a Swedish psychologist, Prof. Ingvar Lundberg, he conducted the first large-scale study of the effects of phoneme awareness training with pre-school children (Lundberg, Frost & Petersen, 1988).

compatible reader by being thematically structured with topics from the students' everyday life – with less attention to the readability or the vocabulary of the texts. All three searchlight readers contained fewer orthographically regular words, fewer frequent words, and had fewer repetitions of words than the phonics reader (based on counts of the first 1000 words in each basal reader). The two modified searchlight readers differed from the standard searchlight reader in a two important ways: they contained a higher proportion regularly spelled words among the first 1000 words (but not as high a proportion as the phonics reader), and they contained about 50 % more text (words). These differences made them somewhat more similar to the phonics reader.

3.2 Survey data

The original BPE study, on which our study is based, comprised roughly 3,500 students (in 124 schools) whose reading and spelling abilities were studied at the beginning and the end of Grade 1 in 1995/96. The participating schools were selected from a large pool of schools that responded to an initial questionnaire in the school year 1994/95 about their use of basal readers. The aim of the initial survey was to establish a representative sample of schools, based on their use of basal readers, school size, and urbanization. Responses were received from 1,220 schools out of a total of approximately 2,200 schools in Denmark. No systematic differences were found between responding and non-responding schools in terms of school size or demographics. From the 1,220 schools, 124 schools were selected in a balanced, random manner. The selection comprised schools that would use one of the five most common basal readers (mentioned above) with their first graders the following school year, and the selection was carried out in a way to ensure an equal distribution of the five basal readers across geographical regions and school demographics. However, within these constraints, the selection was randomized.

At the beginning of the 1995/96 school year, all participating classroom teachers (N = 202) were interviewed about their age, gender, educational background, teaching experience, and their planned use of basal readers. The teachers were also asked about how much emphasis they placed on the teaching of letter knowledge with first graders. The results indicated that 54 of the teachers had changed plans for their use of basal readers. Of these teachers, the majority (43) would use a different basal reader not among the top five from the previous year, while a few were not planning to use any basal reader. The remaining 148 teachers (and classes) used one of the five basal readers yielding 17 – 39 classes per basal reader.

3.3 Register data

For the present long-term follow-up, we link the data from the original first-grade study with Danish registry data containing information about the students' social background characteristics, i.e., parental education, income, employment, and whether the parents are immigrants or first-generation descendants of immigrants. We link the survey data from the original first-grade study on basal readers with the administrative data for schools. A challenge to the current study is that we cannot directly link the individual students from the original study with the choice of basal reader since we only observe which school they attended when they completed compulsory school at the end of ninth grade – not the school they attended in first grade. We are thus linking the information about the uses of basal readers in grade 1 with the 9th grade exam marks at the end of compulsory schooling for the students who belonged to the grade 1 classes provided they did not change school between grade 1 and 9. This is a reasonable assumption, with the exception of feeder schools, because most students stay at the same school throughout their time in compulsory education in Denmark – in contrast to the much higher mobility within countries such as the USA. Furthermore, we were able to observe

whether the parents moved municipality between the years where the students attended grade 1 to grade 9.[8] We found that this was the case for 10% of the sample, and we subsequently deleted these students from the analysis. Also note that, any changes in student allocation to classes will lead to attenuation bias due to measurement error and thus to statistically *more conservative* estimates of effects of basal readers in grade 1. In effect, we can think of the point estimates being lower bounds for any effect (or association) of basal readers.

From the survey, we know which schools used which basal reader the year before the 1995/96 school year survey. We therefore use data for two school cohorts, that being the students, who started school in the 1994/95 school year and the 1995/96 school year.

A few students attended special education schools and could not be tracked for the long-term follow-up. Also, a number of the original Grade 1 schools were feeder schools that only offered teaching up to 6th or 7th grade, which means that their students were unavailable for this study. When we removed students from schools that were either feeder schools or special education schools, our remaining data set comprised 77 schools and 4,481 students. We use a linear probability model to investigate whether there are systematic differences between the students in the surveyed schools and those who were not. Apart from differences in household income, analyses in Table A2 does not show any difference between students in the current sample and those excluded from the sample.

Our main variable of interest is the dominant basal readers found in the original Grade 1 study. As such, we have the basal readers of particular interest in focus: A *phonics compatible basal reader*, A basal reader with *integrated phoneme awareness*, and two modified basal readers. These four basal readers were compared to the reference basal reader "*The Danish*

[8] For the years in question, there were 276 municipalities in Denmark. On average, each municipality covered 156 square km, with an average population of 19,000, and with about 4 public schools.

Books". As explained above, *The Danish books* represented the traditional searchlight method (or "mixed cue" method) of teaching reading.

Outcomes

The main outcomes in the analyses are exam marks at the school leaving exam at the end of the 9th grade (typically at age 16) and at age 30, the students' highest qualifying education, income, and employment.

Exam marks at the end of compulsory schooling. At the end of the 9th grade, students take three exams in the subject of Danish. These exams comprise (1) an oral analysis and presentation of a short text which each student has 15 minutes to prepare, (2) a written essay in which the student either analyse a text or write an essay on a set subject, and (3) a test of spelling to dictation. At the time, each of these exams were marked by the students' teacher and an external examiner on a ten-point scale (range 0 to 13).

Educational level at age 30. The educational level is measured as the number of months of total education from the beginning of Grade 1.

Income at age 30. The measure taken is the disposable income.[9]

Employment at age 30. The measure of employment is simply whether or not the participant was employed for at least 28 work hours per week.[10]

[9] Adapting an inverse hyperbolic sine transformation, to account for any negative income due to debt.

[10] The threshold of 28 hours is chosen because people working 28 hours or more a week are entitled to unemployment benefits. In this sense, working 28+ hours a week signals a relatively strong labour market attachment.

4. Methodological framework

4.1 Identification

In principle, teaching materials are chosen by the individual teacher in Denmark as a function of the teacher's preferences and availability of materials at the school (dependant on many factors including the economy of the school). The relative freedom of choice of materials means that the assignment of teachers and students to basal readers was not random. In consequence, the choice of basal reader could be seriously confounded by factors linked to progress in reading – such as the teachers' preferred teaching method and the socio-economic status of the school district. Such confounds would invalidate causal interpretations of possible associations between basal readers and student outcomes. While it is impossible to prove the hypothesis of no confounds we are nevertheless able to test whether some of the strongest known confounds were present in the study. Two sources of possible endogeneity appear relevant and we investigate these in turn.

4.1.1 Parents and Basal Reader

First, parents might choose a school based on the schools choice of basal reader.[11] It is possible but unlikely that parents might select a school based on the basal reader used in Grade 1 (unless it is a sign of quality, see 4.1.2). Most parents have little or no knowledge of differences between basal readers and neither do they know which basal reader a given class (and school) will use. While some (very limited) awareness of the importance and difference between basal readers may exist among parents today it would have been non-existent in 1995. In addition, in 1995 parents did not have a free school choice. Rather, the school choice came with the choice of

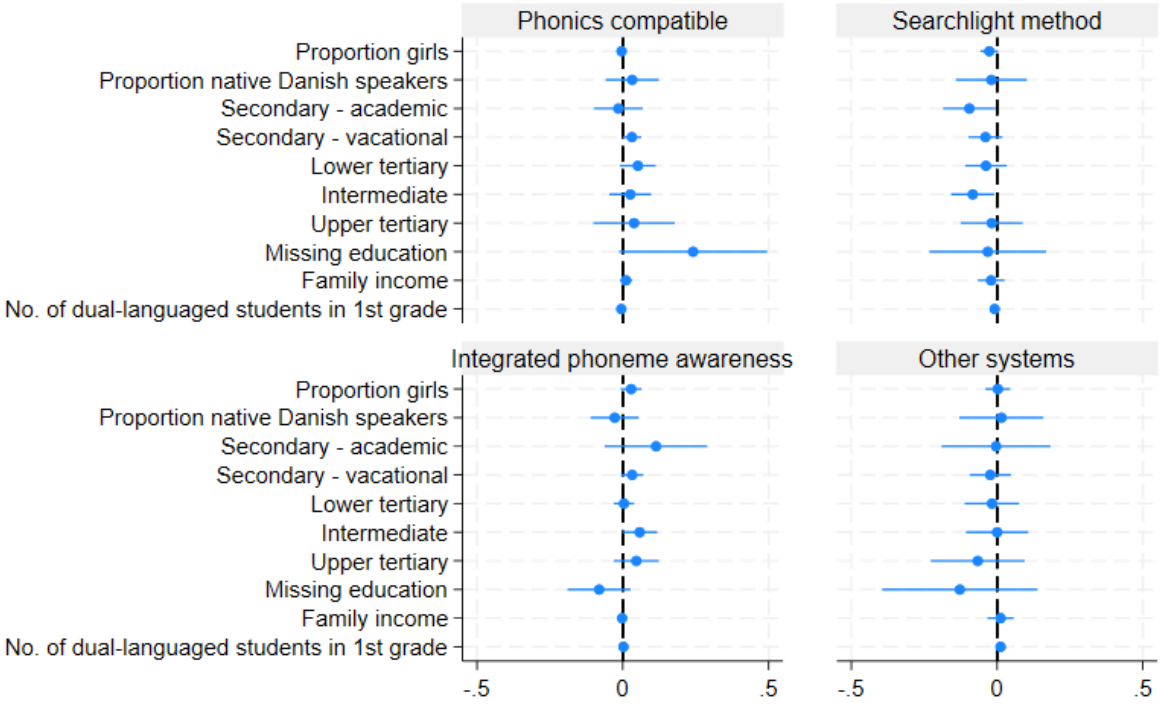
[11] At the time of the study, children in Denmark started school the year they turned seven. When children began school, they were not expected to know the letters or numbers. They are taught the alphabet and numbers in kindergarten while formal reading instruction began in the 1st grade.

residence. While that choice could also be made (in part) due to the local school's basal reader, it seems very unlikely.

However, the availability of the original BPE survey data provides us with an opportunity to test whether this source of endogeneity appear to exist in data by means of balancing tests across several important student characteristics in order to assess possible links between these student characteristics and choice of basal readers.

Our analyses of long-term effects of basal readers rests on the assumption that the use of the different basal readers is uncorrelated with background variables otherwise known to correlate with the outcome measure. In other words, parent SES and other predictors of reading success of the students must be uncorrelated to the choice of basal readers. The balancing test estimates whether any relevant student background characteristics affect the schools' choice of basal readers in 1995, the year that the students entered into 1st grade. Figure 1 shows the results of this analysis.

Figure 1: Balancing test of basal readers on background characteristics



Note: Regression coefficients and 95 % confidence intervals. All models use clustered standard deviations at the school level. The education variables refer to parents highest education.

We find no significant relationship between background characteristics and the choice of basal reader. This supports the assumption that parents would not choose school based on the basic reader.

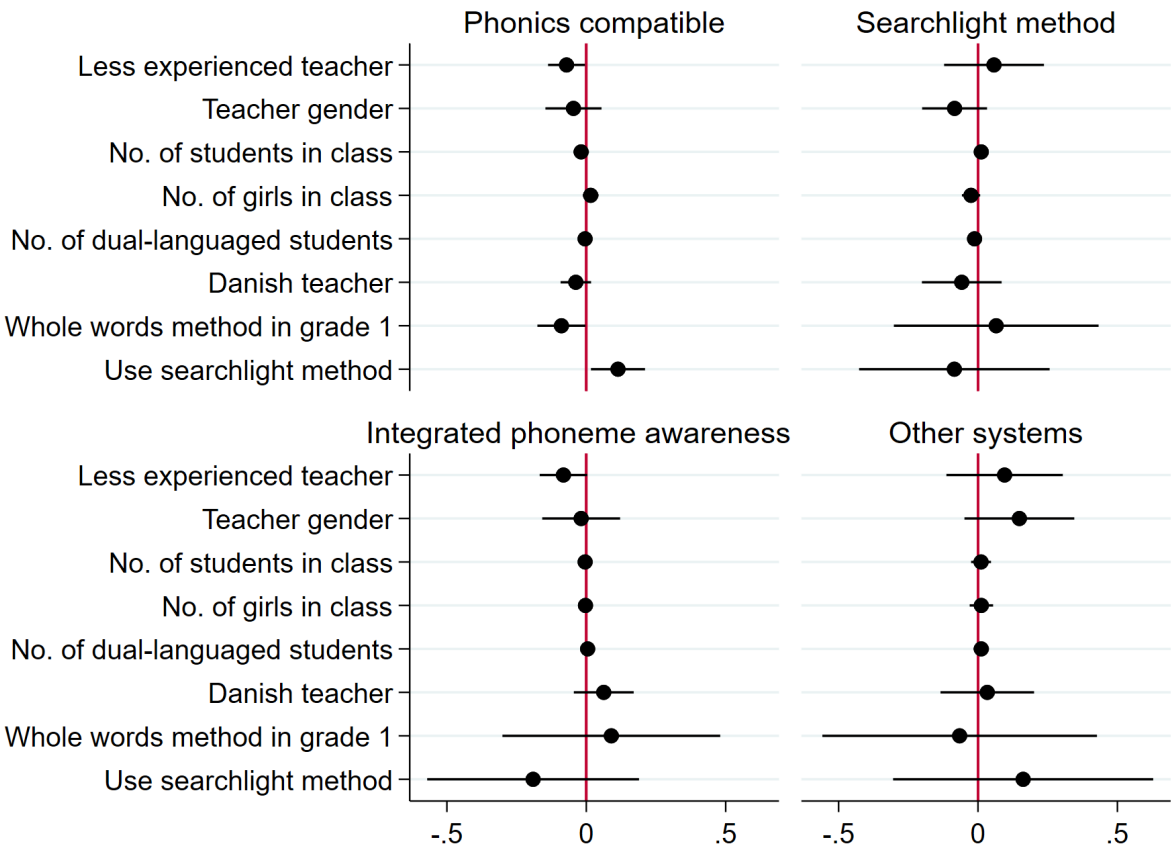
4.1.2 The School, the Teacher, and the Basal Reader

The second concern of endogeneity arises from the possible link between the quality of the school, including school teachers and the school leader, and the school choice of basal reader. While we cannot prove that school quality was not driving the choice of basal reader — a link that appears possible—we can describe why this was unlikely and at most played a very minor role:

Back in the mid-1990s, when the original BPE study was initiated, Danish teachers were generally unaware of the mounting research evidence in favour of phonics instruction. As described in detail above, the dominant method was the searchlight or mixed cue method, and the most phonics-compatible basic reader was viewed as very old-fashioned in Denmark. There was no knowledge available to teachers about which basal readers were better compatible with any particular, well-researched method of instruction.

In order to look for other possible confounds between teacher characteristics and choice of basal reader, we carried out another balancing tests. All participating teachers were interviewed about their age, gender, educational background, teaching experience, and their planned use of basal readers. The balancing test applied here uses the same methodology as the test above, but employed Grade 1 teacher characteristics comprising teacher gender, experience, Danish teacher education, class size, proportion of girls in class, proportion of bilingual students in class.

Figure 2: Balancing test of basal readers on teacher and classroom observable characteristics



Note: Regression coefficients and 95 % confidence intervals. Standard errors are clustered at the school level.

Figure 2 shows an absence of associations between choice of basal readers and teacher and school characteristics – with the one surprising exception that the searchlight method was (just) statistically significantly *more* preferred among teachers using the phonics approach. These results suggest that there was no endogeneity problem.

Obviously, there is some affinity between teachers' beliefs and knowledge and the materials that they select for their teaching. However, multiple constraints limit the teachers' choices of materials, not least the economic incentive of most schools to wear out older materials before new ones are acquired (BPE, 1999: 122). In the original study, teachers were asked directly

about their reasons to choose a particular basal reader. The majority (66 %) of the teachers responded that it was the only basal reader available at their school.

As mentioned, the searchlight method was the dominant way of teaching reading in Denmark in the early 1990s. There was no general knowledge among teachers or teacher educators about the strengths and weaknesses of different ways of teaching beginning reading. This fact is reflected in responses made by grade 1 teachers to a question about their preferred method of teaching reading, see Table 1.

Table 1. The chosen basal reader and teachers' preferred method of teaching beginning reading.

	Chosen Basal Reader			
Preferred teaching method	Searchlight (reference)	Modified searchlight	Integrated phonological awareness	Phonics
Letter name	2.9 %	6.0 %	0.0 %	13.3 %
Letter sound	26.5 %	31.3%	35.3%	13.3%
Whole word	2.9 %	1.5%	5.9%	0.0%
Searchlight	67.6 %	61.2%	58.8%	73.3%
Total	100.0 %	100.0 %	100.0 %	100.0 %

The results in Table 1 supports the claim that a searchlight method was by far the most prevalent across teachers. Somewhat surprisingly, teachers using the phonics compatible basal reader were stronger adherents (73 %) to the searchlight method than other teachers were. On the other hand, letter-sound teaching was a slightly more popular method among teachers who used a basal reader that integrated phonological awareness than among other teachers. Given the results in Figure 2 and Table 1, it appears reasonable to assume that possible links between

choice of basal reader and different teaching methods were very weak – and did not support any particular link between methodological preference and choice of basal reader. Given the general lack of knowledge about the importance of phonics at the time (the mid1990s) this comes as no surprise.

In conclusion, we believe that both possible sources of endogeneity outlined above can be considered very unlikely. These facts make the situation in Denmark in the early 1990s a particularly well-suited natural laboratory for the study of the possible causal effects of different basal readers on reading – and ultimately on long term educational and socio-economic outcomes.

4.2 Average and quantile regression models

Our baseline approach to analyse potential long-term effect of basal readers is to estimate linear models by ordinary least squares (OLS) with cluster robust standard errors at the school-level. This provides mean estimates of the long-term effects of basal readers on the outcomes.

In addition, it is of special interest to analyze whether the choice of basal reader has different impact at different points of the distribution, especially whether economically or academically low achievers benefitted more from the phonics-compatible basal readers. In order to investigate such heterogeneous effects of the basal readers, we estimate unconditional quantile models, see Firpo, Fortin and Lemieux (2009). To compare estimates across models, we standardize the outcomes in order for the estimates to be interpreted in terms of standard deviations. This also facilitates comparisons over time and across countries.[12]

[12] The unconditional quantile regression method developed by Firpo et al. (2009) is estimated with fixed effects. One advantage of the unconditional quantile regression model is that the quantiles are defined preregression, and, therefore, the model is not influenced by any right-hand-side variables (Killewald & Bearak 2014). In unconditional quantile regression, including fixed effects to adjust for selection bias is done without redefining the quantiles.

5. Results

5.1 Estimates of average outcomes

First, we estimate the average effect of choice of basal readers on exam grades in Danish at the end of 9th grade, see Table 2. Exams in Danish included an Oral, Written, and Spelling (dictate) exam. A GPA in Danish was also computed and used as outcome.

Across all four columns of output we find that students who were taught by the phonics basal reader system had significantly higher exam grades than students taught with the search light method (the reference category). Similar statistical differences are not found for the other systems (with one minor exception).

These results are consistent with the hypothesis that the phonics approach to the initial teaching of reading has a positive effect on the development of literacy that lasts to the end of compulsory schooling, i.e. the end of the ninth grade.

Table 2. Average Effects of Basal Readers on Exam Grades in Danish at the end of 9th grade

(reference = searchlight)	**(1) Oral**	**(2) Written**	**(3) Spelling**	**(4) GPA - Danish**	**(5) Math**
Phonics	0.151*	0.122	0.303**	0.207**	0.134
	(0.086)	(0.093)	(0.128)	(0.102)	(0.252)
Integrated phoneme awareness	0.083	0.027	0.080	0.070	-0.091
	(0.092)	(0.090)	(0.092)	(0.090)	(0.167)
Modified searchlight	0.066	0.003	0.038	0.046	0.073
	(0.041)	(0.060)	(0.076)	(0.057)	(0.131)
Observations	4,444	4,479	4,481	4,481	4,454

Note: All models include control variables for individual, family, teacher and school characteristics. See Table A.1 in the appendix for all parameter estimates. Regression coefficients and standard errors in parenthesis. Standard errors clustered at school level. $*$ p < 0.10, $**$ p < 0.05, $***$ p < 0.01.

Next, we estimate the effect of first grade basal reader on long-term labour market outcomes measured at age 30. We here use income, length of education, and an indicator variable for full-time employment as outcome measures.

The average effects of the choice of first grade basal reader on labour market outcomes are generally insignificant. One exception is the estimates for the impact on length of education where the reference basal reader (*Searchlight* compatible) comes out below the other reading systems, including the phonics compatible reader.

Table 3. Average Effects of Basal Readers on Labour Market Outcomes at Age 30

(reference = searchlight)	(1) Income at age 30	(2) Length of education	(3) Full-time employment
Phonics	0.107	0.186*	0.018
	(0.105)	(0.099)	(0.036)
Integrated phoneme awareness	0.053	0.120	0.000
	(0.047)	(0.086)	(0.024)
Modified searchlight	0.076	0.135***	0.020
	(0.051)	(0.046)	(0.021)
Observations	3,769	4,373	3,919

Note: All models include control variables for individual, family, teacher and school characteristics. See Table A.1 in the appendix for all parameter estimates. Regression coefficients and standard errors in parenthesis. Standard errors clustered at school level. ∗ p < 0.10, ∗∗ p < 0.05, ∗∗∗ p < 0.01.

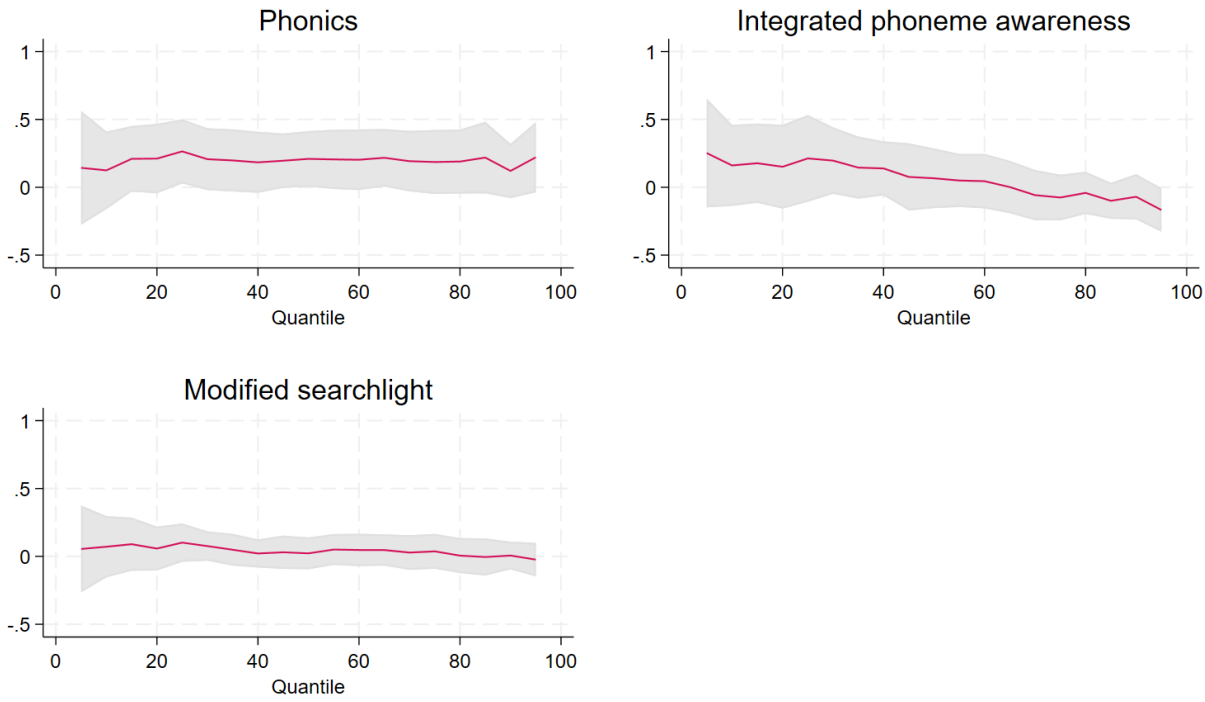
5.2 Quantile estimates

In order to evaluate the relative effectiveness of various basal readers for students of different abilities we estimate unconditional quantile regressions on GPA in Danish exams in grade nine (Figure 3) and income and length of education at age 30 (Figures 4 and 5). Again, with the Searchlight compatible reader as reference.

Even in the absence of average effects there may very well be long-term effects of basal readers elsewhere in the distribution and, in principle, such effects may have opposite signs, which may explain zero-findings on average.

For the 9th grade GPA results, the quantile regression approach reveals that the effect of the phonics-compatible basal reader is significantly positive for grade 9 exam outcomes across the entire distribution of students. The quantile parameter estimates for the integrated phoneme awareness basal reader are significantly positive around the 40th percentile. This corroborates the corresponding average effect, which was significant at the 10% level for Danish GPA. Any effects of modified searchlight basal readers are not statistically significantly different from zero.

Figure 3: Unconditional Quantile Regression Estimates of the Effects of Basal Readers in Grade 1 on GPA in Danish at the end of 9th Grade, (reference = Search light)

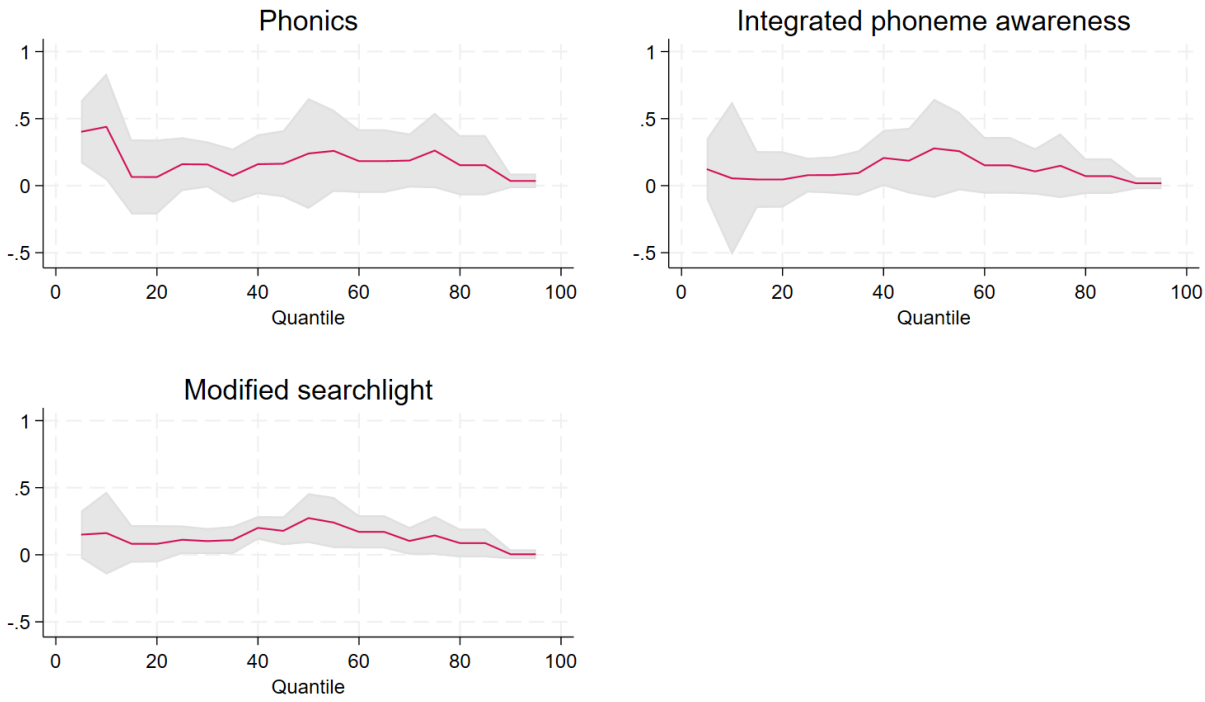


Note: 95% confidence interval in grey. Point estimates are given by the red lines.

The average effect of basal reader on length of education at age 30 was significant for the Phonics and Modified Searchlight basal readers (as compared to the reference Searchlight basal reader). The quantile estimates reveal that the phonics-compatible basal reader had a positive impact on length of education across most of the distribution. The point estimate is relatively high -- almost 0.5 standard deviation at the bottom -- and statistically significant for the bottom 15% only. This corresponds exactly to the share usually deemed "functional illiterate".

For the modified searchlight readers, there are generally also positive and partly significant effects albeit smaller and less significant than for the phonics compatible basal reader. And insignificant in the lower end of the distribution. For the basal reader with integrated phoneme awareness, the quantile estimates are statistically insignificant across the distribution.

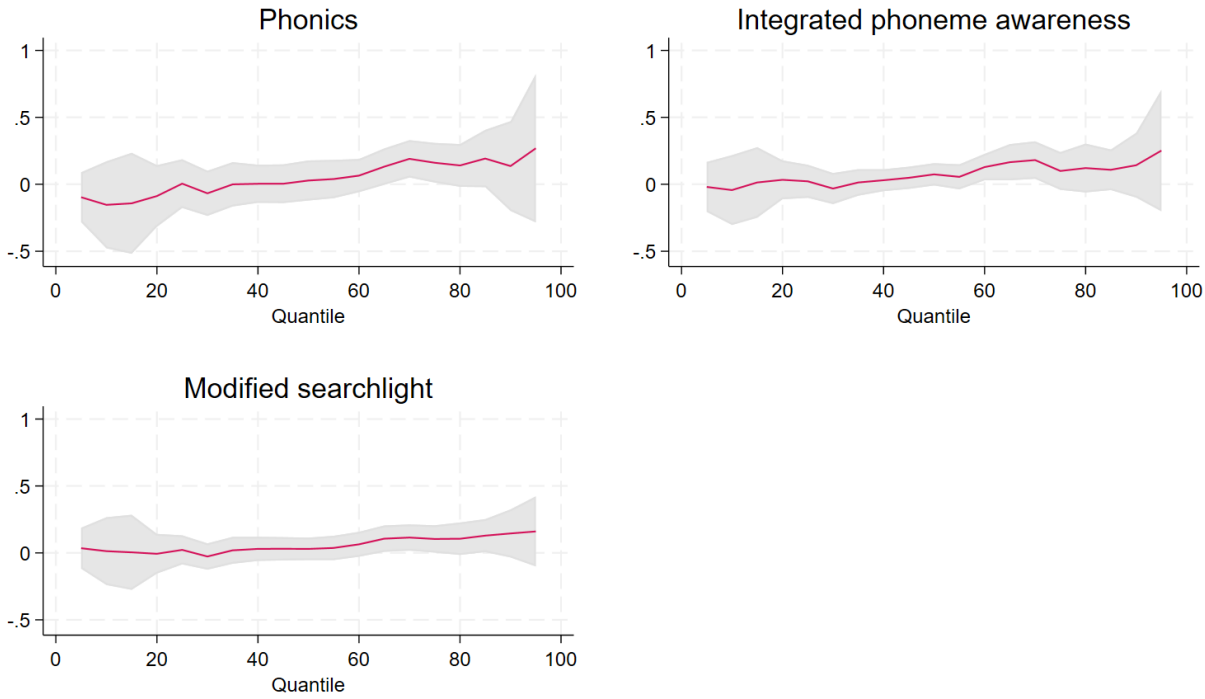
Figure 4: Unconditional Quantile Regression Estimates of the Effect of Basal Reader in Grade 1 for Length of Education at Age 30



Note: 95% confidence interval in grey. Point estimates are given by the red lines.

Estimates of the average effect of basal readers on income at age 30 were generally found to be statistically insignificant. Looking at the unconditional quantile distribution we generally also find insignificant effects of all three basal readers on income (with a small exception around the 70th percentile where all three perform better than the reference).

Figure 5: Unconditional Quantile Regression Estimates of the Effect of Basal Reader in Grade 1 for Income at Age 30



Note: 95% confidence interval in grey. Point estimates are given by the red lines.

6. Robustness analysis

The estimates of the quantile regressions on length of education are strongly positive in the lower end of the distribution, and the estimates for income are negative, but insignificant in the lower end. A possible explanation is that the extra years of education move higher wages further into the future and that more years of education will start to pay off later. To investigate this, we therefore estimate the quantile regressions models on education and income at the age of 25. The results are shown in Figure 6 and Figure 7.

Figure 6: Unconditional Quantile Regression Estimates of the Effect of Basal Reader in Grade 1 for Length of Education at Age 25

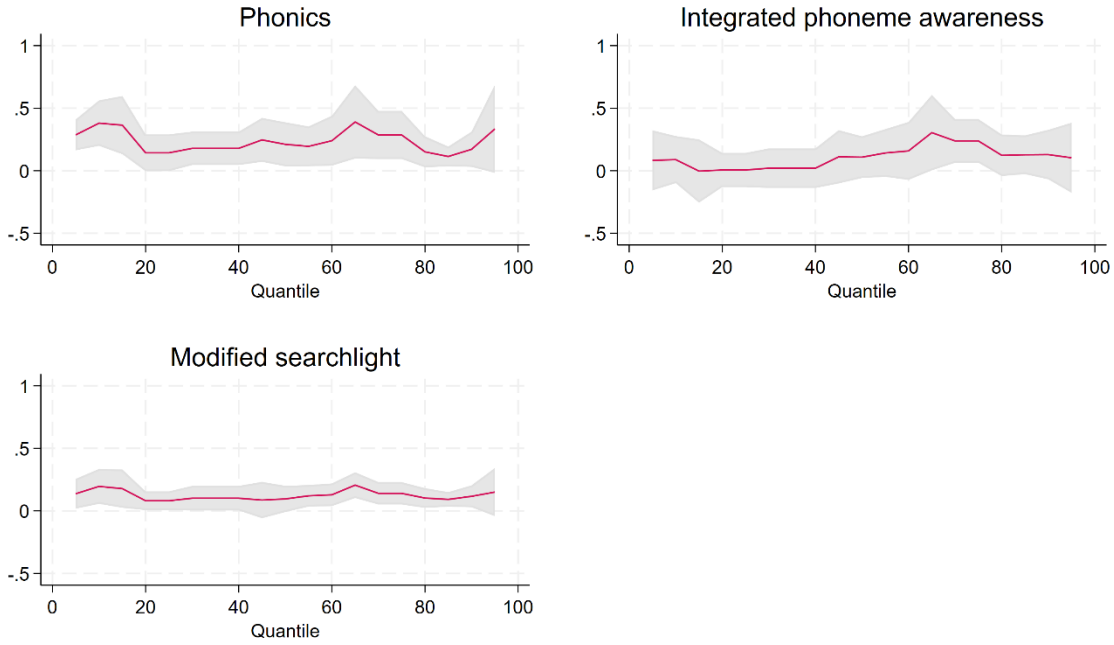
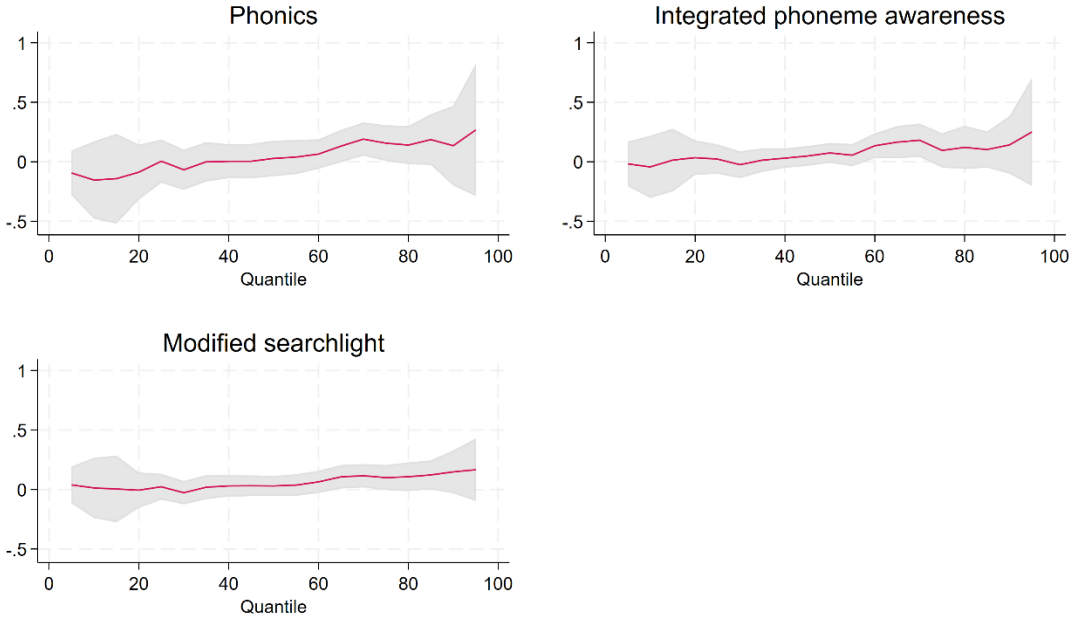


Figure 7: Unconditional Quantile Regression Estimates of the Effect of Basal Reader in Grade 1 for Income at Age 25



The figures show the same tendencies in the estimates of length of education and income. While still statistically insignificant for income, the larger negative effect of phonics basal readers in the lower end of income at age 25 and the larger positive effect on length of education, is very likely to be caused by the fact that young adults in education have a relatively low income.

Could the results be driven by omitted variable bias? Following Oster (2019), we investigate whether the effect of the phonics basal reader – instead of being generated by the students' reading behavior – is caused by omitted variables.

We use the value of δ, which is the ratio of selection on unobservables to selection on observables, that would be necessary for the effect estimates to be fully explained by omitted variable bias.

To do this, we estimate a model in which we only include a dummy variable for the phonics basal reader (uncontrolled model) and compare this to a model in which we also include control variables, as those used in our main specifications (controlled model). A larger value of δ means less bias due to omitted variables and a $\delta>1$ means that unobservables are very unlikely to explain the results.[13] The controlled model serves a benchmark, and we follow Oster (2019) in setting the R_{max} to 1.3 times the R^2 obtained in the controlled model, in order to estimate the δ. The results of these analyses are shown in Table B1. The coefficients are of the same sign for the controlled and uncontrolled models.

The controlled and uncontrolled coefficients, along with the R^2-statistic, show large values of δ. For instance, the δ of the effect of the phonics basal reader on spelling is 5.677, which means that selection on unobservables should be 5.7 times larger than the effect we capture with our

[13] A value of $\delta = 2$, for example, would suggest that the unobservables would need to be twice as important as the observables to produce a treatment effect of zero, Oster (2019).

observables for any of the effects to be zero. This Oster-analysis suggest that there is (very) little bias from omitted variables in our results.

7. Discussion and Conclusion

Many important studies have reported very long-term effects of improvements in early child-hood education and development. The STAR project, Perry Pre School, Abecedarian, and HEAD Start to mention some of the most important programs and interventions. But studies of the very long term importance of a phonics approach in learning how to read are rare. As mentioned, a positive long-term impact (from age 5 to 11) for a phonics approach was found by (Machin et al. 2018).

The purpose of the present study was to add to this literature by investigating the potential long-term effects of the phonics approach adapted in the basal readers in grade 1. We find a relatively strong positive association between the use of a phonics-compatible basal reader in Grade 1 – and Grade 9 exam marks (at age 16) and long-term education completed at age 30. These results are consistent with the basic hypothesis – that a basal reader which lends itself well to a phonics approach to the initial teaching of reading can have positive effects on the development of literacy that may last all the way through compulsory schooling and beyond. The further hypothesis – that students with relatively poor prerequisite abilities would profit the most from a phonics-compatible basal reader - received some support in that the students at the lower end of the ability distribution tended to profit the most with respect to completed length of education for the bottom 15%.

Is causality credible? As described in section 4.1, teachers and school leaders were generally unaware of the importance of a phonics approach in the mid-1990s when the foundation of the current study was laid down. This fits well with the finding that most teachers favoured a

searchlight method in initial teaching of reading no matter which basal reader they used. Combining survey and register data, we searched for but failed to find any systematic differences between students, teachers, and schools that might explain away differences in the effects of basal readers.

Finally, observing that the majority of the positive outcomes are driven by students at the lower part of the distribution is not what one would expect if the significant association between a phonics-compatible basal reader and long-term outcomes was driven by unobservables – for instance, that good teachers and/or good schools chose phonics-compatible readers relatively often. In that case, positive effects would probably have appeared across the entire distribution.

Which channel would render such a modestly looking choice of basal readers to have high long-term impact for selected groups? One simple answer is that nothing succeeds like success. Children who find that early independent reading is relatively easy are much more likely to develop a life-long habit of reading – and they will favour the contexts that associate with reading, such as schools and further education. Another important channel may be class peers. The entire class is treated here and this has proven important in avoiding initial effects to vanish over time. List & Uchida (2024:1) argue that "with enough treated peers the classic fade-out is muted". Our results corroborate their finding.

The policy implication of our study is clear. Choosing a phonics-compatible basal reader and a corresponding teaching approach is an initiative that can be implemented at very low costs in comparison to initiatives such as smaller class size, two teachers in the classroom, or increased instruction time. Provided that the phonics approach has a positive long-term impact into adulthood and labour market outcomes, it will inevitably be extremely cost-effective. At least

in Denmark, published basal readers continue to vary hugely in terms of compatibility with phonics approaches to teaching.

One caveat that merit some consideration is how this may change the locus of control of local government that traditionally decide (and pay) for public municipal schools. Also, school leaders and teachers alike might find a centralized planning of curriculum including implementation of phonics approaches a direct attack on school teachers otherwise highly celebrated autonomy. It would be difficult to maintain a highly decentralized system found in many countries today, including Denmark. However, teacher autonomy is a complex concept related to a multitude of areas. Based on a review of the literature on teacher autonomy, teachers' exercise of control can be found in four different school domains: Teaching and assessment, curriculum development, school functioning, and professional development, Usma-Wilches (2006). Even if the choice of basal reader and teaching approach is somehow restricted it does not necessarily imply that teacher autonomy will be greatly diminished. With stakes this high – with a chance of increasing equality in education and bringing about an improvement especially for the most needy – national strategies on phonics approaches, like the ones found in USA and UK, merit serious consideration in countries that choose a more eclectic approach to the art of teaching reading.

8. References

Andersen, Simon Calmar, Maria Knoth Humlum, and Anne Brink Nandrup. 2016. "Increasing Instruction Time in School Does Increase Learning." *Proceedings of the National Academy of Sciences of the United States of America* 113(27):7481–84. doi: 10.1073/pnas.1516686113.

Andersen, S. C., Beuchert, Louise. V., Nielsen, Helena S., & Thomsen, Mette K. (2020). The Effect of Teacher's Aides in the Classroom: Evidence from a Randomized Trial. *Journal of the*

European Economic Association, **18**(1): 469–505. https://doi.org/10.1093/jeea/jvy048

Andersen, Simon Calmar, Beuchert, Louise. V., Nielsen, Helena. S. (2025). The Effect of Teacher's Aides on Students with Special Needs. Forthcoming, *Scandinavian Journal of Eonomics.*

Angrist, J.D. & V. Lavy (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics* **114**(2): 533-575.

Baulos, A. W., Jj L. García & J. J. Heckman (2024). Perry Preschool at 50: What Lessons Should be Drawn and Which Criticisms Ignored? Working Paper **32972**.

Bjerg, Jens, and Mogens Jansen. 1961. *Huset Højbo (the House Hojbo)*. Copenhagen: Munksgaard.

Borstrøm, I., D. K. Petersen, and C. Elbro. 1999. *Hvordan Kommer Børn Bedst i Gang Med at Læse,* ('How do children best learn to begin to read?'). Center for læseforskning, Københavns Universitet.

Brudholm, M. (1987). *Trip Trap Træsko*. Copenhagen: Nyt nordisk forlag Arnold Busch

Carneiro, Pedro, and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." American Economic Journal: Economic Policy 6 (4): 135–73.

Cherry, Gemma and Anna Vignoles 2020. What is the economic value of literacy and numeracy? IZA world of Labor 229v2.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR. Quarterly Journal of Economics 126 (4): 1593–660.

Dalby, M. A., Elbro, C., Jansen, M., Krogh, T., & Christensen, W. P. (1983). *Bogen om læsning I - forudsætninger og status*. København: Munksgaard/Danmarks pædagogiske Institut.

Dalby, M. A., Elbro, C., Jansen, M., & Krogh, T. (1989). *Bogen om læsning II. Set i et udviklingsperspektiv*. Munksgaard og Danmarks Pædagogiske Institut.

Deming, D. (2009). Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start. *American Economic Journal: Applied Economics*, **1**(3): 111–134.

Elbro, Carsten, and Dorthe Klint Petersen. 2004. "Long-Term Effects of Phoneme Awareness and Letter Sound Training: An Intervention Study With Children at Risk for Dyslexia." *Journal of Educational Psychology* 96(4):660–70. doi: 10.1037/0022-0663.96.4.660.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77(3):953–73. doi: 10.3982/ecta6822.

Fredriksson, P., B. Öckert, & H. Oosterbeek (2013). Long-term effects of class size. *Quarterly Journal of Economics*, **128**(1): 249-285.

Fryer, Roland G. 2014. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence From Field Experiments." *The Quarterly Journal of Economics* 1355–1407. doi: 10.1093/qje/qju011.Advance.

Garcia, J. L. & J. J. Heckman (2023). Parenting Promotes Social Mobility Within and Across Generations . *Annu. Rev. Econ*. 15:349–88.Heinesen, E. (2010). Estimating class-size effects using within-school variation in subject-specific classes. Economic Journal 120: 737-760.

Hermansen, Knud, and Ejvind Jensen. 1954. *Søren Og Mette (Soren and Mette)*. Copenhagen: Gad.

Hoynes, H., D. Whitmore Schanzenbach & D. Almond (2016). Long-Run Impacts of Childhood Access to the Safety Net. *American Economic Review* , **106**(4): 903-934.

Ireson, Juditii, Susan Hallam, and Clare Hurley. 2005. "What Are the Effects of Ability Grouping on GCSE Attainment ?" 31(4):443–58. doi: 10.1080/01411920500148663.

Jansen, Mogens, Bo Jacobsen, and Poul Erik Jensen. 1978. *The Teaching of Reading – Without Really Any Method : An Analysis of Reading Instruction in Denmark*. Copenhagen: Humanities Press.

Johnston, Rhona, and Joyce Watson. 2005. The Effects of Synthetic Phonics Teaching on Reading and Spelling Attainment: A Seven Year Longitudinal Study. Scottish Executive Education Department. Edinburgh, February.

Killewald, A. & J. Bearak (2014). Is the Motherhood Penalty Larger for Low-Wage Women? A Comment on Quantile Regression. *American Sociological Review*, **79**(2): 350-357.

Kjeldsen, Ann-Christina, Silja K. Saarento-Zaprudin, and Pekka O. Niemi. 2019. "Kindergarten Training in Phonological Awareness: Fluency and Comprehension Gains Are Greatest for Readers at Risk in Grades 1 through 9." *Journal of Learning Disabilities* 52(5):366–82.

Kokborg, I. & Rosenberg, P. (1984) *Dansk i 1. klasse ('Danish in Grade 1')*. Copenhagen: Gyldendal.

Krueger, A.B. (1999). Experimental estimates of educational production functions. *Quarterly Journal of Economics*, **114**(2): 497-532.

Lau, Jytte. 1987. *Der Sker Jo Så Meget. Danskbøgerne 1. Klasse ('Why, a Lot Is Happening. The Books for Teaching Danish Grade 1')*. edited by J. Lau, T. K. Heiberg, and P. E. Pagaard. Copenhagen: Munksgaard.

Lavy, V. (2015). Do Differences in Schools' Instruction Time Explain International Achievement Gaps? Evidence from Developed and Developing Countries, The Economic Journal 125(588): F397–F424.

List, John. A. & Haruka Uchida (2024). Here Today, Gone Tomorrow? Toward An Understanding of Fade-Out In Early Childhood Education Programs. NBER Working Paper 33027 (Oct. 2024).

Lundberg, I., Frost, J., & Petersen, O. P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263-284.

Machin, Stephen, Sandra McNally, & Martina Viarengo. 2018. "Changing How Literacy Is Taught: Evidence on Synthetic Phonics." *American Economic Journal: Economic Policy* 10(2):217–41. doi: 10.1257/pol.20160514.

Oster, E. (2019). Unobservable Selection and Coefficient Stability: Theory and Evidence. *Journal of Business & Economic Statistics*, 37(2): 187–204. https://doi.org/10.1080/07350015.2016.1227711

Ramey, C. T. & F. A. Campbell (1984). Preventive education for high-risk children: Cognitive consequences of the Carolina Abecedarian project. *American Journal of Mental Deficiency*, **88**(5): 515-523.

Rivkin, S. G. and J.C. Schiman (2015). Instruction Time, Classroom Quality, and Academic Achievement, The Economic Journal 125, 588:F425–F448.

Rose, J. 2006. *Independent Review of the Teaching of Early Reading*. London: Department for Education and Skills.

Schanzenbach, Diane Whitmore (2006). What Have Researchers Learned from Project STAR? *Brookings Papers on Education Policy*, 205-228.

Schwartz, A.E., B.G. Hopkins, and L. Stiefel (2021). The Effects of Special Education on the Academic Performance of Students with Learning Disabilities. *Journal of Policy Analysis and Management*, **40**(2): 480–520.

Usma-Wilches, Jaime A. 2006. Teacher Autonomy : a review of the research literature. University of Northern Iowa, Graduate Research Papers. 1634.

Vignoles, Anna, Augustin De Coulon, and Oscar Marcenaro-Gutierrez 2011. The Value of Skills in the British Labour Market. *Oxford Economic Papers* , **63**(1): 27-48.

Appendix A: Mean and standard deviations of background variables

Table A1. Mean and standard deviations of background variables

	Mean	SD
Outcome		
Oral	6.93	3.06
Written	6.15	2.65
Spelling	5.80	2.84
GPA - Danish	6.23	1.98
Length of education	181.50	28.51
Annual Income	202164.94	100815.60
Full time employment	0.77	0.42
Basal reader		
Phonics	0.08	0.27
Danskbøgerne	0.15	0.36
Integrated phoneme awareness	0.08	0.27
Modified searchlight	0.65	0.48
Individual characteristics		
Proportion girls	0.49	0.50
Proportion native Danish speakers	0.93	0.25
Family characteristics - 1995		
Primary	0.11	0.31
High School	0.02	0.13
Vocational	0.43	0.50
College	0.34	0.47
Missing education	0.10	0.30
Disposable income	277387.21	99652.75
No. of siblings under 17	1.36	10.63
Family characteristics - 2002 and 2003		
Primary	0.09	0.28
High School	0.02	0.15
Vocational	0.43	0.50
College	0.38	0.48
Missing education	0.08	0.27
Disposable income	394863.97	163480.87
No. of siblings under 17	0.77	0.85
School characteristics		
Changed basal reader between Grade 1 and Grade 2	0.12	0.33
No. of dual-languaged students in 1st grade	3.31	6.20
Teacher educated 1990 or later	0.16	0.37
Observations	4,481	

Table A2: Probability of being in a sampled school. Linear probability model.

	(1)
Proportion girls	-0.003
	(0.002)
Proportion native Danish speakers	-0.005
	(0.012)
Parents' Highest education	
Primary	ref.
High School	0.006
	(0.011)
Vocational	0.000
	(0.004)
College	-0.004
	(0.006)
Missing education	-0.003
	(0.005)
Family income	0.002***
	(0.001)
No. of siblings under 17	-0.000
	(0.000)
No. of students in school-year	0.000
	(0.000)
Observations	46,075

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix B: Supplementary tables

Table B.1 Average estimates including all parameters

	(1) Oral	(2) Written	(3) Spelling	(4) GPA - Danish	(5) Income at age 30	(6) Length of education	(7) Full time employment	(8) Math
Phonics	0.151*	0.122	0.303**	0.207**	0.107	0.186*	0.018	0.134
	(0.086)	(0.093)	(0.128)	(0.102)	(0.105)	(0.099)	(0.036)	(0.252)
Integrated phoneme awareness	0.083	0.027	0.080	0.070	0.053	0.120	0.000	-0.091
	(0.092)	(0.090)	(0.092)	(0.090)	(0.047)	(0.086)	(0.024)	(0.167)
Modified searchlight	0.066	0.003	0.038	0.046	0.076	0.135***	0.020	0.073
	(0.041)	(0.060)	(0.076)	(0.057)	(0.051)	(0.046)	(0.021)	(0.131)
No system	-0.066	-0.007	-0.146*	-0.073	-0.046	0.112	0.058	-0.262
	(0.068)	(0.090)	(0.083)	(0.087)	(0.065)	(0.070)	(0.043)	(0.190)
Shift between 1. and 2. Grade	0.090	0.054	0.180***	0.115*	0.032	0.002	0.003	0.285*
	(0.079)	(0.067)	(0.053)	(0.058)	(0.046)	(0.072)	(0.026)	(0.145)
Proportion girls	0.294***	0.517***	0.336***	0.492***	-0.406***	0.201***	-0.027*	0.242***
	(0.029)	(0.037)	(0.031)	(0.036)	(0.033)	(0.030)	(0.015)	(0.071)
Proportion native Danish speakers	0.150***	0.509***	0.326***	0.375***	0.057	0.004	-0.023	0.805***
	(0.057)	(0.070)	(0.083)	(0.072)	(0.097)	(0.079)	(0.037)	(0.153)
Highest parental education Primary (ref.)								
High School	0.394***	0.409***	0.354***	0.458***	0.370***	0.433***	0.034	1.148***
	(0.107)	(0.097)	(0.093)	(0.100)	(0.107)	(0.115)	(0.054)	(0.241)

Vocational	0.168***	0.185***	0.171***	0.227***	0.179***	0.331***	0.063**	0.458***
	(0.050)	(0.054)	(0.056)	(0.054)	(0.048)	(0.054)	(0.026)	(0.123)
College	0.597***	0.670***	0.589***	0.737***	0.285***	0.797***	0.069**	1.435***
	(0.051)	(0.061)	(0.056)	(0.056)	(0.057)	(0.054)	(0.029)	(0.142)
Missing education	-0.023	0.190**	0.201**	0.170*	0.069	0.275***	0.001	0.177
	(0.099)	(0.091)	(0.100)	(0.094)	(0.113)	(0.093)	(0.054)	(0.178)
Family income	0.010	0.029**	0.028***	0.026**	0.022**	0.027**	0.004	0.052**
	(0.011)	(0.014)	(0.010)	(0.011)	(0.011)	(0.011)	(0.007)	(0.023)
No. of siblings under 17	0.004	-0.005	-0.010	-0.005	-0.009	0.026	-0.008	0.089**
	(0.020)	(0.017)	(0.018)	(0.018)	(0.018)	(0.019)	(0.008)	(0.040)
No. of dual-languaged students in 1st grade	0.001	-0.006*	-0.012***	-0.005*	-0.001	-0.015***	-0.002*	-0.015
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)	(0.010)
Teacher educated 1990 or later	-0.010	-0.014	-0.023	-0.034	-0.064	0.034	-0.015	0.015
	(0.046)	(0.092)	(0.067)	(0.080)	(0.049)	(0.041)	(0.020)	(0.132)
Observations	4,444	4,479	4,481	4,481	3,769	4,373	3,919	4,454

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B2: Estimates of controlled and uncontrolled regressions. Coefficients, R-squared and delta-values.

	Oral		Written		Spelling		GPA - Danish		Income at age 30		Length of education		Full time employment		Math	
	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2	Coef.	R^2
Uncontrolled	0.129	0.001	0.181	0.002	0.318	0.007	0.214	0.003	0.081	0.001	0.130	0.001	0.011	0.001	0.202	0.001
Controlled	0.151	0.084	0.122	0.171	0.303	0.118	0.188	0.168	0.107	0.053	0. 186	0.106	0.018	0.006	0.133	0.096
Delta (δ)	14.017		2.921		5.677		5.082		36.235		-447.128		-18.729		2.825	