Readme file for "The Effect of Schooling on Cognitive Skills"

by Gordon Dahl, Magnus Carlsson, Björn Öckert and Dan-Olof Rooth
Published in the *Review of Economics and Statistics*

A. Data Availability

Our data contains sensitive information on individual variables and hence, Swedish law requires users of the data to hold a permission from an Ethical Review Board ("Regionala Etikprövningsnämnden", in our case at Linköping University). This implies that we are unable to make our data directly available online. However, researchers interested in obtaining our type of data could themselves apply for permission from the Ethical Review Board (ERB) at http://www.epn.se/en/start/startpage/. The cost of applying for permission is currently 5000 SEK (about 550 Euros or 750 dollars). After having obtained an ERB approval to use the data, the next step is to order/buy the data from Statistics Sweden and the Swedish Military Archives (the costs for this are computed on a case-by-case basis). In this process Statistics Sweden assigns each individual an identification number unrelated to the actual personal identification numbers used in Sweden. This identifier then makes it possible to match data from different sources/administrative registers described below. In a final step the identifier should be sent to the Swedish Military Archives in order to get information from the enlistment process. All of these steps are carried out by Statistics Sweden. The data set obtained from Statistics Sweden does not contain personal information such as name and address, but does contain information on region of residence and birth date.

B. Data

The different data sources being used in the empirical work of the article are listed below as well as the different variables being used in the analysis. The main data set being used to arrive at the outcome variables is the Enlistment register for the period 1980-1994. These data are collected by the National Service Administration (NSA). We also use information from the education, income and school registers to create explanatory variables. The population being analyzed is extracted from the 1999 Swedish population register (Registret över totalbefolkningen - RTB). We are then able to match these different data sets by using both the personal identifier mentioned above as well as parental identifiers. The parental identifier is matched to RTB using the Multi-generation register which then facilitates the matching with other administrative registers containing information on parental and family characteristics.

B1. Data sets being used (ambiguous names also presented in Swedish)

Name	Source	Years
Registret över totalbefolkningen, RTB (16-64 year olds)	Statistics Sweden (SCB)	1999
Multi-generation register	Statistics Sweden (SCB)	2009
Education register (in Swedish: Utbildningsregistret)	Statistics Sweden (SCB)	1999, 2003
School register (in Swedish: Grundskolan: slutbetyg)	Statistics Sweden (SCB)	1999

Income register	Statistics Sweden (SCB)	1980
Enlistment records etc.	National Service Admin.	1980-1994
Teacher register (in Swedish: Lärarregistret)	Statistics Sweden (SCB)	1978-94
High school register (in Swedish: Skolregistret)	Statistics Sweden (SCB)	1978-94

B2. A list of all variables being used in the analysis and in which register these are found

Outcome variables

Cognitive test score	Enlistment register, NSA 1980-1994
Non-cognitive test score	Enlistment register, NSA 1980-1994

Explanatory variables or variables used to create explanatory variables

Birthdate	Register of the total population, RTB 1999
Enlistment date	Enlistment register, NSA 1980-1994
Parish	Enlistment register, NSA 1980-1994
School program	Enlistment register, NSA 1980-1994
Years of schooling	Education register, 2003
Father's income in 1980	Income register, SCB 1980
Birth year of parents (mother/father)	Register of the total population, RTB 1999
Family identifier (mother/father)	Multigeneration register 2009
Family size	Multigeneration register 2009
Mother and father's years of schooling	Education register, 1999
Grades in math/Swedish	School register, 1999
Country of birth	Register of the total population, RTB 1999
Teacher-student ratio	Teacher and High school register 1978-94
Teacher experience and education	Teacher register 1978-94

C. Programming files

The Stata do file schooldays.do creates all the tables and figures in the article, as well as the appendix tables and figures which are available online. It uses the data described above as an input.