Plan Overview

A Data Management Plan created using DMPonline

Title: En gymnasieutbildning för alla? Hur påverkar gymnasieutbildningens utformning risken för avhopp och arbetslöshet?

Creator: Jonas Larsson Taghizadeh

Principal Investigator: Jonas Larsson Larsson Taghizadeh, Marcus Österman

Data Manager: Jonas Larsson Larsson Taghizadeh, Marcus Österman

Project Administrator: Jonas Larsson Larsson Taghizadeh, Marcus Österman

Affiliation: Uppsala University

Funder: Swedish Research Council

Template: Swedish Research Council Template

Project abstract:

Dropping out of upper secondary education is associated with a significantly higher risk of unemployment and living in poverty both in Sweden and in the rest of the EU. Increasing the proportion of students who completes the upper secondary level is therefore high on the policy agenda. Reforms aimed at this have mainly focused on introducing vocational programs. However, it is unclear in the literature whether such reforms can really reduce school drop-outs and youth unemployment. Only a handful of international studies have in credible ways managed to isolate the effects of different upper secondary education programs on these outcomes. No studies have yet investigated the latest major reform of the Swedish upper secondary school (Gy11), despite its aim of reducing dropouts by reforming the vocational programs.

The aim of our project is therefore to study how the design and content of upper secondary education affect the risk of students dropping out. We will also study the risk of unemployment among drop-outs. Among other things, we intend to study the effects of attending general and vocational programs before and after Gy11. We will use so-called discontinuities in the admission to the upper secondary school to handle a difficult method problem: separating the effects of the educational orientation from the students' background. We will also study how the outcomes differ depending on gender and the parents' background. Two researchers will carry out the project 2022-25.

ID: 88067

Start date: 01-01-2022

End date: 31-12-2026

Last modified: 08-11-2021

Grant number / URL: 04226

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

En gymnasieutbildning för alla? Hur påverkar gymnasieutbildningens utformning risken för avhopp och arbetslöshet?

General Information

Project Title

En gymnasieutbildning för alla? Hur påverkar gymnasieutbildningens utformning risken för avhopp och arbetslöshet?

Project Leader

Jonas Larsson Taghizadeh

Registration number/corresponding, date and version of the data management plan

2021-04226

Version

1

Date

211108

Description of data – reuse of existing data and/or production of new data

How will data be collected, created or reused?

Vi kommer i projektet inte att själva hämta in ny data utan vi avser att bearbeta avidentifierade registerdata på individnivå från SCB.

What types of data will be created and/or collected, in terms of data format and amount/volume of data?

Ingen ny data kommer att skapas eller samlas in utan vi avser att använda ett urval av redan existerande registerdata från SCB och bearbeta dessa. Bearbetning går i huvudsak ut på att, för det första, identifiera elever som precis kom in på sitt förstahandsval i gymnasiet eller inte gjorde det. På så sätt skapas en variation i vilken utbildningsinriktning som individer går som inte hänger ihop med bakomliggande faktorer. Detta är en s.k. regression discontinuity design (RDD) och bygger på idén att när en skarp gräns avgör utfallet för individer kan vi jämföra dem som ligger precis över gränsen med som ligger precis under för att studera effekten av det som gränsen påverkar. För det andra kommer vi, när väl eleverna är identifierade, att lägga på data om elevernas socioekonomiska bakgrund samt framtida utfall vad gäller karriärer från andra register.

Documentation and data quality

How will the material be documented and described, with associated metadata relating to structure, standards and format for descriptions of the content, collection method, etc.?

I och med att vi använder befintliga data finns datadokumentation tillgänlig via SCB. Vår bearbetning av data kommer att ske genom skript (do-filer) i statstikprogram, vilket på så vis också kan återges. Vi kommer också skapa ett dokument (metadata) där innebörden av samtliga variabler beskrivs och där det framkommer vilket register som variablerna härstammar från.

How will data quality be safeguarded and documented (for example repeated measurements, validation of data input, etc.)?

Datakvaliteten säkerställs i grunden genom att den samlas in från offentliga svenska databaser som SCB administrerar. Dessa databaser håller en internationellt sett mycket hög datastandard. Alla originalversioner av variablerna från registren kommer att bevaras i en orginalfil och datafiler med eventuella ändringar/transformeringar/indexeringar av dessa variabler kommer att sparas som nya filer.

Storage and backup

How is storage and backup of data and metadata safeguarded during the research process?

Vi kommer att utgå från SCBs orginaldata och skapa våra dataset genom skrivna script (do-filer). Det innebär att våra dataset alltid enkelt kan återskapas från SCBs orginaldata. All data kommer att ligga på SCBs servrar som också inkluderar backup.

How is data security and controlled access to data safeguarded, in relation to the handling of sensitive data and personal data, for example?

I vårt projekt kommer vi att använda oss av anonymiserade registerdata som kommer att ligga på en säker server hos SCB vilken bara projektdeltagarna har tillgång till.

Legal and ethical aspects

How is data handling according to legal requirements safeguarded, e.g. in terms of handling of personal data, confidentiality and intellectual property rights?

I vårt projekt kommer vi att använda oss av anonymiserade registerdata som kommer att ligga på en säker server hos SCB vilken bara projektdeltagarna har tillgång till. Enligt den Centrala Etikprövningsnämndens praxis utgör anonyma registeruppgifter av detta slag ett nästan obefintligt brott mot den personliga integriteten (se till exempel Dnr Ö 21-2013).

SCBs har också strikta rutiner vid beställning av data för att säkerställa att aktuella lagar och riktlinjer följs. Vi kommer självklart helt och hållet agera i enlighet med dessa bestämmelser och SCBs förhållningsregler.

How is correct data handling according to ethical aspects safeguarded?

Projektet kommer omfattas av en etikprövning. Antingen genom en särskild ansökan eller genom tilläggsansökan till en befintlig etikprövning för ett snarlikt projekt. Som nämnts ovan är användning av registerdata på detta sätt i enlighet med existerande praxis

Accessibility and long-term storage

How, when and where will research data or information about data (metadata) be made accessible? Are there any conditions, embargoes and limitations on the access to and reuse of data to be considered?

Eftersom vi använder oss av registerdata på individnivå finns det inte möjlighet att direkt tillgängliggöra data. De data vi analyserar är dock i huvudsak möjliga för alla forskare att ansöka om hos SCB. Vi kommer också beskriva hur våra analyser är genomförda vilket innebär att den med tillgång data har möjlighet att replikera våra analyser.

In what way is long-term storage safeguarded, and by whom? How will the selection of data for long-term storage be made?

Datasetet kommer att bevaras på lösenordsskyddade servrar hos SCB
så länge som etikprövningen ger oss tillgång till datan. Metadatan (do-filen) kommer finnas bevarade på lösenordsskyddade servrar på statsvetenskapliga institutionen
så länge som någon av de två forskarna i projektet är aktiva på institutionen

Will specific systems, software, source code or other types of services be necessary in order to understand, partake of or use/analyse data in the long term?

Vi avser att sammanställa och analysera datan i databehandlingsprogrammet Stata 15.

How will the use of unique and persistent identifiers, such as a Digital Object Identifier (DOI), be safeguarded?

Eftersom datasetet inte kommer att göras tillgängligt för andra forskare i syfte att säkerställa forskningspersonernas integritet kommer vi inte att använda sådana identifierare.

Responsibility and resources

Who is responsible for data management and (possibly) supports the work with this while the research project is in progress? Who is responsible for data management, ongoing management and long-term storage after the research project has ended?

Ansvaret för datahantering både under projektet gång och efter projektets avslut kommer att ligga på projektets två medlemmar, Marcus Österman och Jonas Larsson Taghizadeh.

What resources (costs, labour input or other) will be required for data management (including storage, back-up, provision of access and processing for long-term storage)? What resources will be needed to ensure that data fulfil the FAIR principles?

Ekonomiskt ansvar för datahantering kommer under studiens genomförande ligga på projektet och dess deltagare.