

Index:

- I. Introduction
- II. Sources
- III. Variable definition

I. Introduction

The paper tracks the sample of PSU takers from high school graduation until they are several years into the labor market.

The main analysis is performed for three student cohorts, graduating from high school in 2006, 2007, and 2008, for which I observe complete educational histories and longer labor market participation. For these cohorts, the data allows for the observation of outcomes up to 2016 for their academic outcomes and 2017 for labor market outcomes. In addition, younger cohorts will be briefly used to have a broader view of the evolution of the impacts.

II. Sources

A. Educational data

1. PSU scores and socioeconomics at high school graduation
Source: DEMRE/CRUCH
2. Enrollment history and completion
Source: SIES
3. Financial aid assignment
Source: SIES and INGRESA

B. Labor market data

1. Unemployment insurance (UI) [main data]
Source: MinTrab
2. Publicly available earnings information about public employees
Source: Web scrape

A.1. PSU score:

The running variable is the PSU score that comes from administrative data on PSU collected for CRUCH universities by DEMRE. The data includes test takers for the high-school cohorts from 2006 onward. Beyond PSU scores, includes high school GPA, demographics (birth date, gender), and details about household socioeconomic backgrounds, such as household size, parental education, and parents' labor status.

The data with the public mask “mrun” can be accessed with a formal request from SIES.

Contact: estadisticas@mineduc.cl

A.2. Enrollment and completion

SIES provide data with the enrollment records from the Ministry of Education. The information is publicly available and with easy access as explained in below. The data contains yearly enrollment and graduation for all programs in the country.

The data with the public mask “mrun” can be accessed directly from SIES (see links below).

The great majority of data can be downloaded with **mrun** from SIES:

https://datos.gob.cl/organization/subsecretaria_de_educacion or

<https://datosabiertos.mineduc.cl/>

Data Include:

- i. All higher education enrollment
Link: [Matrícula en Educación Superior](#)
- ii. All higher education completion,
Link: [Titulados en Educación Superior](#)
- iii. Application to FUAS
Link: [Postulaciones Formulario Único de Acreditación Socioeconómica \(FUAS\)](#)
- iv. Assignment to benefits
Link: [Asignaciones Formulario Único de Acreditación Socioeconómica \(FUAS\)](#)
- v. Test scores for the newest tests since 2021.
Link: [Pruebas de admisión a la educación superior](#) (PDT2021, PDT2022, PAES2023 AND PAES2024)

For replication purposes, the public data is downloaded to a Github repository:

https://github.com/AlexSolis11/Data_JOLE_Loans_Chile

A.3. Loans and financial aid assignment

SIES also provide application to financial aid through the FUAS data set. The key variable is the family income quintile that determines eligibility. The data also contains assignment to financial aid and loan take-up of the TUL program.

The SGL assignment comes from the INGRESA commission,18 containing the assignment and take-up for the SGL for the first year of enrollment.

B.1. Unemployment insurance data.

The primary data source is the mandatory unemployment insurance (UI) recorded by the Ministry of Labor of Chile. The data is proprietary and can be accessed through a research application.

To apply for the data contact Fernando Carreño from the Studies department:

fcarreño@mintrab.gob.cl

The UI data contain individual-level information that is merge using the national ID number.

B.2. The public sector.

To extend coverage to the public sector, I obtained wages in 2017 through scraping public websites, including 500 different organizations within the Chilean government (e.g., ministries, municipalities, and public hospitals). The information is merged with the PSU data set through full names.

All these governmental dependencies are mandated by law to publish detailed payroll information for all (civilian) employees in the public sector.

The web appendix provides a detailed description of the scraping procedure and the several tests I performed for accuracy and consistency.

The data scrapper and the data can be downloaded from:

https://github.com/AlexSolis11/public_sector_data_chile

III. Variable definition

1. Variable for balance of covariates

Variables	: Description
female	: sex indicator (woman=1).
qqt1	: Income quintile classification in t=1.
agepsu	: Age at the PSU in t=1.
prom_em	: High school GPA.
muni	: Graduated from municipal high school.
fated	: Father years of education.
moted	: Mother years of education.
single	: Students not married at PSU.
work	: indicator whether student work at PSU.
hh_no	: Household size.
fat_work	: Indicator whether students father work at PSU.
motcollege	: Indicator whether mother has college education.
fatcollege	: Indicator whether father has college education.
Housewife	: Indicator whether mother do not work.

2. Educational variables

Educational Attainment

Variables	: Description
ever_grad_u	: Ever graduated from a university program.
ever_grad_ucruch	: Ever graduated from a CRUCH university program.
ever_grad_upriva	: Ever graduated from a Private university program.
ever_grad_voca	: Ever graduated from a vocational program.
ever_grad_anyies	: Ever graduated from any higher education program.
years_college	: Years of enrollment in university programs.
years_vocational	: Years of enrollment in vocational programs.

Education Quality:

Variables	: Description
accredited_prog	: Indicator whether the program enrolled is accredited.
acredita_ies_v2	: Indicator whether the institution enrolled is accredited.
jce_per_student2	: Number of Full-time-equivalent teachers.
doctor_per_student2	: Number of Full-time-equivalent teachers holding a PhD degree.
Tuition	: Tuition cost in the enrolled program
dur_total_carr	: Nominal length of the enrolled program

3. Loan and financial aid Eligibility

Variables	: Description
everelig8	: Ever eligible for University loans (measured 8 years after high school)
ever_elig_voc_loan	: Ever eligible for Vocational loans
ever_anyLoan8	: Ever eligible for any loans (University or vocational)
everelig_BC8	: Ever eligible for Bicentenario grant.
ever_bbic8	: Ever received Bicentenario grant
have_SGL_Vocationalt1	: Indicator for vocational loan take up in t=1
total_loan1	: Debt amount from any loans
total_CollegeLoant1	: Debt amount from University loans in t=1
total_fscu1	: Debt amount from TUL loan in t=1
total_SGL_Colleget1	: Debt amount from SGL-university loan in t=1
total_SGL_Vocationalt1	: Debt amount from SGL-vocational loan in t=1
total_bbic1	: Bicentenario Grant amount received in t=1

4. Labor market variables

Variables	: Description
partLM_t9	: (indicator) Participation in the labor market in t=9
partLM_t11	: (indicator) Participation in the labor market in t=11
experience9	: Accumulated months of experience since high school until t=9.
experience11	: Accumulated months of experience since high school until t=11.
log_suma_salary9	: Yearly earnings in t=9 in logs.

log_suma_salary11 : Yearly earnings in t=11 in logs.