Widening Participation in Higher Education: Policies and Institutional Settings

Cross-Country Perspectives and an Empirical Analysis of Chile

Daniel A. Uribe Jorquera
UCL Institute of Education
Department of Social Science

Thesis for the degree of Doctor of Philosophy
February 2017

Dedicated to my late father Leonardo. My achievements are also his

DECLARATION

I, Daniel Uribe, hereby declare that the work presented in this thesis is entirely my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis accordingly

Word count (exclusive of appendices and bibliography):

55,073 words

ABSTRACT

In the last three decades, Higher Education (HE) has experienced an unprecedented expansion worldwide. In many countries, governments have transferred the cost of HE from taxpayers to individuals and households as a means of increasing the provision on a financially sustainable basis. Most policies have attempted to address the issue of low-income students' participation by setting student aid policies for those unable to afford HE costs. Nonetheless, the starting point of this thesis is that the goal of equity in HE should not begin with, or be confined to, HE policy but must address school education as well. I investigate the effect of the socioeconomic distribution of school achievement on HE enrolment rates in a cross-country framework. I find a mild but statistically significant negative association suggesting that the more school achievement is determined by socioeconomic factors, the less participation in HE is observed. Next, I evaluate the impact of a reform to the student aid system in Chile using household surveys and regression-based and differences-in-differences evaluation techniques. I find the reform increased the probability of access of low-income students to HE by 6 percentage points, or 20 per cent in proportional terms.

After having researched the effects of inequality of school achievement, I focus on the design of student aid and its effect on persistence and dropout. In particular, I investigate the level of harshness of different aid programmes and its effect on students' persistence, completion, and dropout rates. By specifying a logistic multinomial model, I compare the effect of two loan programmes, an income-contingent loan and a mortgage-type, bank-managed, government-guaranteed loan. The harsher, mortgage-type loan was associated with increased persistence and higher completion rates but no difference in dropout rates. Nonetheless, this association was only observable for low-income students; loan harshness made no difference in completion rates for better-off students. In other words, harsher loans seem to be a deterrent only for poor students. This introduces an ethical dilemma: although harsher aid may be more effective, should student aid be disproportionately putting pressure on the poorest students? However, this may in turn reflect poor student's relative higher ability rather than a differential deterrent effect.

ACKNOWLEDGEMENTS

My interest in education comes from my own experience of growing up in a teacher's family. My mother, sister, and other relatives have embraced education in their lives and professional careers. So have I, but through research and policymaking instead of school classrooms.

Doing a PhD would have never been possible without the support and determination of Lorena, my wife, who persuaded me to take on the opportunity and embark on a journey that did not convince me. After long conversations on almost every aspect implied, she made me realise that we had to take the chance. I owe my daughters, Isidora and Florencia, the great joyfulness they have brought into our lives in London and the privilege of being so close to them while they grow. Had it not been for these three ladies, this thesis would have never been written.

I would like to thank my supervisors Kirstine Hansen and Gillian Wyness for their constant support, prompt feedback, and involvement in my research. I express my gratitude to my fellow PhD researchers Diego Carrasco, Rodrigo Torres, Rose Cook, and Christopher Hein. Their valuable feedback helped me detect some weaknesses and make my arguments clearer. I am also very grateful to Jamie Craver, who reviewed key sections of the final draft and helped me leave it written in proper English.

Luis Eduardo González, Pilar Armanet, and José Joaquín Brunner have been hugely influential throughout my professional career. I have learnt the complex relationships between research and policymaking with each of them. I also thank the Social Division at the Chilean Ministry of Social Development for providing CASEN socioeconomic surveys, and Alejandra Contreras and Felipe Gutiérrez from *Comisión* INGRESA for allowing me to use and exploit their databases.

Of course, all the contents of this thesis are my sole responsibility.

The Chilean Government, through CONICYT's Advanced Human Capital Programme, Becas Chile, Scholarship #72110230, sponsored my doctoral studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AFD	Direct Block Grant
AFI	Indirect Block Grant
CAE	State Guaranteed Loan
CASEN	National Socioeconomic Characterisation Survey
CFT	Technical Training Centre
CNED	National Council of Education
CRUCH	Council of Rectors of Chilean Universities
ESCS	OECD Economic, Social and Cultural Status Index
FSCU	Income Contingent Student Loan
FUAS	Student Support Application System
IADB	Inter-American Development Bank
ILO	International Labour Organization
INGRESA	Administrative Commission of the Higher Education Loan System
IP	Professional Institute
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
MDS	Ministry of Social Development (formerly MIDEPLAN)
MIDEPLAN	Ministry of Planning
MINEDUC	Ministry of Education of Chile
OECD	Organisation for Economic Co-operation and Development
PAA	Scholastic Assessment Test (until 2004)
PISA	Programme for International Student Assessment
PSU	University Entry Test (since 2005)

SIMCE	Education Quality Measurement System
UIS	UNESCO Institute of Statistics
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
WDI	World Development Indicators

1. EQUITY IN HIGHER EDUCATION: CORE ISSUES

1.1. Introduction

In 1983, Burton R. Clark published his seminal work, 'The Higher Education System', (Clark 1983), which has been considered a milestone in sociology of higher education (HE). Amongst several contributions, Clark proposed a comprehensive model to analyse the HE system's coordination mechanisms, the so-called 'Clark's coordination triangle'. The model approaches the HE field as shaped by three major forces: the state, the market, and the academic oligarchies. The state intervention in HE; the competition for students, academic staff, and funding (the market); and the influence of the organised academic profession have operated in any one HE system, and a particular combination of all three elements has been a feature of national HE systems historically.

Clark's model relies on an extensive historical and comparative analysis of HE, the main focus of which was to disentangle the fundamental features of one of the oldest and most resilient social institutions in the Western World: the University. Although

a model developed in 1983 would not be entirely appropriate nowadays (Brennan 2010), it is important to remember that the presence of the market in HE is neither a recent nor an unique feature of US-like models, but a longstanding driving force in HE in general. Indeed, market mechanisms (*e.g.*, academic staff recruitment, student selection, and competition for funding) not only influence privatised HE systems but are present in publicly financed and coordinated systems as well.

Since the 1980s, HE coordination has moved from the state to the market. The increasing reliance of policy on market mechanisms is evident when analysing a number of changes in HE. Firstly, market coordination has replaced political negotiation mechanisms between the state and academic oligarchies. Take, for instance, the introduction of performance agreements between governments and higher education institutions (HEIs). Although terms might be negotiated between governments and HEIs, the room for political manoeuvre is limited because the state negotiates with individual institutions rather than a 'sector' or the 'academic corporation'. Secondly, demand-side funding has become more important than institutional funding in many countries. In effect, governments have cut back or frozen block grants, forcing HE funding sources to diversify through the introduction or liberalisation of tuition fees as well as the involvement of the private sector in HE. In general, the trend has been to transfer costs from the state to individuals (though some countries still rely on general taxation to fund HE or, like Germany, have reinforced free HE).

Thirdly, managerialism has become evident in HE. Indeed, actions such as setting academic output goals, imposing scientific productivity standards on academic staff, and moving student services towards meeting customers' demands show how HE has adapted to a new environment deeply penetrated by the market and competition for resources. Although Clark's original view stressed the contradiction between managerialism and the disciplinary nature of the academic authority, HEIs are currently managed in a business-like fashion: there is fierce competition for funding, and external pressures, rather than strictly academic concerns, have influenced academic priorities.

Finally, universities now attract students by using marketing strategies such as brand positioning, product differentiation, and strategic management techniques. The admission process starts taking place well in advance as marketing, outreach, and recruitment strategies are deployed earlier. As students and their families became customers, universities mimicked firms, and governments loosened their direct regulation over HE, handing it over to market forces through 'soft' or 'distance' regulations.

Marketisation has arrived together with massification. The universalisation of secondary education, the diversification of the student body in educational and socioeconomic terms, and the promise of social mobility through HE has resulted in the highest HE participation levels ever seen. The political discourse on the need for countries to boost innovation and growth based on enhancing the skills of their labour forces and accumulating human capital has deemed HE as a central actor.

Apart from efficiency considerations, marketisation relies, in part, on an equity argument (Dill 1997); that is, privatisation of HE's provision and funding would allow for HE's expansion without diverting scarce resources to those who truly benefit from HE – namely, those who are better-off to begin with (Psacharopoulos 2008). However, marketisation has raised concerns with regard to equity of access. Capital markets are often not prepared to lend money to the number and diversity of students sought, and students may not consider the whole benefit of HE when applying for aid either because they lack information or are risk averse (Barr 2012; Goodman and Kaplan 2003).

Provided that the costs of HE have been transferred to individuals: How do HE systems deal with financial barriers to access? Which policies are effective? How does HE policy make sure that student aid policies – specifically loans – do not deter low-income students from entering HE? To what extent does financial aid affect not only access but also persistence and course completion? The policy agenda and the political discourse stress the need for widening participation; marketisation seems – at the very

least – problematic to achieving that goal equitably. All of these concerns are common in literature and policymaking, as I shall discuss later.

In this thesis, I shall examine several aspects affecting equity in HE with a special focus on student aid as the equity policy *par excellence*. While the general context described above refers mostly to changes in HE systems and policies, addressing the problem requires an approach that looks beyond the field of HE, as access to HE is also determined by prior education and social background, which are not affected by HE policies.

Although most equity policies have focused on helping poor students overcome financial barriers to HE access, I argue that such policies would not have a relevant equalising effect unless student aid policies move beyond individual characteristics. There are, of course, individual characteristics such as ability, motivation, and prior academic achievement that may, in part, predict outcomes relevant to HE. Yet, more importantly, there are forces affecting HE participation that are not attributable to individuals: family background, exposure to education in the household, how valued education is in families' day-to-day lives, neighbourhood, and student social environment also model HE participation. HEIs can support students at risk of departure by providing an environment for social and academic integration and by taking measures aimed at enhancing the student experience. Governments play a central role by designing programmes to promote a broader equity agenda, providing resources, targeting support for different social groups, and building the appropriate governance mechanisms concerning regulation and funding. The joint action of variables in the above factors should be taken into account when designing and assessing equity and inclusion policies.

Apart from the above considerations, a central point of this research is that increasing equity and social inclusiveness in HE translates into weakening the relationship between socioeconomic origins and educational outcomes such as access, persistence, employability, or expected earnings. In other words, undermining this deterministic relationship should be a measure of policy effectiveness. Although normative, this

statement has sociological and philosophical implications that I would like to make explicit. Firstly, weakening the predictive power of social background on educational outcomes is at the heart of one of the main promises of HE in the Western World: social mobility. Nevertheless, the optimistic view that education will improve one's quality of life contrasts with a body of research that has found participation in HE has remained highly unequal despite having increased dramatically. Secondly, the philosophical implication is that there needs to be a precise definition of equity in HE. The baseline definition I shall use is the following: two students with the same abilities and preferences should receive the same education, regardless of any other considerations in social, cultural, or economic terms.

Equity is a multidimensional concept that needs to be studied according to its complexities. As such, I shall discuss what equity means and how it applies to HE.

1.2. What is Relevant when Studying Equity in HE?

According to Hansen (1972), equity is seen as normative, yet no analytical tools have been developed to study it. Le Grand (2007) distinguishes among equity of opportunities, processes and outcomes. Equity of opportunities and processes mean that, for instance, access to and permanence in HE should depend on factors other than income, gender, geography, and the like. This concept is similar to what Barr (2012) well defines: two persons with the same abilities and preferences (the only factors that matter) ought to have access to the same education. But Le Grand (1982) made a crucial point: tackling inequality starts by recognising that many differences arise from factors beyond one's control, though the limits of such factors are highly contested -see, for example, the debate about talent and luck (Anderson 1999; Dworkin 1981; Knight 2013).

With a focus on HE, Lemaitre (2005) distinguishes four dimensions of equity. First, 'equity of study opportunities', or whether the institutional settings regulating financial support, admissions, and geographical distribution of places in HE fulfil the requirements of poor students. Second, 'equity of access', which considers the distribution of HE enrolment across socioeconomic groups and the policies designed to

help poor students access HE. Third, 'equity of persistence', or the likelihood of students completing HE according to their socioeconomic characteristics. Finally, 'equity of results', which considers how labour market outcomes are distributed according to socioeconomic origins.

This thesis focuses on policies addressing equity of access and equity of persistence, applying the above definitions.

Policies concerning equity in HE might be seen as importing from, and exporting to, the broader equity and social justice agendas (Brennan and Naidoo 2008). HE imports equity and social justice agendas from the wider society by addressing gender gaps, minorities' participation, and socio-economic exclusion. On the other hand, HE contributes to equity and social justice agendas by improving social cohesion and civic engagement, and by making societies fairer and more equitable (*e.g.*, graduates receive higher quality jobs and their living standards are far better than those of the previous generations).

The more inequitable the socioeconomic composition is in HE, the more HE expansion relies on formerly excluded students. At the same time, HE provides opportunities for social mobility and improvement in living standards in any society. Although, as I shall discuss throughout this thesis, there is much promise left unfulfilled in the above statement, it is incontestable that HE is currently more equitable than three decades ago – anywhere.

Nevertheless, there is much empirical evidence pointing out that there are unfair intergenerational transmission mechanisms affecting equity. One failure of the optimistic premise above is that it underestimates the effect of factors external to individuals and families. For instance, Bowles & Gintis (2002) tested the effect of IQ in the inheritance of socioeconomic status (SES). They found that factors other than IQ – wealth, schooling and race – better explained the correlation between intergenerational economic status. Though these findings do not exactly concern HE, they do express the main rationale behind this thesis: "a policymaker seeking to level

the playing field might use these results to design interventions that would loosen the connection between the economic success of parents and the economic prospects of children" (Bowles and Gintis 2002:22).

Indeed, inequity has long been a concern in sociology of education. HE massification was the main educational phenomenon in the US in the 1960s, and it triggered major concerns regarding access and practices at college level affecting young people's prospects of social mobility (Clark 1973). Walpole (2007) proposed the term 'economically and educationally challenged students', to define low SES, working class, first-generation students. She identified prior educational experiences, SES, parental income, education, and occupation as core variables affecting access and outcomes in HE. In the UK, Gilchrist, Phillips & Ross (2007), on the basis of empirical models and research, hypothesised that low participation of working-class students was due to factors such as lack of information about HE opportunities, perceptions that foregone earnings would not be offset by HE, low entry qualifications, risks involved in financing HE studies, and even perceptions that HE would threaten class identity.

McDonough and Fann (2007), by reviewing 114 journal pieces dealing with inequality of access to HE, divided the literature on college access research into three broad categories: individual, organisational, and field level. Research on college access tends to consider SES as the main influential factor, including family characteristics such as parental involvement, geographical factors, and aspirations. From an organisational standpoint, research identifies academic preparation, characteristics of secondary education, quality of counselling, influence of teachers, university recruitment practices, and diversity of HEIs with special reference to socio-economic segmentation. Interestingly, field-level analysis, by focusing the attention "on the macro-level changes in the institutions, professions, and technology of admissions in order to understand how student perceptions and actions grow from, as well as influence, organizational and institutional perceptions and actions" (McDonough and Fann 2007:77) has a strong explanatory power, as it contributes to a better understanding of the interplay between institutional and individual responses. The main critique from the authors to

the then-current research is the predominant focus on individual characteristics to the neglect of other influential explanatory variables.

In contrast, the literature on student persistence and dropout has always paid attention to factors external to individuals. Student persistence is modelled by individual characteristics, but social and academic integration are crucial to understanding student withdrawal decisions (Tinto 1975, 2006). Other explanations have paid more attention to psychological mechanisms while still others have included economic considerations (Cabrera, Nora, and Castañeda 1992; Stampen and Cabrera 1988), as well as explanations heavily based on choice and economic behaviour (Stratton, O'Toole, and Wetzel 2008). Other studies have highlighted that even though access to HE was equal in socioeconomic terms, it would not translate into equal outcomes such as graduation and dropout: low SES students are more likely to drop out and less likely to obtain a good degree (Crawford 2015). I shall discuss this in more depth in Chapter 5.

In the mid-nineties, Baker and Vélez (1996) reviewed the literature on access to HE of women and minorities and found a shift of the focus: family and class background have lost predictive power, whereas ability, school achievement, and financial aid have become more determinant. Nevertheless, changes in student populations – such as the increasing participation of mature students who are economically independent – have widened the study of social class, ability, and other social characteristics. The emergence of non-traditional students also challenges the traditional conception of student integration, which relies on building social ties between students and matching what HEIs and students expect from each other.

1.3. Thesis Outline

Many HE equity issues are well beyond the field of HE and rest, largely, in school education. The implications for research and policy are as follows. First, there is less space for HE policy to reach equity given that school education is critical for acquiring the necessary skills to succeed in HE and might, therefore, turn into the main source

of inequity. It is hardly possible to have an equitable higher education system and, at the same time, an unequal and segregated school education since high school achievement, the very precondition to HE entry, would be determined in socioeconomic terms. In summary, I seek to establish whether an unequal socioeconomic distribution of achievement in school can be seen as a barrier to HE expansion.

Research has been prolific in studying the effects of student aid on access and persistence in HE, as I shall show in Chapters 4 and 5. Nevertheless, there are unexplored issues that may contribute to a better understanding of the influence of student aid on HE equity and what to expect from it. Firstly, research on student aid evaluation is scarce and hardly addresses issues at national levels. The complexities of student aid systems, regional specificities (which operate at the same time as national policies), and changes in HEIs eligibility for aid programmes make it difficult to assess policies at national levels. As I shall discuss in subsequent chapters, policy changes in tuition fees have triggered research on their impact on low SES students' participation, the main question being whether and how the introduction (or increase) of fees affects the chances of poor students attending HE. Nevertheless, student aid is also in place in publicly funded and no-fee HE systems, so its relevance is not only confined to fee-charging systems. Current approaches fail to explain what happened to countries like South Korea and Chile. In both cases, the private sector led HE expansion and governments had few tools to guarantee access to low SES students. However, inequality in access remains high in Chile but decreased in South Korea. I hypothesise that equity policies in HE are effective so long as there are well-timed policies, institutional mechanisms, and practices that contribute to 'levelling the field'. These institutional mechanisms are to be found in school systems performance and the socioeconomic distribution of school achievement.

Secondly, less attention has been paid to reforms whose aim is to restructure student aid by changing eligibility rules and introducing new designs. In relation to aid design, policies such as re-engineering student loans, changing the balance between grants and loans (*e.g.*, scrapping grants and replacing them with loans), and building aid packages

tailored to the characteristics of different social groups define a new structure of incentives that may trigger behavioural responses.

This thesis seeks to answer the following research questions:

Question 1. In relation to how inequality of school achievement affects HE participation, this thesis asks: Is there a cross-country effect of inequality of school achievement on HE participation or it is a country-specific issue? Does inequality of school achievement limit the expansion of HE systems and, if so, how? Does the explanation hold once controlling for a number of country characteristics, institutional settings, and socioeconomic features?

To answer Question 1, I have consolidated a cross-country panel for 63 countries, using the OECD Programme of International School Assessment (PISA) scores, UNESCO statistics, and World Development Indicators from the World Bank data centre. This dataset includes variables such as enrolment rates, school achievement, and socioeconomic characteristics of countries. This allowed me to use variables informing the general economic and social context and PISA performance, and those indicators dealing with socioeconomic issues, access to education, segregation of school systems, and persistence of SES in school achievement.

Most investigations in the field have focused on studying an individual country or a handful of countries by using existing administrative sources and microdata and have compared, for instance, the effect of school achievement on HE entry as well as the persistence of SES on HE access. A key issue is that micro data and administrative records availability are mainly restricted to a small group of developed countries and this does not allow conclusions to be drawn beyond the countries analysed and their own contexts.

On the other hand, it is common to find relevant research relying on macro data in the field of economics, for instance, with international comparisons of returns to education (Psacharopoulos 1988; Psacharopoulos and Patrinos 2004) and the

relationship between education and income inequality (De Gregorio and Lee 2002). However, there is not a widespread use of international data to predict educational outcomes as seen in economic research. Macro analysis may make significant contributions since it is easy to streamline its results with policymakers' interests.

Question 2. In relation to student aid, this thesis carries out an impact evaluation exercise of a massive student aid reform that took place in Chile in 2005. The research question I address is: To what extent does widening access to financial aid enhance the chances of low-income students accessing HE?

To answer Question 2, I review research on the effect of student aid on access to HE. Empirically, I study the Chilean reform to student aid in 2005, where a massive financial effort alongside a deep re-engineering took place. Nevertheless, Chile lacks long-standing administrative records, so linking HE access, school education, and socio-economic characteristics is not yet possible. In turn, I use comparable socioeconomic surveys for the period 1990 to 2013, the National Socioeconomic Characterisation Survey (CASEN), which represents the entire Chilean population. In order to evaluate the impact of student aid, I undertake a difference-in-difference (D-I-D) analysis.

The attractiveness of D-I-D is that it allows the impact of student aid to be evaluated over time by setting different aid packages. In 2003, aid was restricted to a small group of universities; after 2005, the government made more HEIs eligible for student aid, thus making it available to more students. The government also introduced a new loan scheme, relaxed the academic requirements for grants, and increased amounts in order to finance a higher proportion of tuition fees.

Question 3. In relation to aid structure, I analyse the effect of aid composition on persistence and dropout. I specifically ask: How does debt structure (amount and composition) affect the probability of dropping out/ course completion? Do bank or harsher loans increase the probability of completing studies (or deter students from dropping out) more effectively than other student aid mechanisms? Does loan

structure have different effects for certain type of students (depending on income levels, prior achievement, or type of HEI)?

To answer Question 3, I assess the composition of student aid on student persistence and dropout. I use administrative data containing information on students who have received a combination of loans and grants but due to a massive administrative error that ended up offering a new loan scheme (*Crédito con Aval del Estado*, CAE) to almost all students meeting the academic requirements without considering their socioeconomic situations. This provides a unique opportunity to study the effects of student aid on persistence and dropout because students from the entire socioeconomic spectrum took up loans not normally available to them. CAE students, in general, represent the new profile of Chilean students resulting from the universalisation of secondary education in the nineties and the sustained and expansive student aid policy since then. A number of students now come from low-performing schools, score lower in the university entry test (PSU), and are the first generation in their families in HE. I also linked records to school marks, PSU scores, previous schooling, and parental education.

The thesis is structured as follows:

I start, in Chapter 2, by explaining why Chile presents a worthy case study. Chile introduced market-oriented reforms in 1981, where HE consisted of a few publicly funded universities and enrolment rates reached 5 per cent. In the last 25 years, Chile has quadruplicated its undergraduate enrolment and increased access to low-income students, although access remains highly unequal in socioeconomic terms.

Chapter 3 is an empirical cross-country study aimed at studying the relationship between the distribution of school achievement and participation in HE. Its purpose is to test whether socioeconomic inequality of school achievement is an obstacle for countries seeking to expand HE. I have compiled data from a diversity of international organisations and development agencies such as World Bank, UNESCO and OECD.

I do find a mild, negative but statistically significant association between inequality of school achievement and HE enrolment rates.

In Chapter 4, I estimate the impact of the student aid reform undertaken by the Chilean government in 2005. I use both observational and experimental approaches. Given the nature of the data, I obtain a relevant and significant estimate of the effect of student aid reform.

Chapter 5 examines the effects of credit harshness on persistence, course completion, and dropout, featuring an original contribution I make to the study of aid in HE. I find no significant differences between students being totally funded with a state-guaranteed mortgage-type bank loan (CAE) and those who used CAE to complement other forms of student aid. Nevertheless, when it comes to low-income students, privately funded students are significantly more likely to complete studies. This introduces an additional complexity to policymaking since private loans seem to deter only poor students from dropping out and/or delaying study completion. Is it fair to help the poorest students by virtually forcing them to complete their course?

Chapter 6 outlines the main conclusions of the thesis, policy implications, and a future research agenda in the subject.

The value added of this thesis lies in its following aspects. First, it addresses the relationship between inequality in school systems and learning outcomes, and the consequences for HE. This is at the centre of the policy debate since HE expansion crucially depends on the access of low-income and non-traditional students. Second, I present a new perspective on student aid by regarding aid packaging and aid composition as triggering different behavioural responses, which may be due to poor students being risk averse. Risk aversion needs to be studied beyond loan take up, but also as a determinant variable shaping the decision of persisting or dropping out. I undertake the first attempt to assess those factors at the country level and rely on an unprecedented approach to review the complexities of student aid design. Third, I open a new line of enquiry, as the natural continuation of this work involves other

disciplines such as economic psychology and behavioural economics in order to reach a better understanding of the mechanisms behind risk aversion and behavioural responses of low SES students.

2. CHILE AS A CASE OF STUDY

2.1.Introduction

Chilean HE has some unique characteristics: a high degree of privatisation, extended cost sharing, a high level of tuition fees relative to living standards, and a high persistence of socioeconomic background on school achievement and access. The country has experienced an explosive expansion of HE since the early nineties. Also, more than two-thirds of HE students are the first generation in their families to access HE, and many have even been the first generation to complete secondary education.

Since the early 1980s, when the military dictatorship decided to restructure HE in depth, Chile has been seen as a paradigmatic example of privatisation, cost sharing, and increasing access to HE. The expansion of private provision was timid in the 1980s but has become the main driving force in HE growth since the 1990s. Low regulatory barriers to establishing new HEIs favoured the emergence and proliferation of independent private HEIs. On the other hand, non-university HEIs were created to

meet the demand for occupationally oriented courses, while the relative importance of public HEIs and their political influence decreased.

Funding mechanisms moved their focus from supply to demand, with student aid becoming the most important component of public expenditure on HE. In fact, without considering R&D expenditure, student aid represents more than 40 per cent of total public expenditure in HE. The number of students being supported through loans and grants rose from 130,000 in 2005 to over half a million in 2015, whereas total undergraduate enrolment increased from 650,000 students in 2005 to 1,165,000 in 2015. This progress may not have been possible without the universalisation of secondary education that took place in the 1990s (MDS 2013). HE's gross enrolment rate (GER) of the poorest household income quintile increased from 4.5 per cent in 1990 to 20 per cent in 2009, reaching 27 per cent in 2013 (MDS 2013). In other words, for the poorest young people, the chances of being enrolled in HE are currently 4.5 times higher than in 1990.

Nonetheless, the Chilean school system is highly segregated and achievement is strongly correlated with social origins. Better-off students take most of the places at elite universities and courses leading to liberal professions, while low SES students are confined to low prestige and vocational HEIs, many of which are low quality. Although the rapid increase in coverage is often presented as a very successful policy outcome, many students attend university courses without the academic skills required in HE.

A sensitive issue is that most academically disadvantaged students are supported by government-guaranteed loans and grants so, in fact, the government's support is contributing to low-quality HEIs. At the same time, non-selective, second-tier HEIs present high dropout rates and an important financial dependence from student aid. This translates into a high degree of uncertainty for these HEIs because they are at constant risk of losing their accreditations and thus being unable to receive new supported students. Part of the late expansion of HE in Chile is due to a sharp increase of enrolment at 'aid dependent' HEIs.

2.2.HE Reforms in the 1980s

In 1980, the Chilean dictatorship allowed universities to charge fees and incentivised the private sector to create new HEIs. At that time, Chilean HE was comprised of eight publicly financed universities, two of which were state-owned and six of which were private (though relying on public funding). The two public, national universities, -*Universidad de Chile* (UCH) and *Universidad Técnica del Estado* (UTE)- were divided into a number of new universities – currently 16-, which along with the pre-existing private universities, constitute the Council of Rectors of Chilean Universities (CRUCH). In 1980, 119,000 students attended Chilean universities, representing a quite modest GER of 7.4 per cent of the 18-24 year-old cohort (Bernasconi and Rojas 2004).

There were economic, political, and ideological reasons behind the reform (Bernasconi and Rojas 2004). Economically, the reforms sought efficiency and equity by targeting public resources to school education and diversifying HE funding sources. University students came from better-off backgrounds and subsidies were considered unfair, thus charging fees was seen as justified on both equity and efficiency bases. On the other hand, private returns to HE were and still are quite high: a university graduate earns 3.8 times more than a secondary school graduate, which is far higher than in any OECD country. Moreover, marginal rates of return coming from Mincer's equations are as high as 20 per cent for an additional year of HE in comparison to secondary school graduates (Mizala and Romaguera 2004). Certainly, this is an average, but HE in Chile is still one of the most profitable investments an individual can make.

In political terms, the government sought to atomise the main public universities, which were seen as housing the opposition's political activism, by dividing them into many small, regional universities. The ideology behind this came from a combination of moral and political conservatism and the Chicago economic doctrine. Freedom, choice, and entrepreneurship were seen as a way to counter activism, protest, and debate.

The reforms granted a privileged role to the private sector, as it was allowed to establish new HEIs while public HEIs were not expanded. The Government also promoted the creation of a non-university sector; namely, Professional Institutes (IPs) and Technical Training Centres (CFTs). CFTs were expected to offer vocational-oriented, two-year programmes while IPs offered four-year professional programmes. Initially, three public IPs and two Teacher's Training Academies were created by reorganising sections and regional branches of UCH and UTE. All public, non-university HEIs turned into universities in the early 1990s. Thus, currently, the entire non-university sector is private.

Under the dictatorship, new private institutions required a political 'assessment' given by the Chilean Home Office and a 'technical' authorisation from the Ministry of Education, thus assuring a certain ideological and political homogeneity and control. Quality assurance mechanisms were limited and set during the licensing period. Ordinarily, new universities were examined by existing ones, and after a probationary period, institutional autonomy was granted. In 1990, private HEIs could choose between this examination process or a new licensing procedure carried out by a public agency, which became the most popular.

The reforms transformed Chilean HE into a multi-tier system: the university sector (made up of universities belonging to CRUCH or those newly created by private sector) and the non-university sector (IPs and CFTs). In addition, the reforms diversified funding, moving from purely block-grant funding according to historical considerations, institutional size and complexity, to a combination of block grants (AFD), indirect grants (AFI) depending on new entrants' performances in the National Entry Test (PAA, currently PSU), and student loans (CF). As shown below, maybe the most important shift was the change from a purely supply-side funding to include demand-side financing (loans) as well as the introduction of competitive mechanisms (AFI). Private funding (through tuition fees) and a series of competitive mechanisms to finance research were also set. Interestingly, the original reforms undertaken by the military regime would have never considered public spending on the new independent HEIs they allowed to be created and promoted.

Following Clark's (1983) framework introduced in the previous chapter, the Chilean HE system moved from a coordination scheme dominated by the state and the academic oligarchies to one based on the market (Brunner 1993; Brunner and Uribe 2007). Nevertheless, the military government was never able to fulfil its own goals for HE funding (Arriagada 1989; Castañeda 1990; Lehmann 1990) as the 1980s' economic crisis affected funding projections critically. The Government expected to increase CRUCH universities' funding by 50 per cent in the period from 1980 to 1986: AFD would be cut by half, but AFI and CF would reach the same relevance as the first, as shown in Table 2.1.

Table 2.1. Relative share by public funding mechanisms *(projected).* 1980=100

Type of Public funding	1980	1986
AFD	*100*	*44.35 (50)*
AFI	*0*	*7.3 (50)*
CF	*0*	*16.45 (50)*
Total	*100*	*68.1 (150)*

Source: Arriagada (1989)

Projected funding was far from being fulfilled. This had an obvious impact in the medium and long term, as CRUCH universities became underfunded and student loans were, in fact, scarce. Consistently, Desormeaux & Koljatic (1990) found public contributions to HE fell by 41 per cent in real terms during the 1980s, while public expenditure in HE in relation to GDP plummeted dramatically (Arriagada 1989).

2.3.The 1990s and 2000s

The most relevant feature of higher education policy in the 1990s was the expansion and consolidation of the system. Most of the new private universities were established between 1989 and 1991 because the private sector feared that the new democratic government (*Concertación* coalition) would seek to undermine the private sector's involvement in HE (Uribe 2004). However, these assumptions did not hold. The first *Concertación,* centre-left government, which took office in March 1990, prioritised resourcing to CRUCH universities by increasing funding as well as creating new funding mechanisms to foster investment in infrastructure. Concerning the student aid

system, which had been left heavily underfunded by the military, the government set an income-contingent loan scheme in 1994 instead of the mortgage-type one operating since the early 1980s, and established a grant programme targeting the poorest students. No aid was targeted to students attending new HEIs until the 2000s, despite the fact that a relevant proportion of low SES students were already attending these HEIs.

Chilean HE became one of the most privatised systems in the world, whether considering enrolment or expenditure. Indeed, today, 78 per cent of students attend private independent institutions and 65 per cent of the total expenditure in HE comes from private sources, mostly tuition fees (OECD 2015a). As shown below in Table 2.2, Chile has concentrated on demand-side funding; institutional funding has increased at a very slow pace. In this scenario, student aid systems, defined as policies aimed at supporting young people to meet tuition and living costs, are critical because the government needs to devise cost-effective mechanisms. The evidence shows that Chilean student aid has contributed to enhancing access, but the mechanisms to achieve that access are not robust enough, as I shall show later.

2.4.Shifting HE Funding towards Demand Side

A new trend started in the 1990s, as an increasing share of governmental expenditure in HE went to student funding to the detriment of institutional funding. In fact, in 1990, 74 per cent of public expenditure went to AFD and AFI. Later in 2000, the share decreased to 69 per cent, whereas institutional funding fell to 39 per cent by 2010. This trend has been confirmed since 2000, as more generous financial assistance was set (through grants and loans), reaching few private institutions since 2001 but expanding massively after 2005, when student aid was opened to students enrolled at independent private HEIs, as shown in Table 2.2.

Table 2.2. Public expenditure in Chilean HE. Institutional and demand-side funding
(Million CLP 2010)

Type of Funding	1990	1995	2000	2005	2010	2015
AFD	82,059	116,284	135,568	144,422	177,936	212,170
AFI	25,133	26,581	23,631	22,992	25,001	24,563
Student Support (Grants and Income Contingent Loan)	6,368	22,088	54,791	82,916	234,200	554,755
Institutional Development Fund			23,011	25,812	32,264	55,045
State Guaranteed Loan (*)					277,292	415,951
Other Institutional Funding	2,742	12,981	10,654	11,736	46,934	16,330
Total	116,302	177,935	247,655	287,877	793,628	1,278,814

(*) This item is to repurchase portfolio and to pay guarantees. It should not be accounted as expenditure but as an asset but official statistics do include the item.
Source: Ministry of Education (2015)

Enrolment steadily increased from 245,000 students in 1990 to 1,265,000 in 2015. This expansion put pressure on public expenditure, not only due to the dramatic changes in size but, more importantly, because there were many more low SES completing secondary education and, in consequence, meeting the formal requirements to enter HE (Armanet and Uribe 2005; Espinoza, González, and Uribe 2009).

The generous student support policy in the 1900s and 2000s, an increasing demand due to the universalisation of secondary education, and the creation of a state-guaranteed loan system in 2005 created the conditions to make HE accessible to the very poor. Accordingly, and as expected, coverage rates in HE for the poorest students have consistently increased in the last 15 years. Participation, however, differs throughout the system, as student aid eligibility used to be completely restricted to CRUCH universities. Therefore, the private sector, especially universities, served comparatively more affluent students as no support was provided until 2006.

When the government undertook an ambitious reform of HE student funding in 2005, there were two major developments. The first involved the introduction of a new student loan scheme, CAE, which was conceived to support students attending accredited private institutions (though students attending public and publicly

subsidised institutions were also eligible). The second consisted of a re-engineering of aid mechanisms for CRUCH universities. The government set student aid packages for those attending public and publicly subsidised HEIs, which consisted of a mixture of loans and grants, including a maintenance component for the poorest students: the poorer the student, the larger the grant component and the smaller the loan component. Horizontal equity improved significantly, as before the reform, each HEI managed its own budget to allocate loans. The most important change, however, was that student aid evolved from being strictly residual to a scheme of guaranteed support according to household income.

Equity improved as student support systems were extended to private, independent HEIs. In practice, in the non-university segment, which claims most of low SES student population, the opportunities for access have been improving since the introduction of a grant favouring vocational courses at IPs and CFTs in 2001, before CAE.

2.5. Unintended Consequences

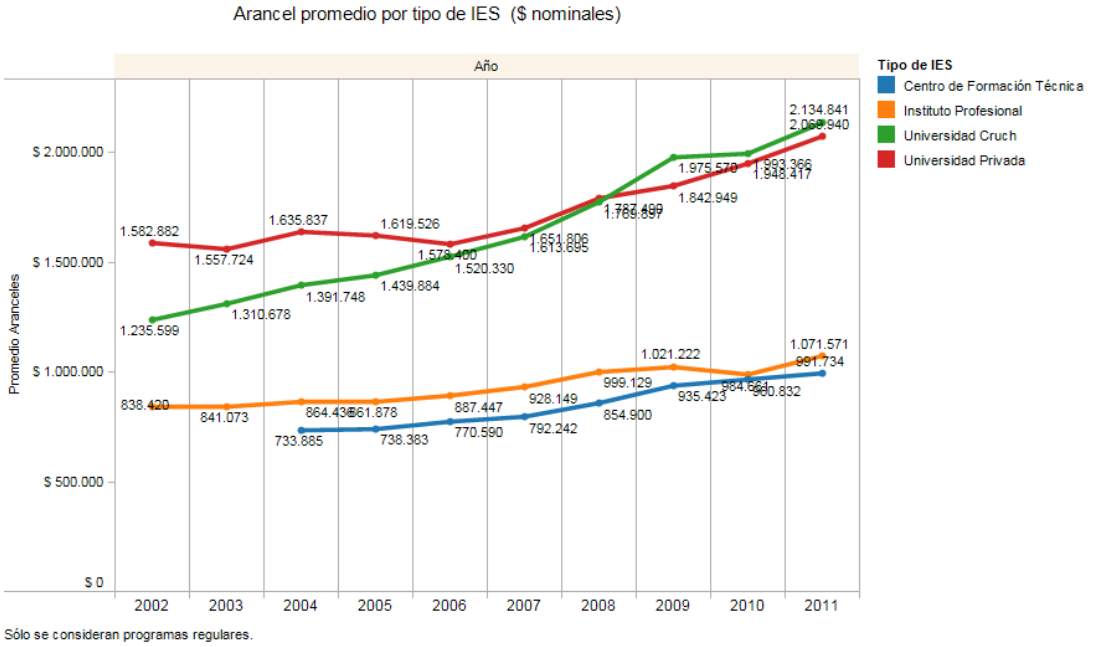
Earnings of HE graduates are still significantly higher than those of secondary school graduates, even though entry salaries may have decreased in the last decade. In Chile, a university graduate earns, on average, 3.8 times more than a secondary school graduate; the figure for vocational ISCED level 5B is 1.8. Although OLS marginal returns estimates for HE are high, the main problem is now the debt burden on CAE students. For income-contingent loans, the situation is different: after 12 to 15 years of repayment – according to the amount owed at the time of starting repayment – any outstanding balance is written off. On the other hand, interest rates also differ: for income-contingent loans, the rate is 2 per cent real; for CAE, it is 5.5 to 6 per cent. This represents a major disadvantage for those with CAE loans; however, the government has introduced an income-sensitivity component and started subsidising interest rates (more details in chapters 4 and 5).

Both mechanisms show design failures. For income-contingent loans, repayment rates are about 50 per cent, forcing the state to provide funding for the system to keep operating. Criticism focuses on the low level of income contingency (5 per cent of gross annual income), short repayment period, subsidised interest rates, and increasing tuition fees. On the other hand, a study commissioned to the World Bank (2011) by the government raised concerns on high default and high debt-to-income ratios in CAE, which endangers the system's financial sustainability.

Despite that most students are better off with these policies, students and the general public are not satisfied. Chileans cite indebtedness of CAE borrowers affecting especially the low and lower-middle classes, rising tuition fees, perceptions of unfair access due to three quarters of the school population attending low-performing secondary schools, increasing enrolment in low quality HEI, and the perceived reluctance of the government to more closely supervise and regulate as critical issues that, to their minds, are not being addressed properly.

As I noted earlier, an important part of the problem is due to the especially low public expenditure in Chilean HE. Even public universities rely on private funding for about 80 per cent of their operational budgets, which is anomalous. Furthermore, in relation to GDP per capita, Chilean tuition fees are amongst the highest in the world, representing 27.9 per cent of GDP per capita in public institutions and 32.0 per cent in the private ones (OECD and The World Bank 2009). Chilean HEIs charge tuition fees similar to many developed countries but have a third of their GDP per capita. This explains why students taking up loans end up with a high financial burden relative to their expected earnings. Moreover, tuition fees have increased quickly, as shown in graph 2.1., representing a major issue for student support policies. It also deters poor students from borrowing money to take long courses, as their debt may become as massive as the value of a house.

Graph 2.1. Evolution of fees by type of HEI (current CLP)



Source: Own estimates based on CNED databases (2012)

This problem affects mainly the lower-middle and middle classes, who, because of a low-quality school education and a highly segregated school system, are hardly able to enter selective institutions. Nonetheless, no relevant research has been carried out on graduates' labour market segmentation, which requires breaking down figures by HEI attended. Some government data suggests that, for the same courses, earnings may differ significantly according to which HEI graduates come from[1]. Labour market segmentation is a common trend in many countries. For instance, UK evidence shows an important variation in earnings by university attended as well as subject studied, where Russell Group HEIs lead in many fields (Ramsey 2008). A similar situation is evident in the US, where Ivy League Universities, again the most socially and academically selective ones, lead earnings in most fields (Gopal 2008).

Although paradoxical it is not uncommon for the least well-off to face the harshest conditions. To solve this paradox, in 2011, Chile's Congress approved a bill that drops the interest rates of CAE to 2 per cent and introduces income sensitivity for

[1] These data are not entirely transparent because the government publishes just earning's ranges for many courses according to HEI attended, but it is not possible to know anything about earning distributions as earning categories do not fit with real earning distributions. Although there are available data, the Government does not publish more precise information due to technical issues (lack of a relevant number of observations), and institutional and political pressures.

repayments, fixing a maximum repayment of 10 per cent of earnings and subsidising amounts remaining. This is, of course, an improvement, but other issues well beyond HE policy remain, as I shall show in the next section.

2.6. The Big Issue: The Chilean School System

Until 1981, the Chilean school system was comprised of public schools, private subsidised schools, and private independent schools. The system was strongly centralised, coordinated by the Ministry of Education, and teachers were public servants. The reforms transferred school's supervision to provincial bodies, public schools' administration to local governments (municipalities), and set a plain voucher system to public and non-fee paying private schools on the basis of pupils' attendance. Additionally, in 1982, the Government set a national standardised test to measure school performance (SIMCE). This quasi-market setting was supposed to increase coverage, foster competition among schools, introduce choice, and increase the quality of education. Unfortunately, the main assumptions required for a quasi-market to work do not hold in this system (Mizala 2007).

In the 1990s, centre-left governments roughly maintained the system structure but made strong investments in infrastructure and increased funding, which the dictatorship had left at historical minimums. Massive programmes to improve quality and equity of the school system, a re-engineering of the teaching career, as well as targeted interventions for the poorest and lowest-performing schools were the distinctive characteristics of that period.

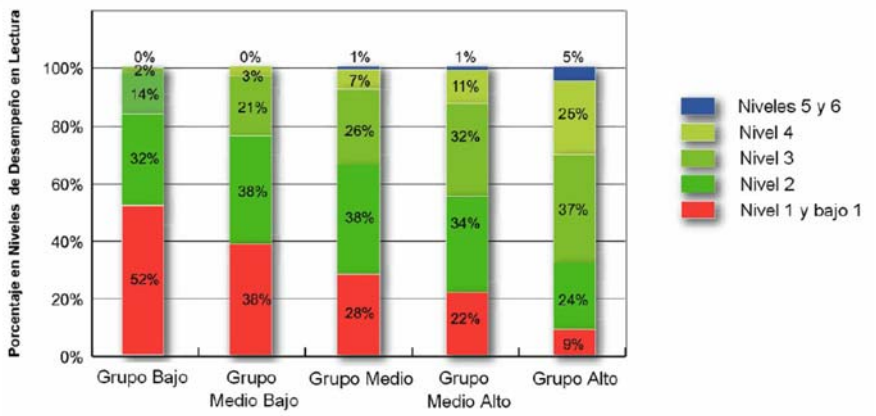
The Chilean school system has become increasingly privatised since the 1981 reform. In 1981, 15 per cent of students attended subsidised private schools. In the 1981 to 1986 period, more than 1,000 new schools entered the market, and by 1990, 31 per cent of school children attended subsidised private schools. Private participation increased during the 2000s, reaching 47 per cent of total enrolment in 2008. Most of this growth has been at the expense of public school enrolment (Elacqua, 2009), which represent a modest 37 per cent of school enrolment nowadays.

A high-impact policy was implemented in 1994. The government allowed private subsidised schools to charge fees without losing voucher financing. Many private schools began charging fees at the maximum allowed in order to keep vouchers to maximise funding. This policy is perhaps the most controversial in the Chilean education system introduced in the last 25 years, and is deemed as highly segregating.

The core issue concerning school education in Chile is that it is one of the most socially segregated in the world. According to Elacqua (2009), more vulnerable students attend public schools that are more socially integrated than private voucher ones. As private schools select students by academic performance or other characteristics, public schools end up receiving most of the academically disadvantaged students.

On the other hand, PISA results show an almost perfect stratification according to SES. Only a 3 per cent of examinees from low SES score in level 4 or above compared to 30 per cent for the upper group. The graph below is unsurprising, but convincing, as the correlation between SES and performance is very high.

Graph 2.2. PISA score levels in reading by SES (%)

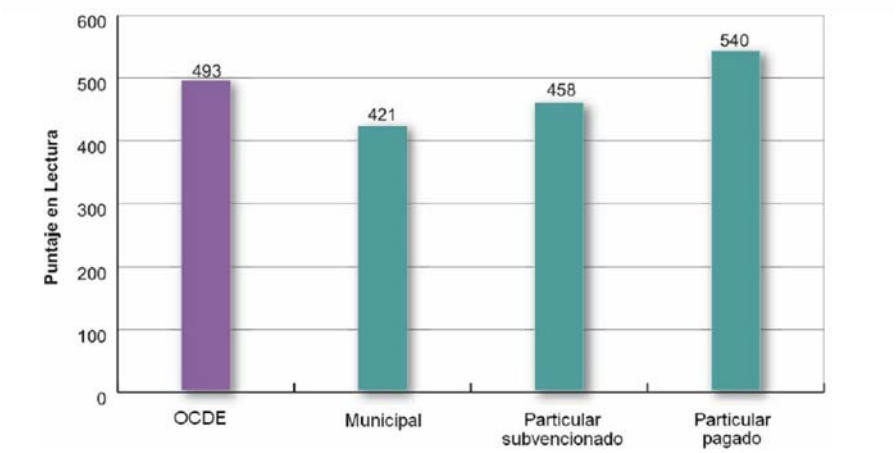


Source: MINEDUC (2010)

A very similar situation can be seen when breaking down PISA scores by school type. Differences in average scores are significant, with public school (*Municipal*) students being the most disadvantaged and private independent school (*Particular Pagado*)

students scoring as high as those in high-performance educational systems (see Graph 2.3).

Graph 2.3 PISA reading score average by type of school



Source: (MINEDUC 2011)

Regarding access to HE, both low SES students and those who have attended public schools perform lower in PSU, as shown in Table 2.3. Only 9.4 per cent of public school students score 600 or more points in PSU, compared with 14.9 per cent from private voucher schools and 53.9 per cent from private independent schools. This means that the majority of students coming from subsidised schools do not score enough to be admitted to a selective HEI, which is feasible only with scores of 600 points or above (see Table 2.3 below).

Table 2.3. PSU scores statistics by type of school (*)

Score	Public	Private Subsidised	Private Independent
Less than 200	23	15	8
200-449.5	38,749	40,869	2,131
450-599.5	39,431	75,380	10,102
600-850	8,157	20,521	14,284
Total	86,360	136,785	26,525
Mean	469.3	501.7	598.5
StDev	93.2	92.4	95.5

(*) PSU Scores have an average of 500 points and a 100 points SD.
Source: DEMRE, 2015.

The consequence of low performance in PSU is that poorer students end up attending low-quality institutions. This has an impact on student persistence, studies completion, and income expectations, which ends up reproducing inequalities coming

from schools. The correlation between either PSU or SIMCE score and social background is high.

A relevant relationship between SES and student performance is unsurprising from an international perspective. In fact, according to OECD (2010), in Chile, a near 20 per cent of variation in PISA reading scores is explained by socio-economic background, close to that of countries such as Germany and the US; far from OECD's average, which is 14 per cent, and even farther from countries like Finland (8 per cent), Japan (9 per cent), and Canada (9 per cent).

The high level of inequity in the school system leaves HE policy no room to manoeuvre beyond remediation. Though HE participation of low SES students has improved significantly in Chile, the distribution of educational opportunities remains unfair. The implications of these levels of inequity drive attention to the performance and development of the school system, as policies aiming at levelling the field need to be implemented at that level.

In the next section, I shall carry out a study testing whether and how inequality of school achievement affects HE participation.

3. HOW DOES INEQUALITY OF SCHOOL ACHIEVEMENT AFFECT HIGHER EDUCATION PARTICIPATION? A CROSS-COUNTRY PERSPECTIVE

In section 1, I pointed out that some studies have shown that students' socioeconomic background has lost predictive power when it comes to HE participation, with ability and achievement becoming more important. Nevertheless, the extent to which school achievement is socioeconomically determined affects low-income students' access to HE. Even assuming that there are no financial barriers to HE access, that school achievement remains highly determined by social background may constitute a barrier for increasing HE participation. This is the main hypothesis I deal with in this chapter.

3.1.Introduction

According to economic theory, a more skilled labour force translates into gains in innovation and productivity, leading to increasing economic growth and competitiveness, higher wages, and a series of non-monetary benefits that determine

prosperity and well-being. Consequently, developing and boosting human capital are seen as crucial tasks for governments and countries. By widening access to education to as many citizens as possible, countries improve and expand the skills of their citizens; barriers to access actually play against countries' interests by impeding the optimal provision of education.

Over the last three decades, HE has been the main driving force of educational expansion. Nevertheless, SES still exerts an enormous influence with regards to access. Despite the global massification of HE, the proportion of low SES students remains small. Student support is inadequate and, in many countries, public money goes mostly to subsidise better-off students (Barr 2005).

Nevertheless, a body of research highlights the importance of prior achievement in ensuring access to HE, thus involving school education policy. Consequently, the socioeconomic distribution of prior achievement becomes crucial because it delimits HE's potential expansion. A positive relationship between SES and school achievement is an obstacle for HE expansion, as poor students are the main source of new HE students. The relationship between inequality and access to HE has been studied in individual or handfuls of countries but there is no research pursuing the identification of a general relationship in a cross-country setting. For instance, there is no research testing whether a negative relationship between inequality of school achievement and participation holds cross-country. This turns relevant in a context where education policy has turned into a global issue and one additional reason for competition amongst countries.

International and development agencies have contributed to the global debate on HE by promoting a wide range of education policies. More important, however, is the fact that these policies have been applied in several countries to bring more resources to the education sector, this way protecting education from fiscal constraints, increasing HE participation, relieving public finances, and meeting an increasing demand for more education and qualifications. I maintain that there is a global policy recipe in

HE but, paradoxically, empirical research has not focused enough on studying variables and relationships likely to be regarded as 'cross-country effects'.

Is there an average effect of inequality of school results on HE access that holds true across countries?

I argue that the interest of the result itself is amplified by its implications for policymaking. In these terms, the most sensitive issue for policymaking is that the extent to which HE participation relies on school achievement and its distribution indicates the extent to which HE policy by itself actually influences HE access for low SES students.

In this chapter, I shall analyse a dataset containing country-level indicators on HE, school systems performance, and countries' socioeconomic characteristics, measured over time. The main objective of the chapter is to find a general relationship between the socioeconomic distribution of school achievement and HE access: Is there a cross-country effect of inequality of school achievement on HE participation or is it a country-specific issue? Does inequality of school achievement limit the expansion of HE systems and, if so, how? Does the result hold once controlling for a number of country characteristics, institutional settings, and socioeconomic features?

The chapter is structured as follows. In section 3.2., I describe the policy and research contexts and discuss the main literature on access to higher education. First, I show the reasons why education and its reforms are relevant in the field of economic theory, as well as the main motivations behind government policies. Next, I discuss the main research on barriers to access to HE, with a special focus on SES persistence. Finally, I review the core literature linking school achievement to HE participation.

In section 3.3., I discuss the advantages and disadvantages of using a cross-country dataset. Then I discuss and describe the main issues of the dataset, specify the analytic model, and discuss the reasons why each variable is included in the model. In section 3.4, I estimate the association between inequality of school achievement and HE

participation and introduce control variables in several steps. I also try different specifications dealing with variables measuring school achievement (PISA scores) and investment in education (expenditure in secondary and HE as a proportion of GDP). In the last section (3.5), I draw some conclusions and identify further issues and future research in the field.

Given data constraints and the nature of the chapter, estimates cannot be considered as causal effects but just as *associations between variables*. Notwithstanding I use several control variables and different empirical approaches to deal with both observable and unobservable confounders, I suggest that this chapter should be regarded as halfway between descriptive and explanatory.

3.2.Policy and Research Context

The fact that governments seek to increase and accumulate human capital as a means of boosting productivity and economic output is now common sense. In this context, confronting barriers to HE access turns into a key issue, as they involve efficiency and equity concerns. Competitive and market mechanisms have been devised in order to optimise and diversify funding and to increasingly target taxpayer's money to school education. Barriers to access are relevant in both economic and social terms, as they threaten the optimal allocation of investment and impede a full and fair use of skills present in societies. Financial barriers not only matter, but represent a major social reproduction mechanism that keeps low SES students outside of HE.

3.2.1. Education, growth and the shifting focus of educational investment

The benefits of education for individuals and economies have been at the centre of the debate for more than 50 years. One of the most prevalent theories giving an account of the benefits of education is that of human capital. The seminal works of Mincer (1958) and Becker (1993) stressed the relationship between education and training, the individual income distribution, and the effects on productivity and growth. In the nineties, endogenous growth theories contested the neoclassical foundations of decreasing returns, giving instead a crucial role to innovation and knowledge as the

driving forces for sustained economic growth. The accumulation of advanced human capital in this context is seen as the driver. Thus, it became common sense that a good education system is a means for a country to assure economic growth and competitiveness.

The interest in studying the relationship between education and earnings resurged in the nineties due to its importance for economic growth and the rapid massification of post-secondary education. A main concern focused on the cost and benefits for those who were new to HE (Card 1999, 2000). Hundreds of studies have been done to measure the value of education using Mincer's equation, and a web of research has been carried out to enquire whether the model still fits given the increasing availability and quality of data along with a number of statistical refinements to the classical model (Heckman 1976; Heckman, Lochner, and Todd 2003).

But formal education is not determinative when measuring the value of education. Cognitive and non-cognitive skills are formed throughout the life cycle, especially in early years' education. Many education gaps are determined in childhood, and those gaps do not seem to narrow in the long term. Early intervention has a much higher return than traditional education and training policies (Carneiro and Heckman 2003) and "reduce[s] the inequality associated with the accident of birth and at the same time raises the productivity of society at large" (Heckman and Masterov 2007:2).

Some empirical literature also stresses the relationship between the wider social and economic context and education. Indeed, many investigations relate economic growth with educational attainment (Barro 2001; Hanushek and Kimko 2000; OECD 2012a; Psacharopoulos 1988). The reasoning has been the following: the more resources allocated to education, the more educational attainment of the workforce, and consequently the more productivity, which translates into higher wages, better jobs and household income, and economic growth.

Although many agree that investing in education has unquestionable benefits for countries and citizens, policies promoted by international agencies have been seen by

some researchers as an 'interested complement' to structural adjustment policies carried out or supported by themselves in the eighties (Espinoza 2008). In the developing world, the World Bank (WB) has promoted the introduction of market mechanisms in school and HE as a standard response to structural adjustment policies (World Bank 2000). Whether or not this assumption and the critique are correct, governments in developing countries have made efforts to devote more and sometimes too scarce resources to education. In addition to that, a chief motivation for bureaucrats and politicians is that citizens trust education as a means of improving opportunities for those of the following generation and, accordingly, voters favour governmental programmes promising educational improvement. This way, the educational debate becomes politically sensitive.

3.2.2. Higher education finance and barriers to access

In developing countries, the scarcity of resources to be devoted to education has been dealt with by focusing public investment on school education, as promoted by international agencies. In fact, since the early eighties, international organisations such as the World Bank (WB) and the Inter-American Bank for Development (IABD) have promoted the use of private funding for HE so that countries can target taxpayer's money to school education. Changes in this direction have been implemented with international support in Brazil, Chile, Mexico, Colombia, some African countries, and former socialist countries in Eastern Europe and Central Asia (Johnstone and Marcucci 2010; World Bank n.d.). This has meant that subsidies for HEIs have been either frozen, cut back, or allocated via competitive schemes, giving priority to demand-side funding mechanisms (Sanyal and Johnstone 2011). In this scenario, the cost of HE has been shared/transferred with/to students and households (Johnstone 2004), and governments have devised student funding mechanisms for poor students to overcome financial barriers.

In relation to the last point, research has been carried out since the late eighties, on the effects of student aid policies on access and persistence. Some evaluation exercises have measured the impact of specific instruments – controlling or not for social

background and other relevant variables – such as guaranteed loans, income contingent/sensitive loans, and scholarships and bursaries (Canton and Blom 2004; Chapman and Ryan 2005; Dynarski 2003; St John and Noell 1989). These investigations have normally found positive effects of aid on access, featuring the 'optimistic' side of research in the field.

Research has also been conducted in order to establish which variables, apart from those informing the wider socio-economic context, relate to higher access to HE. For instance, the impact of socioeconomic background of students on HE access has been studied, amongst other countries, in Chile (Espinoza et al. 2009; Torche 2005), the United States (Alon 2009), Canada (Finnie 2012), and Spain (Mora 1997). Other work in a comparative context has also investigated social stratification in HE (Shavit et al. 2007). Most investigations highlight the persistent but diminishing effect of background on access to HE, then making the point for either reproductionist theories (Bourdieu 1979; Bourdieu and Passeron 1986; Bowles and Gintis 1976) or the Weberian social stratification tradition (Goldthorpe 2010).

3.2.3. Reproduction, prior achievement, and HE access

Reproductionists conceive education as a social process aiming at perpetuating and legitimating domination structures and the division of labour in capitalist societies. Bowles and Gintis (1976, 2002) stressed the role of education as reproducing the stratified structure of labour markets, and Bourdieu argues that selection and elimination mechanisms are more powerful than the pretended meritocracy: even the highest performers amongst the poor are disadvantaged in comparison to their wealthier counterparts. Stratification takes place even within HE through disciplines as low SES students are confined to less recognised disciplines, while better-off students undertake prestigious courses leading to the most socially recognised professions, which offer the best jobs and the highest wages upon graduation (Bourdieu 1979).

I contend that these arguments are not necessarily inconsistent with the fact that HE access for low SES students has increased worldwide and that returns from HE have

remained high. From the reproduction theory, I argue that the important issue to address is the intensity with which school systems reproduce and legitimate domination structures. In any case, the "progressive" and reproductionist approaches to educational phenomena should not be seen as completely contradictive perspectives.

Although the level of determination exerted by socio-economic background on access to HE is consistent and persistent, the most important requirements for HE admission are secondary education completion and good school achievement. Countries (and regions and states in the US, for instance) measure school achievement using a variety of tests and exams. Some standardised tests check the fulfilment of curricular goals at different levels whilst some are properly exit tests leading to a certification or diploma (for instance, A-levels in the UK and Baccalaureat in France). For these reasons, researchers have used international tests–PISA, TIMMS or PIRLS–to carry out comparative research.

A number of investigations use standardised tests such as PISA as a dependent variable, the predictors being socioeconomic level, school practises, geographical and demographic characteristics, immigration status, teacher salaries, and family practices at home that feature children's learning environment and assets (Duru-Bellat and Suchaut 2005; Fischbach et al. 2013; Fuchs and Woessmann 2006). On the other hand, cross-country research has related scores in international tests–as they measure skills quality–to economic growth. The evidence suggests a positive relationship (Hanushek and Woessmann 2007; Jamison, Jamison, and Hanushek 2007).

There also exists research relating school performance to access to HE. Studies in the UK show the importance of prior achievement in narrowing the HE participation gap in socioeconomic terms. Indeed, amongst the highest quintile of school performance, the socioeconomic participation gap is narrower than raw gaps (Chowdry et al. 2013). Additionally, Jerrim (2012) cites evidence suggesting that the socioeconomic gap in school achievement has narrowed in the bottom part of the achievement distribution, due, in part, to an important governmental investment. In Latin American countries, higher levels of inequality of school outcomes are found in comparison to OECD

countries. In fact, Gamboa and Waltemberg (2012), using data from PISA 2006 and 2009, found high levels of inequality of opportunity when considering parental education, gender, and type of school attended. When using parental education, Chile, Mexico, and Uruguay – the highest performers in Latin America – showed the widest achievement gaps. Nonetheless, according to UIS-UNESCO indicators, some Latin American countries show quick improvements in GERs since the late nineties.

The effect of school achievement, as measured by PISA scores, on access to HE has been studied by linking achievement with further educational trajectories. For instance, in Canada (Murdoch, Kamanzi, and Doray 2011), by using longitudinal surveys and following up students, researchers established that PISA literacy scores, schooling, and social factors appear to have an impact on access to HE and persistence, though the greater impact was found on access rather than persistence. In the same country, Finnie (2012) found that cultural variables are of the highest importance and argued that access policies should get away from the 'old' financial barriers issue. In comparative research, Jerrim, Vignoles and Finnie (2012) have studied access for disadvantaged children controlling for educational achievement in four English-speaking countries. They found that socioeconomic differences in HE access are more pronounced in England and Canada than in the US and Australia, and suggested that policies boosting school achievement would be more pertinent than traditional policies focused on aid.

The above findings confirm that school achievement gaps, when taking SES into account, may vary importantly among countries. The rationale here is evident: the wider the socioeconomic achievement gap, the higher the inequality of educational opportunity. Country-level evidence also underlines the importance of prior achievement in the chances of low SES for attending HE. Comparative research shows different patterns for countries dealing with inequality of HE access. There is, however, an operational issue to address. Country level or comparative research using microdata, such as administrative records or longitudinal surveys tailored to examinees populations (as in studies using PISA), are mainly restricted to developed

countries, whereas this sort of dataset is seldom available in developing or middle-income countries.

Nevertheless, international agencies' indicators are widely available for almost any country and data collection has been undertaken for decades and periodically, suggesting one should favour using country-level information instead of microdata as in the studies cited above. Looking at the country level may have some advantages, as including developing and middle-income countries, while controlling for a number of social and economic indicators not considered at the micro level, would allow investigating what happens with the relationship between inequality of school achievement and HE participation and whether it holds once country-level controls are introduced. For instance, a cross-country data set would allow addressing questions such as whether the inequality of school achievement affects HE participation once controlling for macro-level measures such as GDP and labour force qualifications.

The above measures show how different social and economic contexts, institutional settings, and policymaking are, which is not likely to be addressed by using microdata. On the other hand, including developing and middle-income countries, which show higher inequality, lower achievement, and a more modest investment in education, adds by itself an important source of variability. Moreover, in developing countries and transitional economies, massive market-oriented reforms and privatisation in HE have taken place.

Nonetheless, there are many caveats in cross-country research, the most important being data-measurement error. Jerrim and Micklewright (2013) studied measurement error in international tests with regards to SES, as it is reported by children. They concluded that parental occupation and education seem robust, but the number of books at home is problematic. This should be kept in mind, as I use this kind of SES score in my estimates later.

On the other hand, using country means for indicators implies that generalisations on the basis of aggregated data may not hold at lower levels of aggregation and lead to

misleading conclusions, which is known as the "ecological fallacy". A number of other operational issues regarding cross-country data are explained in section 3.

From a policy point of view, measuring the effect of inequality of school achievement on HE access, having controlled for key variables, is of high interest. More important, though, is that having an empirical account of this would allow researchers to test the relevance of HE policy. If it were found, for example, that the most relevant predictors of access are outside of HE policy, it would contest many traditional HE policy devices that are used to equalise access: for example, student aid as some research suggest (Finnie 2012; Jerrim et al. 2012), and more crucially, cost sharing politics: that of cost transfer and privatisation being the solution for the poor to enter HE. Of course in such a situation, addressing the access issue should rely on school education policy rather than HE policy.

Next sections address the lack of cross-country research intending to explain educational outcomes. Specifically, this chapter will ask: Does HE access decrease as the socioeconomic gap in school achievement widens? Does it hold independently from a number of country characteristics, institutional settings, and socioeconomic features?

The above questions assume implicitly that HE access is a demand-side issue. Whether due to a political decision to keep HE enrolment within certain numbers or the lack of enough critical mass to support HE expansion -personnel, infrastructure, funding, etc.-, there might be constraints to the provision of more places in HE. Indeed, public provided/funded HE systems, as they rely upon public expenditure, may find it difficult to grow due to fiscal constraints or a political decision -sometimes with deep institutional and historical roots- in order to keep the numbers within certain manageable range. In some African countries, despite The World Bank has promoted student loans and private participation in HE, the scarcity of resources imposes restrictions to HE expansion.

However, there are HE systems able to respond to the increasing demand quickly, as for instance Chile, Colombia, and the US with for-profit colleges, where lax regulations

to establish new HEIs make adaptation to an increasing demand easier and the private sector has led the expansion. As in the case of Chile I shall study in chapters 4 and 5, important increases in HE participation, alongside improved access conditions to low SES due to student aid (as it is also the case of American for-profit colleges), without policies addressing HE personnel development and the absence of enough quality controls on the new provision, the benefits (or the value) of the new provision/providers is questionable thus the efforts may not pay off. Although important, investigating the value of the new provision/providers is beyond the scope of this chapter.

Countries included in the sample are mostly developed and middle income. Unless politically defined, supply side constrains should not be as severe as to represent a major impediment for HE to grow. It would not be risky to assume that in most countries HEIs would cope with an increasing demand. HE systems have responded from the supply side in one or more ways as for instance by 'upgrading' vocational institutions to HEIs (e.g., UK's post 1992 universities) and allowing the private sector to participate (South Korea, Colombia, and Chile, for instance).

Although I recognise that many more variables are important, I will focus on data provided by PISA, namely scores and the proportion of scores variability explained by SES, and control for variables operating in the wider economic context, school education, and investment in education. For instance, when countries are the unit of analysis it is possible to determine whether school achievement as measured by PISA scores still has a significant effect on HE participation once social inclusiveness indicators such as the proportion of scores variance explained by within-school characteristics as well as variables informing the wider socioeconomic context are included in models. The same rationale applies when including more controls.

Hence, the main contribution of this chapter is to explain the relationship between inequality of school achievement and HE participation rates by using cross-country data.

3.3.Data and Methods

3.3.1. The case for a macro analysis

The use of microdata and the linkage of several administrative sources offer appealing opportunities for research. I recognise that the mechanisms linking school achievement and access to HE might be regarded as country specific to an important extent. Admission systems in HE, the presence of one or several school tiers and a series of cultural characteristics such as attitudes towards education are specific to countries. In effect, most investigations in the field, such as those reported in the above section, have focused on studying individual countries or a handful of them by using existing administrative sources and microdata and have compared, for instance, the effect of school achievement on HE entry as well as the persistence of SES on HE access. A key issue is that the availability of microdata and administrative records is mainly restricted to a small group of developed countries and this does not allow to draw conclusions beyond the countries analysed and their own contexts. More critically, developing countries, which hardly rely on quality microdata, have experienced the most explosive increase in education participation at all levels in the last decades, this being especially true for HE.

Notwithstanding, it is not uncommon to find relevant research relying on macro data in the field of economics, as for instance international comparisons on returns to education (Psacharopoulos 1988; Psacharopoulos and Patrinos 2004) and the relationship between education and income inequality (De Gregorio and Lee 2002). In fact, one of the more recognised works linking education to economic output is that of Barro (1991), which relies on country-level data for 98 countries in order to assess the neoclassical framework on growth on an empirical basis. However, there is not such a widespread use of cross-country data to predict educational outcomes in as seen with economic research in predicting output and growth. Despite the limitations of international data in relation to measurement, harmonising indicator definitions, and crucial statistical shortcomings like reverse causation and endogeneity, a macro approach provides answers to the research questions considered within this chapter.

Therefore, from a methodological point of view, the key issue in relation to the value added of this chapter should be whether obtaining an average coefficient measuring the relationship between school inequality of achievement and HE participation offsets the disadvantages of losing individual and school-level variation. From a strictly technical point of view, the answer is not clear as statistical shortcomings impose limitations to methods, thus the research may end up introducing more complexity instead of certainty. However, macro analysis may still make important contributions since it is easy to streamline its results with policymakers' thinking. In fact, knowing that there is positive cross-country average association between achievement and HE participation, for instance, may contribute to shifting the focus from traditional access policies such as student aid to improving school performance, allowing student aid to become a complementary policy rather than the focus. If the result was that inequality of school achievement is associated negatively with HE participation, policies attempting to tackle these inequalities would also remain within the school education field.

Although a series of country-specific analyses may lead to similar conclusions with a deeper understanding of individual and school factors, the advantage of the macro approach I use in this chapter is that it may find that, in general, and keeping many country characteristics constant, there is an average relationship between inequality of school achievement and HE participation that is independent from individual countries' peculiarities.

In the last two decades, international agencies like the OECD have exerted an increasing influence on national education policies (Baird et al. 2011). Every three years, governments look forward to announcing good news to the public and highlighting how effective their own policies have been and, moreover, blaming former governments when good news is not forthcoming. PISA test scores are seen by many governments as the most authoritative measure of results, thus becoming a sort of assessment of governmental policies.

Despite international league tables and benchmarking exercises that are not questionable by themselves, this sort of information is taken so seriously by some educational authorities that they end up assessing, revising, dismantling, or imitating foreign policies on the basis of very simple associations. Nonetheless, I argue that educational research needs a mixture of strategies in order to gain more presence in the always problematic relationship between evidence and policy, and that cross-country research can help uncover relevant relationships that may eventually translate into policy measures. Regarding this research, the conclusion is evident: finding a general relationship between school achievement, its distribution, and HE participation would contribute to making HE more socially accessible by making critical decisions at the school education level. The main consequence in policy terms would be that there is strong enough evidence to shift the focus of the debate from HE funding and cost transfer to school education, as some studies suggest (Chowdry et al. 2013; Finnie 2012).

3.3.2. The dataset

As I noted earlier, I collected data including variables from the wider economic context, school education characteristics, and investment in education in order to build a model to predict enrolment rates in HE. Those variables have to be widely available for as many countries as possible. Comparative datasets face a number of issues in regards to accuracy, national data sources' quality, and availability across time. Changes in periodicity and the focus of data collection may end up leaving many gaps. In education, national systems differ from each other in many respects, for instance, in starting ages, compulsory years of education, and the very definition of post-secondary, higher tertiary, or higher non-tertiary post-secondary level. Indicators such as coverage or enrolment rates harmonise definitions to allow between country comparisons. Moreover, the most precise and more sophisticated indicators do not always allow for better estimates because some countries are unable to collect them. In addition, policy priorities of international organisations may affect data collection; for example, giving more importance to certain factors in a period only to prioritize others a decade later may generate discontinued series, which may have otherwise

been of the highest value for research. Consequently, the operational criteria I used to gather indicators are (i) availability, (ii) continuity, (iii) simplicity, and (iv) comparability. In respect to comparability, I note the fact that comparability might well be fulfilled due to (i) and (iii) and the fact that a single organisation collects the data, like OECD PISA database.

A dataset has been built taking into account the reflections above. I have used several sources to consolidate the dataset (OECD n.d.; UNESCO n.d.; World Bank n.d.). I took data from 63 countries, namely countries participating in PISA 2009 test, using World Development Indicators (World Bank) and UIS-UNESCO indicators. I considered the 2000 to 2009 period for the predictors and the 2003 to 2012 period for the outcome, thus allowing independent variables to be lagged by three or six years in relation to the outcome, depending on the specification.

PISA has included more developing countries every round so it currently offers a more balanced country profile than 10 years ago. This improves representativeness but also introduce an issue with missing data I shall discuss later.

3.3.3. The outcome

The outcome variable I use in this chapter is the GER, which is defined as the ratio between the number of students attending HEIs and the 20-24 year-old population. I do not use the net enrolment rate, which only considers 20-24 year-olds currently enrolled in HE so giving a more precise measure, due to the lack of continuity in datasets and the fact that many countries do not report the net rate. The source from which the outcome variable was taken is World Development Indicators (WDI). There was also the option of using education attainment; namely the proportion of the 25-29 year-old population with higher education, taken from Barro & Lee (2013) database.

I chose GERs as the outcome variable for the following reasons. HE attainment is reported in quinquennial periods, while HE GERs are collected, in general, on an

annual basis. Using Barro and Lee's database implies just using a half of the information from predictors, whereas HE GER allows using the full dataset. Apart from these rather practical reasons, there are substantive ones: (i) as PISA children were 15 years old when tested, the requirement is to count on attainment data 10 to 15 years later, which means that for PISA 2000 examinees, in order to make sure models test the 'same' cohort, attainment levels for 2010 to 2015 are needed; and (ii) changes in predictor variables may take a long time to crystallise as a relevant change in attainment ratios, while variations in the predictors are more likely to have had an impact on enrolment within a short period of three or six years. For instance, a sharp increase of expenditure in HE through more grants and loans to poor students may have an immediate impact on enrolment rates, as GER are measured for the 20-24 year-old population[2]; whereas an improvement of educational attainment due to this policy may take a decade to become noticeable. In other words, though educational attainment may give a clearer picture of how a country has 'progressed' in educational terms, a more immediate effect of explanatory variables is found in GER.

Other plausible outcomes I did not consider because of operational reasons are entry rates (ER) and graduation rates (GR). Either a policy shock or a sharp improvement in school achievement would affect ERs quicker than GERs. On the other hand, GRs make more sense from the equity of outcome point of view. Indeed, high GERs may not necessarily imply that a country is successful in providing opportunities for low SES students as dropout rates tend to be higher for poor background and minority students (Bean 1980; Herzog 2005; Tinto 1975)). Therefore, improvements in GER may well mean the problem has been postponed. In fact, high GER and low graduation rates are common in Latin American countries. For instance, in Chile, which doubled HE enrolment in the last 10 years, just 50 per cent of students graduate (MINEDUC, 2013). In Argentina and Mexico, rates are even lower, with individual universities reporting as many as 600,000 students at *Universidad de Buenos Aires*, UBA. Thus, graduation rates are cleaner if one thinks about how 'robust' are the provision of educational opportunities. Unfortunately, ERs and GRs are not as widely available as

[2] As GERs is the ratio between HE enrolment and the 20-24 year-old population, values may be higher than 100%.

GERs and attainment measures in international databases, and the number of countries in the sample (already small) would have dropped dramatically had I used ERs or GRs.

An important data issue to deal with is that predictors and outcome should not be measured at the same time. Therefore, I have also compiled data on outcomes for the last available figure in the 2010 to 2012 period. Estimates were drawn by lagging the predictors by three years, or three and six years, depending on the specification. Lagging PISA scores by three years should reflect an impact on GERs since 15 year-old students by t-3 are the new HE entry cohort in year t. Something similar happens with lagging only PISA predictors by six years and the other ones by three years.

3.3.4. Predictors and control variables

The most important predictor in this research is *inequality of education achievement*, which is measured as the proportion of reading scores variance explained by socioeconomic background, namely the OECD Economic, Social and Cultural Status index (ESCS). The variable has been transformed so that it follows a normal distribution with a mean of zero and a standard deviation of 1. In the rest of the chapter, I refer to it as *inequality of school achievement*. The reason why I use this variable instead of more traditional measures is as follows.

In order to deal with a measure of inequality, Gini coefficient is the most traditional measure. Income Gini coefficients are widely available for most countries but databases are not continuous and many countries report Gini every five to ten years. Instead, I use an uncommon indicator such as the proportion of scores variance explained by ESCS. This variable shows how determinant the student's socioeconomic background is for achievement, *i.e.* how 'predictable' school achievement is, given ESCS in a country.

The measure shows an important variability between countries and this may reflect a broader conception of educational systems. Indeed, the extreme values for the model just containing ESCS as a predictor of PISA reading scores fluctuate between 1.8 and

27 per cent of the explained variance in PISA 2009 (1.8 to 22 percent are the values for PISA 2006). In fact, the variable accounts for the persistence of ESCS on scores but not necessarily just as a school education characteristic, but also as for how the country deals with providing equity of opportunities for school students. The more 'predictable' the country, the stronger the social reproduction that takes place. Therefore, I can also use this indicator as a proxy of how strong social reproduction mechanisms through education are for countries.

Alternatively, I may have considered other measures of inequality also coming from PISA. For instance, I could have calculated the ratio between the scores of the top quartile (or any other percentile) of ESCS and the bottom one, this way obtaining a relative measure of inequality of achievement. Other alternative measure might be to build a GINI-like index based on the distribution of achievement.

Nevertheless, such a measure would not necessarily reflect the persistence of SES on achievement and would only be sensitive to variations in country specific socioeconomic distributions of school achievement rather than the persistence of socioeconomic status. Despite the issues that may emerge when considering a rather untraditional indicator, and the fact that coefficients might be difficult to interpret, I maintain that having a measure of inequality and social reproduction through education in a cross-country setting would not be possible by other means.

I also include PISA reading scores as a predictor. The reasons why I selected PISA reading proficiency scores are as follows. PISA measures three proficiency domains: reading, maths and sciences. Every three years, the main subject is properly tested, whereas the remaining subjects are estimates coming from a small set of items as well as other estimation methods. In this case, I took data on reading performance because the reading subject has been the main one twice (2000 and 2009), the others being main subjects just once. For the remaining years (2003 and 2006) PISA scores are estimates of reading performance as the main subjects were, respectively, maths and sciences. There was no other option in order to count on as many test scores as

possible. This way I add a longitudinal component with the advantages it represents as it allows me to control for unobservables.

I include inequality of school achievement and the ESCS for the years 2000, 2003, 2006, and 2009 for all the models in this chapter. ESCS is an indicator including parental education and occupation, cultural assets, and educational resources in the household and it is used as an indicator of socioeconomic status. The base model, thus, considers inequality of school achievement and ESCS. The reason supporting this desicion I also introduce *control variables* related to PISA (OECD 2005, 2010). From there, I use the following variables: (i) and (ii) the percentage of score variance explained by within-school factors, which is sometimes called school inclusiveness index. Both variables consider crucial characteristics of school systems.

ESCS index is a unidimensional measure of SES that OECD has estimated by using the first factor from principal components factor analysis on the basis of the following indicators: (i) the International Socio-Economic Index of Occupational Status (ISEI); (ii) the highest parental education level in years of schooling; (iii) the PISA index of family wealth; (iv) the PISA index of home educational resources; and (v) the PISA index of possessions related to 'classical' culture in the family home (OECD 2012b). The index has a mean of zero -which is the OECD mean, and a standard deviation of 1. A negative value means that the country's ESCS is lower than OECD average and positive values show the opposite. The index is not calculated as a country specific one but as a general index; this being the reason why some countries show means below or above zero (see OECD 2010, 2012b for more details).

In relation to variance composition, within-school variance as a proportion of the total variance has also been included. High within-variance proportion might mean that different schools achieve similar test scores in general, and may be more inclusive in the sense that schools are more likely to serve students with different performance levels. Low within-variance proportion, on the other hand, means that schools differ but are internally more homogeneous, this meaning that school education is segmented as there are low-performance schools that differ too much from high-performance

schools. Finding bad-school, poor-student, low-performance (or the opposite) at the same time would indicate a segregated school system. This indicator intends to measure the first and the third elements of the triad. Following this example, socioeconomic inequality of school achievement as understood in this thesis would account for the second and third elements of the triad.

As a robustness check, I shall also use maths scores and the variance composition indicators related to PISA maths. Maths scores in standardised tests have long been recognised as a better measure of school performance and as having a higher predictive value of future educational outcomes (*e.g.*, university entry, college performance, labour market outcomes, etc.). Nevertheless, including PISA maths scores imposes important shortcomings I shall describe in detail in the results section.

An issue to address is why PISA and why these indicators have been used. Countries (regions and states in the US, for instance) measure school achievement by using a variety of tests and exams. Some tests check the fulfilment of curricular goals at different levels while some are properly exit tests leading to school leaving certifications or diplomas (A-levels in the UK, Baccalaureat in France, for instance). Most are used by HEIs to select students, so they are of critical importance for the student's chances of being offered a place and continuing education. Countries like the US use standardised tests for HE admissions and school results. Since tests are used for different purposes, different aspects are measured. On the other hand, curricula differ amongst countries, thus a measure unaffected by curricula becomes necessary. Although I am aware of the fact that some research takes issue with this, threats to validity and technical critiques to PISA are well beyond this research.

More controls are also included in the models: GDP per capita (PPP), educational attainment of the labour force – as measured by the percentage of the labour force with HE –, the participation in unemployment of people having HE, and GER for secondary education, defined as the ratio enrolment to people in the target population (13 to 17 year-olds in most countries). I have also used variables to measure the investment in secondary and HE as a fraction of GDP per capita. Those indicators

give an account of the wider socio-economic context. Educational attainment of the labour force is important both in terms of productivity and returns to education. Secondary attainment sets the bottom line, as it is the main formal prerequisite to HE. A country with no universal enrolment rate at the secondary level could hardly increase HE GER within a 10-year period. Finally, investment, measured as expenditure in education as a percentage of the GDP per capita is one of the main indicators when focusing on policy. It may reflect countries' priorities as well as the relative cost of specific educational levels when compared to a higher or lower one. In relation to that, it would have been useful to gather data on public-private expenditure as it shows the main features of educational policy: the role of the private sector and cost transfer. Regrettably, that information is only available for a few countries and there is not a systematic pattern in terms of periodicity.

3.3.5. Data issues

There are many gaps in the dataset, especially in data from WDI and UIS- UNESCO where the data are collected every year. I dealt with this by using the last available value for the periods 1998-2000, 2001-2003, 2004-2006, and 2007-2009 to match PISA variables for 2000, 2003, 2006 and 2009, respectively. In other words, I have a collapsed dataset informing four three-year periods.

The final dataset contains the following variables: PISA reading scores, percentage of scores variance found within schools, ESCS, R-squared from ESCS, GDP PPP, percentage of the labour force with HE, participation of HE attainers in total unemployment, GER secondary education, expenditure in secondary education as a percentage of GDP per capita, and expenditure in HE as a percentage of GDP per capita. Every variable is reported (when possible) for the last year of the periods mentioned above. There are data for 63 countries, which corresponds to countries participating in PISA 2009.

Nonetheless, one of the main concerns, especially when working with comparative data, is incomplete information for some variables. I considered imputing predictor

values by using two criteria. Firstly, I replaced missing values with fitted values for each country given by:

$$[3.1] \quad X'_i = \alpha + \beta_1 X + \beta_2 Per + \boldsymbol{DCtry} + \varepsilon$$

Where X' is the predicted value of an independent variable X, given by a linear combination of the valid Xs, the period Per and a vector of country dummies $\boldsymbol{DCtry}$. This is equivalent to fit a fixed effects model and replace the missing values for their fitted values obtained from [3.1].

The second criteria consisted in computing a within-country mean for each country. For each country i and time t, missing values have been replaced by the following expression:

$$[3.2] \quad X'_{it} = \bar{X}_i$$

Where $\bar{X}_i$ is the observed mean value of independent variables throughout the whole period of study.

With either strategy, imputations were made only if at least two observed values were present for the variable. If, for a country, an independent variable had presented only one observation, no imputations would have been made for the independent variable/country.

There are, of course, more imputation alternatives (*e.g.*, more complex regression-based techniques such as chained equations). It is difficult to find variables with no or few missing values in the dataset that may be used as predictors as, with the exception of identification variables, they face the same missingness issues as the variables to impute, then there are low chances of building a good imputation model. Even though the data allows multiple imputation, estimates would not differ substantially from those imputing using [3.1] due to the lack of auxiliary variables. This is not an appropriate setting to try multiple imputation techniques as the missing

at random assumption (MAR) may not hold as missingness is not related to non-response patterns but to countries' strengths and weaknesses affecting data collection carried out by international agencies, as well as countries' policies regarding the generation, production, and dissemination of indicators[3]. For instance, the first round of PISA included OECD countries and a few middle-income and developing countries; whereas students from 63 countries were measured in 2009. This means that participating in PISA is not a random mechanism but conditional to countries' income and development. In addition to that, some countries participating in PISA 2000 did not participate in the following version on purpose, this way allowing policies to crystallise or political reasons. Anyway, multiple imputation estimates, using chained equations, are also shown when assessing robustness and in more detail in Appendix A. However, I prevent that, due to the reasons given, these imputation models might not be particularly robust.

All the models reported in the chapter include missing dummies. In section 3.4, I report estimates for models using [3.1] and [3.2] techniques, lagged predictors as well as models only considering full case data. I am aware of the fact that imputation using averages may reduce the variance critically but chances are limited

3.3.6. Model specification

The model will measure the effect of the socioeconomic distribution of school achievement on access to HE. I start by establishing the access to HE as a dependent variable being a function f of a series of variables as shown below.

$$[3.3]\ GER_i = f(SEdChs, WEC, Inv)$$

In [3.3] GER_i is the outcome variable, namely HE GER; $SEdChs$ is a series of variables informing school education characteristics such as equity of school achievement, performance, and social inclusiveness/segmentation; WEC are variables giving account

[3] For instance, the outcome variable may have many gaps as in some countries enrolment rates are not necessarily calculated by using administrative records but household surveys applied with very different periodicity.

of the wider socio-economic context such as GDP, school attainment of the adult population, and educational profile of the labour force; Inv are variables accounting for investment in education at secondary and higher levels as a fraction of the GDP.

Firstly, I use an OLS estimation. Let *Char* be a series of m country-level variables other than inequality of school achievement as in [3.4]. Formally, the OLS model is:

$$[3.4] \quad GER_i = \alpha + \beta_1 Ineq_i + \sum_{j=1}^{m} \beta_{2j} Char_{ji} + \varepsilon_i$$

For the *i-th* country: GER_i is the HE GER, α is the intercept, $Char_{ji}$ are m variables to be used as controls, and ε_i is the residual. The main coefficient of interest is β_1, though some of β_{2j} might be of interest too, especially those regarding PISA-related variables.

Given the constraints and the nature of the data, I do not intend to establish causal relationships, but associations between variables. Moreover, there are issues like reverse causation and endogeneity that might bias the results. The main question here is whether β_1 is actually measuring the returns to inequality of school achievement or the returns to other unobserved confounders.

Next, I use all available data in order to deal with observed confounders but I take a different approach to deal, in part, with omitted variable bias. One way of specifying the model is to add as many dummy variables D_i as countrites $Country_i$ to [3.4] above. As the time dimension is included, the dependent variables, predictors, and residuals are added to the subscript t, this meaning that now estimates are for the *i-th* country in time t.

$$[3.5] \quad GER_{it} = \alpha + \beta_1 Ineq_{it} + \sum_{j=1}^{m} \beta_{2j} Char_{jit} + \sum_{i=1}^{n} D_i \, Country_i + \varepsilon_{it}$$

In [3.5] all the effects of unobserved time-invariant characteristics are absorbed by the dummies. Another specification is the use of the 'within estimator', which measures the deviations from the mean within each entity – in this case, countries – so deviations of time-invariant unobserved characteristics equals to 0. I specify the model including dummies instead of the demeaned formulation as it is straightforward that controlling for dummies captures countries fixed effects.

The fixed-effects model still has its own caveats. Firstly, omitted variable bias still remains as the fixed effect model can only control for unobservable confounders that are constant across time. Secondly, since the between-subject variance is taken out, variability is reduced, this meaning a loss of 'signal' at the same time that more 'noise' is likely. With a short panel, as the one I used, low variability might be a critical issue because in the fixed effect model all the variation comes from within-subject. Moreover, given this low variability scenario, problems associated with measurement error may distort the analysis and the estimates. This is especially sensitive to some variables in the dataset; for example, those reporting SES based on ESCS. However, the measurement error in this dataset is unlikely to be systematic. Finally, there are many imputed values so it is also necessary to check the robustness by comparing estimates obtained with different imputation techniques as well as complete case analysis (see relevant tables in Appendix A for more detail on missingness).

I shall start specifying the estimation models by establishing a base model containing inequality of school achievement, PISA reading scores and ESCS. Additional controls will be added to the base specification in further steps, starting from within/between school variance composition, then I introduce indicators measuring the wider socioeconomic context, and finally I add variables on educational expenditure. The reason behind this sequence is that I mostly deal with the 'interaction' between ESCS and reading scores so any variables directly involved in the measure of socioeconomic inequality of school achievement are shown from the outset. I expect the inequality measure to attenuate as control variables are introduced.

3.3.7. Power and sample size

The small sample size in this study (228 observations for 61 countries) is also problematic as the power of the estimates is affected. In this context, 'power' means the probability of obtaining a true non-null effect, and depends on sample size, effect size, or both. In other words, a small sample size, as it reduces power, undermines the chances of a statistically significant result being a true effect. It may also overestimate the magnitude of effects unless effect sizes are truly large. Button et al. (2013) illustrate the consequences of low power on the validity of the estimates; namely, low power means low replicability. They highlight the importance of a sound discussion on sample size and statistical power as an essential scientific practice.

This is a threat to the validity of the estimates, so I shall also discuss power estimates in the results section. Power is formally defined as the probability of not making a Type II error, often referred to as β. It is commonly expressed as $1-\beta$. I use a formula that corresponds to the inverse of that used by Dupont and Plummer (1998:599) to assess the power of the estimates:

$$[3.6] \quad t_{power\ df} = {}^{\delta}/_{SE(\delta)} - t_{\frac{\alpha}{2}df}$$

Where $t_{power\ df}$ corresponds to the value of t (two sided) in the t distribution, the first term is to the ratio between δ, which is the absolute value of the regression coefficient, and $SE(\delta)$, which stands for the standard error of the coefficient. The standard error contains the sample size as it is the ratio between the standard deviation and $\sqrt{n}$, where n stands for the sample size. The expression $t_{\frac{\alpha}{2}df}$ denotes the value of t at a given value of $\alpha/2$, with df degrees of freedom. This way, it is straightforward that

$$[3.7] \quad {}^{\delta}/_{SE(\delta)} = t_{\delta\ df}\ ,$$ so replacing the terms, I obtain:

$$[3.8] \quad t_{power\ df} = t_{\delta\ df} - t_{\frac{\alpha}{2}df}$$

It is clear from [3.7] that increasing the sample size leads to smaller standard errors so $t_{\delta\, df}$ is bigger.

With over 100 observations, normal and t distributions are similar, so one could argue that $t_{power\, df} \cong z_{power}$. At $\alpha = .05$, the z value (two sided) is 1.96, then substituting in the equation:

$$[3.9] \quad Z_{power} = {}^{\delta}/_{SE(\delta)} - 1.96$$

If the confidence level were to increase, statistical power drops.

To obtain the probability of finding a true effect, I use the normal cumulative distribution Φ, thus:

$$[3.10] \; Power = \Phi\left(Z_{power}\right),$$

where Φ is the normal cumulative distribution function.

A different perspective is that proposed by Cohen (1988) -the one implemented in GPower software- to study the relationship between power and size effect. In a context of multiple regression, Cohen suggests using f^2, which is the ratio between the variance explained by the model and the residual, as a measure effect size:

$$[3.11] \; f^2 = \frac{R^2}{1 - R^2}$$

The effect size of a given coefficient can be measured as a change in R^2 comparing to the model including the variable with another not including it. In bivariate regression, R^2 is simply the square of the standardised regression coefficient r, which is the Pearson's correlation coefficient, whereas in multiple regression the contribution of a specific variable to the model is measured through partial correlations, then the variable contribution is a partial-R^2, which is defined as:

$$[3.12]\ R^2{}_{YX,X'} = {t^2}/{(t^2 + N - (k+1))}\ ,$$

where Y stands for the dependent variable, X is the relevant predictor, and X' is a set of predictors other than X, t^2 stands for the t-statistic corresponding to the coefficient. $R^2{}_{YX,X'}$ already controls for the effect of any X' variables. Hence, by substituting in [3.11], the calculation of f^2 becomes:

$$[3.13]\ f^2 = \frac{{t^2}/{t^2+N-(k+1)}}{1-{t^2}/{t^2+N-(k+1)}}\ ,$$

Amplifying both the numerator and the denominator by $(t^2 + N - (k+1))$, the expression becomes:

$$[3.14]\ f^2 = \frac{t^2}{t^2+N-(k+1)-t^2} = \frac{t^2}{N-(k+1)}$$

which is the measure of the effect size I use. The parameter f^2 follows the F distribution so, for a particular F-value, given the sample size and the number of parameters the model estimates, the statistical power will be the area under the F curve up to the F-value corresponding to the effect size and the relevant degrees of freedom (for a further explanation, see Nakagawa and Cuthill 2007:598) .

Another common problem with small samples is that the complexity of the model may exceed what the data are able to deliver. For instance, for a small sample with too many parameters, it is likely that the model is overfit. In other words, an overfit model is likely to be tailored to a particular sample so results are hardly replicated in out-of-sample estimations. This is also termed as the model parameters being 'idiosyncratic' to a specific sample. There exist several approaches to test model overfit, but I prefer parsimony as a general methodological principle.

In consequence, I shall report Akaike's and Schwartz's Bayesian information criteria (AIC and BIC, respectively) on the estimates tables as an indirect way of assessing overfit. The advantage of using either information criteria, although there is no test,

is that they penalise the inclusion of additional parameters while at the same time assessing goodness of fit. Beyond their limitations, the use of information criteria is intuitive and researchers use them as conventions to select models (Kass and Raftery 1995).

3.4.Data Analysis and Results

3.4.1. Descriptive statistics

The most recent values of the outcome, 3-year lagged predictors and control variables, are reported in Table 3.1, where an important amount of variability can be found. Table 3.2 shows that GERs in HE have steadily increased on average from 49 to 63 per cent from 2003 to 2012, this representing a growth of 28.6 per cent. Lower HE participation countries (those at P25) show a more rapid increase of 54 per cent, from 33.9 to 52.4 per cent, whereas countries with the highest enrolment rates rose from 64.9 to 74.4 per cent, this representing a 14.6 percent increase in HE participation.

Statistics in Table 3.1 also reflect a diversity of countries characteristics. For instance, the interquartile range for GDP per capita is about US$ 11,000, 80 points for PISA reading scores, 7 percentage points for the proportion of variation in scores explained by ESCS, 12 percentage points for the expenditure in secondary education as a proportion of GDP per capita, and 17 percentage points for HE expenditure as a proportion of GDP per capita.

Relevant to this analysis is the fact that, when predictors are lagged by three years, as inequality of school achievement decreases, GER increases. This negative association is reversed in the last period, as both mean and median ESCS R2 increase in 2009 (see Graph 3.1 and Table 3.2).

No major changes to ESCS methodology were made in PISA 2009, but the inclusion of nine additional countries might have modified the relationship. It may have also been possible that, in 2006, the indicator of inequality had been affected by the scores' estimation procedures although it is not documented anywhere.

Table 3.1. Descriptive statistics. 2009 or last available value (2012 for gross enrolment rate, higher education)

Statistic	Gross enrolment rate (%), higher education	Variance in reading explained by ESCS (%)	PISA Reading score	ESCS (1)	Variance in reading within schools (%)	GDP per capita PPP ($05)	% of Labour force with HE	Unemployed with HE as % of total unemployment	Gross Enrolment rate secondary education	Expenditure in secondary education as % of GDP percápita	Expenditure in higher education as % of GDP percápita
p10	35.20	7.80	390.00	-1.16	39.77	8267.57	15.45	6.70	84.08	14.79	14.15
p25	52.38	9.00	421.00	-0.62	48.69	11649.73	20.10	11.85	89.69	16.94	19.55
p50	62.70	13.30	478.00	-0.13	58.36	23191.64	26.40	17.20	96.69	23.81	26.02
mean	62.76	13.10	462.37	-0.25	59.68	23667.02	28.32	18.10	96.84	23.22	28.39
p75	74.39	16.00	500.00	0.15	71.26	32469.17	35.00	21.75	101.94	29.20	37.68
p90	86.02	19.00	520.00	0.34	78.90	41187.66	40.30	33.20	114.93	32.92	46.14
sd	20.12	4.80	50.90	0.53	14.83	14568.25	11.01	9.80	12.17	7.38	12.02
N	57	63	63	63	62	62	50	56	60	51	49

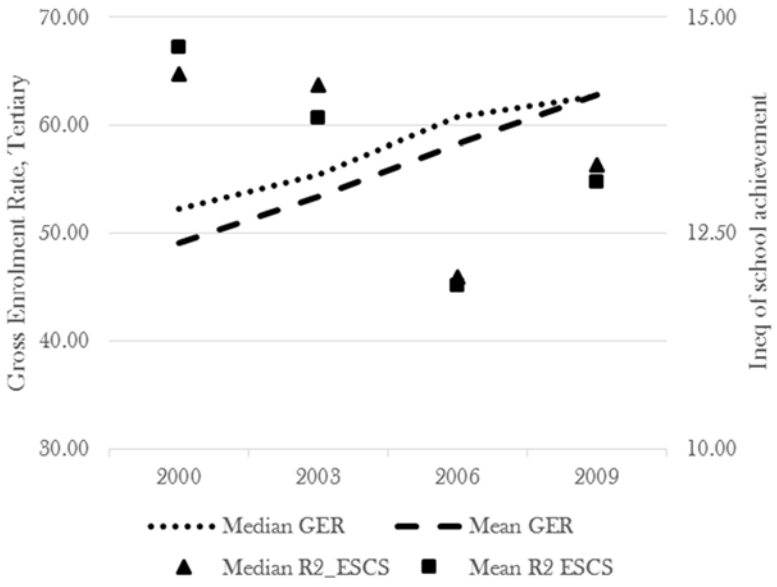
(1) ESCS refers to the PISA index for economic, social, and cultural status.

(2) Cross-country average is 493 for OECD countries. Mean reported here is lower as many countries performing below OECD have been incorporated in the last measurements. There is also an effect from the fact that these statistics are average values of country means instead of the average scores of PISA testees.

Table 3.2. Descriptive statistics for the main predictor and the outcome

	2000 --> 2003		2003 --> 2006		2006 --> 2009		2009 --> 2012	
Statistic	%Reading ESCS	GER HE	%Reading ESCS	GER HE	%Reading ESCS	GER HE	%Reading ESCS	GER HE
p25	10.18	33.91	9.92	37.75	8.48	46.06	9.00	52.38
p50	14.34	52.29	14.22	55.41	11.99	60.81	13.30	62.70
Mean	14.65	49.09	13.83	53.47	11.90	58.29	13.10	62.76
p75	18.88	64.87	17.89	68.00	14.53	71.01	16.00	74.39
Sd	5.12	20.83	5.36	21.46	4.19	19.60	4.80	20.12
N	31	57	39	58	54	56	63	57

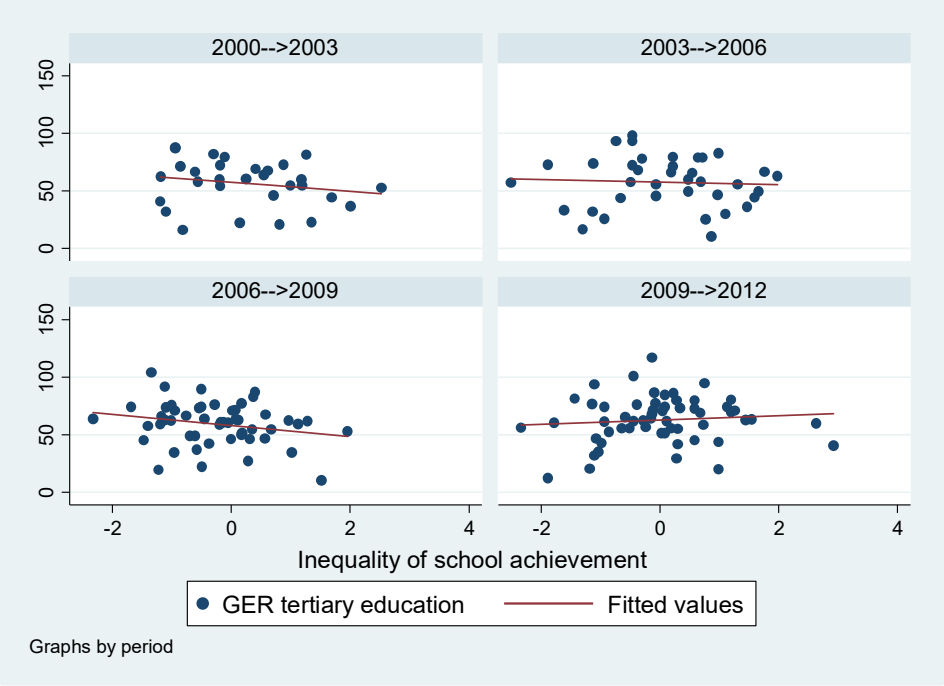
Graph 3.1. Gross enrolment rate (2003-2012) and inequality of school achievement (2000-2009)



A positive association between inequality of school achievement and GER is detected for the last period (the predictor in 2009, the outcome in 2012). A slightly positive slope is shown in the scatterplots (Graph 3.2), while the slope has been negative for previous years. Although this result might be surprising, I would like to hypothesise the reasons.

A plausible explanation might be to assume this as a consequence of the 2008 crisis. A positive association between inequality and enrolment rate might be due to better-off students postponing their entry to the labour market, this way relying on education as the crisis passed. Another hypothesis is that student aid policies may have been cut back by some governments, this way affecting low SES students' opportunities, or credit agencies may have faced severe difficulties, as it happened with Fannie Mae bankruptcy in the US. Nevertheless, the scatterplots below should be taken with caution as they only show simple correlations between variables and might be subject to observable or unobservable omitted variable bias.

Graph 3.2. Scatterplots for predictor and outcome, by period



Through different specifications, I shall analyse the relationship between GER and inequality of school achievement. I start by specifying a multiple linear regression model. Then I introduce the controls specified above.

3.4.2. Basic OLS model

I start the analysis by specifying OLS multiple regressions. The predictor and control variables have been lagged by three years. At a first stage, I use the base specification containing inequality of school achievement, ESCS and PISA reading scores. Then control variables are introduced in blocks to capture effects otherwise in the error term, as in the third model, which includes PISA variance composition.

Table 3.3. OLS regressions on HE gross enrolment rates (robust SE) (+)

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-0.567	0.827	-2.368**	-2.850**	-2.253*	-2.535**
	-1.099	(1.233)	(1.200)	(1.257)	(1.151)	(1.170)
ESCS	8.318***	7.650***	10.204***	7.305**	9.034***	5.110*
	-2.718	(2.597)	(2.849)	(2.820)	(2.667)	(2.612)
PISA Reading (z scores)	8.379***	8.736***	10.509***	9.754***	9.876***	9.376***
	-1.556	(1.541)	(1.564)	(1.618)	(1.564)	(1.597)
% within variance		0.145*	-0.094	-0.116	-0.092	-0.109
		(0.075)	(0.080)	(0.079)	(0.076)	(0.072)
ln (GDP per capita PPP)			-11.706***	-11.695***	-8.902***	-8.216***
			(2.557)	(2.389)	(2.475)	(2.199)
% Lab force with HE			0.255*	0.240*	0.174	0.125
			(0.132)	(0.130)	(0.134)	(0.132)
HE Unemp (% total unemployment)			0.465***	0.488***	0.405**	0.438**
			(0.170)	(0.171)	(0.172)	(0.171)
GER Secondary			0.224**	0.233***	0.219***	0.232***
			(0.088)	(0.082)	(0.079)	(0.071)
Exp in secondary (%GDP)				0.546***		0.644***
				(0.151)		(0.151)
Exp in tertiary (%GDP)					-0.217***	-0.266***
					(0.051)	(0.058)
Constant	61.006***	52.624***	148.193***	136.134***	130.924***	110.816***
	-1.487	(5.111)	(24.205)	(22.526)	(23.249)	(20.927)
N Observations	228	228	228	228	228	228
R2 Adjusted	0.396	0.441	0.592	0.606	0.624	0.646
AIC	1927.9	1911.9	1848.0	1841.7	1830.9	1819.1
BIC	1951.9	1942.8	1906.3	1906.9	1896.0	1891.1

* p<.10, **p<.05, *** p<.01

(+) All models include missing dummies

In these models, inequality of school achievement is not statistically different from zero. By adding variables accounting for the wider economic context, estimates start changing and becoming statistically significant. This remains upon introducing controls accounting for expenditure. Indeed, inequality of school achievement becomes negative and statistically significant. Depending on the specification, a one standard deviation increase in inequality is associated with a decrease of 2.4-2.9 percentage points in GER, as shown in estimates between the third and sixth column in Table 3.3. By adding controls, I still find a significant association, suggesting that the effect is robust to the introduction of more control variables. With regards to the magnitude,

the effects seem mild. However, if the context is considered, a value of 2.4-2.9 points is definitely relevant as the mean increase in GER in the 2003 to 2012 period has been around 4 percentage points every three years (see Table 3.2). Nevertheless, given the sample size, the power of estimates (*i.e.* the probability of finding a true effect), is rather modest, in the range of 0.50 and 0.62 (a threshold of 0.80 is commonly considered as a rule of thumb).

Estimates of PISA reading scores are consistent across models and the effect size seems relevant. An increase of one standard deviation in PISA reading scores would lead to an expansion of 8.4-10.5 percentage points in GER. The coefficient decreases with the introduction of expenditure controls but remains massive. What this result indicates is that the absolute value of school achievement does matter for HE participation and that, given the effect size, countries with high HE participation are those with high PISA reading scores.

Models in Table 3.3 were estimated using interpolation to replace missing values in the independent variables. As a robustness check, I ran the regression using only full case (i.e., cases with complete information for each model). Differences in estimates are not relevant but coefficients with full case show slightly higher estimates (see Table A.3 and Table A.4 in Appendix A). Indeed, an increase of one standard deviation in inequality of school achievement is associated with an increase in GER ranging from 2 to 5 percentage points, against 2.9 points as the highest estimate in Table 3.3.

Results so far tell that inequality of school achievement and PISA test scores matter. Specifically, an increase of one standard deviation is associated with about 10 percentage points in HE enrolment rate – a half of a standard deviation in GER in other words. Could this suggest countries improve PISA results, thus school performance, as the most effective means of increasing participation in HE? Does this tell countries that even though inequality of school achievement affects GER negatively, increasing PISA scores would be more effective in increasing HE access?

There are also some counterintuitive results such as the negative and significant coefficients associated to tertiary education expenditure. Does tertiary expenditure affect negatively HE participation? The underlying reason might be either the lack of better control variables or that OLS is not the appropriate technique. In fact, there are many variables I need to control for that are not measured in the models as, for example, those related to the general institutional characteristics of education systems (for instance, whether early selection takes place), cultural attitudes towards education, or historical factors which may indicate that OLS is not the best approach. In turn, fixed effect estimates may in part address the omitted variable bias issue that is seemingly affecting OLS estimates.

3.4.3. Fixed-effects estimates

Estimates reported above might have been affected by a number of issues concerning OLS, the main being omitted variable bias. A sensitive issue is the lack of information on the main institutional settings in secondary and HE as well as cultural attitudes towards education, which are likely to remain constant across time.

As the dataset contains information for four periods, I take advantage of the data structure. Variables dealing with the wider institutional settings of educational systems and structural characteristics at a societal level are not likely to change over a 10-year period, although educational policies might well change in short time periods. Omitted variable bias could, in part, be addressed by using country fixed effects. A fixed effects approach will show how robust OLS results remain when dealing with omitted variable bias coming from time-invariant unobserved variables.

Variables have been introduced in blocks as in OLS models. Results are presented in Table 3.4. The basic specification only includes inequality of school achievement and the coefficient is statistically significant. An increase of one standard deviation in inequality

Table 3.4. Fixed effects regressions for higher education's GER (+)

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-3.645**	-3.767**	-2.142	-2.482*	-2.451*	-2.738**
	(1.403)	(1.445)	(1.400)	(1.257)	(1.314)	(1.290)
ESCS	4.441	5.033	5.597	2.961	6.534	3.374
	(7.231)	(7.040)	(6.331)	(6.878)	(5.927)	(6.527)
PISA Reading (z scores)	4.824	4.951	1.797	1.750	1.907	1.478
	(3.579)	(3.592)	(3.180)	(3.215)	(3.194)	(3.161)
% within variance		-0.012	-0.024	-0.085	-0.007	-0.062
		(0.083)	(0.068)	(0.065)	(0.065)	(0.068)
ln (GDP per capita PPP)			15.102**	11.490*	18.986***	14.863***
			(6.650)	(6.485)	(6.011)	(5.482)
% Lab force with HE			0.347***	0.260**	0.330***	0.256**
			(0.126)	(0.119)	(0.119)	(0.118)
HE Unemp (% total unemployment)			0.234	0.157	0.254*	0.164
			(0.155)	(0.168)	(0.142)	(0.159)
GER Secondary			0.141	0.245***	0.143*	0.235***
			(0.087)	(0.082)	(0.086)	(0.084)
Exp in secondary (%GDP)				0.718***		0.651***
				(0.198)		(0.200)
Exp in tertiary (%GDP)					0.192	0.116
					(0.149)	(0.126)
Constant	59.884***	60.693***	-112.313*	-95.779	-156.673**	-131.067**
	(1.918)	(5.506)	(66.142)	(65.024)	(60.219)	(55.684)
N Observations	228	228	228	228	228	228
N Groups	61	61	61	61	61	61
R2 within	0.236	0.239	0.408	0.447	0.427	0.459
Rho	0.827	0.827	0.881	0.869	0.912	0.897
AIC	1501.8	1505.1	1461.6	1450.0	1458.3	1449.2
BIC	1522.4	1532.5	1513.0	1508.3	1516.6	1514.4

(+) All models control for missing dummies

When controlling for PISA variance composition, the coefficient rises to 3.8. By adding more control variables, the coefficient for inequality of school achievement drops to 2.1-2.8 and still remains significant, but at a lower significance level (p=0.10), except for the specification including both secondary and HE expenditure, which remains significant at p=0.05. In comparison to OLS estimates, which remained unchanged as control variables were introduced, fixed effects estimates for inequality of school

achievement become smaller. Common to both approaches, nonetheless, is the fact that it seems that inequality of school achievement does affect HE participation.

Fixed effects and OLS models also differ in effect sizes of some control variables. In fixed effects models, coefficients for PISA reading scores are no longer significant. This means that PISA scores' OLS estimates were affected by omitted variable bias as most of the effect has been taken by time-invariant cultural, historical, and institutional characteristics for instance. Equally important is that PISA coefficients effect size is massively reduced. Estimates for tertiary education expenditure are no longer negative nor statistically significant, as in OLS estimates.

Power estimates are high enough for the first two fixed-effects models, in the range 0.8-0.85, whereas I find power estimates in the range of 0.49-0.60 in the more complex models as in the last three columns of Table 3.4. This may crucially affect the reliability of the results because there is a high probability of the last estimates not being a true effect. Thus, results have to be taken with the necessary caution.

With regards to models' appropriateness, unlike OLS models, the introduction of more variables in fixed-effects estimates does not improve goodness of fit – as assessed by BIC – but the opposite, so there is ground, if not to consider overfit to some extent, to contend more complex models.

However, in more substantive terms, I could still underline that there is no significant relationship between a cross-country absolute value of school achievement, as measured in PISA, and GER in HE. Indeed, there are, in fact, countries with comparatively poor performance in PISA reading with high GER as well as high PISA performing countries with lower GER in HE. This is what fixed-effect estimates show when comparing to OLS's. Thus, for a country, the absolute value of PISA is not relevant for HE participation, although the relative school performance within that country seems to be relevant in order to access HE. Therefore, the relationship between prior achievement and access to HE should be regarded as idiosyncratic or country specific.

Nevertheless, the distribution of school achievement seemingly matters at a cross-country level and affects HE participation negatively. The more the persistence of SES on school achievement, the fewer chances of increasing participation in HE. Although this indicator has not been used by researchers and may be hard to interpret, my estimates show that weakening the association between SES and school achievement might be of key importance for countries to increase participation in HE. Moreover, a finding like this highlights the importance of counting on variables measuring social selection mechanisms in education systems, as their influence might be crucial in order to study social reproduction mechanisms through education. More importantly, however, is that this sort of finding may alert countries designing interventions aimed at achieving equity of HE access to look outside HE policy.

I have checked my results' robustness by running the models with different imputation methods and several alternatives of lagged periods as well as including complete case analysis. These estimates are shown in Table 3.5. Column 1 shows the same coefficients for inequality of school achievement shown in Table 3.4.

Estimates do not vary dramatically, but coefficients are bigger in columns 2 and 4 (PISA variables lagged by 6 years and full case, respectively). Full case estimates are similar to those reported in Table 3.4 (column 1), so are estimates using within-country means as imputation technique (column 3).

Table 3.5. Fixed effects coefficients for inequality of school achievement

	1	2	3	4	5
	b	b	b	b	b
Base	-3.645***	-4.394***	-3.386**	-3.061*	-3.206**
Add PISA variance components	-3.767**	-4.323***	-3.553**	-2.689*	-3.210**
Add Socio economic variables	-2.142	-3.888***	-2.280	-1.431	-2.174
Add Secondary expenditure	-2.482*	-3.545**	-2.362	-2.688**	-2.809**
Add HE expenditure	-2.451*	-3.644***	-2.474*	-3.299**	-2.468*
Add Sec & Tert expenditure	-2.738**	-3.301**	-2.699*	-2.673**	-2.968**

* p<.10, ** p<.05, *** p<.01

1: Variable of interest and all controls lagged by 3 years from outcome; missing values interpolated

2: Variable of interest and PISA variables lagged by 6 years, all the remaining variables lagged by 3 years from outcome; missing values interpolated

3: Variable of interest and all controls lagged by 3 years from outcome; missing values imputed by using within-country means

4: Variable of interest and all controls lagged by 3 years from outcome; full case

5: Variable of interest and all controls lagged by 3 years from outcome; multiple imputation estimates

Coefficients in column 2 are estimated from a smaller sample as lagging PISA variables by 6 years means missing one observation point so the total variance is reduced. Nevertheless, the effect size is bigger and statistical power stays in the 0.80 area. Detailed tables reporting estimates for imputed data and full case are presented in Appendix A, as well as the full outputs for the estimates shown in Table 3.5.

In summary, results suggest that an increase of one standard deviation in inequality of school achievement is associated with a decrease of 2.5 to 4 percentage points in HE GER. The coefficient value is relevant because, as noted above, this range of values represents between a half and three-quarters of the average GER increase in the last three years.

However, I introduce an additional robustness check by re-running the models using inequality of school achievement but using PISA maths instead. Although the proportion of PISA maths variance explained by ESCS is not available for PISA 2009, I used 2000, 2003, and 2006 tests to re-run the estimates in Table 3.4. In Table 3.6., I compare the estimates for the same specifications in Table 3.4, but this time using the predictor values for 2000, 2003, and 2006 and the outcome value for 2003, 2006, and 2009.

Table 3.6. Estimates of inequality of school achievement using reading and maths PISA scores

	Reading b	Maths b
Base (Only R-sq ESCS and ESCS)	-3.465***	0.971
Add PISA variables	-3.414***	0.715
Add Socio economic variables	-2.152	-0.202
Add Secondary expenditure	-2.791*	-1.219
Add HE expenditure	-2.188	-0.368
Add Sec & HE expenditure	-2.692*	-1.046

* p<.10, ** p<.05, *** p<.01

I find quite different estimates conditional to which subject score is used. There are good reasons either supporting the use of maths or reading when it comes to measuring school achievement. Firstly, maths skills might be regarded as not being culturally specific, as mathematical concepts might be universal so measures of different

countries should be more transparent, easier, and more comparable than reading. In other words, there are substantive and methodological reasons supporting math scores. Secondly, one may also conjecture that the distribution of maths skills might also be less socially determined than reading. Thus, obtaining coefficients not significantly different from zero for the variable measuring inequality of school achievement is unsurprising.

Nevertheless, there are also good reasons favouring reading scores. First, during the period covered in this study, reading has been the principal subject in 2000 and 2009 whereas maths was the principal subject only in 2003. Scores of non-principal subjects are estimated on the basis of a reduced set of items, so they may be less accurate. Second, the comparison shown in Table 3.5 is made from data having less variance, which may induce unstable coefficients. Third, there are some pieces of research in the UK showing a higher relative importance of reading skills when it comes to access to university (Aucejo and James 2015). In the US some research on academic under preparedness of freshmen suggests that a lack of literacy skills is harder to address with remediation courses than a lack of maths skills (Adelman 2004). Although this work might reflect the fact that a higher proportion of university courses require reading over maths skills, it may also reflect the fact that literacy skills are more likely to be determined by social class than maths skills. This is well in line with classical developments of socio-linguistics because, as Bernstein (2003) noted, working-class children, unlike their middle and upper-class counterparts, were socialised within a more restricted linguistic code.

3.5.Discussion and Conclusions

The above results suggest that inequality of school performance is negatively associated to HE GER. Coefficients are consistent and significant for a relevant number of specifications regardless of methods (OLS or fixed effects), imputation technique, or timing between predictor and outcome variables. Besides the criticisms to PISA-driven education policies, my results show that at a cross-country level, keeping a number of variables constant, inequality of school achievement is negatively associated to HE growth.

The main caveat is the lack of statistical power of the models, including the full set of control variables. This is certainly an issue when the effects are mild as the ones reported in this chapter since small sample sizes make it harder to pick small effects. Moreover, it is also possible that the most complex models are overfit so long as BIC does not improve. I just have 10 observations per parameter in the most complex model (228/21, as missing dummies are included), which means the size/parameters ratio is in the lower bound of the conventionally recommended.

In countries where tests results are more 'predictable', conditional to ESCS, HE GERs drop. In other words, higher inequality of learning outcomes, conditional to social background, does affect HE systems. The main policy implication is that governments should not only focus on improving achievement, but on policies seeking to break socioeconomic gaps in learning outcomes. This might suggest that policies looking to increase HE participation should focus on school education instead of relying on HE student aid.

My results support the claim that the main mechanisms determining access to HE are likely to be country specific as there is no statistically significant cross-country average effect between test results and GERs, otherwise, HE GERs would be streamlined with PISA scores. This is consistent with the fact that, in some countries, HE has been accessible to students sometimes lacking the necessary skills to succeed in HE through the development of second-tier HEIs. Examples of this are for-profit universities in the US and the rapid emergence of private, non-selective HE institution. Both are included in enrolment figures, but the real value of the credentials offered is highly contested.

For more robust estimates, a larger sample – or at least feasible proxies – is required, especially on the composition of expenditure and HE policy mechanisms giving an account of cost sharing and the balance between supply-and-demand-side funding. Obtaining such data for the same number of countries for several time periods seems unfeasible. Fast growing HE systems are more likely to correspond to developing countries that have prioritised financing school education out of public funds; thus, it

does not seem to be other chances than fostering private participation and cost transfer in HE.

The fixed-effects approach undertaken in this chapter is not able to deal with sensitive issues affecting HE policy, such as the creation of new student aid mechanisms, abrupt changes in cost sharing/transfer structures, and fees policies, as they cannot be assumed as time invariant. This is probably the key weakness of the approach undertaken in this chapter: a policy shock may influence enrolment rates immediately, as evident in Latin American countries (Chile, Colombia, and Brazil, to a lesser extent). Notwithstanding, I included several control variables and dealt with time-invariant unobserved variables. It is necessary to find better controls that account for policy changes and the private/public balance in terms of funding and enrolment.

As I shall show in the next chapter, according to the literature, cost sharing/transfer mechanisms are harmless to equity of access so long as student aid policies are supportive enough and also streamlined with the characteristics and the specific needs of the target population. I shall present a case where an aggressive reform to student aid contributed in a great deal to a sharp boost to low SES students' participation in HE.

It is also desirable to have data on the socioeconomic composition of enrolment; for instance, the HE enrolment rates for different levels of income. In fact, in order to study access to HE with more depth, enrolment rates broken down by household income levels would be a key outcome to investigate. This information is available for a few countries, but is neither systematically nor periodically collected by international organisations. A short-term research agenda should include the collection of more measures of outcome and including additional variables in order to have more accurate predictors, as well as using alternative methodological approaches.

School performance and main social reproduction mechanisms play a key role as predictors of HE participation. Equity driven HE policies may have no effect if they do not recognise inequalities coming from school education. Governments have set a

series of policies to overcome barriers to access, but inequity of access still remains high.

This leads to the next Chapter, where I analyse the case of Chile. The country has expanded HE quickly and has also made progress in school education outcomes as measured by PISA scores. Chile has also provided opportunities to poor students by devoting 40 per cent of its public expenditure in HE to student aid. Notwithstanding, Chile has one of the most segregated school systems globally as well as high inequality levels, the second in Latin America behind Brazil. Despite showing impressive economic results in the last 25 years, old, sensitive, and persistent issues remain, which feature Chile as a paradigmatic case of study: most educational issues I have mentioned in the introduction and this chapter are taking place there.

4. INCREASING PARTICIPATION IN HIGHER EDUCATION THROUGH STUDENT AID: AN IMPACT EVALUATION EXERCISE OF CHILE'S 2005 REFORM

4.1.Introduction

Student financial support mechanisms have long been seen as the most prominent set of policies to reach equity in HE. Governments, regardless how public/private HE systems are, have devoted an increasing share of their expenditure to student aid. Indeed, on average, 22 per cent of the total public expenditure in HE in OECD countries corresponds to student aid. Apart from aiming at overcoming financial barriers to access, student aid may also boost student persistence and shorten course completion. This chapter focuses on the effect of student aid on access to HE through an impact evaluation of a change in the student aid system in Chile. Data limitations and the complexity of the policy do not allow an impact evaluation of the whole policy but that of grants, as I shall explain. However, the importance of evaluating grants lies in the fact that the rules have normally reserved them to the poorest students. I

evaluate the impact of the grants policy on the probability of access to higher education for the poorest students.

In 2005, the Chilean government introduced a major reform to student aid. The changes were critical because Chilean HE relies heavily on student financing. Indeed, in 2014, student aid represented a 38.2 per cent of the country's total public expenditure in HE; well above the OECD average. Nevertheless, there were two major issues that the government have to address before the reform. Firstly, there was no support for students attending private independent HEIs[4], which accounted for two-thirds of the undergraduate enrolment in 2005. Private independent HEIs in Chile started in 1981 without any direct government support except for a few –and marginal- student grants since 2001 (for more detail of the 1980's reform, see Chapter 2). However, as enrolment and participation of private independent HEIs became relevant, equity implications and the social pressure to support poor students compelled the government to take action. Secondly, multiple and often inconsistent aid programmes used to support only CRUCH universities students. Both of these issues affected HE equity as a large number of low-income students were excluded from government support and were, therefore, unable to afford HE.

The reform addressed the lack of equity by introducing loans and grants for students attending private independent universities. The main purpose was to reach horizontal equity. In other words, the aim was to realise the ambition that two students with the same skills, preferences and socioeconomic need should receive the same support from the government, regardless of the HEI they attend.

Although addressing the lack of support for students at independent private HEIs was the main focus, the government also sought to reorganise the pre-existent student aid system serving CRUCH universities students. Indeed, the reform also benefited

[4] In what follows I refer to 'private independent' as fee charging HEIs which do not receive institutional grants from the government. Public HEIs stand for state owned institutions, whereas 'private subsidised' or 'private maintained' HEIs stand for private HEIs which receive government grants. 'Subsidised HEIs', in consequence, refer to both public HEIs and private maintained HEIs. I have not strictly followed OECD (2004) definitions but adapted as much as possible to the Chilean context

CRUCH students by increasing the level of support, though the effect was milder than that on private independent HEIs.

The reform also introduced a major change in administrative terms. By implementing a single window system - instead of one application for each programme as in the old system -, the application procedure becomes simpler. This way, the risk of students not applying for aid is minimised and, consequently, the risk of not getting support provided that the poorest students have a limited access to information on the array of benefits available to them.

The government determination to guarantee aid packages according to socioeconomic need and academic merit was critical regarding equity of access. The aspiration was that every student, being aware of their economic situation and academic merit, knew in advance what kind of support was available to them before making any enrolment decision. Hence, the system became more transparent so that students could make enrolment decisions on the basis of straightforward information.

The main policy change consisted in creating a new loan system, CAE, a private loan programme, with the state guarantee as collateral. CAE was designed to support students enrolled at private independent HEIs, on the basis of socioeconomic need, but also as a 'top up' aid for CRUCH universities students, which were in part supported[5] by the existing mechanisms. On the other hand, grants, which were previously available to CRUCH students only, gradually reached the poorest students attending private independent HEIs. The policy also delivered a sharp increase of maintenance allowances - cash and food stamps - benefiting the poorest students, no matter HEI type attended.

In this chapter, I carry out an impact evaluation exercise the reform to student financing in Chile, where a massive demand shock took place in 2005. As I mentioned in Chapter 2, Chile's HE features a paradigmatic combination of: (i) a high tuition

[5] The then current aid mechanisms for CRUCH students were often not enough to cover university fees so the gap may well be covered with CAE.

fee/instruction costs ratio, (ii) consequently, a low proportion of public contributions to HE funding with public expenditure in HE representing less than a third of the total expenditure in the sector, (iii) high proportion of private enrolment, (iv) quick expansion of the total enrolment in the last two decades, and (v) a funding structure massively dominated by demand-side mechanisms.

The research question I shall address is *to what extent does widening access to financial aid enhance the chances of low-income students accessing HE?* To provide an answer, I use a difference-in-difference (D-I-D) approach as the evaluation technique. I use D-I-D which is appropriate for observational and repeated cross-section data. I found a positive and significant effect of the policy on the probability of access to HE for the lowest income students.

The chapter is organised as follows. Firstly, I start in section 4.2. by discussing the global policy context of cost sharing and cost transfer in HE. Then, I show why Chile's HE features a unique case. Secondly, in section 4.3, I discuss the literature on access to HE with a special focus on the empirical literature discussing student financing. Next, in section 4.4, I explain what the reform implications are for different socioeconomic groups. I also give an account of the main developments of the Chilean HE system, especially student aid policies between 1990 and the 2005 reform. In section 4.5, I discuss the data and outline the relevant evaluation techniques, with a special focus on selecting relevant treatment and control groups. Additionally, in section 4.6, I analyse the results and discuss their main implications. I specify D-I-D models in order to estimate the effect of the reform on the probability of access to HE for low SES students. The chapter finishes in section 4.7, where I outline the main conclusions, discuss further research gaps, and the policy agenda.

4.2.Cost Sharing and Cost Transfer in HE: The Global Policy Context

Diversification of provision, new funding arrangements, an increased focus on accountability, and a more heterogeneous student body in socio-economic, educational,

and demographic terms are some core issues HE policy has to deal with (OECD 2008b). These issues are closely interrelated to each other.

As HE has become a mass system in the developed world - and increasingly in developing countries -, there emerged pressures that governments need to address. Rising costs of instruction, infrastructure, and equipment, along with a changing student body have contested the traditional public funding formula. Many countries have adopted cost-sharing as a means of relieving public finances and increasing the resources available to HE on efficiency and equity grounds. On the efficiency side, some authors disagree that HE should continue to be funded out of general taxation because fiscal constraints may translate into sub-optimal investment in HEIs (Barr 2001; Johnstone 2003), whereas on the equity side publicly funded systems have long been deemed as biased towards the affluent students (as in the seminal paper by Glennerster, Merrett, and Wilson 1968). In developing countries, funding HE out of taxation is regarded as inefficient and inequitable, as public resources are often diverted from school education to HE, thereby favouring students that would be able to pay for HE (Psacharopoulos 1988; Psacharopoulos and Papakonstantinou 2005). Nevertheless, the same kind of reasons would also apply to developed countries according to mainstream economic theory.

Since the 1980s, International development agencies such as the World Bank (WB) and the Inter-American Bank of Development (IABD) have exerted a decisive influence in shaping HE cost transfer policies in part of the developing world. They have championed the introduction of fees and student loans, the replacement of institutional block grants with competitive funding, and the private sector participation in HE provision, financing, and coordination. In summary, these policies have reinforced market mechanisms as a driving force in HE coordination. In Latin America, an important area of influence of development agencies, these sort of measures have been introduced in Argentina, Brazil, Colombia, and Mexico through student loans and the development of a dynamic sector of fee-charging private HEIs, but to a lesser extent than the 1980s' Chilean market reforms.

Notwithstanding, similar trends have also taken place in developed countries, such as Australia with the introduction of fees in 1989, or the UK with the successive lifts of fee caps. Even in countries with no tuition fees, living expenses represent important costs for students and families. Thus student aid is not limited to help students pay fees but also supports students to afford increasing living expenses. The Nordic countries, for instance, commonly seen as a paradigm of public and free HE, have also developed extensive and generous loan programmes, either publicly or privately originated, to help students afford living costs (OECD 2008b, 2014).

Henceforth, with the term cost sharing, I shall refer to the fact that there is always a specific proportion of the total costs of HE that are borne by students and their families, on the one hand, and the state, on the other. In contrast, the term cost transfer refers to the process by which students and their families assume costs no longer borne by the state. The magnitude and the direction of cost transfer may operate from the state to individual/families, as with fee introduction, or the opposite, from individuals/families to the state, as with the abolition of fees in Germany, or the state paying HEIs uition fees, as SAAS in Scotland.

Besides the above counterexamples, transferring costs from governments to students and families has been the most common trend in HE funding in the last decades as well as the main means of increasing HE investment. Cost transfer has operated in many countries by either: (i) transferring costs in taxpayer funded systems with low or no students/families contribution towards the cost of HE by introducing tuition fees or any kind of direct contribution, (ii) increasing the individual/families contribution in HE systems already relying on relevant non-governmental funding sources, as for instance policies dealing with lifting or releasing fee caps.

Cost transfer may also make tuition fees the main metric in HE funding, even more so when tuition levels are closer to HE provision costs. The structure of HE funding is also affected. In fact, traditional block grants from governments to HEIs, based on historical quotas or funding formulas, are no longer the most relevant mechanism but demand-side funding.

In this context of cost transfer from the state and taxpayers to students and their families, low SES students who cannot afford tuition fees is a matter of concern so the key question that research has tried to address is to what extent policies, namely student aid systems, contribute to a more egalitarian distribution of HE access opportunities. HE policy is supposed to implement a number of funding mechanisms which should be effective enough to guarantee a fair access without deterring low SES students (Barr 2003).

From an economic point of view, it has been argued that credit markets have imperfections that justify state intervention and credit provision (Barr 2004), with income contingent loans as a popular means of minimising negative impacts on equity, provided that they are the most able to address risk aversion and minimise the impact of debts over the life cycle (Chapman 2006).

The evidence in a number of countries shows that cost transfer has not negatively affected the participation of disadvantaged students as long as student aid mechanisms are supportive enough and are well advertised amongst the target population. I discuss this in further detail in the following section.

Notwithstanding, reporting positive effects of aid may overlook the context in which policies have been introduced. Indeed, establishing student loans - or allowing HEIs to charge fees- differs from expanding and diversifying student funding in order to boost access of low SES students in a system with well-established fee-charging HEIs. In the first case, aid is intended as a means of offsetting the consequences of cost transfer, whereas the other seeks a more socially inclusive HE system. There might also be the case that policies seek to transfer costs to students and households and at the same time expand and enhance student aid due to social inclusion considerations.

Nevertheless, from a policy-making perspective, the effects of cost transfer on access to HE just show one side of the issue, that of how to deal with the increasing demand for HE in an efficient way without affecting the chances of low SES students entering HE. Nevertheless, to arrive at an explanation of the effects of student aid, it is also

necessary to know what other factors affect the chances of access to HE at individual, family, and institutional levels, as well as within the broader social and economic context.

Chile has been regarded as a forerunner in the radical introduction of market coordination mechanisms and the prominent role that the private sector has played (Levy 1986). Nevertheless, there are other characteristics which may make Chile and 2005 reforms a worth studying case. Firstly, it is a country which, before the reform took place, had already relied on demand-side funding. Secondly, as I mentioned in Chapter 2, since 1980 when the military dictatorship diversified funding sources, rolled back (or froze) the enrolment of public universities, and delegated much of the HE development to the private sector, and all Chilean HEIs, either public or private, charge fees representing a high proportion of instruction costs. Thirdly, the reform consisted of three major measures: (i) a re-engineering of the previously existent student aid whose main focus were public and public subsidised HEIs, (ii) an important increase of grants, (iii) the introduction of a state-guaranteed loan scheme aimed at increasing the support for low SES students attending private independent HEIs.

Before the reform, there was virtually no support for students attending private independent HEIs. After the reform, students attending private independent HEIs had the chance of getting CAE and grants. This would require an important degree of responsiveness from HEIs in order to meet the demand for extra places. Accordingly, I make the same assumption as in Chapter 3: HEIs qualifying to participate in government sponsored student aid have no constraints when it comes to fulfil the demand for extra places resulting from student aid increases.

4.3.State of Art. The Role of Student Aid on Access to HE

A more equitable access to HE implies weakening the relationship between socio-economic background and someone's present or prospective situation, as I have already pointed out (see Chapter 1). In that sense, the extent to which access to HE depends on the socioeconomic background is a good measure of effectiveness of equity

policies. More generally, equity policy is a set of measures aiming at weakening the relationship between socioeconomic background and a given educational outcome - access, persistence, employability, etc. Definitions of equity and the influence of inequalities in the school system on access to HE have already been presented earlier. Student aid, as I have also pointed out, is one amongst many factors affecting access to HE. In fact, the effect of aid is likely to be confounded with those of external factors such as the characteristics of school systems, variations in schools' quality and performance, the responsiveness of HE providers to demand, and the policy environment.

Evidence from a number of countries shows that cost sharing strategies have not reduced the participation of disadvantaged students so long as student aid mechanisms are well designed and supportive enough. In other words, to be effective, they should guarantee an appropriate level of financial support. On the other hand, information on benefits should be made widely available, this way making it easier for the target population.

Studies tend to aggregate student aid without making distinctions in relation to aid composition so the effect of aid design and packaging may be neglected (Chen and Zerquera 2011). Nevertheless, studies addressing aid composition, packaging and targeting are more focused on outcomes such as persistence, dropout and completion than access. I shall give an account of this in Chapter 5, which deals with the effect of aid on completion and dropout.

Although aid is often treated as an aggregate variable, a high proportion of the literature is devoted to the effect of several grants and scholarships programmes. In contrast, the effect of loans has not been as extensively studied (Page and Scott-Clayton 2016). Since student aid is normally made up of several instruments (loans, grants, allowances, tuition fee rebates, etc.), measuring the true causal effect of a particular programme would require good quality and wide coverage of data.

In general, empirical research has concluded that aid affects positively the chances of low SES students entering HE. In Australia, after introducing income contingent loans to meet newly introduced tuition fees instead of a no fees approach, the socio-economic distribution of HE became more equitable (Chapman and Ryan 2005). In the US, a country with a long tradition of paid HE, student aid has contributed to increasing college entry for low SES and ethnic minority students (Dynarski 2003; St John and Noell 1989). Canton & Blom (2004, 2010) found similar evidence when evaluating a student loan programme targeted to low SES students attending private –and expensive- universities in Mexico. In Colombia, positive effects of loans on low-income students participation have also been found (Melguizo, Sanchez, and Velasco 2016). In Chile, as I shall show below, the expansion of student aid has also been associated with an increase in HE participation for low SES students.

Although aid has contributed to expanding access, it has also helped the development of a second tier HE, as in the case of the for-profit sector in the US, which targets students eligible for aid. This contests the value for money of the education received (Cellini 2010), as well as the pertinence of a sector growing mostly at the expense of aid (Darolia 2013). Student funding has also had a differential impact on enrolment decisions across socioeconomic groups, aid being more effective to trigger enrolment decisions of less well-off students (Braunstein, Mcgrath, and Pescatrice 1999).

Although the evidence suggests that transferring the costs to students should not deter the poor from entering HE, risk aversion of poor students might deter them from taking up loans if aid is not well designed (Barr 2003). Cultural factors such as the configuration of a pro-HE environment within families through, for instance, having a relative or close friend in HE, may also model the decision of entering HE (Finnie 2012). The same rationale applies to aid packages design: to tackle risk aversion or any other issues affecting equity of access, a combination of loans, grants and maintenance allowances should be tailored to different social groups (DesJardins and McCall 2010; Stampen and Cabrera 1988). However, positive effects of aid packaging go beyond enrolment decisions but are also important for increasing persistence and reducing time to degree (DesJardins and McCall 2010).

Risk aversion may lead poor students to make bad decisions. For instance, in the Netherlands, risk aversion has led to low rates of loan take up. Low-income students prefer working part time -or accepting unqualified jobs- to taking up student loans. The most pervasive consequences of this have been longer courses, as students are distracted from their studies; an underused aid system; and a less than optimal enrolment (Oosterbeek and van den Broek 2009). These issues have consequences on equity and efficiency. So does completing HE studies become harder for low-income students as the risk of dropping out increases: if students' circumstances change, there will be no guarantee of having the resources to afford HE. On the other hand, an underused aid system is inefficient because resources that do not suit the needs of low-income students may have had other uses instead.

Information also plays a central role in loans take up. As affluent students have more information, equity implications are evident. Nevertheless, this common assumption has been contested in some empirical research, for example, Booij, Leuven and Oosterbeek (2012) conducted a randomised trial and found no causal effect of information in take-up rates in the Netherlands. An additional problem relates the aid system complexity. More complex aid systems have countered the impact of aid on college enrolment and persistence since a they are less transparent so not everyone have a clear idea of the true availability of support, this affecting low-income students more critically (Dynarski and Scott-Clayton 2013).

Research on student aid in Chile is scarce. However, the sharp increase of resources to aid, the 2005 reform, the development of high-quality administrative information systems on aid recipients, and the fact that the government started collecting individual records of HE attendants in 2008, has contributed to configuring a research body assessing the impact of aid on HE outcomes. Solís (2013), using a regression discontinuity design found that the participation of students with loans was around 20 per cent higher for students just above cut-off point (PSU score) in comparison to those just below the threshold. Urzúa and Rau (2012) found a higher effect of CAE on persistence in comparison to other aid, as did Horn, Santelices and Catalán (2014)

and Santelices, Catalán, Krueger and Horn (2015). The results are consistent with other studies: a positive and significant effect of student aid on low-income students' access probabilities.

The above studies have taken full advantage of newly available good data and used quasi-experimental evaluation techniques. Further developments should be expected as data on mature cohorts become available and different administrative sources are linked. Nonetheless, there are no investigations on the effects of the 2005 reform.

Regretfully, it is not possible to use microdata to evaluate the impact of the 2005 reform as linking aid recipients and enrolment administrative records is only possible from 2008 onwards. Consequently, as I shall explain in section 4.4, I use a series of household surveys to estimate the impact of the 2005 reform on the probability of access to HE for low-income students.

4.4. Chilean HE Policy

4.4.1. Historical development

Until 1980, Chile's HE was made up of eight publicly funded and non-tuition charging universities, two of them were state owned and operated at a national scale, whereas the other were private owned but relied on public funding. The system was funded out of general taxation and resources were mostly allocated through funding formulas based on enrolment and historical criteria. Since 1980, when the military dictatorship carried out its reform, Chilean HE has experienced dramatic changes. The government introduced tuition fees and allowed the private sector to set up new universities and non-university HEIs. Other relevant institutional transformations have already been outlined in section 2.2 (Chapter 2).

The regime also changed HE funding structure by diversifying financial sources, passing from purely institutional block grant funding, according to historical considerations, institutional size and complexity, to a combination of block grants, student loans, private funding (through tuition payments) and a series of competitive

mechanisms to finance research activities. The government promised to transit from block grant funding to an even combination of block grants, student loans and competitive mechanisms. The 1980s crisis battered the Chilean economy and, consequently, the projected funding was not fulfilled.

The most relevant feature of HE policy in the 1990s was the expansion and consolidation of the system. Many new private universities were established between 1989 and 1991, because of a fear that the new democratic government (*Concertación* centre-left coalition) would undertake a major reform against the private sector involvement in HE (Uribe 2004). The first *Concertación* governments gave the greatest resources to CRUCH universities by increasing funding as well as creating new funding mechanisms to foster investment in infrastructure. Concerning student aid system, in 1994 the government set an income contingent loan system, instead of the mortgage-type operating since the early 1980s, and grants for poor background students[6].

Since the 1990s, an increasing share of governmental expenditure in HE has been funding student aid, reaching almost 40 per cent in 2014 of government contributions to HE (OECD 2014). A major milestone in student funding took place 2001, when aid started reaching students from private independent HEIs for the first time. In political terms, it made the case for extending student loans to private HEIs, especially considering that the reform would not have been endorsed by a left of centre government unless the evidence had been convincing. The reform's leitmotif was that it would make aid available based on students socioeconomic need instead of the HEI they attend: two siblings must receive the same support regardless HEIs they attend.

Enrolment steadily grew from 245,000 students in 1990 to 650.000 in 2006 and reached near 1.2 million students in 2015. Relevant causes for this quick expansion are the

[6] The first scholarship programme was created in 1991. Until then, the state did not use to provide any type of grants to HE students.

universalisation of secondary education[7] which resulted in low-income students completing secondary education; and a massive increase in student aid from 2000 onwards. Accordingly, and as expected, participation rates in HE of the poorest students have consistently increased in the last 15 years. In fact, participation rates of the poorest household income quintile rose from 4.5 per cent in 1990 to 17 per cent in 2006 and 27 per cent in 2013. In other words, for the very poorest young people the chances of being enrolled in HE in 2006 were twice those of 1990. Compared to 1990, a Chilean student from the poorest income quintile was five times more likely to attend HE. These figures on participation, however, differ throughout the system, as student aid eligibility used to be completely restricted to CRUCH universities, then the private sector served comparatively more affluent students as no government support was available.

Inequality of opportunities will diminish so long as student support systems are extended to private institutions. In practice, in the non-university segment, which concentrates an important proportion of low SES students, the opportunities of access have improved since the introduction of a grant favouring vocational courses in Technical Training Centres (*Centros de Formación Técnica*, CFTs) and Professional Institutes (*Institutos Profesionales*, IP) in 2001, and the reform in 2005.

4.4.2. The reform in 2005

In late 2005 the government undertook an ambitious reform of HE student funding. There were three major changes. The first was the introduction of a new student loan scheme originating in private financial institutions but with the state guarantee, CAE, which was conceived as a means of extending the support to all students attending accredited HEIs, this including CRUCH and private independent HEIs' students. Second, the government undertook a re-engineering of the then current student aid system for CRUCH universities. Third, the government set student aid packages consisting of a mixture of loan and grants, including a maintenance component for the

[7] For example, in 1990 a 26 per cent of 20-24 years old in the poorest income quintile had completed secondary education. In 2003, the proportion was 62 per cent. Something similar happened with next quintile (MIDEPLAN 2007).

poorest students. The poorer the student, the larger the grant component, and the smaller the loan component.

Nonetheless, the reform still considered two parallel aid systems: one for CRUCH universities and another for private independent HEIs. For CRUCH universities the government set a policy based on meeting funding goals rather than a strictly residual policy targeting the very poor students. In fact, the government would finance a 100 per cent of tuition fees to students coming from the lowest three household income quintiles through a combination of the pre-existing income contingent loan (FSCU) and grants. For the fourth and fifth, the percentage of aid decreased as income increased but CRUCH students were also able to use CAE to top up until 100 per cent of fees. This scheme was criticised because in practice students with the same socio-economic difficulties are treated differently according to the type of HEI attended, this meaning a lack of horizontal equity.

Private independent HEIs students, which were in practice marginalised from student aid before the reform, could use CAE to cover fees as well as scholarships for the very poor. However, aid is not as generous as CRUCH's, as CAE funded students are charged market interest rates, at fixed instalments instead of subsidised interest and income contingent repayment as in FSCU. CAE was able to cover the three poorest income quintiles and an increasing proportion of the fourth quintile. By 2011, CAE even reached some students from the richest quintile, whereas in 2014 was able to offer support to students from any socioeconomic background. Nonetheless, for the aims of this chapter, I shall consider the period 2005-2009 as the *evaluation period*, although in the discussion I shall also refer to further developments.

Despite the reform, there still exists a lack of horizontal equity, but the fact that the government started guaranteeing a determined level of support according to income actually represented a qualitative change in comparison to the residual pre-reform scenario. The main changes to Chilean HE aid rules are shown in Table 4.1.

Table 4.1. Aid packages before and after the reform by household income quintiles

		Pre-Reform		Post-Reform	
		Loans	Grants	Loan	Grants
Quintiles 1 and 2	CRUCH Universities	YES, FSCU if PSU score>=475	YES, if PSU score>=600	YES, FSCU if PSU score >=475 CAE (although marginal[8])	YES, if PSU score >=550 Maintenance Grant (food and cash)
	Private independent HEIs	NO	YES, some partial grants for vocational courses at non-universities HEIs	YES, CAE if PSU>=475	YES, a few but increasing subject to PSU>=550 Maintenance grant (food and cash). Grants for students taking technical-vocational courses
Quintile 3	CRUCH Universities	YES, FSCU if PSU score>=475	NO	YES, FSCU if PSU score >=475 CAE (although marginal)	NO
	Private independent HEIs	NO	NO	YES, CAE if PSU score>=475	NO
Quintile 4	CRUCH Universities	FSCU (partial and decreasing)	NO	Partial FSCU + CAE for some to top up fees	NO
	Private independent HEIs	NO	NO	CAE	NO
Quintile 5	CRUCH Universities	FSCU but negligible	NO	NO	NO
	Private independent HEIs	NO	NO	NO	NO

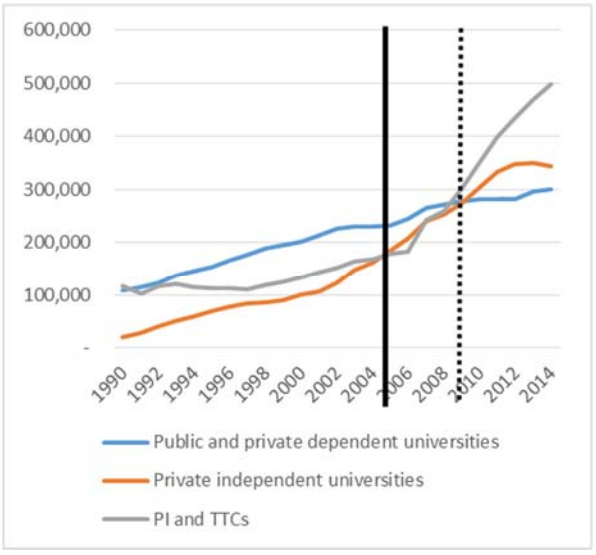
It is noticeable that CRUCH students are still better off in terms of support but Table 4.1 gives a very clear account of how the situation improved for anyone meeting the standard requirements. There is no doubt that the reform contributed to improving horizontal equity, as I shall show below, but HEIs were also able to respond to a

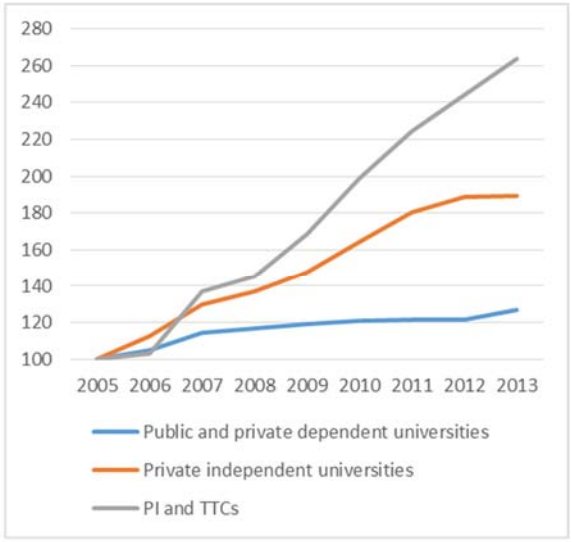
[8] Students from quintiles 1 and 2 at CRUCH universities should not have received CAE but a few were offered the loan because did not applied for other student aid. On the other hand, students in their second year or higher were able to apply for CAE. Although unlikely, there might have been unsupported students who applied for CAE and were offered the loan. The same applies for quintile 3.

demand shock by providing the extra places needed. This responsiveness, nevertheless, would not have taken place without the government setting CAE and sharply increasing the number of grant for students attending technical-vocational courses.

In fact, HE enrolment increased from 595,000 undergraduate students in 2005 to about 850,000 in 2009 (see the period between the solid and dotted vertical line in Graph 4.1). Within the period, private independent HEIs increased their enrolment by 47 per cent, whereas public and public subsidised HEIs increased their enrolment by only a 19 per cent, a rate well in line with which the government had agreed with public HEIs: an annual enrolment growth of up to 6 per cent. Non-university HEIs were the most affected sector: a steep increase of 67 per cent in enrolment took place (see Graph 4.2).

Graph 4.1.Undergraduate enrolment type of HEI Graph 4.2.Enrolment growth(2005=100)





The reform led to an increase in participation of low SES students but most new poor students were absorbed by the private sector, specifically the non-university sector, since the modest expansion of public and private subsidised HEIs and the fact that aid was already available for them. In terms of the government sponsored benefits allocated to students, the figures are clear enough in order to understand what the government intended with the policy (see Table 4.2).

Table 4.2. Number of grants awarded and loans taken up

	2005	2006	2007	2008	2009
Grants	43,061	50,532	62,992	86,998	123,144
CAE		21,263	54,447	90,722	148,416
FSCU	122,779	119,084	133,000	119,928	111,467

Source: Mineduc (2015)

The bulk of the reform consisted in implementing CAE and increasing the number of students awarded with grants. CRUCH universities income contingent loan did not benefit more students as students' numbers were frozen and even cut back by the end of the period. Nevertheless, the ratio FSCU/Tuition fees did improve as the government guaranteed full tuition coverage for students from the three poorest household income quintile. Between 2005 and 2009, the number of grants triplicated and, after just 4 years, CAE became the most important type of aid in terms of the number of students served.

The reform Chile undertook in 2005 is a rather unusual one. As the global trend has been to transfer the costs of HE from the state to individuals, Chilean HE, since the early 2000's and definitely with the reform, has gone in the opposite way as HE policy on access has heavily relied on public expenditure. Indeed, Chilean public expenditure in HE as a proportion of the total expenditure rose from 16 per cent 2005 to 35 per cent in 2012 (OECD 2008a, 2015b). The scale of the change, in a short period of seven years, is striking. Nevertheless, there are internal developments that may explain the uniqueness of this reform. As Chile used to have the lowest participation of public resources in HE funding in the world, despite the scope of the transformation, the post-reform proportion of public expenditure became similar to that of countries with highly privatised HE, such as South Korea and Japan. In addition to that, besides the government made a massive financial effort, the way loans are recorded in the national accounts system turns highly relevant. Whereas, in general, loans do not count as expenditure, in the case of the Chilean system, due to the design of the state guaranteed loan (CAE), at least a part of the resources count as expenditure. As the law prohibits the government to allocate loans in the credit system, the way it intervenes is through repurchasing the portfolio that commercial banks are not willing to finance, which are in turn financed by the government through banks by paying a

surcharge. Although portfolio repurchasing should not count as expenditure but assets, official statistics do consider repurchasing as expenditure, this way introducing a distortion. Nevertheless, the noticeable increase of grants and the fact that the government does devote public money to CAE reflect a major financial effort.

The peculiarity of the Chilean case lies in the radical reform undertaken by the military dictatorship, when major changes in HE structure, funding, and institutions were made at a very quick pace, which did not improve the situation for low income students due to the lack of resources but more critically to the high dropout and low completion rates in secondary education for low income students. The military left education -at all levels- underfunded by historical minimums without doing relevant investments. In 1990, most universities operated in the same conditions as 1973, and even with a lower expenditure per student in real terms.

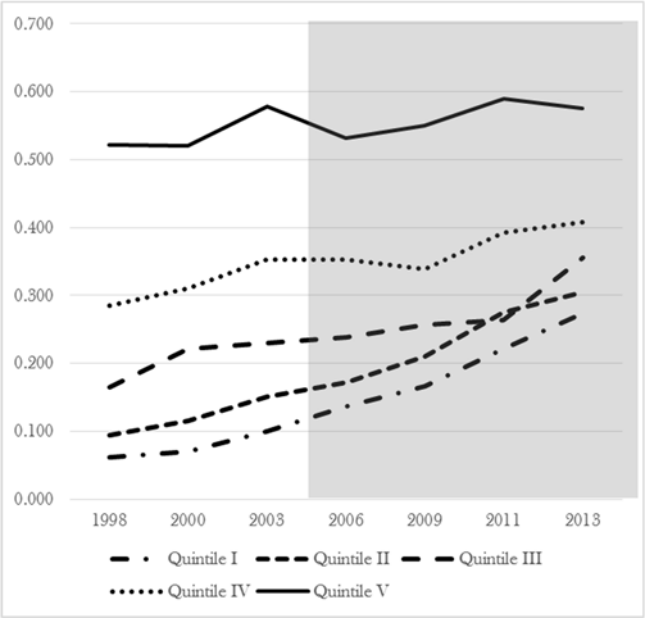
In some way, the public effort in HE might be seen as compensating the depreciated situation of HE (and education in general) in the early nineties rather than reversing the global trend towards more involvement of the private sector in funding and provision. Hence the reform might be seen as part of a structural process of putting things in place but without threatening the basis of Chilean HE: highly privatised, loosely regulated, with increasing public investment which comes to consolidate, with more state control, mixed provision model not very different from the one imposed by the military. Indeed, though important efforts have been made in institutional funding, the bulk of public investment has increasingly gone to demand-side funding.

The socioeconomic distribution of the Chilean student body, as expected, saw relevant changes during the evaluation period. Indeed, the net enrolment rate for quintile 1 increased from 0.10 to 0.16 between 2003 and 2009. This also happened with quintile 2 (from 0.15 to 0.21) and to a lesser extent with quintile 3 (from 0.23 to 0.26).

In Graph 4.3, a steep increase of enrolment rates for the poorest quintiles is noticeable in the shaded area, which corresponds to post-reform. I would like to highlight the extent to which access to HE has increased for low income students in a relatively

short period. From 2011 onwards, the rise of low-income students' participation was even sharper, due to policies implemented that year which are out of the scope of this chapter (as I shall explain in the next section).

Graph 4.3. Net enrolment rates by household income quintile 1998-2013



Source: CASEN survey

4.5.Data and Methods

4.5.1. Sources and data requirements

When the reform took place, there were neither exhaustive nor centralised administrative records of student aid but HEIs used to carry out specific processes. The reform also implied a major change in the way student aid was managed by both the government and HEIs. A proper system allowing linking enrolment and aid records was just developed in 2008 and is not useful in order to evaluate the impact of the 2005 reform since there is no pre-reform data. The Ministry of Education (MINEDUC) generates socioeconomic statistics based on administrative data as a by-product of processes of application to student aid, but not everyone does apply to financial support. Furthermore, administrative records of applicants do not contain exhaustive information on background characteristics, with the exception of income, which determines eligibility to support, as well as parental education.

To address this chapter's research question, I use a series of socioeconomic surveys, namely CASEN (*Encuesta de Caracterización Socioeconómica Nacional*, National Socioeconomic Characterisation Survey), which is the main socioeconomic survey applied in Chile. It is a cross-sectional survey applied every two or three years depending on the period but it is currently biannual. The series started in 1987 and the last version is that of 2015 which has not yet made public. I shall use the surveys from 2003, 2006 and 2009. I use 2003 as a pre-reform survey whereas I evaluate the policy impact in 2006 and 2009. I could have used surveys beyond 2009 but this raises some methodological concerns I shall explain in more depth later.

The surveys allow a precise identification of HE students and former students as well as a series of socioeconomic variables at the individual level as well as household and its member's information. As CASEN does not ask retrospective questions, a key issue is to record household variables vis-à-vis student variables. Surveys are consistent over time, there are no major issues when dealing with varying coding structures and questions because the core nuclei of questions has remained across time.

In order to assess the impact of the policy change in the probability of low SES student accessing HE, I need information on similar people at different points across time. This is guaranteed by the fact that the survey corresponds to a random sample of the Chilean population in every version so, in fact, I am using repeated cross-sections. Along with making sure that I have repeated measures of a similar population, it would also be necessary to establish both treatment and control groups and include enough control variables so that I can treat the data as coming from a natural experiment. I can also introduce control variables in order to test robustness as well as deal with the non-random assignment to treatment and control groups to a limited extent.

4.5.2. Data structure

CASEN is applied to a random, stratified and multistage households sample (geographical sample units have changed over time). Every present adult is interviewed face to face whereas the head of household responds on behalf of absent

adults and children. Sample sizes were of about 50 000 households in 1998 and 70 000 in 2013, this meaning a number of individuals in the sample have fluctuated between 210 000 to 250 000. The sample design is complex and has been subject to several changes with regards to improving the efficiency and allowing reliable estimates for smaller geographical areas but not in terms of target population. Therefore, it should not represent a threat to representativeness, since each cross-section is a sample of the same population. As I do not observe the same subjects over time, it is convenient to rule out a major change in CASEN sampling procedures so the samples in different cross-sections are likely to be 'identical' with each other.

Sampling design has changed over time so it is not possible to replicate the sample structure when data from several years are merged into a unique dataset because primary sampling units and strata are different so the use of estimates accounting for complex sample design (as the svy: prefix in STATA) cannot be used. Nevertheless, it is still possible to use sample weights but as if it were just a proportional sample.

The survey has always included modules on residents, income, employment, housing, health and education, while modules on assets, ICT penetration, energy usage and disability have also been introduced in subsequent surveys. CASEN also allows identifying each member of the household and their relationship with the head of household and, consequently, most of the relationships between each member of the household. Nevertheless, I made a series of decisions in order to get the most appropriate data and choose the correct individuals which are as follows.

The first measure I took concerning data was to consider 18-24 year-old individuals as my target population. This is consistent with the fact that in Chile students typically enter HE at the age of 18 and that the average duration of courses is 5 or 6 years. The age criterion also allows including both current students and former students (18-24 year-old graduates and dropouts) and is sensitive to the fact that student aid supports students on their first degrees so students over 24 years old, although a relevant proportion of the student body, are unlikely to be affected by the reform. As I am interested in measuring the impact of the student aid reform on access to HE,

the outcome variable is the probability of having *ever* entered HE. Therefore, I need to identify anyone that could have been supported with student aid so dropouts and graduates must also be included. I use this variable instead of enrolment because it is more inclusive. If I had used enrolment, people who finished HE –for example two or four-year programmes in IPs or CFTs- would have been excluded and treated as if they have never attended HE, even though they had studied their whole courses within the evaluation period. The same could have happened with students who defer or leave their courses.

The survey does contain a variable dealing with student aid and even the amount and coverage of tuition costs so it is possible to establish for an individual which kind of aid was awarded (FSCU, CAE or grants) and, at least in theory, what proportion of tuition fees student aid represents. Regretfully, the variable does not work well as there are major inconsistencies in the number of students being supported with grants, a crucial component of aid for the poorest students, in comparison to administrative data. In effect, FSCU is overestimated by 25 per cent in 2003, whereas grants are underestimated by 30.2 per cent in 2009. Although some degree of inconsistency between administrative sources and household surveys is not surprising, having overestimation is some years and underestimation in others seems problematic. Although CASEN has experienced several changes in sample design, when it comes to follow a specific population, sample variation is likely to affect aid recipient estimates rather than changes in sample design. For instance, CASEN used to overestimate HE enrolment by 15% in comparison to official statistics. However, from 2009, the estimated number of students has been much closer to the student population; for instance, in 2009 HE enrolment was 835,247, whereas CASEN estimated 826,345 students, very different from 2003 when the official figure was 542,516 but CASEN estimated 685,264, 26 per cent of overestimation.

The above is not due to pitfalls in CASEN but to changes in data collection and the generation of official statistics. In effect, until 2007, the government used to ask HEIs for enrolment statistics. In contrast, since 2008 the government asks for administrative data using a refined definition of enrolment, which also includes special part-time

programmes targeted to workers that did not use to capture in former data collection processes.

Table 4.3. CASEN estimates and administrative data by type of aid 2003 and 2009

	CASEN estimates		Administrative data		dif %	
Type of aid	2003	2009	2003	2009	2003	2009
FSCU	146,583	95,063	117,019	111,467	25.3%	-14.7%
CAE	-	132,006	-	148,416		-11.1%
Grants (+)	39,335	85,935	38,840	123,144	1.3%	-30.2%

+ Only considers grants supporting tuition fees payment
Sources: CASEN and MINEDUC

Anyway, one has to consider that the more disaggregated the population, as in disaggregated estimates of recipients, the less precise the estimates due to sample variation. Despite the above issues, which affect the estimates in terms of absolute numbers, there is still possible to get good estimates in terms of proportions by income quintiles, which is what I actually do in this chapter. Nevertheless, changes in the questionnaire might have been crucial as CASEN used to ask for state-supported loan and grants in general but currently the question inquires whether the student is the recipient of specific programmes.

The next decision regarding data was that, in order to get the family background and/or household characteristics, I restricted the population to those still living with at least one of their parents/tutors. At least in the Chilean context, this decision seems sensible. As shown in Table 4.4., over a 70 per cent of 18-24 year-old population lived with at least one parent/step parent and this has been constant over time. A similar approach has been used in Canada (Corak, Lipps, and Zhao 2004) although not without criticism as living arrangements may influence HE participation (Finnie and Usher 2006). Nevertheless, there is no chance to track people retrospectively because the survey does not have a longitudinal element and, more importantly, there is no other way of having data containing family and background characteristics.

Table 4.4. Number of 18-24 year-old by relationship with the head of household

Relationship with head of household	1998	2000	2003	2006	2009	2011	2013
Children or stepchildren	1,231,851	1,236,048	1,341,478	1,427,244	1,561,340	1,631,304	1,541,258
Other	537,079	506,875	554,598	550,077	575,882	621,418	610,424
Total 18-24	1,768,930	1,742,923	1,896,076	1,977,321	2,137,222	2,252,722	2,151,682
Proportion of children/stepchildren	0.696	0.709	0.708	0.722	0.731	0.724	0.716

There are some socio-demographic differences between those living with their parents and those who do not. Those living on their own live in partnerships and are employed, head of household and women in a higher proportion than those living with parents. This group is also more concentrated in the lowest income quintiles, with 46 per cent against 38 per cent of those living in the parental house in quintiles I and II. Statistics are shown in Table 4.5.

Table 4.5. Socio-demographic characteristics of 18-24 year-old with at least complete secondary education. 2009

	Not living at parental home	Living at parental home	Total
Marital status			
Married	0.10	0.01	0.03
In partnership	0.33	0.04	0.11
Other	0.58	0.95	0.85
Total	1.00	1.00	1.00
N ('000s)	509.3	1,462.9	1,972.3
Employment			
Employed	0.45	0.34	0.36
Unemployed	0.11	0.12	0.12
Economically Inactive	0.44	0.55	0.52
Total	1.00	1.00	1.00
N ('000s)	509.3	1,462.9	1,972.3
Income quintile			
1	0.20	0.17	0.18
2	0.26	0.21	0.23
3	0.22	0.22	0.22
4	0.18	0.21	0.20
5	0.15	0.18	0.17
Total	1.00	1.00	1.00
N ('000s)	506.6	1,462.9	1,969.5
Gender			
Male	0.45	0.51	0.49
Female	0.55	0.49	0.51
Total	1.00	1.00	1.00
N ('000s)	509.3	1,462.9	1,972.3

I have also introduced an additional restriction considering individuals with at least complete secondary education. This way, I avoid an important distortion as completion rates at the secondary level are higher for better off students so my target population consists only of students which at least have the formal qualifications required in HE so I do not penalise low-income students due to their lower secondary attainment.

The final dataset contains data considering 2003, 2006, and 2009 surveys, the total sample size reaching 27,600 records with 9,000 to 11,000 cases for each respective year. The dependent variable is a dummy taking the value of 1 if the subject has ever attended HE and 0 if not, which is henceforth referred to as *access*. There are many identification variables such as age, sex, marital status and whether or not in employment. Background characteristics include per capita household income quintiles, household size, the number of siblings, whether it is urban or rural, sex of the head of household, years of schooling of the head of household, mother's education attainment, and the head of household occupation (ISCO 2008 but aggregated).

4.5.3. Model identification

I shall start with a logistic regression approach to compare the probability of access to HE before and after the reform. One way of doing it would have been to run specific regressions for each year and compare the effect sizes of the variable measuring income and test whether the differences between the effects are statistically significant. However, this approach is not robust because there are unobserved variables such as school performance and quality of prior education, affecting both predictors and the outcome. Whether or not unobserved variables are fixed over time, the problem arises when effect sizes are considered. Comparing effect sizes of regression models including the same variables measured at different time periods is not equivalent to measuring policy impact but might be an approximation to assess the main objective of student aid: weakening the effect of socioeconomic background on the probability of access to HE. I shall provide a comparison in the results section but the estimates must be taken cautiously.

A more experimentalist approach would be to build a counterfactual simulating what would have happened post-reform had the conditions of student aid, the influence of social background and other relevant variables remained the same as in the pre-reform period. This is equivalent to build a forecasting model which is as follows.

To achieve this, I specify a logistic regression model given by the logistic function:

$$\Pr(Y = 1) = \frac{1}{1 + e^{-z}} \qquad [4.1],$$

where Y is the outcome variable and z is a linear combination

$$Z = \beta_0 + \boldsymbol{\gamma}\mathbf{X} \quad [4.2],$$

$\mathbf{X}$ is a vector of predictor variables. An interesting feature of the logistic model is that the exponential of $\boldsymbol{\gamma}$ can be interpreted as odd ratios, the ratio by which the probability of access to HE increases/decreases when $\mathbf{X}$ increases one unit or when the attribute takes the value of one, in the case of dummy variables. Values greater than 1 mean that probability increases, whereas the opposite happens with values lower than 1.

The counterfactual is simulated by regressing the probability of access on $\mathbf{X}$ in 2003. Next, these estimates are applied to 2006 and 2009 surveys and the probability of access is forecasted. The logic behind this is straightforward as the model forecasts a counterfactual: how the post-reform probabilities of access would have been had the effects of all variables in the model remained constant. After that, logistic regressions, containing the same variables, are run for each post-reform year (2006 and 2009) so the impact measure for each income quintile would be the difference between the average forecasted probability by income quintile and the actual mean predicted probabilities given by year-specific logistic models. I do not use sample probabilities for the comparison as the metric is different. The fitted probability is a continuous variable whereas the sample one is binary so standard errors are not useful to run

significance tests. The above technique is an approach to an experimental setting but lacks a control group so it may be subject to biases and inconsistencies affecting the estimates. The identification problem to solve is that the differences between forecasted and actual probabilities are not yet an approximation to the true causal effect of the policy. Comparing intra-quintiles probabilities would not establish the difference between affected and unaffected students which is the true causal effect of the policy to estimate. Although the above approach does build a counterfactual, the policy evaluation toolbox has always been used to compare groups which are (un)affected by the policy.

I specify a differences-in-differences (D-I-D) model in order to evaluate the impact of the student aid reform. The crucial step is to choose an adequate control group, unaffected by the policy and at the same time supports the common trend assumption, the key identification one in D-I-D models. Had not the reform taken place, the differences in access between control and treatment group would have been the same in pre and post reform period (the counterfactual). The most straightforward and intuitive control group, as it is easy to deduct from Table 4.1., would be to choose the highest quintile since it is not sensitive to the policy change (there is no aid for quintile 5 in pre and post reform years). As I have control variables, I could deal with pre-existent differences in household variables and personal characteristics.

Although there is no direct way of testing the key identifying assumption, it is clear from Table 4.5. that the pre-reform trajectory of quintile 5 is one of stabilisation of enrolment rate instead of the increasing trend observed in the poorer income quintiles. On the other hand, participation rates, the type of previous schooling -with a much higher incidence of private schools-, and PSU entry scores features quintile 5 as hugely different. On the one hand, there is a group showing enrolment rates similar to the developed world, whereas on the other extreme enrolment rates rather correspond to those of the third world.

Table 4.6. Net Enrolment Rates by Household Income Quintile

Year	I	II	III	IV	V
1998	0.062	0.094	0.165	0.285	0.521
2000	0.070	0.116	0.222	0.310	0.520
2003	0.100	0.150	0.229	0.352	0.578
2006	0.136	0.172	0.238	0.353	0.531
2009	0.166	0.210	0.256	0.338	0.550
2011	0.221	0.275	0.263	0.392	0.590
2013	0.274	0.305	0.355	0.408	0.575

Source: CASEN survey

On the basis above, I ruled out quintile 5 as a control group. The alternative is to choose groups that have been affected by the reform in a very different fashion. I have slightly modified Table 4.1 showing pre and post reform situations but instead of considering the two parallel aid systems, I compare the pre and post reform situation for the HE system as a whole. In Table 4.6, instead of explaining the rules and conditions for aid eligibility, I show how every income quintile is affected by the policy.

From both tables 4.1 and 4.7, it is easy to note that the poorest two income quintiles were affected in the same way, from partial to full access to loans and an increasing coverage of grants either those aimed at paying fees or the maintenance ones. Access to loan was partial because there was no loan programme for students attending private independent HEIs and access to loans for CRUCH students in quintiles I and II, although having a high coverage, was not guaranteed and also subject to budgetary constraints and specific issues affecting HEIs management

Table 4.7. Effect of the New Policy by Income Quintiles

	Pre-Reform		Post-Reform	
	Loans	Grants	Loan	Grants
Quintile 1 and 2	Partial coverage subject to PSU>=475	Low coverage, pay fees Subject to PSU>=600	Full coverage, subject to PSU>=475	Increasing coverage, fees and maintenance expenditures PSU>=550
Quintile 3	Partial coverage subject to PSU>=475	n/a	Full coverage subject to PSU>=475	n/a
Quintile 4	Partial coverage subject to PSU>=475	n/a	Partial coverage subject to PSU>=475	n/a
Quintile 5	n/a	n/a	n/a	n/a

The situation differs when comparing to quintile 3, which also became fully covered, subject to the same entry score requirements because quintile 3 students are not entitled to grants. Indeed, there was no CAE before the reform and the coverage for CRUCH students was partial and not guaranteed. Nevertheless, FSCU coverage for quintile 3 students was lower than that of quintiles 1 and 2 so this may represent a comparability issue when contrasting quintiles 1 and 2 vis-à-vis quintile 3 in the pre-reform scenario. However, as the bulk of the post-reform new loans corresponds to CAE, changes in FSCU, in terms of changing the probability of access to HE for low and middle-income quintiles, should have had a modest impact.

For quintile 4, there is partial CAE coverage because it is restricted by the budget allocated by the government to guarantee debts. After 2009, CAE has increased its coverage so has even been offered to any applicant of any socioeconomic condition. In the case of quintile 5, it has already been ruled out due to common trends assumption hardly holds.

Using quintile 3 as a control group is approximately equivalent to measuring the joint impact of (i) the extension of the grants programme through lowering the score requirement from 600 to 550 points (equivalent to half a standard deviation), (ii) the government offer in order to guarantee grants to students from the two poorest household quintiles, with the consequent substitution of FSCU with grants. Hence, rather than assessing the impact of the whole reform, I am evaluating the effect of the new grants policy.

The nature of the data features an opportunity for using quasi-experimental evaluation techniques. In fact regression discontinuity (RD) has been used to evaluate the impact of aid (van der Klaauw 2002; Melguizo et al. 2016; Solis 2013). Although it is not difficult to make a case for Regression Discontinuity (RD), the lack of data on PSU scores in CASEN survey does not allow taking advantages from changes in the assignment rule. Even though the reform introduced rules linking amounts of aid to income, not knowing PSU scores, a statutory requirement, would make the exercise misleading.

After choosing treatment and control groups, I then proceed to specify the evaluation technique which is the D-I-D model. Although D-I-D is a non-experimental technique whose main shortcoming is the lack of random assignment of treatment and control groups, it is one of the most common evaluation techniques when using observational data, as in this chapter. The technique has been widely used in a variety of topics in the social sciences and impact evaluation in the US (Dynarski 2003) and recently in the UK (Dearden, Fitzsimons, and Wyness 2014).

Equation [4.3] shows the D-I-D model as an OLS model.

$$Access = \beta_0 + \beta_1 Treatment + \beta_2 After + \beta_3 (Treatment \times After) \quad [4.3],$$

where the dummy *Access* is the dependent variable which takes the value of 1 if the person has ever attended HE, β_1 is the mean difference in probability for the treatment group, in this case quintiles 1 and 2, β_2 represents the mean difference in probability after the intervention, and β_3 is the effect of the interaction term, the D-I-D estimator, which captures the difference in access to HE over time for treatment group compared with the difference over time for the control group.

It is also possible to extend the model by introducing more covariates as control variables, as in [4.4]

$$Access = \beta_0 + \beta_1 Treatment + \beta_2 After + \beta_3 (Treatment \times After) + \gamma \boldsymbol{X} \quad [4.4]$$

Where $\boldsymbol{X}$ is a vector of individual, household and social background characteristics.

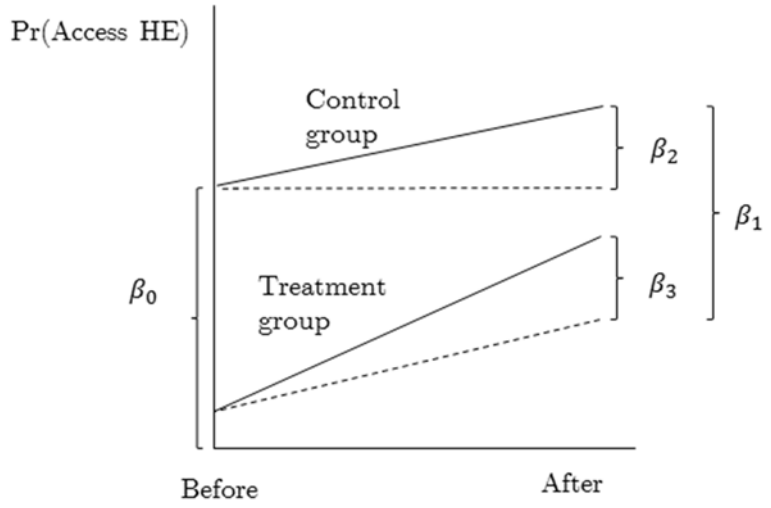
I do not use a non-linear model but the linear probability model due to simplicity and the fact that interaction effects in non-linear models such as probit and logit may not be meaningful and might even change sign or become statistically (non)significant according to the value of the other covariates in the model (Ai & Norton, 2003). Anyway, I provide D-I-D estimates using logit as a link function in Annex B.

A graphical representation of how D-I-D estimator works and what the coefficients measure is provided in Graph 4.4. Nevertheless, as I have more than one post-reform period, I use a slightly different specification to allow estimates for different post-reform periods to capture the D-I-D estimator, as equation [4.5] shows.

$$Access = \beta_0 + \beta_1 Treatment + \beta_2 Year1 + \beta_3 Year2 + \beta_4 (Treatment \times Year1) + \beta_5 (Treatment \times Year2) + \gamma \boldsymbol{X} \qquad [4.5]$$

Where β_1 is the treatment dummy, β_2 and β_3 are coefficients for each post-reform year, and β_4 and β_5 are the D-I-D estimators, which are the main parameters of interest in this chapter.

Graph 4.4. Graphical representation of the differences-in-differences estimator



The main advantages of D-I-D estimators are the following. Firstly, from a practical point of view D-I-D offers an experimental framework in the absence of randomisation so observational data, as either panel or repeated cross sections, is used within an experimental framework. Second, it controls for observables and unobservables time invariant variables in a similar fashion as fixed effects. Third, D-I-D deals with common time effects, such as the state of the economy and changes in the policy environment, across treatment and control groups.

Nevertheless, D-I-D is not robust to time-variant unobservables so even with good time-variant control variables there may still be confounders affecting the estimates. Moreover, there also needs to be data before and after the intervention, which is not always possible. Common trends, *i.e.* that differences between treatment and control group would have remained constant over time in the absence of the intervention, which is the D-I-D key identification assumption, is hard to hold because there is no direct way of testing that non-random assigned groups would be equally (un)affected by time effects. Nevertheless, I provide placebo estimates in the next section which, though indirectly, suggest that holding the common trend assumption would not be problematic.

However, the most important critique to studies using D-I-D is the lack of accuracy when computing standard errors. An influential paper (Bertrand, Duflo, and Mullainathan 2004) found a surprisingly high proportion of papers whose estimates were likely to be false positives due to the underestimation of the standard errors, with serial correlation being the main cause. Serial correlation is likely to affect estimates drawn from data containing many points of time (Bertrand et al. 2004 report a mean of 16.5 time period in the papers they reviewed), whereas collapsing the data in before and after periods is likely to lead to consistent standard errors, which is the case that applies to this chapter.

4.6.Results

4.6.1. Descriptive statistics

In this section, I present and explain the results obtained by using the evaluation techniques above described. Statistics on educational attainment are provided for the whole 18-24 year-old population and those living with at least one parent/guardian for the period 1998-2013 (Table 4.7). There is a noticeable increase in the proportion of students reporting either complete or incomplete HE, rising from 32 per cent in 2003 to 37 per cent in 2009. For those living with their parents, the proportion attaining HE is slightly higher, with 34 per cent in 2003 and 40 per cent in 2009. As I

pointed out in section 3, living arrangements are likely to influence HE participation, especially for students from the lowest socio-economic background.

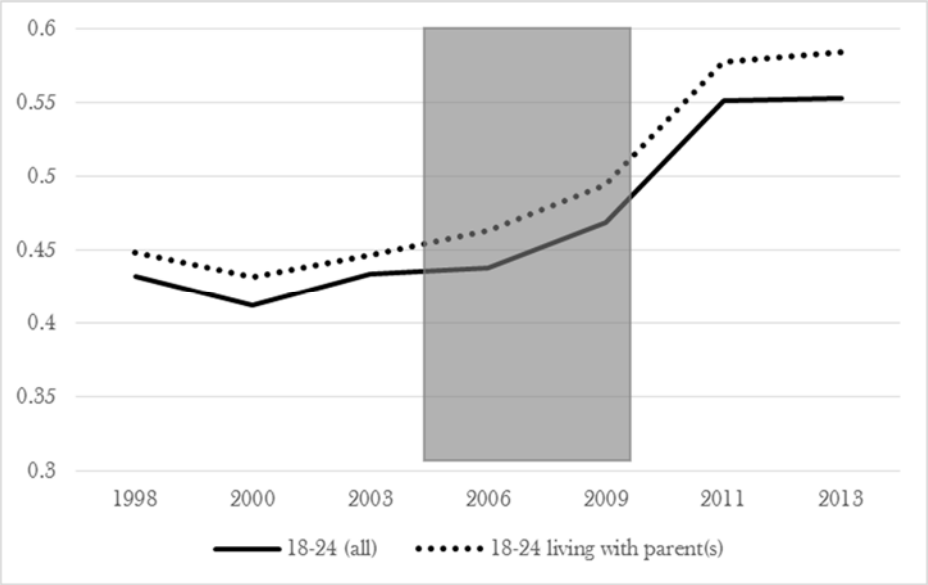
Table 4.8. Proportion (%) of 18-24 year-old by educational attainment (*living with at least one parent*)

Educational attainment	2000	2003	2006	2009	2011	2013
No formal education	0.005 *(0.004)*	0.009 *(0.009)*	0.010 *(0.01)*	0.011 *(0.011)*	0.009 *(0.009)*	0.010 *(0.01)*
Incomplete Primary	0.076 *(0.062)*	0.046 *(0.037)*	0.030 *(0.024)*	0.024 *(0.019)*	0.024 *(0.019)*	0.018 *(0.013)*
Complete Primary	0.068 *(0.054)*	0.056 *(0.04)*	0.044 *(0.033)*	0.042 *(0.032)*	0.033 *(0.025)*	0.038 *(0.028)*
Incomplete Secondary (academic)	0.142 *(0.127)*	0.118 *(0.105)*	0.113 *(0.102)*	0.110 *(0.096)*	0.163 *(0.159)*	0.099 *(0.086)*
Incomplete secondary (technical)	0.043 *(0.043)*	0.041 *(0.035)*	0.031 *(0.028)*	0.027 *(0.024)*	0.045 *(0.047)*	0.018 *(0.017)*
Complete Secondary (academic)	0.244 *(0.25)*	0.248 *(0.254)*	0.277 *(0.272)*	0.284 *(0.279)*	0.224 *(0.211)*	0.247 *(0.234)*
Complete secondary (technical)	0.143 *(0.147)*	0.164 *(0.172)*	0.156 *(0.156)*	0.134 *(0.133)*	0.101 *(0.1)*	0.116 *(0.114)*
Incomplete HE	0.241 *(0.275)*	0.286 *(0.316)*	0.304 *(0.337)*	0.324 *(0.36)*	0.358 *(0.384)*	0.385 *(0.425)*
Complete HE	0.029 *(0.027)*	0.030 *(0.028)*	0.034 *(0.033)*	0.045 *(0.043)*	0.041 *(0.042)*	0.065 *(0.065)*
Not informed	0.009 *(0.006)*	0.001 *(0)*	0.002 *(0)*	- *(0)*	- *(0)*	0.003 *(0.003)*
Total	1 *(1)*	1 *(1)*	1 *(1)*	1 *(1)*	1 *(1)*	1 *(1)*
N Total	1,742,923 *(1,236,048)*	1,896,076 *(1,341,478)*	1,977,321 *(1,428,244)*	2,137,222 *(1,561,340)*	2,252,722 *(1,631,304)*	2,151,682 *(1,541,248)*

If only people with at least secondary education were considered, the differences would be smaller. This gap has slightly widened over time as shown in Graph 4.5 but at a much slower pace than the gap in attainment shown in Table 4.8. Indeed, in 2003, 43.4 per cent of 18-24 years-old had ever attended HE whereas 44.7 of the subgroup

living with at least one parent did so. For 2006-2009 the gap is wider: 43.8 vs 46.6 in 2006 and 46.9 vs 49.4 per cent in 2009 (the shaded area corresponds to the evaluation period).

Graph 4.5. Access to HE for 18-24 year-old having at least completed secondary education



Descriptive statistics and cross tabulations of the dependent variable are presented for the years 2003, 2006 and 2009 in Table 4.9. It is easy to appreciate the increasing pattern of access to HE for the lowest income quintiles, whereas the higher –quintiles 4 and 5- remained constant. However, when the experimental and control groups are considered, both groups experienced an increase in access during the evaluation period. There were also changes in access when gender and area are considered. In effect, female access to HE increased by 7 percentage points whereas access for males remained constant; access improved in both urban and rural areas, where the increase was 9 percentage points. Children of unqualified workers and craftsmen and operators also increased their participation in HE at a faster rate than those from other occupational backgrounds. So did students whose mothers had only attained primary education. In summary, social groups traditionally seen as the most disadvantaged were the ones which made more gains in access to HE during the evaluation period.

Table 4.9. Access to HE by background variables. Weighted sample proportions

	2003		2006		2009	
	Mean	SD	Mean	SD	Mean	SD
Household income quintile						
1	0.263	0.440	0.328	0.470	0.369	0.482
2	0.333	0.471	0.394	0.489	0.437	0.496
3	0.425	0.494	0.446	0.497	0.471	0.499
4	0.594	0.491	0.585	0.493	0.601	0.490
5	0.830	0.375	0.791	0.407	0.818	0.386
Total	0.509	0.500	0.527	0.499	0.550	0.498
Gender						
Male	0.519	0.500	0.511	0.500	0.527	0.499
Female	0.499	0.500	0.544	0.498	0.573	0.495
Total	0.509	0.500	0.527	0.499	0.550	0.498
Area						
Urban	0.528	0.499	0.549	0.498	0.570	0.495
Rural	0.277	0.447	0.301	0.459	0.361	0.480
Total	0.509	0.500	0.527	0.499	0.550	0.498
Occupation Head of Household						
Managers, professionals and technicians	0.807	0.395	0.798	0.402	0.818	0.386
Clerks and sales	0.559	0.497	0.576	0.494	0.605	0.489
Farming and agriculture	0.296	0.457	0.356	0.479	0.338	0.473
Craftmen and operators	0.412	0.492	0.463	0.499	0.501	0.500
Unqualified workers	0.262	0.440	0.324	0.468	0.371	0.483
Unknown	0.425	0.494	0.472	0.499	0.463	0.499
Total	0.509	0.500	0.527	0.499	0.550	0.498
Mother's education						
Primary	0.245	0.430	0.303	0.459	0.334	0.472
Secondary	0.520	0.500	0.561	0.496	0.555	0.497
Higher	0.888	0.316	0.828	0.378	0.857	0.350
Unknown	0.486	0.500	0.470	0.499	0.483	0.500
Total	0.509	0.500	0.527	0.499	0.550	0.498
Total N	11,931		14,260		15,102	

4.6.2. Logistic regression estimates

I start with regression-based impact estimates, for which I provide regression coefficients in Table 4.10. I regress the access dummy on a series of predictor variables measuring individual, socioeconomic background, and household characteristics. Firstly, I estimate the model for 2003, then I use the same estimates to compute fitted probabilities by income quintile for years 2006 and 2009. Next, I estimate the model but this time using years 2006 and 2009. I compute fitted probabilities for 2006 and 2009 and compared them with those predicted by the 2003 model. The logic behind this is that I build a counterfactual by using 2003 estimates and computing fitted probabilities for post-reform years so the differences between fitted values are in fact a measure of impact for each income quintiles.

Table 4.10. Logistic models to predict access to HE

	2003	2006	2009
	exp(b)/se	exp(b)/se	exp(b)/se
Income quintile			
Quintile 2=1	1.503***	1.527***	1.588***
	(0.196)	(0.172)	(0.156)
Quintile 3=1	2.642***	2.124***	1.755***
	(0.342)	(0.233)	(0.174)
Quintile 4=1	3.447***	2.771***	2.490***
	(0.474)	(0.323)	(0.283)
Quintile 5=1	6.108***	4.219***	4.333***
	(1.027)	(0.596)	(0.612)
Age	6.099***	26.573***	22.107***
	(3.250)	(12.854)	(9.757)
Age2	0.963***	0.932***	0.936***
	(0.012)	(0.011)	(0.010)
Female=1	1.061	1.189**	1.221***
	(0.083)	(0.086)	(0.080)
Married=1	0.395***	0.379***	0.455***
	(0.099)	(0.072)	(0.077)
Employed==1	0.172***	0.194***	0.245***
	(0.015)	(0.015)	(0.019)
Head of household woman=1	0.887	0.814**	0.795***
	(0.093)	(0.076)	(0.067)
N Siblings	1.124*	1.209***	1.162***
	(0.070)	(0.071)	(0.060)
Household size	0.790***	0.743***	0.812***
	(0.036)	(0.035)	(0.031)
Urban=1	1.251**	1.311***	1.296**
	(0.118)	(0.100)	(0.135)
Mother's education			
Mother Secondary=1	1.538***	1.513***	1.350***
	(0.149)	(0.135)	(0.110)
Mother HE=1	3.777***	1.908***	2.692***
	(0.676)	(0.324)	(0.392)
Mother Unknown=1	0.591**	0.541***	0.674**
	(0.134)	(0.128)	(0.114)
Head of household occupation			
Clerks and sales=1	0.819	0.739**	1.066
	(0.118)	(0.104)	(0.143)
Farming and Agriculture=1	0.671**	0.672**	0.711**
	(0.117)	(0.113)	(0.123)
Craftmen and operators=1	0.630***	0.675***	0.744**
	(0.083)	(0.086)	(0.092)
Unqualified worker=1	0.591***	0.536***	0.781*
	(0.092)	(0.075)	(0.108)
Unknown=1	0.899	0.969	0.976
	(0.127)	(0.137)	(0.138)
Head of household years of schooling	1.101***	1.119***	1.104***
	(0.015)	(0.015)	(0.013)
Constant	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
N	11894	14231	15102
Pseudo R2	0.32	0.27	0.24
Log likelihood	-426720	-506012	-599199
Log likelihood null model	-625851	-696238	-788303

p<.10, ** p<.05, *** p<.01

The above models, in general, show that the goodness of fit worsens over time. This means that the association between the background variables reported in Table 4.10 and the outcome, access to HE, have weakened over time, this meaning that the outcome becomes less dependent on socioeconomic variables over time. For instance, relative to quintile 1, quintile 5 students were 6.1 times more likely to attend HE in 2003, whereas the odds ratio dropped to 4.3 in 2009.

Accordingly, children of unqualified workers were 41 per cent less likely to attend HE in 2003 but only 22 per cent less likely in 2009. Furthermore, females were 22 per cent more likely to access to HE than males in 2009, up from 6 per cent in 2003. These reflections indicate that the 2005 reform seems to have weakened the link between socioeconomic background and access to HE but I do not provide impact estimations in this instance.

By calculating fitted probabilities for each income quintiles I obtained Table 4.11, which compares access rates predicted with 2003 estimates and actual rates fitted for 2006 and 2009. Standard errors were obtained by calculating the ratio between predicted probabilities' standard deviations and the square root of cell frequencies.

Table 4.11. Impact estimates for income quintiles

	2006			2009		
	Predicted	Actual	Difference	Predicted	Actual	Difference
1	0.2606	0.3274	**0.0668**	0.2999	0.3688	**0.0689**
2	0.3155	0.3916	**0.0761**	0.3458	0.4369	**0.0911**
3	0.4009	0.4450	**0.0441**	0.4575	0.4706	0.0131
4	0.5495	0.5845	**0.0351**	0.5781	0.6007	**0.0226**
5	0.7920	0.7904	-0.0015	0.8131	0.8180	0.0049

Note: Differences in bold are statistically significant at p=.05

According to these estimates, the 2005 reform increased the probability of access to HE by 7 per cent for quintile 1 students in the evaluation period, 9 per cent for quintile 2 students, and also affected quintiles 3 and 4 but to a lesser extent. Those differences, although an approximation, are not measuring the true causal effect of the reform but at least confirm what the descriptive statistics showed. The above estimates try to mimic what would have happened in the absence of the reform, the counterfactual, but do not offer a comparison between affected and unaffected groups.

4.6.3. Differences-in-differences estimates

To address the above issue, I use the D-I-D model to estimate the causal effect of the reform. As I highlighted in the last section, there are two crucial steps. Firstly, I need to define treatment and control groups so the control group is ideally unaffected – or slightly affected - by the policy change. As I argued in the last section, although quintile 3 has also been affected by the policy, the main difference in comparison to quintiles 1 and 2 is that quintile 3 students were not eligible for grants. Apart from other issues presented in the last section, quintile 3 makes a sensible control group. Secondly, I need to provide evidence in order to test, although indirectly, the common trends assumption. I provide placebo estimates to indirectly test common trend.

I use the linear probability model (LPM) as a functional specification. In other words, I use an OLS model instead of link functions such as logit or probit. Besides its shortcomings, especially regarding the prediction of probabilities higher than 1 or lower than 0, LPM is not seen as too problematic when it comes to D-I-D models. There is no technical ground for treating the effects as linear but the alternative of using non-linear models may introduce more serious issues. In fact, a logit model, for instance, would estimate odds ratios. This is even more problematic in non-linear models when interaction effects are concerned, provided that the estimates would vary according to the estimation point at which the function is evaluated. Notwithstanding far from optimal, using the LPM in a D-I-D context seems to be a sensible approximation to the measurement of the true causal effect even though estimates should be taken with the necessary caution.

I provide summary statistics of the outcome and control variable in Table 4.12 for the treatment and the control group, broken down by year. The differences between both groups are straightforward with the control group presenting higher access, more employed people, proportionally less rural population, children of more educated mothers, and a higher proportion of heads of household with managerial and professional occupations. I also present sample means for treatment and control group from 1998 to 2013 in table 4.13.

Table 4.12. Summary statistics. Control and treatment groups

	Control group						Treatment group					
	2003		2006		2009		2003		2006		2009	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Access	0.425	0.494	0.446	0.497	0.471	0.499	0.304	0.460	0.366	0.482	0.408	0.491
Female=1	0.503	0.500	0.480	0.500	0.487	0.500	0.555	0.497	0.529	0.499	0.545	0.498
Age	21.166	1.759	21.139	1.832	21.095	1.835	20.985	1.797	20.954	1.854	20.836	1.805
Married=1	0.045	0.208	0.042	0.201	0.050	0.219	0.059	0.235	0.054	0.225	0.062	0.242
Employed=1	0.523	0.499	0.561	0.496	0.461	0.498	0.332	0.471	0.365	0.481	0.278	0.448
Female H.of HH=1	0.212	0.409	0.245	0.430	0.286	0.452	0.269	0.443	0.303	0.460	0.335	0.472
N siblings	2.583	1.116	2.512	1.054	2.390	0.975	2.847	1.196	2.725	1.240	2.597	1.087
HH size	4.990	1.667	4.812	1.450	4.661	1.450	5.383	1.759	5.189	1.803	4.982	1.658
Rural=1	0.085	0.278	0.098	0.297	0.104	0.306	0.124	0.329	0.150	0.357	0.144	0.351
Mother primary=1	0.330	0.470	0.369	0.483	0.315	0.465	0.460	0.498	0.455	0.498	0.415	0.493
Mother secondary=1	0.559	0.496	0.518	0.500	0.558	0.497	0.471	0.499	0.466	0.499	0.497	0.500
Mother HE=1	0.084	0.278	0.080	0.271	0.097	0.295	0.048	0.214	0.055	0.228	0.053	0.224
Mother unknown ed.	0.026	0.160	0.032	0.177	0.031	0.172	0.021	0.142	0.024	0.152	0.035	0.183
Managers, professionals and technicians=1	0.135	0.341	0.104	0.305	0.105	0.307	0.048	0.213	0.045	0.207	0.052	0.223
Clerks and sales=1	0.165	0.372	0.157	0.364	0.191	0.393	0.102	0.302	0.121	0.326	0.136	0.342
Farming and agriculture=1	0.051	0.219	0.050	0.218	0.040	0.196	0.064	0.245	0.057	0.232	0.048	0.214
Craftmen and operators=1	0.342	0.474	0.333	0.471	0.324	0.468	0.291	0.454	0.268	0.443	0.226	0.418
Unqualified workers=1	0.164	0.370	0.204	0.403	0.184	0.387	0.196	0.397	0.241	0.428	0.251	0.434
Unknown=1	0.144	0.351	0.152	0.359	0.156	0.363	0.299	0.458	0.269	0.443	0.287	0.452
Unweighted N	2,649		3,384		3,687		5,629		6,040		6,602	

Table 4.13. Access rate for control and treatment groups. 1998-2013

	Control Group	Treatment Group
1998	0.415	0.310
2000	0.451	0.319
2003	0.425	0.304
2006	0.446	0.366
2009	0.471	0.408
2011	0.481	0.462
2013	0.603	0.583

D-I-D estimates are presented in Table 4.14. I start with the basic model just containing treatment, time dummies and their interactions - the D-I-D estimators. There is no significant effect for the year 2006. The second model controls for individual level characteristics. After that, I control for household characteristics and the socioeconomic background controls are introduced in the last models. Estimates do not show a significant effect of the policy reform by 2006. This is completely sensible as it was the first year of implementation and the impact estimator may not have captured the effect of the 11 000 new grants (in comparison to 2003), mostly concentrated in the lowest quintiles, which makes the difference in how the policy affected treatment and control groups. The coefficients fluctuate between 2.3 and 4.2 percentage points and only one specification, that of household controls, showed a statistically significant effect but at just 0.10 level.

There are significant effects for the interaction between treatment group and year 2009. Coefficients are stable at around 6 percentage points. The magnitude of the impact estimator keeps upon introducing control variables, which means that D-I-D estimates were unlikely correlated with the error term, at least with the set of control variables I used. In summary, D-I-D estimates indicate that there is no significant effect of the reform of grants policy in 2006 but it did have a significant effect in 2009. Had I estimated the true causal effect, the changes in the grants policy would have increased the probability of access to HE for students from income quintiles 1 and 2 in 6 percentage points.

The size of the participation effect for the poorest students is massive and equivalent to an increase of 14 per cent for students from the poorest 40 per cent of households.

Table 4.14. Differences-in-differences estimates. OLS, LPM

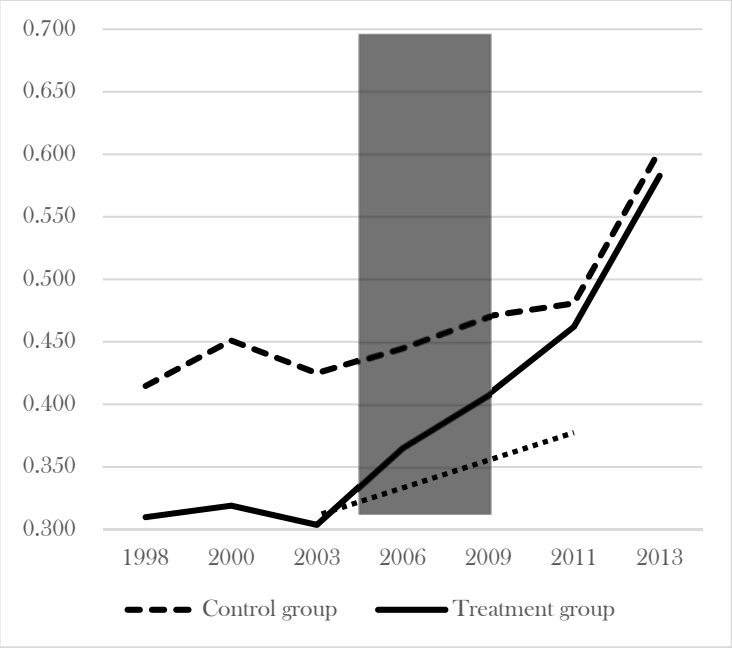
Variable	Base model b/se	Individual controls b/se	Household controls b/se	Background controls b/se
Treatment=1	-0.121***	-0.171***	-0.154***	-0.105***
	(0.020)	(0.018)	(0.018)	(0.017)
Year 2006=1	0.020	0.036*	0.032*	0.043**
	(0.021)	(0.019)	(0.019)	(0.018)
Year 2009=1	0.046**	0.033*	0.028	0.023
	(0.021)	(0.019)	(0.019)	(0.018)
Treatment=1*Year 2006=1	0.042	0.038	0.042*	0.026
	(0.027)	(0.025)	(0.025)	(0.023)
Treatment=1*Year 2009=1	0.059**	0.062**	0.060**	0.057**
	(0.026)	(0.024)	(0.024)	(0.023)
Age		0.416***	0.414***	0.408***
		(0.064)	(0.062)	(0.060)
Age^2		-0.009***	-0.009***	-0.009***
		(0.002)	(0.001)	(0.001)
Female=1		0.010	0.020**	0.023**
		(0.010)	(0.010)	(0.009)
Married=1		-0.251***	-0.153***	-0.151***
		(0.017)	(0.019)	(0.018)
Employed==1		-0.322***	-0.312***	-0.264***
		(0.010)	(0.010)	(0.010)
Head of household woman=1			-0.024**	-0.039***
			(0.011)	(0.011)
N Siblings			0.045***	0.023***
			(0.006)	(0.006)
Household size			-0.055***	-0.043***
			(0.004)	(0.004)
Urban=1			0.136***	0.039***
			(0.008)	(0.009)
Mother's education Secondary=1				0.095***
				(0.011)
Mother's education Higher=1				0.232***
				(0.024)
Mother's education Unknown=1				-0.058**
				(0.029)
H.of Household Clerks and sales=1				-0.060***
				(0.021)
H.of Household Farm and Agriculture=1				-0.116***
				(0.023)
H.of Household Craftsmen and operators=1				-0.108***
				(0.020)
H. of Household Unqualified worker=1				-0.123***
				(0.020)
H.of Household Unknown occupation=1				-0.072***
				(0.020)
Head of household years of schooling				0.018***
				(0.002)
Constant	0.425***	-4.169***	-4.145***	-4.198***
	(0.016)	(0.663)	(0.650)	(0.623)
N	27631	27628	27628	27590
R^2	0.012	0.134	0.156	0.219
Adjusted R^2	0.012	0.134	0.156	0.218

*p<.05, **p<.01, ***p<.001

Common trends checks

I show an approximation to the common trends assumption in Graph 4.6 that displays the probability of access to HE for treatment and control groups, pre and post reform, using CASEN surveys from 1998 to 2013. The shaded area refers to the evaluation period where it is straightforward to appreciate a change in comparison to the pre-reform trend, whereas the dotted line would represent the counterfactual if the common trend assumption held.

Graph 4.6. Probability of access to HE. Treatment and control group (sample means)



I report placebo estimates to indirectly assess the common trends assumption. To assess it, I run D-I-D for the pre reform period to establish whether the interaction terms are significant. This means testing whether the lines representing treatment and control group were in fact parallel before the intervention took place. Placebo estimates are reported in Table 4.15. I use the full model including individual, household and background controls. I found no significant interaction effects at p=.05 although the interaction between treatment and year 1998 is slightly significant at p=.10

Table 4.15. Placebo estimates for D-I-D estimates

Variables	Full model
	b/se
Treatment	-0.109***
	(0.017)
Year=1998	-0.006
	(0.020)
Year=2000	0.002
	(0.022)
Treatment=1*Year=1998	0.044*
	(0.026)
Treatment=1*Year=2000	0.021
	(0.028)
Individual controls	YES
Household controls	YES
Background controls	YES
N	18787
R2	0.236
Adjusted R2	0.235

* p<.10, **p<.05, *** p<.01

4.6.4. Discussion

I have aimed at measuring the true causal effect of the 2005 reform throughout the chapter. My estimates report that there is a significant effect of changes in the grants programme which were not yet reflected in 2006, the first year of implementation, but in 2009. The impact measure is about 6 percentage points. However, there are still issues with the data that may affect the estimates that I discuss now.

One crucial issue is that unobserved variables affecting HE participation may have changed differently for control and treatment groups, such as school achievement. In effect, Chile had already experienced a performance improvement as measured by PISA reading test. In effect, whereas Chilean students scored on average 410 points in 2000, scores reached 449 points in 2009. In 2000, 42 per cent of 15 year-old students performed at PISA level 1 or lower (the most basic level), while in 2009 just 30 per cent scored at level 1. The proportion of students in the lowest performance level fell from 16 per cent to 9 per cent in the highest socioeconomic groups but also fell from 73 per cent to 53 per cent in the lowest socioeconomic group (MINEDUC 2011). In absolute terms, what happened is that low-income students improved faster than higher income students so poor students might take advantage from an enhanced aid system.

Besides differential improvements in PISA, according to socioeconomic level, the violation of the common trends assumption does not seem plausible. In fact, socioeconomic gaps in entry test scores (PSU) have not narrowed. Neither have gaps in the national assessment test (SIMCE). The key identification assumption of D-I-D is unlikely to be under threat.

Anyway, having controlling for observable variables and, to some extent, ruled out differential changes in achievement as a threat to common trends, it is not enough guarantee that the estimates are unbiased. There still remain other unobserved variables that can alter the results as for instance those dealing with the structure of incentives the reform put in place. Amongst these, higher motivation of low income students due to more support being available for them, an anticipation effect since CAE started being discussed in the Congress three years earlier, as well as the fact that the government sought, beyond supporting more students, to guarantee a relevant amount of money according to household income, may have influenced poor student in a very different fashion compared to better-off students. No research has been conducted in that area but the way incentives are designed -and the way students respond to them- might be crucial for my research.

4.7.Final Remarks

Besides the caveats, this evaluation exercise represents a step forward for the study of access to HE in Chile. At least for the short period this chapter has covered, background variables do seem to have become weaker predictors of access to HE. Despite the methodological shortcomings, the evidence gathered in this chapter does suggest that the lowest income group was noticeably benefitted with the grants policy change. The effect of the reform on access rates for low SES students and its effectiveness in closing SES gaps in participation is massive.

An emerging trend of studies linking student aid to educational outcomes is taking place on Chilean HE research. The development of information systems and the accumulation of good quality and rich administrative records, alongside the continued commitment to improving and funding more generous student aid from the

government, will trigger new research. However, this chapter remains as the only research piece that has addressed the 2005 reform so far.

The relevance of 2005 reform is unquestionable: apart from the sharp funding increase, the reform set a new basis for student aid. Firstly, income determines what level of aid would be available to students and the government is committed to guaranteeing a type of support and aid amount instead of the residual approach of the previous system. New developments in student aid policy in Chile have happened on the 2005 reform basis. Expanding income thresholds for loans and grants, introducing an income sensitive repayment structure and a subsidised interest rate for CAE are policies that were introduced in 2011 as a 'natural' continuation of 2005 reform, as well as a response to massive student demonstrations. However, measures taken in 2011 had already been widely discussed since the 2005 reform took place.

The estimates I provided throughout the chapter support the fact that the reform of student aid did favour the poorest students. D-I-D estimates showed that changing the grants policy had an impact of 6 percentage points on the probability of control group students entering HE, between 2003 and 2009. The magnitude of the effect of grants is to be regarded as an important achievement. Considering that the increase in participation for quintiles 1 and 2 was around 10 percentage points between 2003 and 2009, an impact measure of 6 percentage points is certainly relevant. Despite concerns that I am not picking up a 'pure' causal effect of the policy intervention, the use of several approaches to evaluating policy impact proved fairly consistent.

Supply-side changes, which have not been the focus of this chapter, are worth considering as they may confound the reform effect. For instance, one may argue that poor students willing to attend HE were not able to make it due to, for instance, some local HEIs offering limited places. If this were the case, the participation effect may be due to increased supply of places rather than the effect of more governmental support. This seems unlikely in the Chilean case given fee levels as a proportion of household income. In effect, considering that cheapest courses fees -mostly technical- were around US$ 1,800 in 2009, it was equivalent to a quarter of the annual household

income in quintile 2. This would be even more unlikely for someone willing to study an average five-year professional programme (Business, Law or Psychology at an average public university) that should have paid US$4,500 in 2009, which is equivalent to a 62.5 per cent of a quintile 2 annual household income. A high tuition fees/costs of instruction ratio makes very difficult for low income families to pay tuition fees without governmental support. Hence, it is fairly safe to maintain that the effects from my estimates are a sensible approximation to the true causal effect of the reform.

Reaching equity in HE is a real challenge to policy makers, while this chapter has shown that devising appropriate student support mechanisms may improve the situation, there is a long way to meet the goal of 'two people with the same ability and preferences should receive the same education'. Apart from household and family influences, there are a number of institutional and social factors affecting the chances of the poor to reach HE, such as admissions procedures, HEIs' recruitment policies, school practices regarding HE access, as well as the influences of peer groups, teachers and even geographical barriers. The lesson here is that policy makers should address equity issues in HE *beyond* the scope of HE policy. Importing the wider equity agenda to HE without addressing inequalities in the process earlier is a disproportionate and ineffective burden for HE policy.

In order to reach equity, policies need also to look at improving persistence, completion and results in the labour market. Motivations, expectations, the student experience, and the institutional factors are all relevant processes which policy should address. In the next chapter, I shall also take advantages of the 2005 reform to address the issue of aid composition and its influence on persistence, completion and dropout.

5. HOW DO STUDENT LOAN CHARACTERISTICS AFFECT COMPLETION AND DROPOUT IN HIGHER EDUCATION? A MULTILEVEL MULTINOMIAL DISCRETE SURVIVAL APPROACH

5.1.Aims and Research Questions

Student loans are an increasingly popular device to help low SES students afford tuition fees and/or living expenses in HE. Financial aid has contributed to either improve access (Canton and Blom 2004; Chapman and Ryan 2005; Dynarski 2003; St John and Noell 1989) and persistence in HE (Chen and DesJardins 2008, 2010; Dowd and Coury 2006). On the other hand, research in persistence and dropout also shows that students who drop out have higher risks of defaulting students loans than those completing their courses (Dynarski 1994; Hillman 2014; Woo 2002).

I shall argue that, as students progress with their studies, the debt burden increases so that they are compelled to persist and/or finish their courses on time, thus preventing either over-indebtedness or repaying a loan for an unfinished course. Nonetheless, the effectivity of these incentives may vary depending on loans characteristics. Students may borrow from several sources such as the government, as

in most income contingent/ deferred payment schemes; HEIs, as in some loan schemes in the US; and the private sector, by either contracting conventional bank loans with a private collateral or state guaranteed loans.

The main hypothesis behind this chapter is that different credit schemes imply a variable levels of harshness, which may trigger a variety of behavioural responses. For instance, students who find out they lack academic skills or simply do not like the course they chose have to make a tough decision, especially if the course is financed by a loan: to persist until the studies are finished or drop out. By persisting, the student avoids being indebted for a 'no-good' as there are no half-engineers or three-quarter doctors, while by dropping out they can stop the debt growing, but at the cost of wasting the investment. Moreover, when having to decide between persisting and dropping out, students weigh, amongst other factors, their academic fit, academic/professional preferences, income prospects, the value of the credentials, and critically, their ability to repay loans. Soft loans may make easier to drop out as the consequences of defaulting are perceived as not severe in the medium/long term. Indeed, not repaying a government provided loan may be seen as less problematic than not repaying a bank loan because the latter would usually imply bad credit records and may affect students' future more severely.

The aim of this chapter is to investigate how student loan characteristics affect study completion and dropout, having controlled for a number of household, institutional and individual characteristics. Three questions will be addressed in this chapter: *(i) How does the debt structure (amount and composition) affect the probabilities of dropping out/ course completion? (ii) Do bank or harsher loans increase the probability of completing studies (or deter students from dropping out) more effectively than other student aid mechanisms? (iii) Does loan structure have differential effects for certain type of students (income levels, prior achievement, or type of HEI)?*
To answer these questions, I carry out a survival analysis of Chilean university students participating in CAE, whose repayment process operates through commercial banks. I take advantage of a policy failure occurred during the first year of the programme implementation (2006). The failure consisted in offering a loan to rich

students instead of the poor due to a random software failure. The government was forced to increase the budget so that poor students were also offered loans. Take up ratios were similar for different income groups so the failure features a unique opportunity to study the entire socioeconomic distribution being supported with aid.

The implication of the failure for this chapter is that it allows me to observe students from the whole socioeconomic distribution. Had the failure not occurred, I would have been able to get information for only the poorest students. The failure, in consequence, does not provide an experimental setting but allows observing students that would not have been supported with the loan had the assignment procedure been applied correctly.

The chapter is organised as follows. First, I discuss the main theoretical and disciplinary approaches in the literature on persistence and dropout (section 5.2). After that, in section 5.3, I describe the Chilean student aid system, with an emphasis on the crucial role it has played in providing opportunities to the less well-off students over the last two decades. Next, in section 5.4, I describe the dataset and its main variables, discuss the methods commonly used in empirical research, with a special emphasis on survival analysis and random effects (frailty) models methods, and data constraints. Then, I specify multinomial logistic models for university students entering five-year undergraduate courses in 2006 and supported by CAE loans. I analyse the results in two stages: (i) the effects of debt-related variables on the outcome (completion/dropout) for five-year course entrants and (ii) the effects of debt-related variables for specific groups by interacting variables. I also introduce random effects as a means of dealing with unobserved heterogeneity (section 5.4). The chapter finishes with the main conclusions and issues for future research and policy agendas (section 5.5).

5.2.State of Art

The effect of student aid on access of low SES students, minorities, and non-traditional students to HE has been widely studied (Bettinger 2004; Canton and Blom 2004; Chapman and Ryan 2005; Dynarski 2003; Finnie and Usher 2006; Johnstone 2004; St

John and Noell 1989). Most investigations have shown a positive effect of aid on access to HE and some of them have suggested specific policy mechanisms to improve equity of access beyond overcoming financial barriers (Finnie 2012).

Interestingly, the study of dropout/persistence processes goes far beyond socio-economic and background variables as psychological, economic, social, institutional, and academic variables are regarded as affecting the process. Following Cabrera, Nora and Castañeda (1992), two types of research on dropout/ persistence can be identified[9]. Firstly, the literature trying to understand the structural processes underlying dropout processes. This research seeks to discuss/review theoretical frameworks and conceptual approaches, along with testing empirically the adequacy of particular theories. On the other hand, there is a body of research focusing on measuring the effects of a series of variables, critically student aid, on dropout and persistence.

Theoretical approaches to dropout from HE have been developed since the 1970s, especially in the US. The main conceptual developments are the integration model (Tinto 1975, 2010), the attrition model, developed by Bean (1980, 1985) and the 'choice' model. These approaches to persistence and dropout arise from a variety of disciplines such as psychology, sociology and economics.

The first theoretical framework, that of the integration, pays attention to social and academic integration levels and focuses on students adaptation to the academic environment and variables which boost it, such as SES, personal attributes, family context and prior academic performance. Tinto's model features a variation of Durkheimian sociology which emphasises social integration and normative components (Spady 1970), where dropout behaviour is treated in a similar way as Durkheim dealt with suicide (Durkheim 2006). The integration model relates dropout to the mismatch between students and institutions, where factors explaining persistence such as motivation, students' skills, and academic and social characteristics of HEIs model and, at the same time, provide feedback to the student commitment to their academic

[9] An illustrative synthesis of earlier theoretical developments may be found in Bean (1982).

goals and studies. The more important academic goals and commitment are, the higher the probability of persisting (Cabrera, Nora, and Castaneda 1993). The integration model has sometimes been regarded as the most influential and the one which more empirical research has produced (Cabrera et al. 1993; Donoso and Schiefelbein 2007).

The second model -that of Bean (1980, 1981)-, introduces external factors such as job opportunities as well as considers the intention of persisting/dropping out as the main predictor. This model incorporates non-cognitive factors such as attitudes, motivation and interests as well as environmental and organisational factors. In fact, Bean and Vesper (1990) suggested that these factors along with family approval play a central role (see also Bean 1982:27). Nevertheless, Bean's major progress in the field is the fact that he intended to model the intention to leave, which has a direct impact on institutional research.

The third model establishes a connection between student's academic choices and persistence. Dropout follows a three stages process: (i) socio-economic factors and academic skills model the predisposition of attending the university; (ii) students balances costs and benefits of attending a specific HEI; (iii) the academic experience models perceptions on economic and non-economic benefits of persisting. Here student aid is important because it affects costs that student should face. Stratton, O' Toole and Wetzel (2008) state that dropout is a rational response to the changes of the probabilities of getting a degree and/or costs and benefits associated with it.

Cabrera et al (1993) point out that an important part of empirical research is based on integration and attrition models. Results show that both theories complement each other and that factors external to students are more important in comparison to what the integration model may predict. By integrating theoretical views, Stratton et al (2008) suggest that a major weakness in the literature is the fact that there has not been a clear distinction between dropping out and leave HE temporarily (stopping out). I would agree with this view but also argue that studying stopout relies too much on the availability of quality data covering periods long enough to make sure that stop outs can be properly detected. More crucial, however, is the fact that stopout is

difficult to operationalise as there is not a precise definition of it so it is keen to definitions *ad hoc* to the available data-

Chen and St. John (2011) concluded that integration is one of the most stable predictors of persistence. Students attending selective institutions show lower dropout rates, student aid coverage as a proportion of tuition fees has a significant impact and, in consequence, there should be a coordination between policies affecting the level of fees and student aid packages. Nevertheless, the authors concluded that non-need-based aid had a smaller effect on persistence than need-based aid. This relates to the incentive structures behind the funding instruments and their related behavioural response.

In relation to research on the effect of student aid on dropout rates, nevertheless, results are mixed rather than concluding. Some studies find significant effects on persistence, whereas other examine the importance of matching aid policies according to student characteristics and the temporal dimension of the educational process. However, there are also investigations finding no effects of student aid in dropout/persistence.

For instance, Chen and Desjardins (2008) noted that research carried out on the basis of dominant theoretical frameworks has not done important contributions as it lacks a focus on the interaction of student aid and income levels when explaining dropout. For this reason, they used a longitudinal approach and estimated a coefficient measuring the interaction between student aid and income level. They found heterogeneous effects of aid on dropout according to income level, the most important effect being found for low and middle-income students. In another study, the same authors arrived at similar conclusions but this time by testing differential effects by racial and ethnic characteristics (Chen and DesJardins 2010).

On the other hand, Stratton et al (2008) suggest that getting a job at the university and being granted a scholarship is associated with a lower probability of dropping out in comparison to those receiving loans.

Other research pieces have tested whether aid packages design leads to differential outcomes. Herzog (2005) points out that the return of student aid is much less important than the student academic experience, but at the same time, the relevance of the level of aid in relation to fees is well established: middle-income students with unmet economic needs are the keenest to drop out, while students with good academic skills show a higher probability of moving to other HEI. Aid structure has also been studied together with timing –temporal effects-, i.e. how aid affects departure decisions in different stages of the educational process (DesJardins, Ahlburg, and McCall 2002b). They found that moving from loans to scholarships, as Princeton University did, had a positive effect on persistence. Other studies obtained important conclusions for policymaking as temporal effects may help package aid by both differentiating aids according to how long the student has been enrolled and targeting aid packages to those at higher risk of leaving before graduating (DesJardins, Ahlburg, and McCall 2002a).

Nonetheless, studies such as Dowd and Coury's (2006), which focused on Community Colleges in the US, found no effect of student loans on persistence and warned that even though allocating resources through loans seems to be correct conceptually, the evidence they found contests this assumption.

This disparity of results is due to the scope -institutional, state and national levels-, data limitations, and the lack of consistent theories. Interestingly, the main theory frameworks are contended by empirical research but the validity of attrition and integration models, the most comprehensive ones, have hardly been questioned.

In the particular case of Chile, research on the subject is recent and underdeveloped. However, there is a comprehensive theoretical discussion (Díaz 2008; Donoso and Schiefelbein 2007; Himmel 2002). In turn, earlier empirical research has measured neither factors affecting dropout nor the impact of student aid but rather qualitative approaches (De los Ríos y Canales, 2007; Centro de Microdatos, 2008). There are also studies measuring dropout globally through small samples (Centro de Microdatos,

2008) or by developing proxies on the basis of official statistics (González and Uribe 2003; Uribe 2004).

Only two pioneering studies have made use of administrative records to link persistence/ dropout with student aid. Urzúa and Rau (2012) used administrative records from CAE and HE enrolment database. They found 2006's CAE students outperformed non-CAE students with lower dropout rates. That paper uses a multiple choice approach and only tracks students one year after the first enrolment and after 5 years in order to link the data with unemployment insurance records to obtain data on wages (all formal employees contribute to unemployment insurance based on their earnings). Nevertheless, there are many flaws regarding some particular characteristics of the programme, the effects of the administrative mistake and the special conditions that HEIs established for their first CAE student cohort (I shall return to this in the data and methods section). The other study (Horn et al. 2014), investigated the differential effect between state-provided and institutionally provided aid. Nevertheless, this study used data from the most selective Chilean university, the Catholic University of Chile (PUC), thus the variability of some parameters of interest, such as prior achievement, is highly restricted so that the estimates might be problematic. Additionally, this restriction threatens the generalisability of the results. Research discussed above focuses on measuring the effects of aid and/or aid packages targeted to specific populations. The results confirm the positive effect of aid on persistence as well as the effectiveness of aid packages in targeting specific groups, mostly low SES students. Nonetheless, there are not direct comparisons between financial instruments introducing different incentive structures that may trigger a variety of behavioural responses according to the level of harshness implied.

5.3. The Chilean Student Aid System

5.3.1. General features

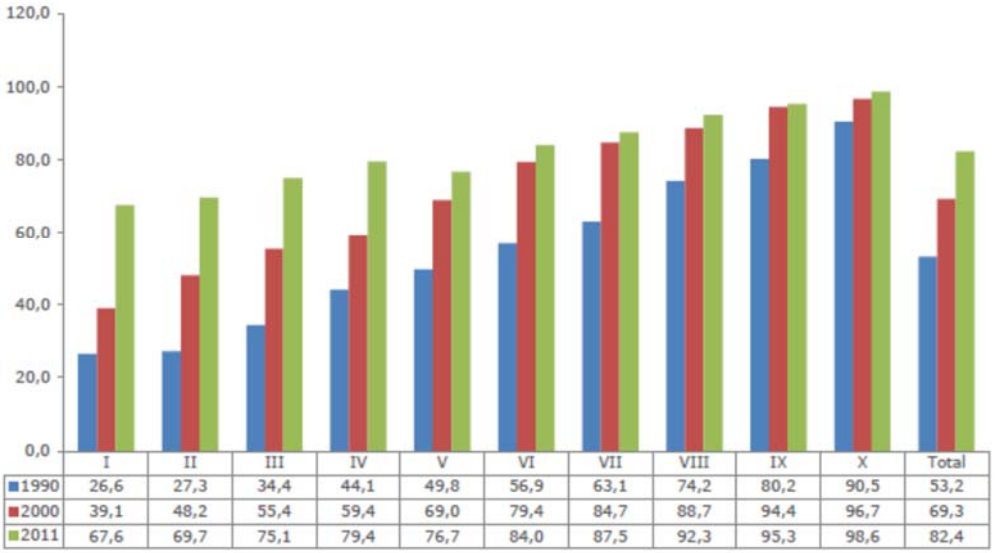
Currently, 1.2 million students attend HE in Chile at the undergraduate level, this representing four times the total enrolment in 1990. This is well in line with global

trends. In fact, in the last 20 years, HE enrolment has been growing by 5 per cent a globally (British Council 2012; OECD 2008b).

Even though the tuition fees introduction in Chile dates back to 1981, it is a rather recent trend in advanced countries. Cost sharing policies have been accompanied by financial mechanisms allowing low SES students to access HE (Chapman 2006; Johnstone 2004; Teixeira et al. 2008). In the Chilean case, two major processes explain the steady increase in HE participation, which took place since the early 2000s: (i) the universalisation of the coverage of secondary education and (ii) an important governmental effort to enhance demand-side funding through student aid (Armanet and Uribe 2005; MDS 2012).

Reaching universal coverage in secondary education had an obvious impact on the demand for HE. The secondary education attainment of the poorest decile's 20-24 year-old population rose from 39 per cent to 68 per cent between 2000 and 2011, while in the next decile the figures improved from 48 percent to 69 per cent in the same period, as shown in Graph 5.1. Poorest young people attaining secondary education became the most important source of demand for HE and most of its growth prospects relied on it.

Graph 5.1. Share of the 20-24-years old population with at least complete secondary education by household income decile.



	I	II	III	IV	V	VI	VII	VIII	IX	X	Total
1990	26,6	27,3	34,4	44,1	49,8	56,9	63,1	74,2	80,2	90,5	53,2
2000	39,1	48,2	55,4	59,4	69,0	79,4	84,7	88,7	94,4	96,7	69,3
2011	67,6	69,7	75,1	79,4	76,7	84,0	87,5	92,3	95,3	98,6	82,4

Source: MDS (2012)

Reaching universal coverage in secondary education had an obvious impact on the demand for HE. The secondary education attainment of the poorest decile's 20-24 year-old population rose from 39 per cent to 68 per cent between 2000 and 2011, while in the next decile the figures improved from 48 percent to 69 per cent in the same period, as shown in Graph 5.1. Poorest young people attaining secondary education became the most important source of demand for HE and most of its growth prospects relied on it.

5.3.2. The Chilean aid system. Policies and developments

Chilean student aid system is made up of a variety of programmes, including loans, grants and maintenance allowances. Students apply to the whole student aid system before the university applications process takes place. Student aid is allocated to students upon PSU scores and university application outcomes are released. Aid is granted on the basis of academic performance, namely GPA, PSU scores, and SES. Once PSU and admission results are published, students are required to enrol in HEIs but currently they know in advance which kind of support is available to their choices. It is important to note that, by far, the main selection criteria used by HEIs is PSU score, this meaning that having good test results assures a place in the most selective

universities. For the twenty-five public and private dependent universities (CRUCH) selection relies entirely on PSU scores and secondary GPAs whilst an increasing number of private independent universities do use PSU as the main selection criteria. Non-selective universities, however, do not pay much attention to PSU scores with selection purposes. In practice, for CRUCH universities, the most selective HEIs, students are eligible to income contingent loans (FSCU) and a generous grant scheme, whereas, for independent private HEIs students, aid relies mainly on CAE. The next table (5.1) shows eligibility rules for student aid.

Table 5.1. Student aid eligibility criteria in for new entrants (as at 2006)

Type of HEI	SES (household income quintile)	Academic requirements	Benefit package	HEIs eligibility
CRUCH universities (public and private dependent)	I and II	PSU>=550	Grant to cover 100% of tuition fees. Food and maintenance allowance	None
		550>PSU>=475	FSCU up to 100% of fees Food allowance	None
	III	PSU >= 475	FSCU up to 100% of fees	None
	IV	PSU >= 475	FSCU from 20% to 80% of fees. CAE subject to budget constraints	None, except that HEIs ought to participate in CAE if applicable
Private Universities	I and II	PSU >= 475	CAE, grant, food and/or maintenance allowance also apply	Accredited and/or CAE participant
	III and IV	PSU >= 475	CAE	CAE Participant
Professional Institutes (IPs)		GPA>=5.3	CAE, Grant, allowances	CAE Participant
		GPA>=5.0	Grant, allowances	Determined by MINEDUC
	III and IV	GPA>=5.3	CAE	CAE Participant
Technical Training Centres (CFTs)	I and II	GPA>=5.3	CAE, Grant, allowances	CAE Participant
		GPA>=5.0	Grant, allowances	Determined for MINEDUC
	III and IV	GPA>=5.3	CAE	CAE Participant

Top performing students prefer selective CRUCH universities and a handful of private independent institutions on the basis of academic indicators, prestige, and the kind of aid available to them.

As noted above, the Chilean student aid supports students on the basis of SES and academic merit. There are several packages, the most attractive ones favouring students entering CRUCH universities. Moreover, the government guarantees aid packages to anyone meeting the requirements as well as budgets enough resources to meet students' needs on the basis of enrolment and tuition fees estimates. Nevertheless, it is important to prevent that CAE coverage has evolved according to the budget allocated by the government on annual basis. In fact, in the first years of CAE operation, loans were offered to students with per capita household income up to the third income quintile. Nevertheless, the government has increased resources to CAE and is currently (as at 2015) able to offer loans to anyone fulfilling academic requirements.

5.3.3. The 2005 reform and the introduction of CAE

In late 2005, a major reform of student aid took place. Three important changes were critical: (i) the creation CAE, which has leveraged resources from private banks and had benefited 600,000 students by 2014, (ii) a re-engineering of student aid affecting CRUCH universities, and (iii) an increase of the number and coverage of grants alongside the extension of maintenance allowances. The government established aid packages according to student academic and socioeconomic characteristics (see MINEDUC 2007; OECD and The World Bank 2009 for more details) including grants, loans, and maintenance support targeted to the poorest students (see Table 5.1 on last page). Student aid started being allocated exclusively on the basis of per capita household income. Moreover, subsequent aid expenditures started being budgeted by forecasting aid demand but this time explicit parameters such as enrolment growth and fee increases are considered.

The new loan scheme has concentrated the biggest share of the aid increase. In fact, 21,000 students were supported by CAE in 2006 and roughly 600,000 have been favoured until March 2014. At the same time, the number of current students under the support of CAE increased from 21,263 in 2006 to 356,574 in 2014, this accounting for about one-third of the Chilean HE enrolment (INGRESA 2015).

Participation levels by SES, although still unequal, increased for all socioeconomic groups with the exception of the richest income quintile, which had also reached an almost universal coverage.

However, policymakers' focus has shifted from access to the mechanisms contributing to increasing retention of poor students as well as securing the financial sustainability of aid policies. There is a wide interest in explaining which different settings of student aid are more effective according to socioeconomic characteristics along with a recognition of the heterogeneity of the impact of aid by SES. This turns relevant given that student population in Chile has experienced an important diversification (UNDP 2005). The student body is no longer the homogeneous group with good school background entering HE straight away from the school, but a much diverse one.

These concerns are a consequence of both the rapid transition from an elite to a mass HE, which relied on the inclusion of otherwise marginalised students and the increase in the number of HEIs from the mid-nineties onwards. The current student population also comprises low SES background young people, workers seeking a second qualification, mature students, etc. The risk of dropping out has become a major concern, as many new students lack academic skills and differ from the once normal well-prepared and academically oriented student seen before.

The Chilean HE policy debate has also paid attention to a fundamental issue: students' debt burden, which has questioned systemic foundations such as cost transfer and a comparatively low share of public funding in HE total expenditure. However, it is a matter of concern that relevant facts are neither investigated nor well documented: dropout rates of CAE supported students are much lower than the

average. For Chilean standards, a modest 12 per cent of CAE students have dropped out in the first three years of study, in sharp contrast with the four times higher system level dropout rate (González and Uribe 2003; Rolando, Salamanca, and Lara 2010). The main question here is why CAE dropout rates are lower, even though the prospects are high income to debt ratio (World Bank 2011), which is often associated with low retention and higher risks of default (Volkwein et al. 1998; Woo 2002) and harsher repayment conditions in comparison to FSCU supported students.

CAE loans started in 2005 to finance students entering in 2006. Loans are generated by private financial institutions while the main mechanism to leverage funds is the state guarantee. Crucially, the government faces a major constraint: the law does not allow the government to originate CAE loans directly, which is what defines the CAE financial model where the government must repurchase the credits that banks are not willing to finance. Once students take up loans, the government splits the portfolio into equally sized and homogeneous loan packages which are subsequently tendered. Bank bids are selected based on the following parameters: (i) the highest proportion of students to be financed by the bank and (ii) the lowest surcharge to portfolio repurchase

Another particular feature of CAE is that both the state and HEIs share the risk: HEIs guarantee a decreasing proportion of loans until the student graduates, which is completely transferred to the state upon graduation. HEIs participation in CAE, in consequence, is not mandatory and they are allowed to cap the number of CAE-supported students according to their institutional practices. HEIs are also able to cap the number of CAE-supported students they accept and establish academic requirements on top of legal requirements (either scoring 475 in PSU or reaching a GPA of 5.3[10]). The way HEIs responded to the incentives in the first year of CAE operation is rather mixed. Some HEIs were conservative in capping numbers and set demanding academic requirements to minimise dropouts so taking a very low risk, others did not cap the number of CAE-supported students but used CAE to substitute their own mechanisms of student support -institutional loans, grants, and bursaries-,

[10] Chilean marks normally range between 1.0 and 7.0, 4.0 being the pass threshold.

while for other HEIs CAE represented an opportunity to expansion (CAE driven HEIs). Nonetheless, most HEIs ended up sticking to the legal requirements and relaxing caps in subsequent years.

At least for this analysis, as only 2006 entry cohort -and the first- is considered, the incentives to HEIs in order to expand enrolment are rather limited, whereas there might be an incentive to have a special attention on CAE-supported students. Indeed, a number of HEIs devised institutional mechanisms aiming at preventing dropout. Nevertheless, CAE students having lower dropout rates, at least in the first cohort, does not seem to be due to special measures taken by HEIs but more demanding academic requirements set by HEIs. The impact this may have on the estimates of debt composition seems to be rather limited, provided that all students in the sample are CAE-supported and most students are totally CAE-supported.

The policy failure

A major policy failure took place in the first year of CAE operation. CAE regulations stated that loans must be offered on the basis of socioeconomic need, the poorest students having priority. After that, loans would be offered to the extent to which the budget allows so[11]. Due to a very rare software failure, loans were offered to the richest students instead of the poor. This meant a major shock for a new programme which was created to improve HE access and forced a quick and radical government intervention. Originally, the government had intended to offer 14,000 credits but as a consequence of the failure, the budget would be increased enough so that poor students could obtain a loan while keeping the loan offer to better-off students. Finally, 21,263 students took up the loans from which 13,695 corresponded to new entrants. Middle-income groups were the most affected with the failure, as shown in Table 5.2, which instead of the intended pyramidal distribution shows one looking as an hourglass.

[11] The government budget is not intended to originate loans directly but to portfolio repurchase –including banks' mark-up and guarantees payment.

Table 5.2. CAE loans taken up in 2006 by household income quintile. New entrants

Household income quintile	N	%
1	2,924	21.35
2	2,208	16.12
3	2,130	15.55
4	3,277	23.93
5	3,156	23.04
Total	13,695	100

Source: *Comisión* INGRESA

Had the assignment rule been correctly applied –giving priority to those coming from low SES households-, a relevant group of students would not have been offered a loan. That group of students had (upper)middle-class background and a variety of sources of government sponsored support were available to them, as shown in Table 5.1. Some of them, those attending to state subsidised universities, were supported by a mixture of CAE loan and income contingent loan, whereas the other group -those attending independent HEIs- was completely financed by CAE.

This features a unique opportunity for research as those students financed by mistake, many being also supported by other aid mechanisms, would not have been present otherwise. It also helps build a proxy for debt composition, which consists in what proportion –as a percentage- of tuition fee payments corresponds to CAE loan and consequently compare the outcomes across debt composition levels. Indeed, without the mistake, CAE was not meant to cover students from quintile 4. Quintile 4 students at CRUCH universities are able to top up their fees -mostly covered with FSCU- with CAE in a variable degree (from 20 to 80 per cent of fees as shown in Table 5.1), but the mistake benefited them, this way allowing me to have a relevant number of cases to study debt composition. In summary, the effect of the policy failure on my research was twofold: it allowed me to count on observations from the whole socioeconomic spectrum and provided a relevant number of cases with different aid composition. This features the uniqueness of the data I use in this chapter as well as the way I take advantage from the mistake.

I discuss the use of debt composition as a proxy of harshness and the assumptions implied in more depth in section 5.4.

The assignment rule has been applied correctly in the following years and the system has been able to, step by step, offer loans to everyone, thus covering the whole socio-economic spectrum. However, the 2006 cohort features a unique case as it is the only one covering the whole socioeconomic spectrum and, at the same time, allow me to track students throughout the complete duration of their studies. On the other hand, loans take up ratios by income quintiles were very similar- about 75 per cent in the 2006 to 2010 period (World Bank 2011). This makes a good point to support the use of the policy failure as a natural experiment (I shall discuss this further in the data section).

It is important to discuss the consequence of the policy failure on the probability of getting aid. The fact that CAE distribution did not have a pyramidal shape may slightly affect the middle (third) income quintile but does not affect students beyond 2006 process. In that sense, provided that quintile 3 was used as control group in Chapter 4, the D-I-D estimator presented for year 2006 might be slightly overestimated. Nonetheless, given the number of students involved in the reform and the fact that CAE was just starting so it did not represent a relevant share of student aid, the effects on the estimates are not important. Neither are 2009 estimates affected because four years later this very small effect would have vanished.

5.4. Data and Methods

The dataset comes from a merging of several administrative databases from Chilean HE student aid systems. My main predictors are a dummy variable 'fully funded by CAE/partially funded by CAE' and the standardised cumulated debt (z score). By using several methodological approaches, I use a series of individual, institutional, course, and household characteristics as control variables. My aim goes beyond merely understanding the effects of student aid on persistence and dropout but this chapter focus on studying the differential effect of both types of support.

5.4.1. Outcome variable

The data used in this chapter come from the Administrative Commission for the Higher Education Loan System, INGRESA. INGRESA is a public institution in charge of managing the State Guaranteed Loan, which is responsible for setting the budget, application processes, loan allocation, and coordination with the banking system. It is also responsible for collecting information from the HEIs to monitor and supervise the system. As loans are renewable every year, INGRESA checks student statuses which are relevant to keep the system working properly. Three main statuses are of interest: (i) continuing registration, (ii) dropout, and (iii) completed studies.

The status 'continuing registration' is the most straightforward as INGRESA just checks with HEIs whether the student is still registered and meets the academic requirements -70 per cent of completed credits- and then proceeds to renew the loan for the following year if requirements are met and the student requests so. This is important to be taken into account as some students may not need financial support or may have been awarded other type of support in a particular year. When requirements are not met, the loan is not renewed for the following year but the student would be eligible for subsequent years if academic credit completion is met in the future.

In contrast, 'dropout' status is not straightforward. Law # 20,127, which regulates the system operation, defines that a student has dropped out when there is no registration at any participant HEI during the whole last academic year. For instance, suppose a student started studying in year n. Then the student was not registered during year n+1 and consequently, INGRESA recorded the student as 'presumably dropped out'; and finally, if the student did not register at any participant HEI in year n+2, INGRESA deems the student as 'dropout'. The issue at this point is that the student might well have dropped out in year n or n+1, since the status variable refers to the beginning of the academic year. I assumed the student has 'dropped out' in year n and so was it recorded in the dataset.

The status 'completed studies' is slightly simpler. For example, a student has been recorded as 'continuing registration' from year n to year n+4. At the beginning of year n+5, HEIs inform INGRESA that the student has finished and then INGRESA records the student as 'completed studies'. Besides it is recorded as 'completed studies' in year n+5, I assume the student actually finished in year n+4, as all HEIs report status at the beginning of each academic year so the student should have finished at the end of year n+4 at the latest.

The outcome is a categorical variable made up of three categories: Continuing registration (0), dropout (1) and completed studies (2). I shall consider transfers, which account for 14.3 per cent of the sample, as continuing registration due to the following reasons. First, some students appear as transferred when they move to other HEI when in fact there might be a regulated pathway so it would not be a proper transfer. This is not straightforward to quantify as it would require me having every regulated pathway in the dataset, which is not the case. Second, the condition 'transferred' does not exclude other outcomes happening thereafter, which is one of the key conditions for the model I develop in this chapter to hold. Moreover, not always are transferred students easy to identify as the data do not allow to track every possible move into the HE system. Therefore, most transferred students are found within category '0'. By doing so, I am indirectly assuming that the events as coded in the outcome variable are absorbing, i.e. no other event must occur after dropout or completion have taken place. In this setting the status 'transferred' would be a non-event.

In order to obtain a basic profile of transferred students, I present some basic statistics in Table 5.3 from which it is noticeable that the proportions of transferred students by income quintile are rather even; entry scores are significantly lower for transferred students, though depending on the income quintile; the proportion of students having completed their courses by year 6 is significantly lower for transferred students; and there is a higher proportion of dropouts among students who have stayed at their entry HEI compared to transfers, even though this depends on the income quintile

Table 5.3. Characterisation of students according to transfer condition

	Household income quintile					
Transfer condition	1	2	3	4	5	Total
	Distribution					
Stayed at entry HEI	815	687	772	1,304	1,442	5,020
% Stayed	86.3	86.1	84.1	83.9	87.7	85.7
Transferred to another HEI	129	111	146	251	202	839
% Transferred	13.7	13.9	15.9	16.1	12.3	14.3
Total	944	798	918	1,555	1,644	5,859
%	100	100	100	100	100	100
	PSU score means					
Stayed at entry HEI	539.8	542.5	547.7	572.5	601.1	567.5
Transferred to another HEI	539.0	529.8	550.5	557.4	591.8	558.0
Dif	0.8	*12.8*	-2.8	*15.1*	*9.3*	*9.5*
	Proportion completing course by year 6					
Stayed at entry HEI	0.54	0.59	0.58	0.6	0.64	0.599
Tranferred to other HEI	0.22	0.24	0.26	0.29	0.23	0.254
Diff	0.32	*0.34*	0.32	*0.31*	*0.42*	*0.35*
	Proportion dropped out by year 6					
Stayed at entry HEI	0.21	0.18	0.18	0.15	0.14	0.165
Tranferred to other HEI	0.14	0.21	0.10	0.11	0.07	0.116
Diff	*0.07*	-0.03	*0.09*	0.04	*0.07*	*0.05*

Note: Statistically significant differences at p=.05 reported in italics

The above makes sense since transferred students do not leave HE but instead try a different university, whereas a lower proportion of transferred students complete their courses by year 6. In summary, transferred students are not different in socioeconomic terms from those staying at the entry HEI, perform lower in PSU but the difference is of less than 0.1 standard deviation, whereas the relevant difference concerns course completion rates.

5.4.2. Covariates

There are two kinds of covariates in the dataset: time invariant and time variant. Time invariant covariates are background variables as per capita household income at the entry point (measured as household income quintiles), GPA in secondary school and university admissions test scores (PSU scores). Other time-invariant variables are the HEI and type of HEI the student entered and the expected income –as at 2006- upon graduation. Time variant covariates are the cumulated debt (in z scores) and the proportion of CAE debt as a percentage of the fees.

Two debt related variables were built: (i) the proportion of total tuition fees being financed by CAE and (ii) the cumulated CAE debt. These are the main variables of interest, their study featuring the core motivation of this chapter. The proportion of CAE financed tuition fees has been dichotomised, the cut-off point being 80 percent. This takes into account the fact that partial CAE funding starts when FSCU finances 20 per cent of tuition fees. It would have been possible to use debt composition as a continuous variable but the frequency distribution does not allow it as a large majority of observations concentrate at around 100 per cent of CAE funding (consider I use the CAE database).

The cumulated debt variable only considers the total CAE accrued debt since the amount corresponding to other aid instruments is not recorded in the database. I have defined those students whose CAE debt proportion is lower than or equal to 80 per cent as 'partially funded', whereas 'totally funded' refers to students with a CAE proportion higher than 80 per cent. The assumption behind this is that non-CAE financed fees are funded by other aid mechanisms such as income-contingent loans, state and institutional grants, third sector aid, etc.

The dataset consists of a merge of administrative data on CAE borrowers from 2006, the year of the administrative error. The total number of students is 21,163, from which 13,097 entered HEIs for the first time that year. Three databases were used to build the dataset. Firstly, student aid applications database (FUAS), which contains socioeconomic data such as per capita household income, household members' activity, educational background, and income. Income is self-reported and checked by the Chilean tax service (*Servicio de Impuestos Internos*, SII), which returns household income quintile on the basis of members' national ID numbers as provided by applicants. SII just reports per capita income quintiles due to legal constraints safeguarding data privacy, while quintiles cut-off values follow standards from CASEN household survey. Thus, there are two measures of income whose consistency is around 70 per cent, the self-reported income being continuous and the corrected categorical. Although SII adjusts income and this measure should be more reliable, it is the applicant who fills in application forms regarding household members' data. Therefore,

omitting the 'wealthy' household member or adding household members with no income drops the per capita income artificially so it may end up supporting students that should not be benefited. There is some evidence of manipulation of socioeconomic data so the government has devised a number of consistency checking routines. Nonetheless, application checks have turned irrelevant as the income cut-off point has been pushed upwards over the last years. Anyway, I use the household income quintile as adjusted by SII.

National ID numbers (every Chilean is assigned one upon registering birth) make also possible to merge FUAS database with PSU scores, GPA, and the secondary school attended. Secondary schools' identifiers allow merging data on school performance tests (SIMCE) at the school level, type of school (public, private subsidised or private independent), and a school vulnerability score based on the proportion of students entitled to free school meals. Individual standardised test scores are problematic as the measurement system intends assessing performance at the school level rather than individual achievement. Though some studies in Chile have used individual scores, there is not enough technical ground supporting the use of individual scores. I use PSU scores and school GPAs instead.

Other course characteristics are also reported. For instance, the Chilean Ministry of Education (MINEDUC) tracks student employment and wages. The information is delivered by generic course/profession, as for instance medicine, engineering, architecture, etc. There are data on wages and the proportion getting a job in a specific year for 181 courses/professions of which 167 has been matched to CAE borrowers. Employment prospects, namely courses' current average income is used as a proxy of how attractive a course is, thus allowing an extra control variable. Despite the variability in earnings for the same course at different universities may be high, average incomes are a good proxy for prestige and attractiveness.

Finally, an index of university complexity was built as a proxy of quality, selectivity and how demanding courses might be. The index considers research output and academic staff qualifications and productivity for each university as at 2006. By using

factor analysis via principal components factor extraction, I obtained a unique index to be used as control variable, which corresponds to the one with the highest eigenvalue.

5.4.3. Debt composition as a proxy of harshness

A key issue for this chapter is to relate debt composition and harshness. I shall argue that CAE loan is the harshest kind of student aid due to the following reasons. Firstly, loans are originated by private financial institutions, the same applies to debt collection. Secondly, the repayment method is done by fixed instalments which are collected for 10, 15 or 20 years on the basis of the amount accrued. Although an income sensitive element was introduced in 2013 where the government subsidises the gap between the fixed instalment and 10 per cent of the gross monthly income, this does not affect the incentives structures at the time the borrowers were studying. Thirdly, there are not provisions for writing off unless a permanent disability is demonstrated or the borrower's death. Neither do income thresholds apply to CAE. Fourthly, CAE interest rates are fixed by INGRESA on the basis of government borrowing rate plus a 2 per cent spread. For 2006 CAE borrowers, the annual interest rate was 5.5 per cent after accounting for inflation. Interest is also accrued during the study period.

In contrast, FSCU, the income contingent loan used by most government supported HEIs, is written off after 15 years, is repaid at a fixed 5 per cent of the monthly income over an income threshold. The interest rate is 2 per cent real, well below the commercial banks' interest and the government cost of borrowing. Debt collection is done by the universities through some contract specialised firms to support the process. Therefore, it is a much more student-friendly loan and there is much more at stake in the case of defaulting a bank loan.

In consequence, it is safe to maintain that a higher proportion of tuition fees funded by CAE translates into a harsher debt. Nevertheless, due to the data nature and the high proportion of fully CAE funded students found in the dataset (the average is 90

per cent by the sixth year), it was not sensible to use a continuous debt composition variable but a dummy. Despite I could not maintain that 80 per cent of CAE funding does make much difference with 75 per cent, for instance, I do maintain that even a small amount of CAE does make the difference since a bank demanding payment is much more deterrent than the university doing so.

5.4.4. Data structure and hierarchies

HEIs report enrolment by course attended every year and a status variable over time: continuing student, graduated, dropped out, suspended, etc. Specific course descriptors are also reported as course duration and tuition fees for the current academic year. Amounts borrowed are recorded every year so it also allows tracking cumulated debts at any point in time. Additionally, the dataset responds to a hierarchical structure as all the students are identified by HEI and course, thus providing an opportunity to test whether observations are independent through introducing random effects (see section 5.4.5).

Nested data might be crucial in order to deal with unobserved heterogeneity. The reason is that there is a high variability within Chilean universities, where faculties enjoy a high degree of autonomy. Most student services, although centrally organised, operate at this level where social and academic integration mechanisms operate. As the data do not allow reproducing the faculty structure, I use the course-within-university as the level two grouping variable. I shall not estimate random slopes but only random intercepts. Identifying a group to which students belong with acceptable precision is crucial in order to deal with unobserved heterogeneity and intra-class correlation.

5.4.5. Methods

As the main theoretical frameworks consistently emphasise, persistence and dropout are longitudinal processes. Empirical research on dropout uses longitudinal data in different ways, as for instance 'before and after' data as some registration records – when a student enrols and when leaves-; time to person records, where a status

variable tells whether or not an event of interest has occurred at the time 't' for each value of t. These data structures depend on how the time variable is recorded. The shape of the dataset and how time is recorded, on the other hand, critically determine which methods are appropriate. Time to person records, as in my dataset, are more flexible as the introduction of time-varying covariates is allowed, whereas other structures where only starting and event occurrence time are recorded would only allow for time-invariant covariates.

A common issue in longitudinal research is the presence of censoring. Censoring happens when the event(s) of interest might not have happened during the observation period as when there is not an outcome for a group of students, thus the researcher has no information on whether and when the event(s) will take place because observation periods are limited. This is called right censoring. On the other hand, censoring also occurs when researchers find some students which were not tracked during the observation period or when there are no observations for some subjects in some time periods, which is called interval censoring.

Censoring affects critical distributional assumptions leading to biased estimators. Survival analysis, time to event or event history analysis are widely used methods to deal with censoring in several ways, by modelling time within the censored data framework (Hosmer, Lemeshow, and May 2008; Rabe-Hesketh and Skrondal 2012; Singer and Willett 1993, 2003; Willett and Singer 1991). Unlike some approaches that only use complete cases, by imputing censored values or dichotomising event/no event regardless censoring, survival analysis uses all the available information from the dataset (Hosmer and Lemeshow 1999).

Survival analysis has been intensively used in empirical research on time-to-degree and dropout (Bruinsma and Jansen 2009; Denson and Schumacker 1996; DesJardins, Ahlburg, and McCall 1999; DesJardins et al. 2002a; Ishitani and DesJardins 2002; Letkiewicz et al. 2014; Ortiz and Dehon 2013; Paura and Arhipova 2014; Petras et al. 2011; Wao 2010) as well as in studies measuring the effect of student financing on persistence/dropout where most of them found aid as either improving time-to-degree

or reducing the risk of dropout (Chen and DesJardins 2010; DesJardins et al. 2002b; van der Haert et al. 2014; Horn et al. 2014).

Classic and widespread survival analysis techniques such as Cox's regression requires knowing exactly when the event(s) of interest happened. Given the nature of the data, in this chapter, I deal with a discrete time survival analysis. For instance, my database allows knowing the year in which the events of interest occurred but not the exact date. However, students may depart from the university at any time within the year so this violates the core assumptions of the classic continuous time survival approach.

5.4.6. Model specification

In survival analysis, events of interest might be of binary, as withdraw/stay, live/dead, smoker/non-smoker, or multiple outcomes as dropout-graduate-continuing. The last case features the competing risks model –or a multiple absorbing model in the survival analysis jargon-, which is the approach I use in this chapter. The occurrence of completion/dropout would prevent any other event from ever happening. A binary model would treat the other event as censored.

Discrete-time competing risks survival models are an extension of multinomial logistic regression but adding specific time dummies. Instead of odd ratios, the multinomial model reports relative risk ratios (RRR). The interpretation of RRRs is similar to that of odds ratios but refers to a base event. For instance, in a binary model of dropout/non-dropout, an odd ratio of 1.5 for a given predictor would mean that a unit increase on the predictor would increase the odds of dropping out by 50 per cent. Instead, in a multinomial context where the base category is 'continuing studies,' coefficients report the RRR, where a value of 1.5 indicates that an increase on the predictor would increase the odds of dropping out by 50 per cent but relative to 'continuing studies'.

The model specification for a competing risks model is

$$\log\left(\frac{h_i(k,t)}{h_i(0,t)}\right) = \alpha_{k0} + \alpha_{k1}D_{1i} + \alpha_{k2}D_{2i} + \cdots + \alpha_{kT}D_{kT} + \beta_k X_i \quad [5.1]$$

Where, $h_i(k,t)$ is the hazard of experiencing event k, relative to non event $h_i(k,0)$, D_i are time dummies for each time period, α_{ki} capturing the baseline hazard function (see *e.g.* Allison 2014; Rabe-Hesketh and Skrondal 2012; Singer and Willett 2003; Willett and Singer 1991), and X_i being a vector of covariates. Not only should the non-event category be used as reference; it might be any one so long as coefficients are meaningful. I use dropout as base category but also look at the coefficients for completion relative to continuing studies (non-event).

As explained above, I also introduce to [5.1] a random effect μ_{k0j} at course j within university to control for unobserved heterogeneity, thus allowing intercepts to vary.

$$\log\left(\frac{h_i(k,t)}{h_i(0,t)}\right) = \alpha_{k0} + \alpha_{k1}D_{1i} + \alpha_{k2}D_{2i} + \cdots + \alpha_{kT}D_{kT} + \beta_k X_{ij} + \mu_{k0j} \quad [5.2]$$

Although the use of random effects does not address the issue of unobserved variable bias as random effects and individual predictors are supposed to be orthogonal (*i.e.* zero correlation), it does introduce a factor consistent with the theory, as a random intercept at course level might well account for institutional factors, that I am not able to observe, at a level which is seen as critical from the integration model since social interactions and the academic experience occur mostly at faculty/course level. The use of random effects in the context of survival analysis is often referred to as frailty models (Blossfeld and Hamerle 1989; Hougaard 1995; Jones-White et al. 2009; Liu 2014; Steele, Diamond, and Wang 1996; Steele, Goldstein, and Browne 2004). Frailty models have a multiplicative effect (additive in the log-odds scale) on the baseline hazard function given by α_{ki} in [5.1] and [5.2], which estimate the hazard function in the log-odds metric.

5.5.Results

5.5.1. Descriptive statistics

Out of the 13,000 new students entering HE in 2006, I have selected a group of 5,859 students entering 5-year courses at a university, leading to academic oriented professional qualifications, corresponding approximately to ISCED level 5A qualifications. This decision was made on the following basis: (i) five year courses are the most common in Chilean universities, where courses last 4.92 years on average (with a 0.75 year standard deviation) and (ii) the observation period, after the temporal adjustments made in order to match the status variable with more substantive definitions, rather than the bureaucratic ones as used by CAE management, is reduced to 6 years. So it seems reasonable to restrict the population this way. Had I considered longer courses, I would not have allowed for some extra time for completion, which may lead to artificially pessimistic predictions given that course completion goes beyond the formal duration very often. I did not consider shorter courses as they have a more vocational profile and are hardly comparable with five-year courses with regards to academic orientation, the level of skills required to succeed, selection criteria, and entry requirements.

I dealt with missingness by both imputing the mean and regression-based techniques with variables not used as predictors in the estimates. PSU means were imputed for 220 out of 5,859 individuals. The type of school attended was imputed by multinomial logistic regression taking the commune (local authority) as the predictor. The number of individuals with imputed values for the type of school was 117 out of 5,859. The number of imputed values is low enough to make sure they do not bias the variance.

The main descriptive statistics are reported in Table 5.4 which follows. As shown, the consequences of the administrative error on the distribution of students by household income quintiles are more evident in this subpopulation than the whole population of CAE new students (as shown in Table 5.2) as it takes a pyramidal shape rather than the glass hour one. In addition, most of the effect is due to the fact that CRUCH universities better off students were disproportionally benefited in comparison to

private universities students. For instance, 84 per cent of CRUCH CAE students come from the richest quintiles (4 and 5), whereas the figure drops to 41 per cent in private universities.

Table 5.4. Descriptive statistics. 5-year university courses

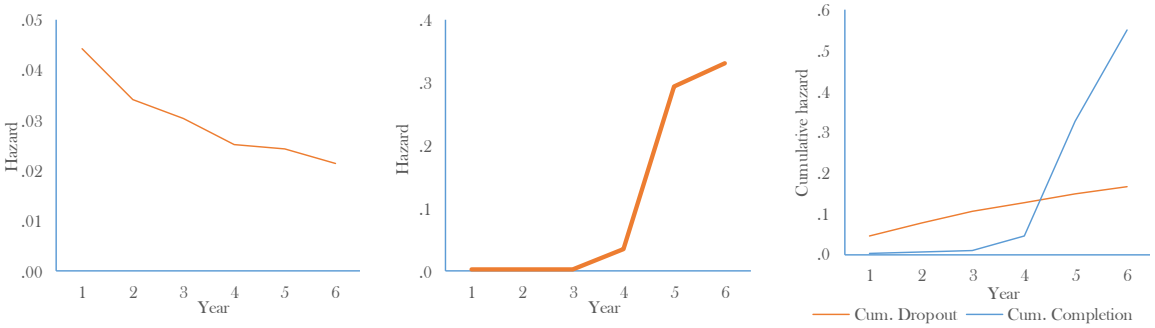
Individual Characteristics		
Predictors	5-year university courses	
	Mean	SD
Demographic characteristics		
Gender (%)		
Male	59.67	*49.06*
Female	40.33	*49.06*
Starting age	19.49	*2.09*
Prior Academic achievement		
GPA	580.58	*54.99*
PSU Score	566.13	*58.67*
Socioeconomic characteristics		
Income quintile (%)		
1	16.11	*36.76*
2	13.62	*34.30*
3	15.67	*36.35*
4	26.54	*44.15*
5	28.06	*44.93*
Type of Secondary School (%)		
State	31.97	*46.64*
Private Maintained	54.84	*49.77*
Private Independent	13.19	*33.84*
Debt Characteristics		
Cumulated CAE debt ('000 CLP$)		
Year 1	1339.77	293.33
Year 2	2671.93	635.59
Year 3	4088.29	1010.42
Year 4	5680.80	1345.81
Year 5	7329.82	1712.72
Year 6	8933.95	2136.78
Proportion of fees financed by CAE loan		
Year 1	94.59	14.78
Year 2	92.19	16.51
Year 3	91.31	17.27
Year 4	91.69	16.07
Year 5	91.42	15.45
Year 6	90.09	16.31
University Characteristics		
	Mean	SD
Type of entry HEI (%)		
CRUCH University	30.82	*46.17*
Private University	69.28	*46.13*
University complexity index	0.08	*1.00*
Profession Characteristics		
	Mean	SD
Expected Income ('000 CLP$)	892.88	*301.26*
Total N	*5859*	

Sample hazard statistics and hazard functions are shown in Table 5.5 and Graphs 5.2, 5.3 and 5.4 below. Dropout probabilities decrease with time whereas completion probabilities increase, which looks reasonable and consistent with available studies. There still remains, by year 6, an important proportion of students who have not yet completed their 5-year programmes. This is not surprising for Chilean standards since dropout rates are high and studies last longer than the formal duration. In fact, for 5-year courses at Chilean universities, completing studies may easily take up to 7 years.

Table 5.5. Sample hazard probabilities. 5-Year courses, universities

		Events		Hazard probabilities (%)			
Year	Risk set	Dropout	Completion	Dropout	Completion	Cum. Dropout	Cum. Completion
1	5,859	259	16	4.42%	0.27%	4.42%	0.27%
2	5,573	190	19	3.41%	0.34%	7.68%	0.61%
3	5,346	162	24	3.03%	0.45%	10.48%	1.06%
4	5,141	130	181	2.53%	3.52%	12.74%	4.54%
5	4,779	116	1,404	2.43%	29.38%	14.86%	32.59%
6	3,194	69	1,058	2.16%	33.12%	16.70%	54.92%

Graph 5.2. Dropout hazard **Graph 5.3. Completion hazard** **Graph 5.4.Cumulated hazard**



Few students have been missed or were not completely tracked during the observation period. This does not represent an issue as the number -164- is too small in order to represent a real threat.

5.5.2. Estimates

Maximum likelihood estimates are provided in Table 5.6, starting with the baseline hazard model which only includes time dummies. The key variable of interest is debt

composition, which is coded as 1 for 'partially CAE funded' students and 2 for 'totally CAE funded' students. The reference category is 'totally CAE funded'.

I took the year of the first enrolment as a base category for time dummies. I also provide Wald's chi-square tests for categorical variables in order to test whether or not their effect equal zero as well as goodness of fit statistics, namely log-likelihood, Akaike's Information Criterion (AIC), Schwartz's Bayesian Information Criterion (BIC)[12], and Mc Fadden's R-squares. Models were estimated by using clustered robust standard errors.

Debt characteristics are the main variables of interest. Having controlled for social background variables, the model shows that for partially funded CAE students the relative risk of completing studies would be expected to decrease by 42 per cent. The effect of the debt composition dummy keeps significant with an important variation after introducing demographic controls, the RRR dropping to 39 per cent, but jumps to .61 once prior achievement and HEIs characteristics are controlled for. When introducing interactions, the coefficient for the main effect of debt composition becomes not significant but interaction effects are highly significant. In one way or another, the effect of debt composition is statistically significant for all the specifications where it is included. Equations testing debt composition in RRR between continuing students and dropouts do not show any statistically significant coefficient. No statistically significant coefficients are found for cumulated debt.

Nevertheless, the effect of the dummy variable ought to be studied for some subpopulations as partially funded students are defined by the policy rules shown in Table 5.1. For this purpose, I introduce interactions between debt, time and socioeconomic variables. The cumulated debt was squared in order to identify concavity. As expected, interactions between debt composition and income quintiles are highly significant for quintiles 4 and 5. The risk ratios for quintiles 4 and 5, taking

[12] To calculate BIC, instead of the number of observations or subjects, I use the number of events, following Raftery (1995) and Singer and Willet (2003) recommendations.

quintile 1 as the reference category adds up to the baseline odds. Nevertheless, interaction in nonlinear models should be taken carefully (Ai and Norton 2003).

Predictive margins provide a clearer picture. What interaction coefficients show is that the relative risk for CAE partially financed students, relative to quintile 1 fully CAE funded students, improves for quintiles 4 and 5, at a higher rate in comparison to fully CAE financed students. By year 6, the probability of completion (hazard probability), for fully CAE supported students in quintile 1 is 2.27 times (.45/.20) as that of partially financed. For quintile 5 the ratio is 1.32 (.37/.28). That means that completion probabilities for partially supported students increase for quintiles 4 and 5 and that the gap between fully and partially CAE supported students narrow as income level increases, as shown below in Graph 5.5. By year 5, I find a similar pattern.

Table 5.6. Relative risk ratio estimates for competing risk models, comparing risks of continuing and course completion, relative to drop out

	Baseline		Add background and debt		Add demographic characteristics		Add academic achievement & HEIs		Add expected income & interactions.	
	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout
Year (*Wald χ^2*)	*29.01****	*656.70****	*18.72***	*284.48****	*17.89***	*308.78****	*17.74***	*304.05****	*19.95***	*202.60****
2	1.31*	1.62	1.21	1.68	1.31*	1.75	1.37*	1.96	1.44*	2.24
3	1.48***	2.40**	1.24	2.58*	1.47*	2.78**	1.60*	3.51**	1.91*	4.29**
4	1.72***	22.54***	1.32	24.95***	1.71*	28.10***	1.95**	40.33***	2.82**	63.02***
5	1.30	195.92***	0.91	224.73***	1.28	264.21***	1.54	443.16***	1.70	575.41***
6	1.39*	248.21***	0.89	293.35***	1.37	363.80***	1.76	715.67***	1.69	734.36***
Debt structure										
Partially CAE funded			1.17	0.58**	1.18	0.61*	0.99	0.39***	0.59	1.04
Cumulative debt			1.18	0.96	1.18	0.97	1.07	0.75	1.08	0.70
Sq_Cumulative Debt									1.10	1.13
Socioeconomic background										
Income Quintile (*Wald χ^2*)			*11.07**	*33.18****	*9.75**	*30.71****	*5.33*	*9.83**	*1.59*	*4.50*
2			1.15	1.15	1.16	1.17	1.18	1.17	1.10	1.27
3			1.20	1.20	1.19	1.22	1.17	1.17	1.10	0.92
4			1.44**	1.66***	1.41**	1.68***	1.32*	1.43*	1.18	0.96
5			1.46**	2.01***	1.41**	2.00***	1.23	1.45**	1.14	0.78
Secondary School (*Wald χ^2*)			*6.52**	*6.01**	*4.98*	*5.24*	*3.30*	*2.47*	*3.22*	*2.76*
Public			0.80	0.76*	0.80*	0.74*	0.86	0.90	0.86	0.89
Maintained			0.95	0.92	0.91	0.87	0.97	1.04	0.97	1.03
Demographic chars.										
Male=1					0.69***	0.56***	0.70***	0.57***	0.71***	0.61***
Age					0.92***	0.95**	0.92***	0.97	0.92***	0.97
Prior achievement										
PSU							1.26***	1.61***	1.28***	1.72***
GPA							1.13***	1.31***	1.13***	1.34***
University Characteristics										
Cruch University=1							1.53*	1.52	1.46*	1.51
Complexity Index							0.82*	0.83	0.82*	0.76*

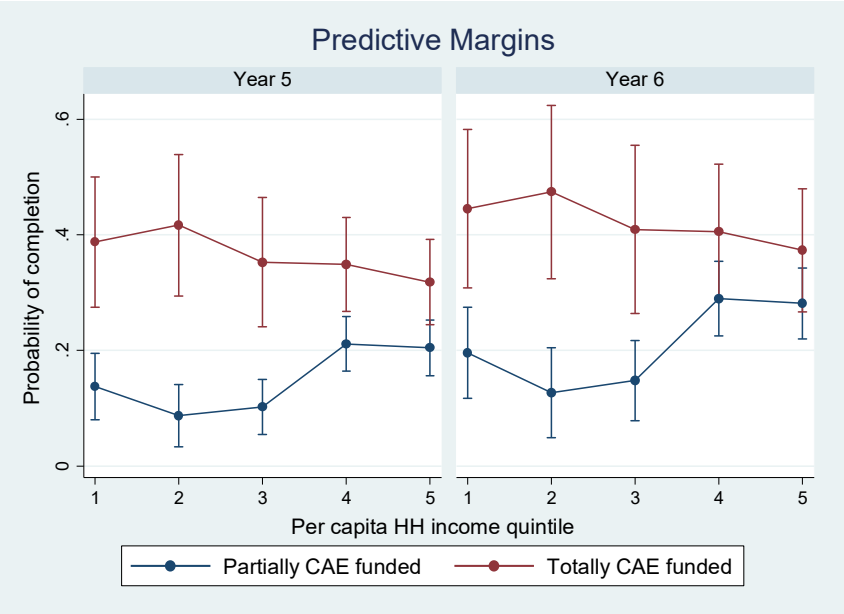
	Baseline		Add background and debt		Add demographic characteristics		Add academic achievement & HEIs		Add expected income & interactions.	
	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout
Year * Partially CAE funded (*Wald* χ^2)									*21.31***	*15.50***
Year=2 * Partially CAE funded									1.64	0.92
Year=3 * Partially CAE funded									0.97	0.52
Year=4 * Partially CAE funded									0.45**	0.15**
Year=5 * Partially CAE funded									1.16	0.16**
Year=6 * Partially CAE funded									1.47	0.24
Income Quintile*PartiallyCAE funded (*Wald* χ^2)									*11.13**	*19.34****
Income Quintile=2 * Partially CAE funded									1.56	0.78
Income Quintile=3 * Partially CAE funded									1.65	1.36
Income Quintile=4 * Partially CAE funded									2.09**	4.21***
Income Quintile=5 * Partially CAE funded									2.10**	4.69***
Expected income									0.93	0.68***
Income quintile * St.Cum.Debt (*Wald* χ^2)									*2.06*	*5.12*
Income Quintile=2 * St. Cum. Debt									0.85	0.77
Income Quintile=3 * St. Cum. Debt									0.92	1.06
Income Quintile=4 * St. Cum. Debt									0.91	1.09
Income Quintile=5 * St. Cum. Debt									0.84	1.21
Intercept	21.56***	0.06***	21.74***	0.05***	147.97***	0.18***	96.53***	0.07***	88.48***	0.06***
N Obs	29892		29892		29892		29892		29892	
N Individuals	5859		5859		5859		5859		5859	
Goodness of fit statistics										
-2LL	20243		20112		20002		19825		19594	
AIC	20267		20168		20066		19905		19734	
BIC	20341		20342		20265		20152		20168	
Pseudo R^2 (+)	0.23		0.23		0.24		0.24		0.25	

* p<0.05, ** p<0.01, *** p<0.001

Note: Wald chi-sq were calculated to test the null hypothesis that the effect of categorical variables equals to zero

(+) Pseudo R-sq= 1 - Log likelihood model/Log likelihood null model

Graph 5.5. Fitted hazard probabilities of completing studies by income quintile (full single level model)



On the other hand, by years 5 and 6, low income fully CAE funded outperformed their richer counterparts. This may suggest that poor students are more compelled to finish their courses on time than the richer ones when the debt is harsher but the differences are not statistically significant when looking at the boxplots above. In fact, within the fully CAE funded group, completion probabilities are similar across the socioeconomic spectrum. In contrast, within the partially CAE funded group, poor students show lower completion probabilities than their richer counterparts. This may have important implications as the model would predict that the harsher debts make the difference among poor students, whereas it would not amongst the better off.

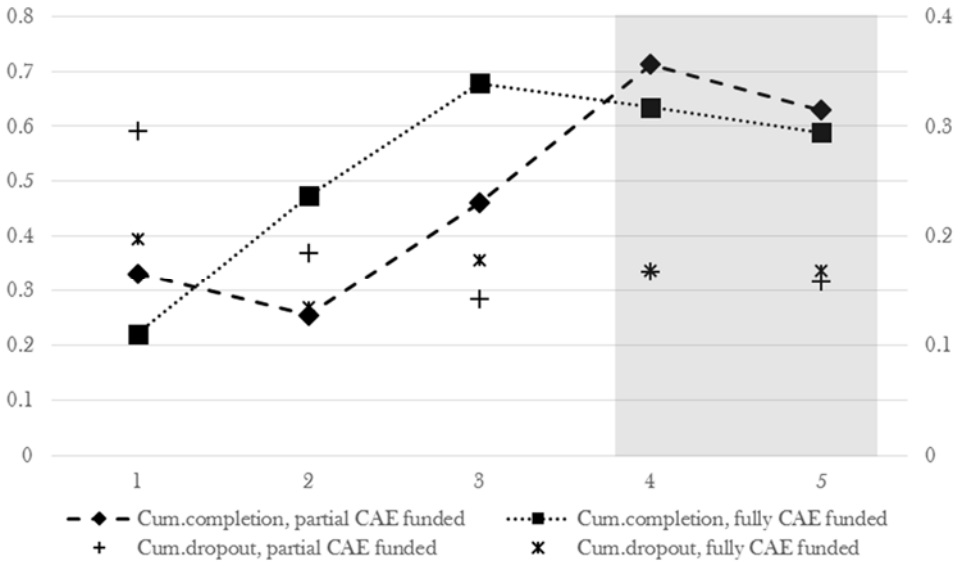
Instead of looking at hazard probabilities, a clearer picture might be obtained by examining the cumulated probabilities of either event for both debt composition groups, as shown in Graph 5.6 (dropout rates on the secondary axis). The cumulated probability $Cum\ h_{(a,T)}$ of an event (excluding the non event) by time T can be defined as follows:

$$Cum\ h_{(a,T)} = 1 - \prod_{t=1}^{T}(1 - h_{(a,t)}) \quad [5.3],$$

where $h_{(a,t)}$ is the hazard probability at time t and T corresponds to the time to which the cumulated probability is calculated.

Fully CAE funded students also present lower dropout rates than partially funded ones, with the most noticeable gap being amongst poor students. Once again, the debt composition seems to make the difference on poor students, namely up to the middle-income quintile, in contrast to students from quintiles 4 and 5 (grey area on Graph 5.6)

Graph 5.6. Cumulated probabilities by household income quintile. Year 6



Although I have introduced controls at individual and institutional levels, there still remains some degree of uncertainty because debt composition, as the available data allow, only considers the proportion of tuition fees being financed by CAE. Due to the Chilean aid system rules, I have assumed that partially CAE funded students are also supported by another sort of aid from the government, HEIs or third sector organisations so that the proportion of CAE loans in relation to tuition fees for each student features a good proxy for debt composition.

The estimates and predicted probabilities shown above may be subject to unobserved heterogeneity. For instance, student A belongs to the group G and student B to G'.

In the absence of heterogeneity, if both students had identical profiles (covariates), their predicted probabilities of dropping out would be the same. Now suppose that students in group G face more demanding academic requirements, a higher level of institutional support, or a higher quality of the facilities and IT than those from G'. All these variables are not observed and may lead to violating the assumption of independent observations: the true probabilities for two individuals with the same profile would not be equal.

Theory on dropout and persistence would not support that the independence assumption holds. As discussed in section 5.3, theoretical developments have paid attention to students' integration processes, which operate at the institutional level. As in the example above, I am not observing these variables but the intuition is that if the theory is correct, there will not be ground for the independent observation assumption to hold. As I have argued, in Chilean universities academic and social integration occur mostly at faculties, which in many cases are very autonomous from the central administration. Faculties structures are not reproduced by the data, thus I use programme/course-within-university as a proxy and use it as the clustering unit of interest.

To deal with unobserved heterogeneity, I include a random intercept (Table 5.7). I also prevent that with the inclusion of a random effect address the unobserved variable bias will not be addressed as random effects are assumed as being orthogonal to fixed effects.

Table 5.7. Relative risk ratio estimates for random effects competing risk models, comparing risks of continuing and course completion, relative to drop out

	Baseline		Add background and debt		Add demographic characteristics		Add academic achievement & HEIs		Add expected income & interactions.	
	RRR Continuing /Dropout	RRR Completion/Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout
Year (*Wald* χ^2)	*27.22****	*723.47****	*19.00***	*305.69****	*17.76***	*310.07****	*17.37***	*302.08****	*21.04****	*218.53****
2	1.30*	1.61	1.20	1.42	1.30*	1.49	1.35*	1.58	1.51*	1.74
3	1.46***	2.38**	1.23	1.83	1.45*	2.03	1.57*	2.28*	2.07**	2.62*
4	1.69***	23.32***	1.30	15.30***	1.67*	18.06***	1.89*	22.03***	3.20***	33.32***
5	1.28	267.94***	0.89	151.34***	1.24	189.88***	1.48	265.80***	1.98	362.52***
6	1.35	438.33***	0.86	213.39***	1.32	286.79***	1.67	464.54***	1.99	546.35***
Debt structure										
Partially CAE funded			1.18	0.72	1.19	0.73	1.01	0.59**	0.59	1.10
Cumulative debt			1.18	1.34	1.19	1.35	1.08	1.14	0.99	1.24
Sq_Cumulative Debt									1.11	1.06
Socioeconomic background										
Income Quintile (*Wald* χ^2)			*16.24***	*27.61****	*12.71**	*24.84***	*6.89*	*13.53***	*2.50*	*3.56*
2			1.16	1.12	1.17	1.13	1.18	1.12	1.10	1.31
3			1.24	1.26	1.21	1.26	1.20	1.20	1.12	1.06
4			1.48**	1.73***	1.44**	1.72***	1.36*	1.58**	1.22	1.44
5			1.50***	1.80***	1.43***	1.75***	1.26*	1.46**	1.16	1.21
Secondary School (*Wald* χ^2)			*10.33***	*5.77*	*7.10**	*3.29*	*5.32*	*2.98*	*5.18*	*2.80*
Public			0.76*	0.89	0.78*	0.90	0.83	1.09	0.82	1.07
Maintained			0.94	1.08	0.91	1.04	0.96	1.21	0.96	1.19
Demographic chars.										
Male=1					0.72***	0.53***	0.73***	0.57***	0.73***	0.58***
Age					0.91***	0.94***	0.92***	0.96*	0.92***	0.96*
Prior achievement										
PSU							1.27***	1.98***	1.28***	2.05***
GPA							1.13***	1.51***	1.13***	1.52***
University Characteristics										
Cruch University=1							1.56*	0.97	1.50*	0.99
Complexity Index							0.83*	0.72**	0.83*	0.70**
Year * Partially CAE funded (*Wald* χ^2)									*21.80****	*13.24**

	Baseline		Add background and debt		Add demographic characteristics		Add academic achievement & HEIs		Add expected income & interactions.	
	RRR Continuing /Dropout	RRR Completion/Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout	RRR Continuing /Dropout	RRR Completion /Dropout
Year=2 * Partially CAE funded									1.63	1.06
Year=3 * Partially CAE funded									0.94	0.72
Year=4 * Partially CAE funded									0.43**	0.22*
Year=5 * Partially CAE funded									1.11	0.21*
Year=6 * Partially CAE funded									1.41	0.31
Income Quintile*Part. Funded (W*ald* χ^2)									*10.12**	*14.26***
Income Quintile=2*Partially CAE funded									1.57	0.87
Income Quintile=3*Partially CAE funded									1.65	1.12
Income Quintile=4*Partially CAE funded									2.03**	3.66**
Income Quintile=5*Partially CAE funded									2.03**	3.56**
Expected income									0.95	0.59***
Income quintile * St.Cum.Debt (W*ald* χ^2)									*2.05*	*3.63*
Income Quintile=2 * St. Cum. Debt									0.85	0.73
Income Quintile=3 * St. Cum. Debt									0.91	0.97
Income Quintile=4 * St. Cum. Debt									0.92	0.86
Income Quintile=5 * St. Cum. Debt									0.85	0.90
Random effects Var / Cov										
$\sigma^2_{\mu_i}$	0.21**	1.03***	0.18**	1.04***	0.14*	0.99***	0.13*	1.10***	0.12	0.92***
$\sigma_{\mu_1\mu_2}$	0.22*		0.20*		0.16*		0.15		0.11	
ρ	0.06	0.24	0.05	0.24	0.04	0.23	0.04	0.30	0.04	0.22
Intercept	23.47***	0.05***	22.97***	0.05***	158.61***	0.23**	99.89***	0.10***	81.91***	0.06***
N Obs	29892		29892		29892		29892		29892	
N Individuals	5859		5859		5859		5859		5859	
-2LL	19397		19273		19170		18918		18778	
AIC	19427		19335		19240		19004		18924	
BIC	19520		19527		19457		19271		19377	
Pseudo R^2 (+)	0.26		0.26		0.27		0.28		0.28	

* p<0.05, ** p<0.01, *** p<0.001

Note: Wald χ^2s were calculated to test the null hypothesis that the effect of categorical variables equals to zero.

(+) Pseudo R-sq= 1 - Log likelihood model/Log likelihood null model.

I have also introduced institutional controls in single and two-level models by including proxies for prestige and complexity. Entry scores, secondary GPA, and type of secondary school attended may also have an institutional correlate. This way, though I cannot make sure that random effects are orthogonal to individual characteristics, I maintain that, given the nature of the control variables used at individual level in the two-level model, it is sensible to state that my individual controls are likely to capture institutional factors which interact with the individual covariates.

Estimates including a random effect are reported in Table 5.7 above. The coefficients look different from those on Table 5.6, but they do not necessarily show the same thing. In fact, when a random effect is included, the estimates refer to subject-specific effects or conditional RRRs, instead of a marginal probability only conditional on the covariates (Rabe-Hesketh and Skrondal 2012:529). In this case, the effects refer to course specific effects and the interpretation is slightly different. For instance, students A and B attend the same course at the same university but A is fully CAE funded and B is partially CAE funded. In the model introducing academic achievement and HEIs characteristics (fifth column on Table 5.7), a relative risk of 0.59 for partially funded means that the odds of completion for B are 41% lower than those of A, relative to dropout, but conditional on students attending a given course within a university.

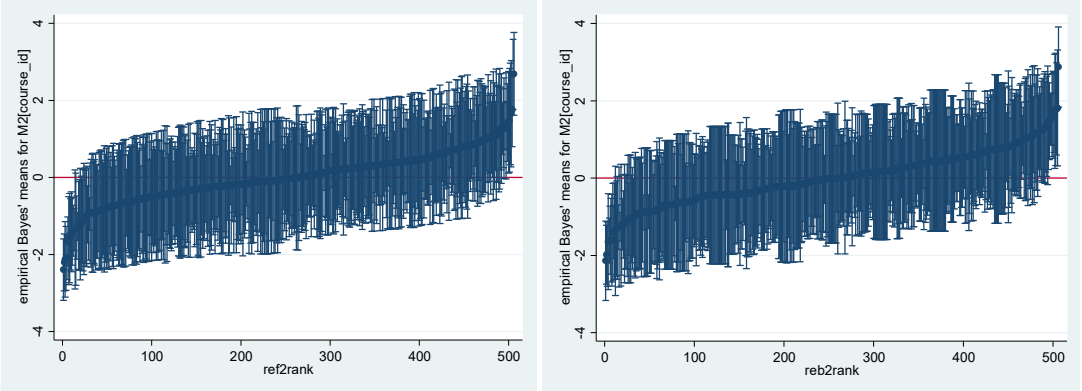
The effect of the debt composition dummy is statistically significant for the model controlling for demographic, prior achievement, socioeconomic and institutional variables, as well as in the interacted model. In fact, in the same pattern of models in Table 5.6, interaction effects add up to baseline odds in a similar fashion while coefficients are less extreme but, as already explained, they are conditional on given values of the random effect.

Random effects are consistently significant for completion, whereas the one for continuation becomes non-significant as more controls are introduced. Intra-class correlation (ICC), in turn, remains within the 0.20-0.30 range in the completion equations. Although ICC is not particularly high, it suggests that the introduction of random effects does improve the models –as reported in goodness of fit statistics in

tables 5.6 and 5.7- and it also introduced the appropriate flexibility to allow for subject (course) specific hazard functions. Furthermore, the level of intra-class correlation found in this study is higher than what it is found in many school achievement models.

The introduction of random effects and their contribution to the models are reported in Graph 5.7. The graph shows the estimated value of random effects with their 95% confidence intervals, taking a caterpillar shape. There is a number of subjects where the random effects are statistically different from zero, which are the boxes overlapping the zero line. Between the baseline model and the full interacted one, plots look very similar but the full model shows slightly more subjects with statistically significant random effects.

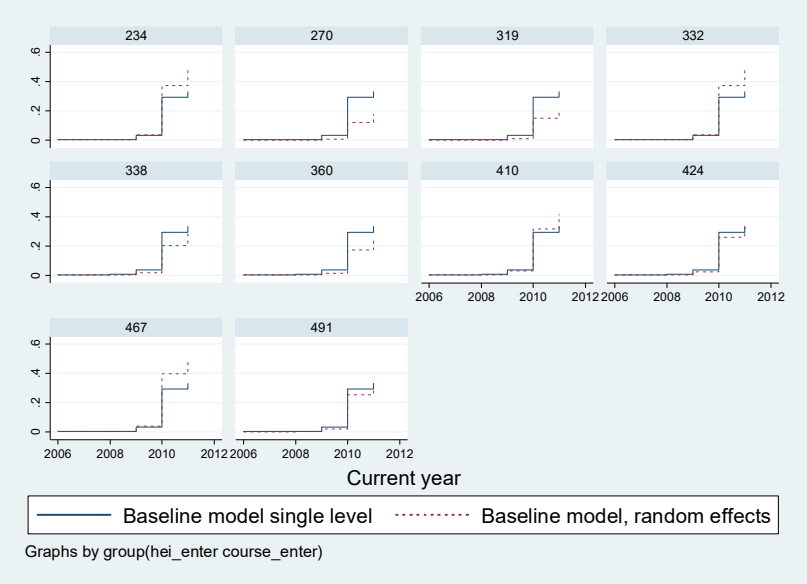
Graph 5.7. Caterpillar plots for random effects for completion. Baseline (left) and full (right) models



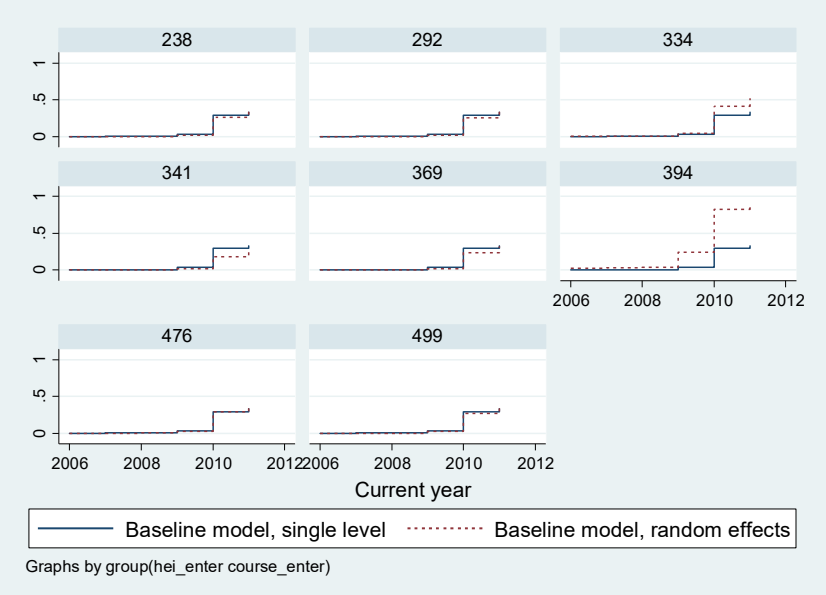
A different way of examining the contribution of random effects is by looking into the variations of the hazard function comparing the baseline single level (constant hazard function) and the baseline random effects models. Graphs 5.8 and 5.9 show examples of how the introduction of random effects results in more flexible hazard functions. Graph 5.8 refers to Law undergraduate programmes, while Graph 5.9 refers to Business programmes. Each cell represents a course within a university, so for instance courses 234 and 467 are Law programmes taught at different universities. The solid line is the constant hazard function and dashed lines represent hazard functions if random effects are introduced. I show the examples by using baseline models instead of full models for simplicity reasons. I have also selected courses in order to make the differences clear, as a number of random effects are not statistically different from zero.

Random intercepts account for a set of characteristics to be found at the course level. Following Tinto's (1975, 2010) student integration model, these characteristics, which I do not observe directly, may make the difference in terms of the student experience and, more critically, the matching between student's and HEI's expectations which is, according to that theory, the core mechanism behind the decision of withdrawing or persisting.

Graph 5.8. Fitted completion hazard probabilities for law courses

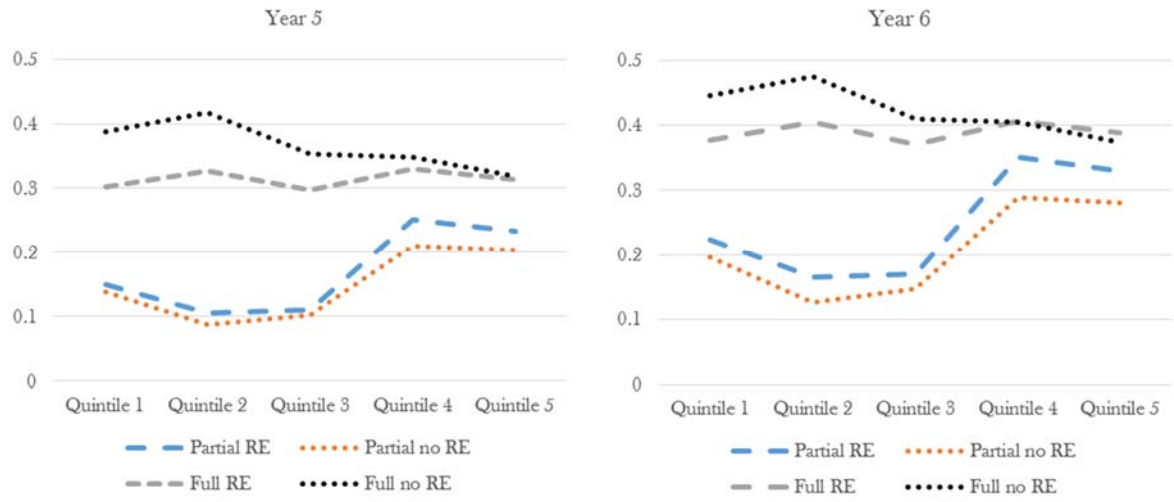


Graph 5.9. Fitted completion hazard probabilities for business courses



Interacted models in Table 5.7 are less extreme than the ones in Table 5.6. This does not necessarily mean that random effects attenuated the gaps for fully and partially CAE funded students but reflects the difference between coefficients referring to marginal and subject-specific effects. Anyway, there still remains the question of whether or not the introduction of random intercepts makes any difference. In order to illustrate this, I compare fitted probabilities between single level and random effects models (Graph 5.10).

Graph 5.10. Fitted hazard probabilities of completion by income quintile; fully and partially CAE funded students. Single level and random effects models



Dotted lines represent single-level fitted probabilities (as in Graph 5.5) whereas dashed lines correspond to random effect estimates. The introduction of random effects attenuates both the probabilities of completing studies for either partially and fully CAE funded students, as well as narrows the gap between both debt composition levels, which is consistent with the fact that coefficients are much less extreme in random effects models (as compared to single level models). These results must be taken with caution as the fitted probabilities shown are estimates at specific values of time and keeping the rest of the values as observed. That means that results would change if other values were used to calculate probabilities instead.

The introduction of random effects, presumably representing a set of variables operating at institutional rather than individual level, slightly alters the estimates by attenuating the coefficients as well as the predicted probabilities. Nevertheless, it is

necessary to insist that controlling for unobserved heterogeneity is not equivalent to deal with confounders at the individual level so it does not solve the issue of omitted variable bias.

Another way of examining the results is taking average marginal effects for the variables. As marginal effects including random effects are computer intensive and extremely time demanding, I use the full single level model and report the effects in Table 5.8. Marginal effects inform the change in the probability when a particular predictor increase in one unit or when switching from one category to another in the case of categorical variables.

Table 5.8. Average marginal effects (Standard errors)

	Outcome					
	Continuing		Dropout		Completion	
	$\frac{dy}{dx}$	SE	$\frac{dy}{dx}$	SE	$\frac{dy}{dx}$	SE
Year						
2	0.017	(0.009)	-0.018	-0.009	0.001	(0.001)
3	0.022	(0.013)	-0.024	(0.012)*	0.002	(0.001)
4	-0.005	(0.017)	-0.031	(0.013)*	0.036	(0.009)***
5	-0.296	(0.040)***	-0.032	(0.015)*	0.327	(0.037)***
6	-0.353	(0.055)***	-0.034	(0.014)*	0.387	(0.053)***
Partially CAE						
funded=1	0.040	(0.009)***	0.003	(0.005)	-0.044	(0.007)***
Cum. Debt	0.011	(0.010)	0.001	(0.005)	-0.012	(0.008)
Income Quintile						
2	0.010	(0.006)	-0.006	(0.004)	-0.003	(0.005)
3	0.007	(0.006)	-0.006	(0.004)	-0.001	(0.006)
4	0.003	(0.006)	-0.010	(0.004)*	0.007	(0.005)
5	-0.004	(0.006)	-0.009	(0.004)*	0.013	(0.005)*
Secondary School						
Public	-0.007	(0.007)	0.005	(0.003)	0.002	(0.006)
Maintained	-0.004	(0.006)	0.001	(0.003)	0.004	(0.005)
Male=1	0.000	(0.005)	0.011	(0.002)***	-0.010	(0.004)*
Age	-0.005	(0.001)***	0.002	(0.000)***	0.003	(0.001)***
PSU	-0.011	(0.004)**	-0.008	(0.002)***	0.019	(0.004)***
GPA	-0.007	(0.002)**	-0.004	(0.001)***	0.011	(0.002)***
Cruch Uni=1	0.008	(0.011)	-0.010	(0.005)*	0.003	(0.011)
Complexity Index	-0.001	(0.005)	0.006	(0.003)*	-0.005	(0.005)
Expected income	0.017	(0.005)***	0.003	(0.001)	-0.020	(0.005)***

* p<0.05, ** p<0.01, *** p<0.001

If a student with given characteristics, being fully CAE supported, switched to partially CAE supported, their completion probabilities would decrease by 4.4 percentage points. For the same case, probabilities of continuation would increase by 4 percentage points. This suggests that for partially funded students their probabilities of completing courses decrease but at the same time their probabilities of being still enrolled are increased. This can be interpreted in different ways. First, it can be said that for students with less harsh finances is not an issue to spend more time before graduating. But a second plausible explanation could be that harsher finances actually encourage students to finish on time. I maintain that the most important mechanism here is a deterrent effect so this result should be seen as a harsher debt pushing students rather than softer debts favouring longer study periods.

The effect of debt composition on dropout is rather negligible. If the debt composition variable were a proxy of harshness, the results presented in this chapter would suggest that harsher student loans affect completion positively. The effect is the opposite when the outcome is continuing studies. Nevertheless, these marginal effects are average effects at observed values and only provide a general picture.

5.6.Final Remarks

The present chapter provides some evidence suggesting that harsher loans are associated with higher probabilities of study completion and that the effect of debt composition is noticeable at low socioeconomic levels. Nevertheless, the positive effect of debt composition on course completion differ across the models specified. Indeed, debt composition was significant in all single level models but only in two random effects models. Data limitations and the extent to which these results explain the underlying mechanisms are powerful enough reasons in order to take these results with caution.

Having seen the above results, the question that still remains is whether it is sensible to conclude that harsher debts boost completion probabilities for poor students, prevent them from dropping out, and that the debt composition does not make differences among the better-off students. It seems feasible to link this finding to the

fact that poor families are more risk averse, which is a well-documented fact in the literature, so that poor students, being fully supported by CAE, face a more distressing experience with regards to indebtedness and, in consequence, are more compelled to complete their studies or persist than partially CAE funded poor students. So the fact that debt composition only makes the difference for poor students seems plausible. I leave this as a hypothesis because I have no direct means of testing it.

Nevertheless, there are other mechanisms operating that needs to be considered. First, fitted probabilities suggest that poor fully funded CAE students are more likely to finish their studies than their richer counterparts. Not only the harsh debt may be more effective to push poor students to complete studies but the explanation might be higher ability. Poor students have arrived at the same place but after a more difficult trajectory. Presumably they were educated at a lower quality school, had less exposition to educational and cultural assets in the family, and lack some soft skills that may be crucial to succeed in an academic environment. Richer students, in turn, have had advantages in school quality and a social environment that help them to succeed in the university. Hence, the higher performance of poor students with harsher loans, in comparison to better-off students might not be due to harsher debts not deterring richer student but reflect a different level of ability.

Considering Chilean aid policy rules, partially CAE funded students are in fact supported by a combination of grants and income contingent loans. This happens to all students with the exception of those in the highest income quintile. Therefore, it is not risky to assume that poor students count on an appropriate level of financial support in order to pay tuition fees so not receiving the appropriate amount of support does not hold as an explanation.

Results so far seem to indicate that debt composition, which I assume as a proxy of harshness, makes a difference on outcomes which is localised on the lower SES students. Nevertheless, the main policy implication should not be a move towards bank loans without first understanding the underlying process, identifying the most affected population –in positive or negative terms- and uncovering subtle behavioural

implications. This of course features a dilemma between a policy that seems to be effective in relation to the outcomes studied here and a deterrent effect unfairly affecting the poorest students. The analysis would also benefit from the introduction of a variable accounting for ability and, more crucially, the interplay between ability and the other variables included in models. How aid composition affects academic outcomes for students with different level of ability should be the question to answer in future research. Although it does not seem easy to find a measure of ability, currently Chilean administrative records allow tracking students back in the school system so that it would be feasible to model school trajectories, progression, rankings and other milestones in the school trajectory that may enrich the analysis.

The introduction of random effects allowed for more flexible estimates but what they actually measure is still a sort of black box. More research is required on student academic support, curricular innovation and new pathways within the Chilean higher education system in order to reach a deeper understanding of what happens within universities. In the same line, it is necessary to understand the perception of indebtedness throughout the socio-economic spectrum and what sort of behavioural response it may trigger.

No administrative dataset can provide the sort of insight required to study behavioural responses but there is an opportunity to open new research fields. However, the perception of indebtedness may also be modelled by specific cultural/national factors that may depend on the very features of HE access, i.e. admissions, funding, support, the perceived value of education, the student experience, and their social and individual returns. There might also be other country-specific characteristics such as a culture of repayment, the exposition of households to indebtedness, and the country income level that contend the generalisability of this chapter's findings.

6. CONCLUSIONS

6.1. Main Findings

I started this thesis by paying attention to three major issues. First, the increasing marketisation of HE, which has transformed one of the most resilient institutions in the western world, the university. Marketisation arrived together with massification and in fact, it was seen in some countries as a means of increasing HE supply without involving substantially more public expenditure. Although it may seem contradictory, rapid growing HE systems have relied on the market and the emergence of private funding mechanisms. The policy response to equity concerns has been student aid, which seeks specifically to overcome financial barriers. The second major issue is that, in terms of equity, a great deal is at stake outside HE. What HE policy and the own HEIs are able to do to assure social equity is more limited that one could think at a first sight. Segregation in the school system, quality of school education according to income, and, critically, an unequal distribution of learning outcomes are factors that impede equitable access to HE. The third issue is that the design of aid policies plays a key role in achieving equity goals in HE. Existing research, as well as this piece, supports tailoring and packaging student aid to a different type of students. For

example, the poorest students might be better supported through scholarships rather than loans because of risk aversion. Nevertheless, as seen in Chapter 5, the policy design matters even in the case of very specific aid programmes. As aid design may trigger unforeseen behavioural responses, it transpires that future research should take on the opportunity of studying attitudes and behavioural responses that student aid design may trigger.

In this thesis, I have worked towards providing an answer to whether there is a cross country effect of inequality of school achievement on HE participation. Does an increase in the socioeconomic inequality of school achievement translate into lower participation in HE once as series of relevant variables are kept constant? Chapter 3 was devoted to addressing this topic. The key results were that (i) inequality of school achievement is an obstacle for HE expansion, (ii) considering countries as the units of analysis, enrolment rates in HE are not related to an absolute level of performance as measured by PISA although what the first results suggested was that PISA performance does matter for increasing HE coverage. This finding was conflicting with the fact that not only do high achievement countries have high enrolment rates in HE but also coverage may be high in countries where school performance is comparatively poor. By controlling for unobserved time-invariant country fixed effects, I obtained estimates that might be seen as challenging. The clear effect of PISA reading score on HE enrolment rate vanished but a mild and significant effect of inequality of school achievement remained. These findings are in line with other research pieces which highlight the relevance of prior achievement in the school for access to HE but my estimates apply at country level. Despite data limitations, the results were robust to the inclusion of control variables as well as different treatments for missing values. Nevertheless, the most important methodological issues are the low statistical power and the likelihood of having obtained overfit estimates.

A relevant amount of data and indicators at countries level are publicly accessible from international organisations but, with a few exceptions, they remain surprisingly underused. Even though macro data might oversimplify and provide less insightful evidence than microdata, providing fairly solid evidence indicating that HE

participation is country specific should be encouraging enough for researchers to make the most of it in the same way as the studies on the effect of education on economic growth. In contrast to the common assumption of cross-country data being low quality and lacking comparability, a careful selection of indicators, a sensible revision of definitions, and a rather simple screening of data patterns in order to identify variables with the most observations made possible to build a panel with which, although limited, I obtained estimates on a substantive topic.

I have also conducted an impact evaluation of a major reform to the student aid system in Chile. The reform was massive in terms of volume of investment and the number of students it benefited. Having appropriate data was the very challenge I had to deal with. Sensibly, I obtained a good approximation to the true causal effect of the 2005 reform. The evidence suggests an important effect on the probability of access to higher education. Between 2003 and 2009, the policy would increase the probability of access to HE for the poorest students by 6 percentage points. Considering that the probability of access for quintiles 1 and 2 was about 30 per cent prior the reform, the policy impact is indeed relevant. An alternative regression-based estimation technique generated consistent results.

Nevertheless, working out the data and choosing the appropriate treatment and control groups required me to make assumptions and simplifications which, though unlikely to affect the results, may have hidden more complex mechanisms. In effect, I had to balance the chances of having better estimates with making sure that the differences between treatment and control groups were meaningful. A key issue with the data was that sample estimates on the number of students having loans and scholarships were inconsistent with official figures so that I had to stick to the rules of student aid as stated in the legislation. Whether this was due to the sample design or an issue with the appropriate questionnaire formulation remains unanswered so far.

Despite the above pitfalls, this is the only piece of research attempting to measure the impact of this massive policy change in Chile. Other investigations have used regression discontinuity approach to evaluate the impact of changes in specific

programmes rules or aid packages but for limited periods, and have struggled to find relevant results because of the difficulty of isolating specific instruments which operate jointly. More precise impact evaluations of the 2005 reform are not possible as there are no administrative records for the pre-reform years. Nevertheless, as enrolment rates by quintile continued to increase after the evaluation period along with more changes in aid policy, the following step is obviously to make use of new data sources.

How aid is designed matters. As shown in the literature, student aid works better when it is tailored and packaged to meet the need and match the characteristics of different groups. Nevertheless, much of the discussion on aid design tries to address the issue of risk aversion of poor students. The main contribution I made with this regards is to link debt composition to harshness in terms of repayment. The originality of this approach lies in a different understanding of deterrent effects, moving the focus from loan take up to persistence. The mechanism operates in a different fashion: instead of just deterring poor students from being indebted, harsh aid may also dissuade students from dropping out in relation to how severe the conditions agreed are perceived. In other words, debt conditions may deter students at different decision points: before and after loan take up. Research methods proved fit to purpose and made it possible to take full advantage of the unique data and circumstances. From a methodological point of view, the main innovation was to include a shared frailty - individual frailties are the most common- in a discrete survival model which allowed me to estimate the variance of random intercepts. The meaning and interpretation of the random effect as a proxy for the concrete student experience and its smoothing effect compared to single level models are consistent with the theory.

6.2. Policy Implications

The findings that this thesis provides are of the highest importance for policymaking. They show that for a given country access to HE might be in part influenced by the socioeconomic distribution of achievement in school education. Some research referenced elsewhere in this thesis points out that low school achievement acts as a barrier for socioeconomic participation gaps to narrow in HE. Cross-country estimates suggest that there is a cross country effect of the socioeconomic distribution of

achievement on access to HE. The policy implication is clear: the claims for meritocratic access to HE, on the basis of individuals' academic achievement disregarding social, cultural and economic considerations, does not hold without an equitable school education. Although necessary this is not a sufficient condition.

The Chilean case illustrates the statement above. Indeed, it is a country with high HE participation regardless the high level of socioeconomic segregation affecting the school system. Chilean HE has experienced a steady growth which most politicians promote as an achievement. Nevertheless, as access opportunities have improved with student aid and so have low-income students' opportunities, inequalities brought from the schools are transmitted to the higher level through the emergence of a residual-second-tier sector serving poor students, which has developed at the expense of financial aid, in a similar fashion to American for-profit colleges. Chilean HE is divided between a bunch of academically and socially selective universities and a non-selective-low quality sector made up of either universities, IPs or CFTs. Non-university HE, without being academically selective, is divided between a handful of big-sized high-quality institutions ones and low-quality ones.

A softly regulated HE with a relevant share of private independent HEIs, alongside a generous student aid, which is the government's preferred channel to fund the system, at least needs an assessment. High tuition fees, which translates into high indebtedness, may not necessarily have a correlate in terms of education quality and, more crucially, valued qualifications. To the extent which the Chilean government does not undertake a substantial reform to the institutional foundations of HE, it may end up -if it has not already- harming poor students by setting unrealistic expectations and incentivising them to contract harsh credits in exchange of low-value qualifications; deceiving them in other words. Although the benefits of expanding student aid to historical levels has had an enormous impact on chances of low income students' participation, there needs to be policies directed to the supply side. For instance, the scale of the enrolment expansion has not had a correlate in terms of academic staff development; the quality assurance system need to be streamlined to the student experience as a whole, beyond management, and set higher standards to

a bunch of second tier HEIs which have relied on student aid. Both policies operating on the institutional side are crucial for the sustainability of the HE system and need to be taken alongside equity policies.

As governments have been effective in making it possible for low SES to enter HE, at the same time it is expectable for HEIs to receive students lacking basic academic skills. There are just small-scale remediation policies and their effectiveness has not been assessed beyond specific HEIs. Yet Chilean HE is producing engineers with severe academic skills gaps in maths and science, school teachers with less than basic reading comprehension levels, and technicians with no relevant skills for the labour market.

More critical, however, is that promising more access to HE without improving quality and equity at the secondary level will not bring good results. The lesson is that demand-side funding might boost participation figures but there needs to be an opportune governmental action beyond student funding in order to make sure that young people acquire relevant qualifications, instead of a broken promise of a life-changing opportunity.

The reform undertaken in 2005 had a significant impact on access to HE of the poorest student. Beyond the critique and the preventions that I made above, the new policy has proven effective in incorporating otherwise excluded students. One might say that Chile has been successful in giving access to HE to low SES students and there is ground for politicians to claim it as an achievement. Nevertheless, the priority needs a shift towards making big decisions and long-term goals. The radical reform introduced by the military dictatorship, without political opposition took about two decades in order to consolidate a new model. Reforms in democracy will not crystallise faster.

6.3. Towards a Future Research Agenda

I have identified the following research prospects on the basis of this as the natural continuation of the research undertaken.

Firstly, there is a potential for further research with cross-country data which are likely to improve their quality and cover more issues on education. There are more data sources to conduct research aimed at finding explanations to educational outcomes. The variety and expansion of international tests to countries other than OECD-developed ones will allow researchers to widen the study the relationships between skills and educational performance with social, economic and educational outcomes.

Secondly, the design of student aid with the consequent behavioural effects it may induce remains as an unexplored field offering a high potential for research. Psychological mechanisms and the behavioural response to harsh and distressful conditions are central to a more thorough understanding of the subtle incentive structures behind aid design which are hardly accessible with quantitative methods so probably requiring a mixed methods approach. I maintain that research must go well beyond assessing the effectiveness of student aid to boost participation and persistence in HE but instead should seek to uncover otherwise hidden unfair mechanisms or even introduce ethical issues.

Last, but not least, the use of random intercepts in chapter 5 should be the first attempt towards using the student experience as a key explanatory variable when using quantitative methods. Developing good proxies by recording and making use of process data such as attendance, use of libraries, borrowed books, IT connection times, study spaces, and participation in non-academic activities should guide data management at HEIs. From the point of view of government data sources, in the case of Chile, there currently exist high-quality administrative data sources in HE. It is only a matter of time to count on mature cohorts allowing the application of the econometric and statistical toolbox.

REFERENCES

Adelman, Clifford. 2004. ‘Principal Indicators of Student Academic Histories in Postsecondary Education, 1972-2000.’ *US Department of Education*. Retrieved 26 June 2016 (http://eric.ed.gov/?id=ED483154).

Ai, Chunrong and Edward C. Norton. 2003. ‘Interaction Terms in Logit and Probit Models’. *Economics Letters* 80(1):123–129.

Allison, Paul D. 2014. *Event History and Survival Analysis*. 2nd edition. Los Angeles: SAGE Publications, Inc.

Alon, Sigal. 2009. ‘The Evolution of Class Inequality in Higher Education Competition, Exclusion, and Adaptation’. *American Sociological Review* 74(5):731–755.

Anderson, Elizabeth S. 1999. ‘What Is the Point of Equality?’ *Ethics* 109(2):287–337.

Armanet, Pilar and Daniel Uribe. 2005. ‘Escenarios futuros de la educación superior chilena: desafíos de medición’. in *Chile en la tarea de superar las brechas de la desigualidad / aspectos conceptuales y de medición*. Santiago de Chile: Instituto Nacional de Estadísticas.

Arriagada, Patricio. 1989. *Financiamiento de La Educación Superior En Chile 1960-1988*. Santiago de Chile: FLACSO.

Aucejo, Esteban M. and Jonathan James. 2015. ‘The Path to College Education: Are Verbal Skills More Important than Math Skills?’ Retrieved 13 March 2016 (https://wpcarey.asu.edu/sites/default/files/uploads/department-economics/college.pdf).

Baird, J. et al. 2011. *Policy Effects of PISA*. Oxford University Centre for Educational Assessment Oxford.

Baker, Therese L. and William Vélez. 1996. ‘Access to and Opportunity in Postsecondary Education in the United States: A Review’. *Sociology of Education* 69:82–101.

Barr, N. A. 2001. *The Welfare State as Piggy Bank: Information, Risk, Uncertainty, and the Role of the State*. Oxford; New York: Oxford University Press.

Barr, Nicholas. 2003. ‘Financing Higher Education: Comparing the Options’. *London School of Economics and Political Science*. Retrieved 1 September 2013 (http://econ.lse.ac.uk/staff/nb/Barr_HE_options030610.pdf).

Barr, Nicholas. 2004. ‘Higher Education Funding’. *Oxford Review of Economic Policy* 20(2):264–83.

Barr, Nicholas. 2005. ‘Financing Higher Education’. *Finance and Development*. Retrieved 9 March 2016 (http://www.imf.org/external/pubs/ft/fandd/2005/06/barr.htm).

Barr, Nicholas. 2012. *Economics of the Welfare State*. Oxford: Oxford University Press.

Barro, Robert J. 1991. 'Economic Growth in a Cross Section of Countries'. *The Quarterly Journal of Economics* 106(2):407–443.

Barro, Robert J. 2001. 'Human Capital and Growth'. *The American Economic Review* 91(2):12–17.

Barro, Robert J. and Jong Wha Lee. 2013. 'A New Data Set of Educational Attainment in the World, 1950–2010'. *Journal of Development Economics* 104:184–98.

Bean, John P. 1980. 'Dropouts and Turnover: The Synthesis and Test of a Causal Model of Student Attrition'. *Research in Higher Education* 12(2):155–87.

Bean, John P. 1981. 'The Synthesis of a Theoretical Model of Student Attrition'. Los Angeles, CA. Retrieved 5 June 2015 (http://eric.ed.gov/?id=ED202444).

Bean, John P. 1982. 'Conceptual Models of Student Attrition: How Theory Can Help the Institutional Researcher'. *New Directions for Institutional Research* 1982(36):17–33.

Bean, John P. 1985. 'Interaction Effects Based on Class Level in an Explanatory Model of College Student Dropout Syndrome'. *American Educational Research Journal* 22(1):35–64.

Becker, Gary Stanley. 1993. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 3rd ed. Chicago ; London: University of Chicago Press.

Bernasconi, Andrés and Fernando Rojas. 2004. *Informe Sobre La Educación Superior En Chile, 1980-2003*. Santiago de Chile: Editorial Universitaria.

Bernstein, Basil. 2003. *Theoretical Studies Towards a Sociology of Language (Class, Codes and Control Vol 1)*. 1 edition. London; New York: Routledge.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. 'How Much Should We Trust Differences-In-Differences Estimates?' *The Quarterly Journal of Economics* 119(1):249–75.

Bettinger, Eric. 2004. *How Financial Aid Affects Persistence*. National Bureau of Economic Research. Retrieved 16 December 2015 (http://www.nber.org/papers/w10242).

Blossfeld, Hans-Peter and Alfred Hamerle. 1989. 'Unobserved Heterogeneity in Hazard Rate Models: A Test and an Illustration from a Study of Career Mobility'. *Quality and Quantity* 23(2):129–141.

Booij, Adam S., Edwin Leuven, and Hessel Oosterbeek. 2012. 'The Role of Information in the Take-up of Student Loans'. *Economics of Education Review* 31(1):33–44.

Bourdieu, Pierre. 1979. *The Inheritors: French Students and Their Relation to Culture*. Chicago ; London: University of Chicago Press.

Bourdieu, Pierre and Jean-Claude Passeron. 1986. *Distinction: A Social Critique of the Judgement of Taste*. London: Routledge.

Bowles, Samuel and Herbert Gintis. 1976. *Schooling in Capitalist America: Educational Reform and the Contradictions of Economic Life*. New York: Routledge & Kegan Paul.

Bowles, Samuel and Herbert Gintis. 2002. 'Schooling in Capitalist America Revisited'. *Sociology of Education* 75(1):1–18.

Braunstein, Andrew, Michael Mcgrath, and Donn Pescatrice. 1999. 'Measuring the Impact of Income and Financial Aid Offers on College Enrollment Decisions'. *Research in Higher Education* 40(3):247–59.

Brennan, John. 2010. 'Burton Clark's The Higher Education System: Academic Organization in Cross-National Perspective'. *London Review of Education* 8(3):229–37.

Brennan, John and Rajani Naidoo. 2008. 'Higher Education and the Achievement (And/Or Prevention) of Equity and Social Justice'. *Higher Education* 56(3):287–302.

British Council. 2012. *The Shape of Things to Come: Higher Education Global Trends and Emerging Opportunities to 2020*. British Council.

Bruinsma, Marjon and Ellen P. W. A. Jansen. 2009. 'When Will I Succeed in My First-year Diploma? Survival Analysis in Dutch Higher Education'. *Higher Education Research & Development* 28(1):99–114.

Brunner, José Joaquín. 1993. 'Chile's Higher Education: Between Market and State'. *Higher Education* 25(1):35–43.

Brunner, José Joaquín and Daniel Uribe. 2007. *Mercados universitarios: el nuevo escenario de la educación superior*. Santiago, Chile: Ediciones Universidad Diego Portales.

Button, Katherine S. et al. 2013. 'Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience'. *Nature Reviews Neuroscience* 14(5):365–76.

Cabrera, Alberto F., Amaury Nora, and Maria B. Castañeda. 1992. 'The Role of Finances in the Persistence Process: A Structural Model'. *Research in Higher Education* 33(5):571–93.

Cabrera, Alberto F., Amaury Nora, and Maria B. Castaneda. 1993. 'College Persistence: Structural Equations Modeling Test of an Integrated Model of Student Retention'. *The Journal of Higher Education* 64(2):123–39.

Canton, Erik and Andreas Blom. 2004. 'Can Student Loans Improve Accessibility to

Higher Education and Student Performance?' *World Bank Policy Research Working Paper* 3425. Retrieved 2 September 2013 (http://www.eldis.org/vfile/upload/1/document/0708/DOC16860.pdf).

Canton, Erik and Andreas Blom. 2010. 'Student Support and Academic Performance: Experiences at Private Universities in Mexico'. *Education Economics* 18(1):49–65.

Card, David. 1999. 'The Causal Effect of Education on Earnings'. *Handbook of Labor Economics* 3:1801–1863.

Card, David. 2000. *Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems*. National Bureau of Economic Research. Retrieved 28 May 2014 (http://www.nber.org/papers/w7769).

Carneiro, Pedro Manuel and James J. Heckman. 2003. *Human Capital Policy*. Rochester, NY: Social Science Research Network. Retrieved 27 May 2014 (http://papers.ssrn.com/abstract=434544).

Castañeda, Tarsicio. 1990. *Para combatir la pobreza: política social y descentralización en Chile durante los '80*. Centro de Estudios Públicos.

Cellini, Stephanie Riegg. 2010. 'Financial Aid and for-Profit Colleges: Does Aid Encourage Entry?' *Journal of Policy Analysis and Management* 29(3):526–52.

Chapman, Bruce. 2006. 'Income Contingent Loans for Higher Education: International Reforms'. Pp. 1435–1503 in *Handbook of the Economics of Education*, vol. 2, edited by E. Hanushek and F. Welch. Elsevier. Retrieved 2 September 2013 (http://www.sciencedirect.com/science/article/pii/S1574069206020253).

Chapman, Bruce and Chris Ryan. 2005. 'The Access Implications of Income-Contingent Charges for Higher Education: Lessons from Australia'. *Economics of Education Review* 24(5):491–512.

Chen, Jin and Desiree Zerquera. 2011. 'A Methodological Review of Studies on Effects of Financial Aid on College Student Success'. in *Annual Conference of the Association for Education Finance and Policy, Seattle, WA*. Retrieved 8 June 2015 (https://aefpweb.org/sites/default/files/webform/Methodological%20Review2.pdf).

Chen, Rong and Stephen L. DesJardins. 2008. 'Exploring the Effects of Financial Aid on the Gap in Student Dropout Risks by Income Level'. *Research in Higher Education* 49(1):1–18.

Chen, Rong and Stephen L. DesJardins. 2010. 'Investigating the Impact of Financial Aid on Student Dropout Risks: Racial and Ethnic Differences'. *The Journal of Higher Education* 81(2):179–208.

Chowdry, Haroon, Claire Crawford, Lorraine Dearden, Alissa Goodman, and Anna

Vignoles. 2013. ‘Widening Participation in Higher Education: Analysis Using Linked Administrative Data’. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176(2):431–457.

Clark, Burton R. 1973. ‘Development of the Sociology of Higher Education’. *Sociology of Education* 46(1):2–14.

Clark, Burton R. 1983. *The Higher Education System: Academic Organization in Cross-National Perspective*. Berkeley: University of California Press.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2 edition. Hillsdale, N.J: Routledge.

Corak, Miles, Garth Lipps, and John Zhao. 2004. *Family Income and Participation in Post-Secondary Education*. St. Louis, United States: Federal Reserve Bank of St Louis. Retrieved 14 June 2016 (http://search.proquest.com.libproxy.ucl.ac.uk/docview/1698642317?rfr_id=info%3Axri%2Fsid%3Aprimo).

Crawford, Claire. 2015. ‘Socio-Economic Differences in University Outcomes in the UK: Drop-Out, Degree Completion and Degree Class’. *Labour* 27:02.

Darolia, Rajeev. 2013. ‘Integrity versus Access? The Effect of Federal Financial Aid Availability on Postsecondary Enrollment’. *Journal of Public Economics* 106:101–14.

De Gregorio, José De and Jong–Wha Lee. 2002. ‘Education and Income Inequality: New Evidence From Cross-Country Data’. *Review of Income and Wealth* 48(3):395–416.

Dearden, Lorraine, Emla Fitzsimons, and Gill Wyness. 2014. ‘Money for Nothing: Estimating the Impact of Student Aid on Participation in Higher Education’. *Economics of Education Review* 43:66–78.

Denson, Katy and Randall E. Schumacker. 1996. ‘Student Choices: Using a Competing Risks Model of Survival Analysis.’ Retrieved 27 January 2015 (http://eric.ed.gov/?id=ED399341).

DesJardins, S. L., D. A. Ahlburg, and B. P. McCall. 1999. ‘An Event History Model of Student Departure’. *Economics of Education Review* 18(3):375–90.

DesJardins, Stephen L., Dennis A. Ahlburg, and Brian P. McCall. 2002a. ‘A Temporal Investigation of Factors Related to Timely Degree Completion’. *The Journal of Higher Education* 73(5):555–81.

DesJardins, Stephen L., Dennis A. Ahlburg, and Brian P. McCall. 2002b. ‘Simulating the Longitudinal Effects of Changes in Financial Aid on Student Departure from College’. *The Journal of Human Resources* 37(3):653–79.

DesJardins, Stephen L. and Brian P. McCall. 2010. ‘Simulating the Effects of Financial Aid Packages on College Student Stopout, Reenrollment Spells, and Graduation Chances’. *Review of Higher Education* 33(4):513–41.

Desormeaux, Jorge and Matko Koljatic. 1990. 'El financiamiento de la educación superior desde una perspectiva libertaria'. in *Antecedentes y tendencias en el sistema de financiamiento de la educación superior chilena*, edited by C. Lehmann. Santiago de Chile: Centro de Estudios Públicos.

Díaz, Christian. 2008. 'Modelo Conceptual Para La Deserción Estudiantil Universitaria Chilena'. *Estudios Pedagógicos* 34(2):65–86.

Dill, David D. 1997. 'Higher Education Markets and Public Policy'. *Higher Education Policy* 10(3–4):167–85.

Donoso, Sebastián and Ernesto Schiefelbein. 2007. 'Análisis de Los Modelos Explicativos de Retención de Estudiantes En La Universidad: Una Visión Desde La Desigualdad Social'. *Estudios Pedagógicos* 33(1):7–27.

Dowd, Alicia C. and Tarek Coury. 2006. 'The Effect of Loans on the Persistence and Attainment of Community College Students'. *Research in Higher Education* 47(1):33–62.

Dupont, William D. and Walton D. Plummer. 1998. 'Power and Sample Size Calculations for Studies Involving Linear Regression'. *Controlled Clinical Trials* 19(6):589–601.

Durkheim, Emil. 2006. *On Suicide*. Third Edition. London: Penguin Classics.

Duru-Bellat, Marie and Bruno Suchaut. 2005. 'Organisation and Context, Efficiency and Equity of Educational Systems: What PISA Tells Us'. *European Educational Research Journal* 4(3):181–94.

Dworkin, Ronald. 1981. 'What Is Equality? Part 2: Equality of Resources'. *Philosophy & Public Affairs* 10(4):283–345.

Dynarski, Mark. 1994. 'Who Defaults on Student Loans? Findings from the National Postsecondary Student Aid Study'. *Economics of Education Review* 13(1):55–68.

Dynarski, Susan M. 2003. 'Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion'. *The American Economic Review* 93(1):279–88.

Dynarski, Susan and Judith Scott-Clayton. 2013. 'Financial Aid Policy: Lessons from Research'. *The Future of Children* 23(1):67–91.

Elacqua, Gregory. 2009. 'The Impact of School Choice and Public Policy on Segregation: Evidence from Chile'. Retrieved 15 August 2016 (http://www.economiaynegocios.uahurtado.cl/wp-content/uploads/2010/08/25-9-09.pdf).

Espinoza, Oscar. 2008. 'Creating (in) Equalities in Access to Higher Education in the Context of Structural Adjustment and Post-Adjustment Policies: The Case of Chile'. *Higher Education* 55(3):269–84.

Espinoza, Oscar, Luis Eduardo González, and Daniel Uribe. 2009. 'Movilidad Social En Chile: El Caso Del Gran Santiago Urbano'. *Revista de Ciencias Sociales* 15(4):586–606.

Finnie, Ross. 2012. 'Access to Post-Secondary Education: The Importance of Culture'. *Children and Youth Services Review* 34(6):1161–70.

Finnie, Ross and Alex Usher. 2006. 'The Canadian Experiment in Cost-Sharing and Its Effects on Access to Higher Education, 1990–2002'. Pp. 159–87 in *Cost-Sharing and Accessibility in Higher Education: A Fairer Deal?*, *Higher Education Dynamics*, edited by P. N. Teixeira, D. B. Johnstone, M. J. Rosa, and H. Vossensteyn. Springer Netherlands. Retrieved 8 September 2013 (http://link.springer.com/chapter/10.1007/978-1-4020-4660-5_7).

Fischbach, Antoine, Ulrich Keller, Franzis Preckel, and Martin Brunner. 2013. 'PISA Proficiency Scores Predict Educational Outcomes'. *Learning and Individual Differences* 24:63–72.

Fuchs, Thomas and Ludger Woessmann. 2006. 'What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data'. *Empirical Economics* 32(2–3):433–64.

Gamboa, Luis Fernando and Fábio D. Waltenberg. 2012. 'Inequality of Opportunity for Educational Achievement in Latin America: Evidence from PISA 2006–2009'. *Economics of Education Review* 31(5):694–708.

Gilchrist, R., D. Phillips, and A. Ross. 2007. 'Participation and Potential Participation in UK Higher Education'. Pp. 75–96 in *Higher Education and Social Class:Issues of Exclusion and Inclusion*. London,UK: Routledge Falmer.

Glennerster, Howard, Stephen Merrett, and Gail Wilson. 1968. 'A Graduate Tax'. *Higher Education Review* 1(1):26–38.

Goldthorpe, John H. 2010. 'Class Analysis and the Reorientation of Class Theory: The Case of Persisting Differentials in Educational Attainment'. *The British Journal of Sociology* 61:311–35.

González, Luis Eduardo and Daniel Uribe. 2003. 'Estimaciones Sobre La Repitencia Y Deserción En La Educación Superior'. *Calidad de La Educación* (Segundo semestre). Retrieved 16 June 2014 (http://www.cned.cl/public/secciones/seccionpublicaciones/doc/35/cse_articulo140.pdf).

Gopal, Prashant Gopal. 2008. 'Which College Grads Earn the Most?' *Bloomberg.com*, August 7. Retrieved 15 August 2016 (http://www.bloomberg.com/news/articles/2008-08-07/which-college-grads-earn-the-most-businessweek-business-news-stock-market-and-financial-advice).

van der Haert, Margaux, Elena Arias Ortiz, Philippe Emplit, Véronique Halloin, and Catherine Dehon. 2014. 'Are Dropout and Degree Completion in Doctoral Study Significantly Dependent on Type of Financial Support and Field of

Research?' *Studies in Higher Education* 39(10):1885–1909.

Hansen, W.Lee. 1972. 'Equity and the Finance of Higher Education'. *Journal of Political Economy* 80(3):S260–73.

Hanushek, Eric A. and Dennis D. Kimko. 2000. 'Schooling, Labor-Force Quality, and the Growth of Nations'. *The American Economic Review* 90(5):1184–1208.

Hanushek, Eric A. and Ludger Woessmann. 2007. *The Role of Education Quality in Economic Growth [Electronic Resource]*. Washington, D.C.: World Bank.

Heckman, James J. 1976. 'The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models'. Pp. 475–492 in *Annals of Economic and Social Measurement, Volume 5, number 4*. NBER. Retrieved 29 May 2014 (http://www.nber.org/chapters/c10491.pdf).

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2003. *Fifty Years of Mincer Earnings Regressions*. National Bureau of Economic Research. Retrieved 29 May 2014 (http://www.nber.org/papers/w9732).

Heckman, James J. and Dimitriy V. Masterov. 2007. *The Productivity Argument for Investing in Young Children*. National Bureau of Economic Research. Retrieved 29 May 2014 (http://www.nber.org/papers/w13016).

Herzog, Serge. 2005. 'Measuring Determinants of Student Return VS. Dropout/Stopout VS. Transfer: A First-to-Second Year Analysis of New Freshmen'. *Research in Higher Education* 46(8):883–928.

Hillman, Nicholas W. 2014. 'College on Credit: A Multilevel Analysis of Student Loan Default'. *Review of Higher Education* 37(2):169–95.

Himmel, Erika. 2002. 'Modelos de Análisis de La Deserción Estudiantil En La Educación Superior'. *Calidad de La Educación* 17:91–108.

Horn, Catherine, María Verónica Santelices, and Ximena Catalán. 2014. 'Modeling the Impacts of National and Institutional Financial Aid Opportunities on Persistence at an Elite Chilean University'. *Higher Education* 68(3):471–88.

Hosmer, David W. and Stanley Lemeshow. 1999. *Applied Survival Analysis: Time-to-Event*. Wiley.

Hosmer, David W., Stanley Lemeshow, and Susanne May. 2008. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. 2 edition. Hoboken, N.J: Wiley-Interscience.

Hougaard, Philip. 1995. 'Frailty Models for Survival Data'. *Lifetime Data Analysis* 1(3):255–73.

INGRESA. 2015. 'Estadísticas | Ingresa'. Retrieved 26 May 2015 (http://portal.ingresa.cl/estadisticas-y-estudios/estadisticas/).

Ishitani, Terry T. and Stephen L. DesJardins. 2002. 'A Longitudinal Investigation of Dropout from College in the United States'. *Journal of College Student Retention: Research, Theory & Practice* 4(2):173–201.

Jamison, Eliot A., Dean T. Jamison, and Eric A. Hanushek. 2007. 'The Effects of Education Quality on Income Growth and Mortality Decline'. *Economics of Education Review* 26(6):771–88.

Jerrim, John. 2012. 'The Socio-Economic Gradient in Teenagers' Reading Skills: How Does England Compare with Other Countries?*'. *Fiscal Studies* 33(2):159–184.

Jerrim, John and John Micklewright. 2013. 'Socioeconomic Gradients in Children's Cognitive Skills: Are Cross-Country Comparisons Robust to Who Reports Family Background?' *GINI Discussion Paper* 65. Retrieved (http://www.gini-research.org/system/uploads/520/original/65.pdf?1380552594).

Jerrim, John, Anna Vignoles, and Ross Finnie. 2012. 'University Access for Disadvantaged Children: A Comparison across English Speaking Countries'.

Johnstone, D.Bruce. 2003. 'Cost Sharing in Higher Education: Tuition, Financial Assistance, and Accessibility in a Comparative Perspective'. *Sociologickỳ Časopis/Czech Sociological Review* 351–374.

Johnstone, D.Bruce. 2004. 'The Economics and Politics of Cost Sharing in Higher Education: Comparative Perspectives'. *Economics of Education Review* 23(4):403–10.

Johnstone, D.Bruce and Pamela Marcucci. 2010. *Financing Higher Education Worldwide: Who Pays? Who Should Pay?* Baltimore: Johns Hopkins University Press.

Jones-White, Daniel R., Peter M. Radcliffe, Ronald L.Huesman Jr, and John P. Kellogg. 2009. 'Redefining Student Success: Applying Different Multinomial Regression Techniques for the Study of Student Graduation Across Institutions of Higher Education'. *Research in Higher Education* 51(2):154–74.

Kass, Robert E. and Adrian E. Raftery. 1995. 'Bayes Factors'. *Journal of the American Statistical Association* 90(430):773–95.

van der Klaauw, Wilbert. 2002. 'Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression–Discontinuity Approach*'. *International Economic Review* 43(4):1249–87.

Knight, Carl. 2013. 'Luck Egalitarianism'. *Philosophy Compass* 8(10):924–34.

Le Grand, Julian. 1982. *The Strategy of Equality: Redistribution and the Social Services*. London, UK: Allen & Unwin.

Le Grand, Julian. 2007. *The Other Invisible Hand: Delivering Public Services through*

Choice and Competition. Princeton, NJ, USA: Princeton University Press.

Lehmann, Carla. 1990. 'Antecedentes Y Tendencias En El Sistema de Financiamiento de La Educación Superior Chilena'. in *Financiamiento de la educación superior: Antecedentes y desafíos*, edited by C. Lehmann. Santiago de Chile: Centro de Estudios Públicos.

Lemaitre, María José. 2005. 'Equidad en la educación superior: un concepto complejo'. Retrieved 12 August 2016 (https://repositorio.uam.es/handle/10486/660776).

Letkiewicz, Jodi et al. 2014. 'The Path to Graduation: Factors Predicting On-Time Graduation Rates'. *Journal of College Student Retention: Research, Theory & Practice* 16(3):351–71.

Levy, Daniel C. 1986. *Higher Education and the State in Latin America*. University of Chicago Press.

Liu, Xian. 2014. 'Survival Models on Unobserved Heterogeneity and Their Applications in Analyzing Large-Scale Survey Data'. *Journal of Biometrics & Biostatistics* 5. Retrieved 2 February 2015 (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4267525/).

McDonough, P. and A. Fann. 2007. 'The Study of Inequality'. Pp. 53–93 in *Sociology of Higher Education. Contributions and their Contexts*, edited by P. Gumport. Baltimore: The Jonhns Hopkins University Press.

MDS. 2012. 'Resultados Educación. CASEN 2011'. Retrieved (http://observatorio.ministeriodesarrollosocial.gob.cl/casen/casen-documentos.php?c=108&a=2011).

MDS. 2013. 'Observatorio Social - Ministerio de Desarrollo Social - Gobierno de Chile'. Retrieved 15 November 2013 (http://observatorio.ministeriodesarrollosocial.gob.cl/casen_obj.php).

Melguizo, Tatiana, Fabio Sanchez, and Tatiana Velasco. 2016. 'Credit for Low-Income Students and Access to and Academic Performance in Higher Education in Colombia: A Regression Discontinuity Approach'. *World Development* 80:61–77.

Mincer, Jacob. 1958. 'Investment in Human Capital and Personal Income Distribution'. *The Journal of Political Economy* 66(4):281–302.

MINEDUC. 2007. *Country Background Report - Chile. OECD Thematic Review of Tertiary Education*. Retrieved 16 June 2014 (http://www.oecd.org/education/skills-beyond-school/41473042.pdf).

MINEDUC. 2011. 'Resumen Resultados PISA 2009. Chile'. Retrieved 15 November 2013 (http://ww2.educarchile.cl/UserFiles/P0001/File/PISA/Resumen_Resultados_PISA_2009_Chile.pdf).

Ministry of Education. 2012. ‘Mi Futuro’. Retrieved 30 September 2013 (http://www.mifuturo.cl/).

Mizala, Alejandra. 2007. *La Economía Política de La Reforma Educacional En Chile*. CIEPLAN. Retrieved 15 August 2016 (http://educahoy.cl/wp/wp-content/uploads/2010/10/mizala-economia-educacional.pdf).

Mizala, Alejandra and Pilar Romaguera. 2004. ‘Remuneraciones Y Tasas de Retorno de Los Profesionales Chilenos’. in *Oferta y demanda de profesionales y técnicos en Chile: El rol de la información pública*, edited by J. J. Brunner and P. Meller. Santiago de Chile: RIL Editores.

Mora, José-Ginés. 1997. ‘Equity in Financing and Access to Higher Education’. *Higher Education in Europe* 22(2):145–54.

Murdoch, Jake, Pierre Kamanzi, and Pierre Doray. 2011. ‘The Influence of PISA Scores, Schooling and Social Factors on Pathways to and within Higher Education in Canada’. *Irish Educational Studies* 30(2):215–35.

Nakagawa, Shinichi and Innes C. Cuthill. 2007. ‘Effect Size, Confidence Interval and Statistical Significance: A Practical Guide for Biologists’. *Biological Reviews* 82(4):591–605.

OECD. 2004. *OECD Handbook for Internationally Comparative Education Statistics: Concepts, Standards, Definitions and Classifications*. Paris: Organisation for Economic Co-operation and Development.

OECD. 2005. *School Factors Related to Quality and Equity: Results from PISA 2000*. Paris: Organisation for Economic Co-operation and Development. Retrieved 13 September 2013 (http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/34668095.pdf).

OECD. 2008a. *Education at a Glance 2007*. Paris: Organisation for Economic Co-operation and Development.

OECD. 2008b. *Tertiary Education for the Knowledge Society*. Paris: Organisation for Economic Co-operation and Development.

OECD. 2010. *PISA 2009 Results: Overcoming Social Background. Equity in Learning Opportunities and Outcomes*. Paris: Organisation for Economic Co-operation and Development.

OECD. 2012a. *Education at a Glance 2012: OECD Indicators*. Paris: Organisation for Economic Co-operation and Develpoment.

OECD. 2012b. *PISA 2009 Technical Report*. Paris: Organisation for Economic Co-operation and Development. Retrieved 22 January 2014 (http://www.oecd.org/edu/school/programmeforinternationalstudentassessmentpisa/pisa2009technicalreport.htm#PISA_2006_Technical_Report).

OECD. 2014. *Education at a Glance 2014*. OECD Publishing. Retrieved 19 May 2016

(http://www.oecd-ilibrary.org/education/education-at-a-glance-2014_eag-2014-en).

OECD. 2015a. *Education at a Glance 2015*. OECD Publishing.

OECD. 2015b. *Education at a Glance 2015*. OECD Publishing. Retrieved 1 December 2015 (http://www.oecd-ilibrary.org/education/education-at-a-glance-2015_eag-2015-en).

OECD. n.d. 'PISA Products - Organisation for Economic Co-Operation and Development'.

OECD and The World Bank. 2009. *Reviews of National Policies for Education: Tertiary Education in Chile 2009*. Paris: Organisation for Economic Co-operation and Development.

Oosterbeek, Hessel and Anja van den Broek. 2009. 'An Empirical Analysis of Borrowing Behaviour of Higher Education Students in the Netherlands'. *Economics of Education Review* 28(2):170–77.

Ortiz, Elena Arias and Catherine Dehon. 2013. 'Roads to Success in the Belgian French Community's Higher Education System: Predictors of Dropout and Degree Completion at the Université Libre de Bruxelles'. *Research in Higher Education* 54(6):693–723.

Page, Lindsay C. and Judith Scott-Clayton. 2016. 'Improving College Access in the United States: Barriers and Policy Responses'. *Economics of Education Review* 51:4–22.

Paura, Liga and Irina Arhipova. 2014. 'Cause Analysis of Students' Dropout Rate in Higher Education Study Program'. *Procedia - Social and Behavioral Sciences* 109:1282–86.

Petras, Hanno, Katherine E. Masyn, Jacquelyn A. Buckley, Nicholas S. Ialongo, and Sheppard Kellam. 2011. 'Who Is Most at Risk for School Removal? A Multilevel Discrete-Time Survival Analysis of Individual- and Context-Level Influences.' *Journal of Educational Psychology* 103(1):223–37.

Psacharopoulos, George. 1988. 'Education and Development: A Review'. *The World Bank Research Observer* 3(1):99–116.

Psacharopoulos, George. 2008. 'Funding Universities for Efficiency and Equity: Research Findings versus Petty Politics'. *Education Economics* 16(3):245–60.

Psacharopoulos, George and George Papakonstantinou. 2005. 'The Real University Cost in a "free" Higher Education Country'. *Economics of Education Review* 24(1):103–8.

Psacharopoulos, George and Harry Anthony Patrinos. 2004. 'Returns to Investment in Education: A Further Update'. *Education Economics* 12(2):111–34.

Rabe-Hesketh, Sophia and Anders Skrondal. 2012. *Multilevel and Longitudinal Modeling Using Stata, Volumes I and II, Third Edition*. 3 edition. College Station, Tex: Stata Press.

Ramsey, Alan. 2008. 'Graduate Earnings. An Econometric Analysis of Returns, Inequality and Derpivation across the UK'. Retrieved 15 August 2016 (https://www.economy-ni.gov.uk/sites/default/files/publications/del/Graduate%20Earnings%20An%20econometric%20analysis%20of%20returns%2C%20inequality%20and%20dep%20main%20report.pdf).

Rolando, Rodrigo, Juan Salamanca, and Alfredo Lara. 2010. *Retención de Primer Año En El Pregrado: Descripción Y Análisis de La Cohorte de Ingreso 2007*. Retrieved 23 June 2014 (http://www.mifuturo.cl/images/Estudios/Estudios_SIES_DIVESUP/retencion_primer_ao_pregrado.pdf).

Santelices, María Verónica, Ximena Catalán, Diana Kruger, and Catherine Horn. 2015. 'Determinants of Persistence and the Role of Financial Aid: Lessons from Chile'. *Higher Education* 71(3):323–42.

Sanyal, Bikas C. and D.Bruce Johnstone. 2011. 'International Trends in the Public and Private Financing of Higher Education'. *Prospects* 41(1):157–75.

Shavit, Yossi, Richard Arum, Adam Gamoran, and Gila Menachem. 2007. *Stratification in Higher Education: A Comparative Study*. Stanford University Press.

Singer, J. D. and J. B. Willett. 1993. 'It's About Time: Using Discrete-Time Survival Analysis to Study Duration and the Timing of Events'. *Journal of Educational and Behavioral Statistics* 18(2):155–95.

Singer, Judith D. and John B. Willett. 2003. *Applied Longitudinal Data Analysis*. Oxford University Press. Retrieved 8 June 2015 (http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195152968.001.0001/acprof-9780195152968).

Solis, Alex. 2013. *Credit Access and College Enrollment*. Working Paper, Department of Economics, Uppsala University. Retrieved 14 June 2016 (http://www.econstor.eu/handle/10419/82577).

Spady, William G. 1970. 'Dropouts from Higher Education: An Interdisciplinary Review and Synthesis'. *Interchange* 1(1):64–85.

St John, Edward P. and Jay Noell. 1989. 'The Effects of Student Financial Aid on Access to Higher Education: An Analysis of Progress with Special Consideration of Minority Enrollment'. *Research in Higher Education* 30(6):563–81.

Stampen, Jacob O. and Alberto F. Cabrera. 1988. 'The Targeting and Packaging of Student Aid and Its Effect on Attrition'. *Economics of Education Review* 7(1):29–46.

Steele, Fiona, Ian Diamond, and Duolao Wang. 1996. 'The Determinants of the Duration of Contraceptive Use in China: A Multilevel Multinomial Discrete-Hazards Mdeling Approach'. *Demography* 33(1):12–23.

Steele, Fiona, Harvey Goldstein, and William Browne. 2004. 'A General Multilevel Multistate Competing Risks Model for Event History Data, with an Application to a Study of Contraceptive Use Dynamics'. *Statistical Modelling* 4(2):145–159.

Stratton, Leslie S., Dennis M. O'Toole, and James N. Wetzel. 2008. 'A Multinomial Logit Model of College Stopout and Dropout Behavior'. *Economics of Education Review* 27(3):319–31.

Teixeira, Pedro N., Johnstone, D. B., Maria J. Rosa, and Hans Vossensteyn. 2008. *Cost-Sharing and Accessibility in Higher Education: A Fairer Deal?* Berlin: Springer.

Tinto, Vincent. 1975. 'Dropout from Higher Education: A Theoretical Synthesis of Recent Research'. *Review of Educational Research* 45(1):89–125.

Tinto, Vincent. 2006. 'Research and Practice of Student Retention: What Next?' *Journal of College Student Retention: Research, Theory & Practice* 8(1):1–19.

Tinto, Vincent. 2010. 'From Theory to Action: Exploring the Institutional Conditions for Student Retention'. Pp. 51–89 in *Higher Education: Handbook of Theory and Research*, vol. 25, edited by J. C. Smart. Dordrecht: Springer Netherlands.

Torche, Florencia. 2005. 'Privatization Reform and Inequality of Educational Opportunity: The Case of Chile'. *Sociology of Education* 78(4):316–43.

UNDP. 2005. *Expansión de La Educación Superior En Chile. Hacia Un Nuevo Enfoque de La Calidad Y Equidad*. United Nations Development Program. Retrieved 23 June 2014 (http://www.desarrollohumano.cl/otraspub/pub10/Ed%20superior.pdf).

UNESCO. n.d. 'UNESCO Institute for Statistics'. Retrieved 22 January 2014 (http://stats.uis.unesco.org/unesco/TableViewer/document.aspx?ReportId=136&IF_Language=eng&BR_Topic=0).

Uribe, Daniel. 2004. 'Oferta Educativa Y Oferta de Graduados de Educación Superior'. in *Oferta y Demanda de Profesionales y Técnicos en Chile*, edited by J. J. Brunner and P. Meller. Santiago de Chile: RIL Editores.

Urzua, Sergio and Tomas Rau. 2012. *Higher Education Dropouts, Access to Credit, and Labor Market Outcomes: Evidence from Chile*. Society for Economic Dynamics. Retrieved 8 June 2015 (https://ideas.repec.org/p/red/sed012/228.html).

Volkwein, J.Fredericks, Bruce P. Szelest, Alberto F. Cabrera, and Michelle R. Napierski-Prancl. 1998. 'Factors Associated with Student Loan Default

among Different Racial and Ethnic Groups'. *The Journal of Higher Education* 69(2):206–37.

Walpole, Marybeth. 2007. 'Economically and Educationally Challenged Students in Higher Education: Access to Outcomes'. *ASHE Higher Education Report* 33(3):1–113.

Wao, Hesborn O. 2010. 'Time to the Doctorate: Multilevel Discrete-Time Hazard Analysis'. *Educational Assessment, Evaluation and Accountability* 22(3):227–47.

Willett, John B. and Judith D. Singer. 1991. 'From Whether to When: New Methods for Studying Student Dropout and Teacher Attrition'. *Review of Educational Research* 61(4):407.

Woo, Jennie H. 2002. 'Factors Affecting the Probability of Default: Student Loans in California'. *Journal of Student Financial Aid* 32(2):5–23.

World Bank. 2000. *Higher Education in Developing Countries: Peril and Promise*. Washington, DC: World Bank.

World Bank. 2011. *Chile's State-Guaranteed Student Loan Program (CAE)*. The World Bank. Retrieved 1 June 2015 (http://documents.worldbank.org/curated/en/2011/03/16406290/chiles-state-guaranteed-student-loan-program-cae).

World Bank. n.d. 'Education - Tertiary Education - Projects'. Retrieved 2 September 2013a (http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTEDUCATION/0,,contentMDK:20528441~menuPK:617592~pagePK:148956~piPK:216618~theSitePK:282386~isCURL:Y,00.html).

World Bank. n.d. 'The World Bank DataBank - Create Widgets or Advanced Reports and Share'. Retrieved 22 January 2014b (http://databank.worldbank.org/data/views/variableSelection/selectvariables.aspx?source=world-development-indicators).

APPENDIX A (CHAPTER 3)

Table A.1. Gross enrolment rates and inequality of school achievement by country and period

	Gross Enrollment Rate (HE)				Inequality of School Achievement (R2 ESCS)			
Country	2003	2006	2009	2012	2000	2003	2006	2009
Albania	15.8	18.4		55.5				10.7
Argentina	64.9	67.1	71.2	80.3			13.5	19.0
Australia	72.7	71.2	75.9	86.3	17.4	14.2	11.8	12.7
Austria	46.1	49.3	60.2	72.4	16.6	21.3	12.9	16.0
Azerbaijan			19.0	20.4			7.2	7.4
Belgium	59.7	62.6	67.5	70.8	18.9	22.8	16.0	19.3
Brazil	22.3	25.6			13.9	8.6	13.7	13.0
Bulgaria	41.1	45.7	53.0	62.7			22.7	20.2
Canada	60.0				12.2	9.9	9.5	8.6
Chile	42.9	46.7	59.2	74.4			18.6	18.7
Colombia	24.7	32.0	37.1	45.0			10.3	16.0
Croatia	39.0	45.1	49.2	61.6			10.2	11.0
Czech Republic	36.7	49.3	60.7	64.2	22.9	15.5	12.5	12.4
Denmark	67.4	78.9	74.4	79.6	16.1	16.3	10.7	14.5
Estonia	64.8	65.8	62.7	76.7			7.7	7.6
Finland	87.1	93.3	91.6	93.7	8.6	9.6	7.7	7.8
France	54.4	55.4	54.5	58.3	19.0	19.5	16.4	16.7
Germany			46.3		23.6	22.5	15.9	17.0
Greece	72.2	93.1	89.4	116.6	12.2	10.9	10.7	12.5
Hong Kong SAR, China	31.5	33.4	57.2	60.1	7.8	5.3	6.3	4.5
Hungary	52.6	66.4	61.7	59.6	25.4	21.7	19.4	26.0
Iceland	62.2	72.7	74.1	81.4	7.4	4.0	4.9	6.2
Indonesia	15.7	16.5	22.4	31.5	9.1	6.8	10.8	7.8
Ireland	54.2	57.7	61.0	71.2	12.3	16.5	12.3	12.6
Israel	56.8	57.7	62.5	67.9			8.2	12.5
Italy	57.8	65.8	66.0	62.5	10.4	14.1	7.5	11.8
Japan	52.0	57.6	59.0	61.5		10.8	7.3	8.6
Jordan	34.4	37.7	41.8	46.6			11.3	7.9
Kazakhstan	44.5	52.8	40.0	56.3				12.0
Korea, Rep.	87.7	97.8	103.9	100.8	8.6	10.9	6.6	11.0
Kyrgyz Republic	40.8	43.5	48.8	41.3			9.8	14.6
Latvia	71.5	73.6	66.1	65.1	9.0	7.7	9.4	10.3
Liechtenstein	20.7	29.8	34.4	42.5	17.1	18.6	18.1	8.4
Lithuania	68.1	76.4	77.4	73.9			14.0	13.6
Luxembourg	12.2	10.3	10.5	19.7		17.4	20.6	18.0
Macao SAR, China	76.6	56.8	63.8	56.1		0.9	1.8	1.8
Mexico	22.8	25.1	27.0	29.0	19.8	16.9	14.5	14.5
Montenegro	16.8	25.8	44.9	55.5			6.0	10.0
Netherlands	55.9	59.7	62.7	77.3		15.5	13.7	12.8
New Zealand	68.9	78.7	82.7	79.8	15.2	16.7	14.9	16.0
Norway	79.1	77.7	73.8	74.1	12.6	11.7	7.8	8.6
Panama	46.5	44.6	44.6	43.5				18.0
Peru	31.8	35.0		40.6				27.4
Poland	60.2	65.5	70.5	73.2	14.3	15.8	13.3	14.8
Portugal	54.6	55.4	62.2	68.9	18.0	12.8	17.9	16.5

	Gross Enrollment Rate (HE)				Inequality of School Achievement (R2 ESCS)			
Country	2003	2006	2009	2012	2000	2003	2006	2009
Qatar	15.2	19.3	10.1	12.1				4.0
Romania	37.6	51.8	63.8	51.6			11.0	13.6
Russian Federation	66.3	72.3	75.9	76.1	10.2	10.9	8.3	11.3
Serbia			49.8	52.4			14.0	9.0
Slovak Republic	33.9	44.3	54.2	55.1		20.9	14.9	14.6
Slovenia	69.3	83.0	86.9	86.0			15.1	14.3
Spain	64.0	68.0	73.2	84.6	15.8	11.4	10.5	13.6
Sweden	81.4	79.3	70.8	70.0	11.7	14.2	8.5	13.4
Switzerland	43.9	46.4	51.5	55.6	21.4	17.9	14.0	14.1
Thailand	40.6	43.8	45.7	51.4	7.3	9.9	14.7	13.3
Trinidad and Tobago	8.4	11.5						9.0
Tunisia	26.8	31.8	34.4	35.2		7.6	8.5	8.1
Turkey	28.7	36.0	45.8	69.4		20.3	13.1	19.0
United Kingdom	61.7	58.9	58.5	61.9			12.2	13.7
United States	81.2	82.6	89.1	94.3	19.3	18.0		16.8
Uruguay	41.2	45.4	63.3	63.2		12.8	13.5	20.7
Total	49.1	53.5	58.3	62.8	14.7	13.8	11.9	13.1

Table A.2. Valid and missing values by variable and country

Country	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	Valid values	Imputed values(*)	% imputed
Albania	3	1	2	2	2	3	1	2	3	0	0	16	14	47%
Argentina	4	2	3	3	3	4	2	4	4	4	4	33	7	18%
Australia	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Austria	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Azerbaijan	2	2	2	2	2	2	1	2	2	1	2	18	2	10%
Belgium	4	4	4	4	4	4	4	4	4	3	3	38	2	5%
Brazil	2	2	2	2	2	2	2	2	2	2	2	20	0	0%
Bulgaria	4	2	3	3	3	4	4	4	4	4	4	35	5	13%
Canada	1	1	1	1	1	1	1	1	1	1	1	10	0	0%
Chile	4	2	3	3	3	4	2	4	4	4	4	33	7	18%
Colombia	4	2	2	2	2	4	1	4	4	4	4	29	11	28%
Croatia	4	2	2	2	2	4	4	4	4	1	3	28	12	30%
Czech Republic	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Denmark	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Estonia	4	2	2	2	2	4	4	4	4	4	4	32	8	20%
Finland	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
France	4	4	4	1	4	4	4	3	4	4	4	36	4	10%
Germany	1	1	1	1	1	1	1	1	1	1	0	9	1	10%
Greece	4	4	4	4	4	4	4	4	4	3	3	38	2	5%
Hong Kong SAR, China	4	4	4	4	4	4	3	4	3	3	3	36	4	10%
Hungary	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Iceland	4	4	4	4	4	4	4	2	4	4	4	38	2	5%
Indonesia	4	4	4	4	4	4	4	4	4	1	1	34	6	15%
Ireland	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Israel	4	2	3	3	3	4	4	4	4	4	4	35	5	13%
Italy	4	4	4	4	4	4	4	4	4	4	4	40	0	0%

Country	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	Valid values	Imputed values(*)	% imputed
Japan	4	3	4	3	3	4	4	4	4	4	4	37	3	8%
Jordan	4	2	2	2	2	4	0	0	4	4	0	20	20	50%
Kazakhstan	4	1	1	1	1	4	2	4	4	0	3	21	19	48%
Korea, Rep.	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Kyrgyz Republic	4	2	2	2	2	4	1	3	4	0	4	24	16	40%
Latvia	4	4	4	4	4	4	4	3	4	4	4	39	1	3%
Liechtenstein	4	4	4	1	4	0	0	0	3	3	3	22	18	45%
Lithuania	4	2	2	2	2	4	4	4	4	3	3	30	10	25%
Luxembourg	4	3	4	3	4	4	4	4	4	3	0	33	7	18%
Macao SAR, China	4	3	3	3	3	4	4	4	4	1	4	33	7	18%
Mexico	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Montenegro	4	2	2	2	2	4	1	1	3	0	0	17	23	58%
Netherlands	4	3	3	3	3	4	4	4	4	4	4	36	4	10%
New Zealand	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Norway	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Panama	4	1	1	1	1	4	4	3	4	4	3	26	14	35%
Peru	3	1	2	2	2	1	1	1	1	1	3	15	15	50%
Poland	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Portugal	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Qatar	4	1	2	2	1	3	1	1	3	1	0	15	25	63%
Romania	4	2	2	2	2	4	4	4	4	3	3	30	10	25%
Russian Federation	4	4	4	4	4	4	4	4	4	0	4	36	4	10%
Serbia	2	2	2	2	2	2	0	1	2	1	1	15	5	25%
Slovak Republic	4	3	3	3	3	4	4	4	4	4	4	36	4	10%
Slovenia	4	2	2	2	2	4	4	4	4	1	3	28	12	30%
Spain	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Sweden	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Switzerland	4	4	4	4	4	4	4	4	4	4	4	40	0	0%
Thailand	4	4	4	4	4	4	0	4	4	4	4	36	4	10%
Trinidad and Tobago	2	1	1	1	1	2	2	2	2	2	2	16	4	20%
Tunisia	4	3	3	3	3	4	0	3	4	4	4	31	9	23%

Country	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	Valid values	Imputed values(*)	% imputed
Turkey	4	3	3	3	3	4	4	4	4	1	3	32	8	20%
United Kingdom	4	2	3	3	4	4	4	4	4	4	4	36	4	10%
United States	4	3	3	3	4	4	4	4	4	4	4	37	3	8%
Uruguay	4	3	3	3	3	4	3	3	4	3	3	32	8	20%
Valid values	228	177	187	179	187	221	185	202	222	179	192	1931	349	15%
Imputed values	0	51	41	49	41	7	43	26	6	49	36	349		
% imputed	0%	22%	18%	21%	18%	3%	19%	11%	3%	21%	16%	15%		

[1] GER HE; [2] R-sq ESCS; [3] PISA Reading score; [4] % variance within; [5] ESCS]; [6] Ln GDP per capita; [7] % Labour force TE al; [8] Unemployment TE % tot; [9] GER Secondary; [10] Expenditure in secondary as % of GDP p/c; [11] Expenditure in HE as % of GDP p/c
(*) Values imputed through interpolation for independent variables.

Table A.3. OLS Estimates with and without Imputed Data

	Imputed data						Only full case					
	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-0.567	0.827	-2.368**	-2.850**	-2.253*	-2.535**	-2.059	-0.669	-3.669***	-5.027***	-1.943*	-2.806**
	-1.099	(1.233)	(1.200)	(1.257)	(1.151)	(1.170)	-1.277	(1.480)	(1.270)	(1.329)	(1.165)	(1.188)
ESCS	8.318***	7.650***	10.204***	7.305**	9.034***	5.110*	10.914***	10.367***	14.581***	13.568***	14.815***	9.014**
	-2.718	(2.597)	(2.849)	(2.820)	(2.667)	(2.612)	-3.41	(3.320)	(3.660)	(4.406)	(3.076)	(3.676)
PISA Reading (z scores)	8.379***	8.736***	10.509***	9.754***	9.876***	9.376***	7.595***	7.813***	12.076***	10.268***	6.517**	5.979**
	-1.556	(1.541)	(1.564)	(1.618)	(1.564)	(1.597)	-2.12	(2.021)	(2.937)	(2.736)	(2.776)	(2.648)
% within variance		0.145*	-0.094	-0.116	-0.092	-0.109		0.106	-0.133	-0.094	0.017	0.023
		(0.075)	(0.080)	(0.079)	(0.076)	(0.072)		(0.082)	(0.091)	(0.076)	(0.076)	(0.072)
ln (GDP per capita PPP)			-11.706***	-11.695***	-8.902***	-8.216***			-14.042***	-23.468***	-2.137	-4.794
			(2.557)	(2.389)	(2.475)	(2.199)			(4.109)	(4.230)	(3.227)	(4.064)
% Lab force with HE			0.255*	0.240*	0.174	0.125			0.003	0.242	0.058	0.157
			(0.132)	(0.130)	(0.134)	(0.132)			(0.254)	(0.251)	(0.165)	(0.193)
HE Unemp (% total unemployment)			0.465***	0.488***	0.405**	0.438**			0.721**	0.585**	0.324	0.315
			(0.170)	(0.171)	(0.172)	(0.171)			(0.289)	(0.283)	(0.204)	(0.243)
GER Secondary			0.224**	0.233***	0.219***	0.232***			0.221**	0.247***	0.139	0.200***
			(0.088)	(0.082)	(0.079)	(0.071)			(0.097)	(0.078)	(0.086)	(0.069)
Exp in secondary (%GDP)				0.546***		0.644***				0.734***		0.717***
				(0.151)		(0.151)				(0.177)		(0.161)
Exp in tertiary (%GDP)					-0.217***	-0.266***					-0.399***	-0.464***
					(0.051)	(0.058)					(0.086)	(0.087)
Constant	61.006***	52.624***	148.193***	136.134***	130.924***	110.816***	61.601***	55.671***	176.927***	246.157***	76.242**	78.844**
	-1.487	(5.111)	(24.205)	(22.526)	(23.249)	(20.927)	-1.598	(5.736)	(40.709)	(41.311)	(32.844)	(39.292)
N Observations	228	228	228	228	228	228	176	170.000	142.000	120.000	127.000	114.000
R2 Adjusted	0.396	0.441	0.592	0.606	0.624	0.646	0.336	0.377	0.480	0.581	0.541	0.608
AIC	1927.9	1911.9	1848.0	1841.7	1830.9	1819.1	1485.2	1423.013	1156.806	937.756	977.047	858.164
BIC	1951.9	1942.8	1906.3	1906.9	1896.0	1891.1	1497.9	1438.692	1183.408	965.631	1005.489	888.262

* p<.10, ** p<.05, *** p<.01

Table A.4. Fixed effects estimates with and without imputed data

	Imputed data						Only full case					
	Base	Add PISA variance composition		Base	Add PISA variance composition		Base	Add PISA variance composition		Base	Add PISA variance composition	
	b/se	b/se		b/se	b/se		b/se	b/se		b/se	b/se	
Ineq Sch Ach	-3.645**	-3.767**	-2.142	-2.482*	-2.451*	-2.738**	-3.061**	-2.689*	-1.431	-2.688**	-3.299**	-2.673**
	-1.403	(1.445)	(1.400)	(1.257)	(1.314)	(1.290)	-1.457	(1.582)	(1.363)	(1.153)	(1.233)	(1.203)
PISA Reading (z scores)	4.441	5.033	5.597	2.961	6.534	3.374	11.094	10.806	8.432	5.722	8.050	5.594
	-7.231	(7.040)	(6.331)	(6.878)	(5.927)	(6.527)	-6.77	(6.828)	(7.131)	(7.207)	(6.492)	(7.362)
% within variance	4.824	4.951	1.797	1.750	1.907	1.478	5.126	4.718	2.998	-1.174	-0.696	-1.250
	-3.579	(3.592)	(3.180)	(3.215)	(3.194)	(3.161)	-4.308	(4.653)	(4.937)	(4.610)	(4.609)	(4.613)
ESCS		-0.012	-0.024	-0.085	-0.007	-0.062		-0.056	-0.010	-0.053	0.079	-0.041
		(0.083)	(0.068)	(0.065)	(0.065)	(0.068)		(0.099)	(0.093)	(0.091)	(0.077)	(0.095)
ln (GDP per capita PPP)			15.102**	11.490*	18.986***	14.863***			26.536***	25.685**	22.102***	25.375**
			(6.650)	(6.485)	(6.011)	(5.482)			(8.256)	(10.380)	(7.993)	(10.578)
% Lab force with HE			0.347***	0.260**	0.330***	0.256**			0.256*	0.187	0.206*	0.104
			(0.126)	(0.119)	(0.119)	(0.118)			(0.136)	(0.119)	(0.115)	(0.125)
HE Unemp (% total unemployment)			0.234	0.157	0.254*	0.164			0.485*	0.153	0.340	0.208
			(0.155)	(0.168)	(0.142)	(0.159)			(0.269)	(0.232)	(0.240)	(0.245)
GER Secondary			0.141	0.245***	0.143*	0.235***			0.110	0.260**	0.139	0.263**
			(0.087)	(0.082)	(0.086)	(0.084)			(0.096)	(0.097)	(0.088)	(0.099)
Exp in secondary (%GDP)				0.718***		0.651***				0.658**		0.706***
				(0.198)		(0.200)				(0.250)		(0.251)
Exp in HE (%GDP)					0.192	0.116					0.111	-0.050
					(0.149)	(0.126)					(0.153)	(0.112)

	Imputed data						Only full case					
	Base	Add PISA variance composition		Base	Add PISA variance composition		Base	Add PISA variance composition		Base	Add PISA variance composition	
	b/se	b/se		b/se	b/se		b/se	b/se		b/se	b/se	
Constant	59.884*** -1.918	60.693*** (5.506)	-112.313* (66.142)	-95.779 (65.024)	-156.673** (60.219)	-131.067** (55.684)	61.744*** -1.508	65.943*** (6.206)	-225.561*** (81.081)	-239.307** (105.488)	-188.175** (80.635)	-233.707** (107.387)
N Observations	228	228	228	228	228	228	176	170	142	120	127	114
N Groups	61	61	61	61	61	61	60	60	49	42	44	38
R2 within	0.236	0.239	0.408	0.447	0.427	0.459	0.087	0.070	0.376	0.500	0.410	0.498
Rho	0.827	0.827	0.881	0.869	0.912	0.897	0.824	0.816	0.933	0.938	0.908	0.904
AIC	1501.8	1505.1	1461.6	1450.0	1458.3	1449.2	1133.7	1099.7	882.1	700.9	757.1	673.4
BIC	1522.4	1532.5	1513.0	1508.3	1516.6	1514.4	1143.2	1112.2	905.7	726.0	782.7	700.8

* p<.10, ** p<.05, *** p<.01

Table A.5. OLS estimates. Missing values imputed by interpolation. All predictor lagged 3 years

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-0.567	0.827	-2.368**	-2.850**	-2.253*	-2.535**
	-1.099	(1.233)	(1.200)	(1.257)	(1.151)	(1.170)
ESCS	8.318***	7.650***	10.204***	7.305**	9.034***	5.110*
	-2.718	(2.597)	(2.849)	(2.820)	(2.667)	(2.612)
PISA Reading (z scores)	8.379***	8.736***	10.509***	9.754***	9.876***	9.376***
	-1.556	(1.541)	(1.564)	(1.618)	(1.564)	(1.597)
% within variance		0.145*	-0.094	-0.116	-0.092	-0.109
		(0.075)	(0.080)	(0.079)	(0.076)	(0.072)
ln (GDP per capita PPP)			-11.706***	-11.695***	-8.902***	-8.216***
			(2.557)	(2.389)	(2.475)	(2.199)
% Lab force with HE			0.255*	0.240*	0.174	0.125
			(0.132)	(0.130)	(0.134)	(0.132)
HE Unemp (% total unemployment)			0.465***	0.488***	0.405**	0.438**
			(0.170)	(0.171)	(0.172)	(0.171)
GER Secondary			0.224**	0.233***	0.219***	0.232***
			(0.088)	(0.082)	(0.079)	(0.071)
Exp in secondary (%GDP)				0.546***		0.644***
				(0.151)		(0.151)
Exp in tertiary (%GDP)					-0.217***	-0.266***
					(0.051)	(0.058)
Constant	61.006***	52.624***	148.193***	136.134***	130.924***	110.816***
	-1.487	(5.111)	(24.205)	(22.526)	(23.249)	(20.927)
N Observations	228	228	228	228	228	228
R2 Adjusted	0.396	0.441	0.592	0.606	0.624	0.646
AIC	1927.9	1911.9	1848.0	1841.7	1830.9	1819.1
BIC	1951.9	1942.8	1906.3	1906.9	1896.0	1891.1

* p<.10, ** p<.05, *** p<.01

Table A.6. Fixed effects estimates. Missing values imputed by interpolation. All predictor lagged 3 years

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-3.645**	-3.767**	-2.142	-2.482*	-2.451*	-2.738**
	(1.403)	(1.445)	(1.400)	(1.257)	(1.314)	(1.290)
ESCS	4.441	5.033	5.597	2.961	6.534	3.374
	(7.231)	(7.040)	(6.331)	(6.878)	(5.927)	(6.527)
PISA Reading (z scores)	4.824	4.951	1.797	1.750	1.907	1.478
	(3.579)	(3.592)	(3.180)	(3.215)	(3.194)	(3.161)
% within variance		-0.012	-0.024	-0.085	-0.007	-0.062
		(0.083)	(0.068)	(0.065)	(0.065)	(0.068)
ln (GDP per capita PPP)			15.102**	11.490*	18.986***	14.863***
			(6.650)	(6.485)	(6.011)	(5.482)
% Lab force with HE			0.347***	0.260**	0.330***	0.256**
			(0.126)	(0.119)	(0.119)	(0.118)
HE Unemp (% total unemployment)			0.234	0.157	0.254*	0.164
			(0.155)	(0.168)	(0.142)	(0.159)
GER Secondary			0.141	0.245***	0.143*	0.235***
			(0.087)	(0.082)	(0.086)	(0.084)
Exp in secondary (%GDP)				0.718***		0.651***
				(0.198)		(0.200)
Exp in tertiary (%GDP)					0.192	0.116
					(0.149)	(0.126)
Constant	59.884***	60.693***	-112.313*	-95.779	-156.673**	-131.067**
	(1.918)	(5.506)	(66.142)	(65.024)	(60.219)	(55.684)
N Observations	228	228	228	228	228	228
N Groups	61	61	61	61	61	61
R2 within	0.236	0.239	0.408	0.447	0.427	0.459
Rho	0.827	0.827	0.881	0.869	0.912	0.897
AIC	1501.8	1505.1	1461.6	1450.0	1458.3	1449.2
BIC	1522.4	1532.5	1513.0	1508.3	1516.6	1514.4

* p<.10, ** p<.05, *** p<.01

Table A.7. OLS estimates. Missing values imputed by interpolation. Predictor and PISA variables lagged 6 years, all the remaining lagged 3 years

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-3.517***	-3.279**	-3.487***	-3.087**	-3.093**	-2.566**
	(1.167)	(1.387)	(1.261)	(1.211)	(1.247)	(1.125)
ESCS	12.594***	12.992***	12.355***	10.381***	12.013***	9.445***
	(2.731)	(2.721)	(3.206)	(3.239)	(3.081)	(3.053)
PISA Reading (z scores)	6.550***	7.151***	9.189***	9.181***	8.390***	8.322***
	(1.436)	(1.499)	(1.773)	(1.819)	(1.766)	(1.892)
% within variance		-0.056	-0.166**	-0.157**	-0.139	-0.125
		(0.079)	(0.083)	(0.077)	(0.086)	(0.076)
ln (GDP per capita PPP)			-10.536***	-10.549***	-7.871***	-7.592***
			(2.273)	(2.240)	(2.399)	(2.264)
% Lab force with HE			0.277**	0.240*	0.172	0.110
			(0.131)	(0.133)	(0.125)	(0.126)
HE Unemp (% total unemployment)			0.483***	0.535***	0.471***	0.539***
			(0.175)	(0.179)	(0.179)	(0.184)
GER Secondary			0.181	0.181*	0.160	0.160*
			(0.110)	(0.104)	(0.101)	(0.092)
Exp in secondary (%GDP)				0.359**		0.460**
				(0.178)		(0.181)
Exp in tertiary (%GDP)					-0.227***	-0.267***
					(0.085)	(0.093)
Constant	64.559***	68.993***	147.530***	138.157***	131.516***	118.023***
	(1.590)	(5.452)	(22.172)	(21.606)	(22.292)	(21.475)
N Observations	170	170.000	170.000	170.000	170.000	170.000
R2 Adjusted	0.438	0.471	0.602	0.607	0.619	0.629
AIC	1420.785	1412.369	1371.203	1370.945	1365.616	1362.653
BIC	1442.736	1440.591	1424.511	1430.525	1425.196	1428.505

* p<.10, ** p<.05, *** p<.01

Table A.8. Fixed effects estimates. Missing values imputed by interpolation. Predictor and PISA variables lagged 6 years, all the remaining lagged 3 years

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-4.394***	-4.323***	-3.888***	-3.545**	-3.644***	-3.301**
	(1.470)	(1.467)	(1.286)	(1.420)	(1.195)	(1.313)
ESCS	2.844	3.085	-5.002	-2.016	-3.817	-0.820
	(6.023)	(6.136)	(6.194)	(6.367)	(6.132)	(6.110)
PISA Reading (z scores)	-2.037	-1.282	-1.747	0.041	-1.304	0.133
	(3.719)	(3.591)	(3.904)	(3.827)	(3.902)	(3.714)
% within variance		0.071	0.057	0.059	0.054	0.053
		(0.087)	(0.087)	(0.085)	(0.086)	(0.083)
ln (GDP per capita PPP)			10.314	10.871	8.381	7.693
			(8.675)	(8.472)	(9.781)	(9.017)
% Lab force with HE			0.300**	0.213	0.301**	0.209
			(0.149)	(0.133)	(0.143)	(0.126)
HE Unemp (% total unemployment)			0.305	0.213	0.309	0.182
			(0.261)	(0.251)	(0.273)	(0.251)
GER Secondary			0.111	0.226**	0.120	0.248**
			(0.106)	(0.102)	(0.108)	(0.108)
Exp in secondary (%GDP)				0.634**		0.752***
				(0.244)		(0.260)
Exp in tertiary (%GDP)					-0.008	-0.196
					(0.147)	(0.133)
Constant	60.447***	56.276***	-69.641	-96.332	-51.414	-63.409
	(1.687)	(5.987)	(84.571)	(85.209)	(96.856)	(90.496)
N Observations	170	170	170	170	170	170
N Groups	60	60	60	60	60	60
R2 within	0.251	0.272	0.344	0.389	0.354	0.398
Rho	0.907	0.901	0.893	0.875	0.891	0.859
AIC	1053.6	1052.6	1049.0	1040.7	1050.4	1042.4
BIC	1072.4	1077.7	1096.1	1094.0	1103.7	1101.9

* p<.10, ** p<.05, *** p<.01

Table A.9. OLS estimates. Missing values imputed by using within-country means. All predictor lagged 3 years

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-0.514	0.884	-2.558**	-3.433**	-1.510	-2.144*
	(1.103)	(1.242)	(1.241)	(1.386)	(1.102)	(1.182)
ESCS	8.248***	7.654***	11.362***	7.913*	13.759***	7.289**
	(2.726)	(2.598)	(3.132)	(4.016)	(2.771)	(3.531)
PISA Reading (z scores)	8.422***	8.777***	11.245***	11.010***	5.363***	5.730***
	(1.569)	(1.558)	(1.835)	(2.164)	(1.993)	(2.123)
% within variance		0.139*	-0.119	-0.147*	-0.017	-0.042
		(0.076)	(0.084)	(0.083)	(0.071)	(0.076)
ln (GDP per capita PPP)			-13.124***	-13.823***	1.638	2.844
			(2.936)	(3.304)	(2.775)	(3.077)
% Lab force with HE			0.222	0.270	-0.020	0.116
			(0.140)	(0.230)	(0.115)	(0.183)
HE Unemp (% total unemployment)			0.534***	0.589**	0.409***	0.302
			(0.182)	(0.276)	(0.153)	(0.227)
GER Secondary			0.231**	0.238***	0.117	0.139*
			(0.092)	(0.091)	(0.075)	(0.074)
Exp in secondary (%GDP)				0.524***		0.468***
				(0.190)		(0.166)
Exp in tertiary (%GDP)					-0.364***	-0.419***
					(0.064)	(0.083)
Constant	60.989***	52.969***	163.058***	156.647***	42.949	18.513
	(1.488)	(5.173)	(28.730)	(32.647)	(27.286)	(30.510)
N Observations	228	228	210	191	194	182
R2 Adjusted	0.396	0.441	0.571	0.568	0.611	0.628
AIC	1927.9	1912.1	1713.8	1563.4	1529.3	1434.2
BIC	1951.9	1943.0	1767.4	1621.9	1588.1	1498.2

* p<.10, ** p<.05, *** p<.01

Table A.10. Fixed effects estimates. Missing values imputed by using within-country means. All predictor lagged 3 years

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-3.386**	-3.533**	-2.280	-2.362	-2.474*	-2.699*
	(1.439)	(1.477)	(1.479)	(1.429)	(1.444)	(1.475)
ESCS	3.572	4.217	6.381	4.976	10.484	5.878
	(7.389)	(7.162)	(6.775)	(7.824)	(6.440)	(7.932)
PISA Reading (z scores)	4.651	4.827	3.277	4.213	3.868	4.688
	(3.597)	(3.600)	(3.524)	(3.588)	(3.855)	(3.962)
% within variance		-0.039	-0.046	-0.114	-0.053	-0.095
		(0.101)	(0.092)	(0.083)	(0.086)	(0.083)
ln (GDP per capita PPP)			15.857**	11.086	18.248***	14.642*
			(6.522)	(7.484)	(6.362)	(7.356)
% Lab force with HE			0.259**	0.293***	0.238**	0.245**
			(0.109)	(0.104)	(0.113)	(0.112)
HE Unemp (% total unemployment)			0.290*	0.307	0.390**	0.407
			(0.168)	(0.244)	(0.160)	(0.272)
GER Secondary			0.134	0.191**	0.100	0.161*
			(0.091)	(0.083)	(0.087)	(0.086)
Exp in secondary (%GDP)				0.630***		0.554**
				(0.221)		(0.211)
Exp in tertiary (%GDP)					0.228	0.160
					(0.168)	(0.118)
Constant	59.648***	62.168***	-116.560*	-87.230	-140.332**	-121.180
	(1.967)	(6.371)	(65.822)	(73.360)	(63.858)	(73.123)
N Observations	228	228	210	191	194	182
N Groups	61	61	56	51	51	48
R2 within	0.23	0.232	0.391	0.403	0.377	0.432
Rho	0.828	0.829	0.885	0.877	0.895	0.87
AIC	1503.8	1506.9	1364.0	1221.6	1248.9	1167.4
BIC	1524.3	1534.4	1414.2	1276.9	1304.4	1228.3

* p<.10, ** p<.05, *** p<.01

Table A.11. OLS estimates. All predictor lagged 3 years. Full case

	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-2.059	-0.669	-3.669***	-5.027***	-1.943*	-2.806**
	(1.277)	(1.480)	(1.270)	(1.329)	(1.165)	(1.188)
ESCS	10.914***	10.367***	14.581***	13.568***	14.815***	9.014**
	(3.410)	(3.320)	(3.660)	(4.406)	(3.076)	(3.676)
PISA Reading (z scores)	7.595***	7.813***	12.076***	10.268***	6.517**	5.979**
	(2.120)	(2.021)	(2.937)	(2.736)	(2.776)	(2.648)
% within variance		0.106	-0.133	-0.094	0.017	0.023
		(0.082)	(0.091)	(0.076)	(0.076)	(0.072)
ln (GDP per capita PPP)			-14.042***	-23.468***	-2.137	-4.794
			(4.109)	(4.230)	(3.227)	(4.064)
% Lab force with HE			0.003	0.242	0.058	0.157
			(0.254)	(0.251)	(0.165)	(0.193)
HE Unemp (% total unemployment)			0.721**	0.585**	0.324	0.315
			(0.289)	(0.283)	(0.204)	(0.243)
GER Secondary			0.221**	0.247***	0.139	0.200***
			(0.097)	(0.078)	(0.086)	(0.069)
Exp in secondary (%GDP)				0.734***		0.717***
				(0.177)		(0.161)
Exp in tertiary (%GDP)					-0.399***	-0.464***
					(0.086)	(0.087)
Constant	61.601***	55.671***	176.927***	246.157***	76.242**	78.844**
	(1.598)	(5.736)	(40.709)	(41.311)	(32.844)	(39.292)
N Observations	176	170	142	120	127	114
R2 Adjusted	0.336	0.377	0.48	0.581	0.541	0.608
AIC	1485.2	1423.0	1156.8	937.8	977.0	858.2
BIC	1497.9	1438.7	1183.4	965.6	1005.5	888.3

* p<.10, ** p<.05, *** p<.01

Table A.12. Fixed effects estimates. All predictor lagged 3 years. Full case

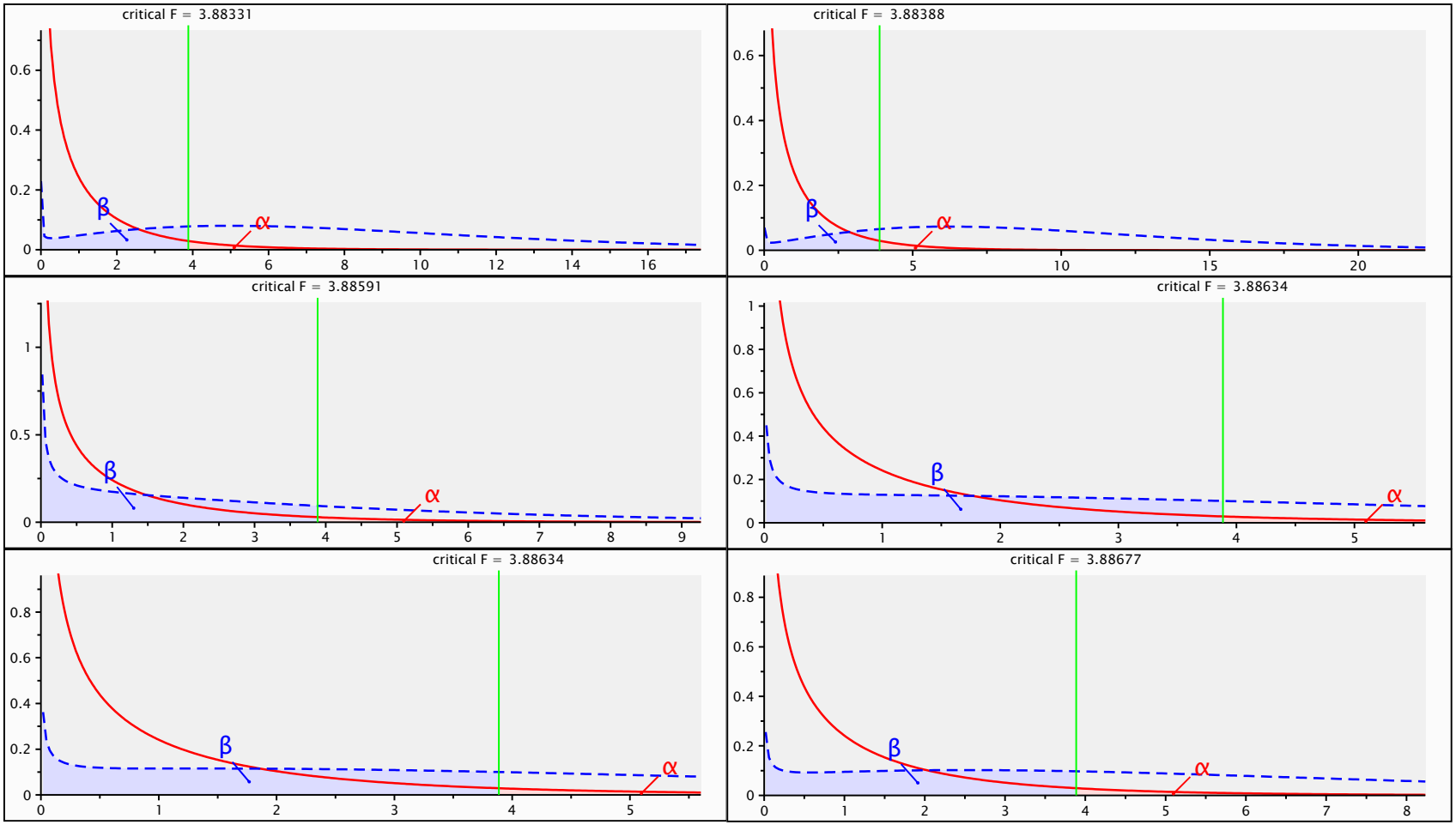
	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-3.061**	-2.689*	-1.431	-2.688**	-3.299**	-2.673**
	(1.457)	(1.582)	(1.363)	(1.153)	(1.233)	(1.203)
ESCS	11.094	10.806	8.432	5.722	8.050	5.594
	(6.770)	(6.828)	(7.131)	(7.207)	(6.492)	(7.362)
PISA Reading (z scores)	5.126	4.718	2.998	-1.174	-0.696	-1.250
	(4.308)	(4.653)	(4.937)	(4.610)	(4.609)	(4.613)
% within variance		-0.056	-0.010	-0.053	0.079	-0.041
		(0.099)	(0.093)	(0.091)	(0.077)	(0.095)
ln (GDP per capita PPP)			26.536***	25.685**	22.102***	25.375**
			(8.256)	(10.380)	(7.993)	(10.578)
% Lab force with HE			0.256*	0.187	0.206*	0.104
			(0.136)	(0.119)	(0.115)	(0.125)
HE Unemp (% total unemployment)			0.485*	0.153	0.340	0.208
			(0.269)	(0.232)	(0.240)	(0.245)
GER Secondary			0.110	0.260**	0.139	0.263**
			(0.096)	(0.097)	(0.088)	(0.099)
Exp in secondary (%GDP)				0.658**		0.706***
				(0.250)		(0.251)
Exp in tertiary (%GDP)					0.111	-0.050
					(0.153)	(0.112)
Constant	61.744***	65.943***	-225.561***	-239.307**	-188.175**	-233.707**
	(1.508)	(6.206)	(81.081)	(105.488)	(80.635)	(107.387)
N Observations	176	170	142	120	127	114
N Groups	60	60	49	42	44	38
R2 within	0.087	0.07	0.376	0.5	0.41	0.498
Rho	0.824	0.816	0.933	0.938	0.908	0.904
AIC	1133.66	1099.702	882.062	700.873	757.14	673.4
BIC	1143.171	1112.246	905.709	725.961	782.737	700.762

* p<.10, ** p<.05, *** p<.01

Table A.13. Fixed effects estimates. All predictor lagged 3 years. Multiple imputation estimates

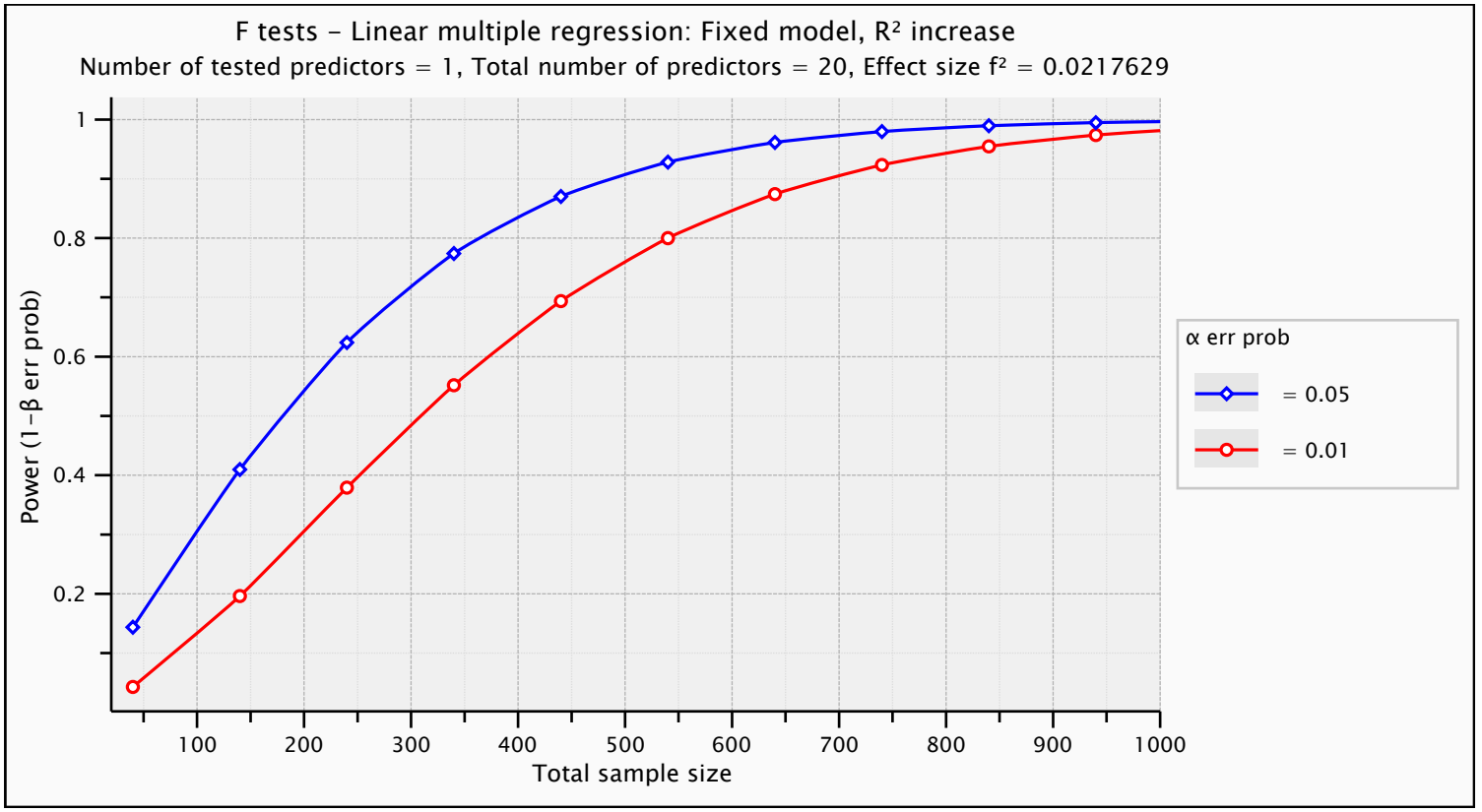
	Base	Add PISA variance composition	Add Socio economic variables	Add Secondary expenditure	Add Tertiary expenditure	Add Sec & Tert expenditure
	b/se	b/se	b/se	b/se	b/se	b/se
Ineq Sch Ach	-3.206**	-3.210**	-2.174	-2.809**	-2.468*	-2.968**
	(1.523)	(1.536)	(1.412)	(1.303)	(1.373)	(1.301)
ESCS	1.354	1.318	2.621	-0.726	4.421	0.443
	(7.827)	(7.884)	(7.534)	(7.010)	(7.278)	(6.999)
PISA Reading (z scores)	6.734*	6.747*	0.291	0.455	1.392	1.040
	(3.653)	(3.680)	(2.957)	(2.869)	(3.116)	(3.092)
% within variance		0.005	-0.025	-0.090	-0.034	-0.095
		(0.087)	(0.086)	(0.087)	(0.082)	(0.084)
ln (GDP per capita PPP)			20.367***	14.721***	22.163***	15.908***
			(5.470)	(5.361)	(5.551)	(5.748)
% Lab force with HE			0.297**	0.220*	0.295**	0.220*
			(0.129)	(0.128)	(0.127)	(0.124)
HE Unemp (% total unemployment)			0.331	0.177	0.352	0.193
			(0.223)	(0.206)	(0.234)	(0.218)
GER Secondary			0.173*	0.316***	0.167*	0.306***
			(0.088)	(0.089)	(0.087)	(0.092)
Exp in secondary (%GDP)				0.836***		0.795***
				(0.259)		(0.274)
Exp in tertiary (%GDP)					0.166	0.081
					(0.152)	(0.154)
Constant	55.793***	55.497***	-170.742***	-138.915***	-191.162***	-150.415**
	(2.152)	(6.002)	(53.572)	(50.946)	(55.448)	(55.082)
N Observations	252	252	252	252	252	252
N Groups	63	63	63	63	63	63

Graph A.1. Power estimates for fixed effects models, 3-year lagged (*)



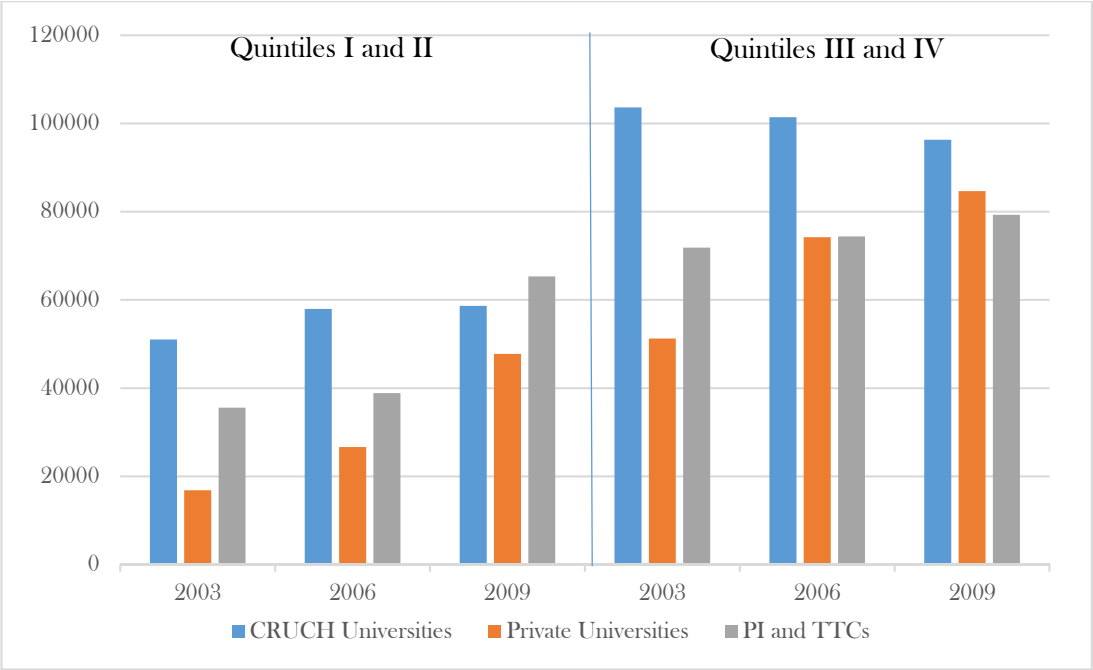
(*) Power estimates (left to right): .7642, .8252, .3532, .5369, .4936, .6015

Graph A.2. Power estimates according to sample size and p-level, full model.

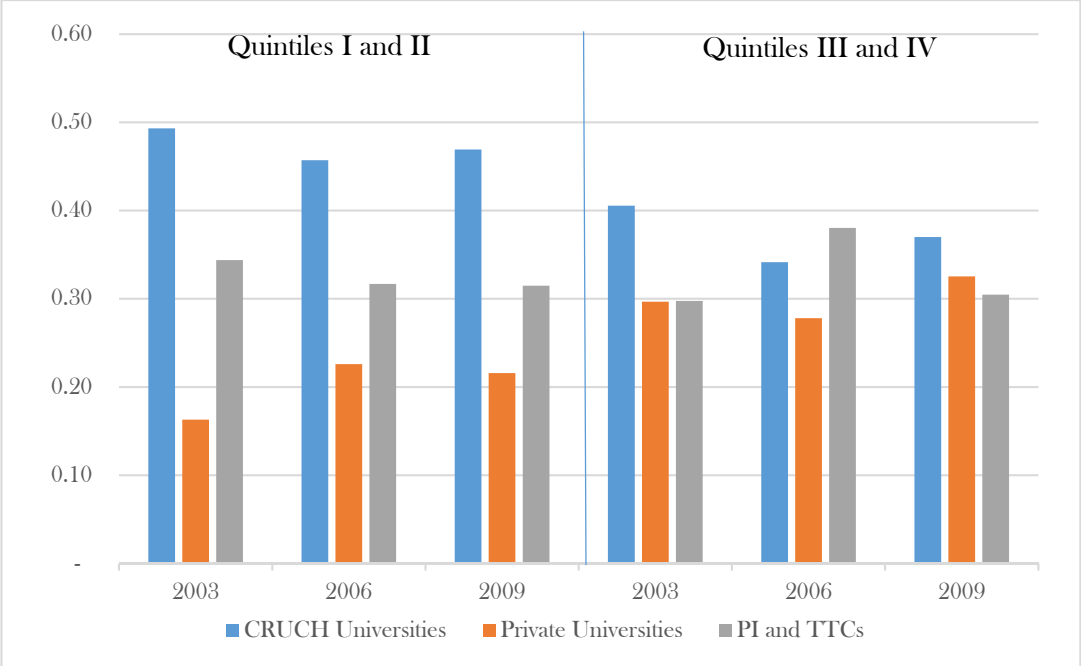


APPENDIX B (CHAPTER 4)

Graph B.1. Enrolment by income quintile and type of HEI, 2003, 2006 and 2009



Graph B.2. Share of undergraduate enrolment by income quintile and type of HEI, 2003, 2006 and 2009



Graph B.4. Number of students receiving student aid and undergraduate enrolment by type of HEI (2005-2009)

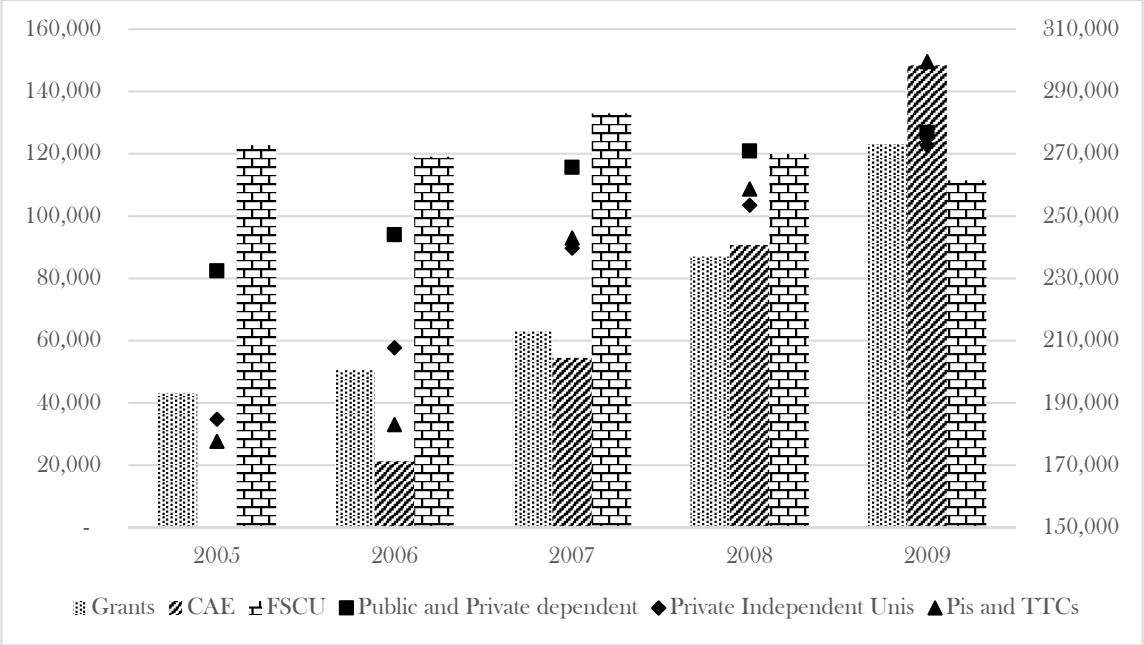


Table B.5. Differences-in-differences estimates. Logit estimates (odds ratio)

	Base model OR/se	Individual controls OR/se	Household controls OR/se	Background controls OR/se
Treatment=1	0.590***	0.426***	0.451***	0.553***
	(0.052)	(0.039)	(0.042)	(0.053)
Year 2006=1	1.087	1.195**	1.171*	1.262**
	(0.094)	(0.108)	(0.108)	(0.122)
Year 2009=1	1.202**	1.161*	1.134	1.109
	(0.101)	(0.102)	(0.102)	(0.105)
Treatment=1*Year 2006=1	1.216*	1.227*	1.264*	1.177
	(0.140)	(0.150)	(0.156)	(0.149)
Treatment=1*Year 2009=1	1.314**	1.383***	1.396***	1.401***
	(0.146)	(0.162)	(0.167)	(0.172)
Age		7.174***	7.313***	8.216***
		(2.259)	(2.322)	(2.697)
Age2		0.959***	0.958***	0.956***
		(0.007)	(0.007)	(0.007)
Female=1		1.068	1.137***	1.163***
		(0.050)	(0.054)	(0.057)
Married=1		0.224***	0.364***	0.374***
		(0.029)	(0.051)	(0.052)
Employed==1		0.214***	0.217***	0.249***
		(0.012)	(0.012)	(0.014)
Head of household woman=1			0.872**	0.797***
			(0.048)	(0.050)
N Siblings			1.283***	1.168***
			(0.046)	(0.044)
Household size			0.740***	0.763***

	Base model OR/se	Individual controls OR/se	Household controls OR/se	Background controls OR/se
			(0.020)	(0.021)
Urban=1			2.040***	1.263***
			(0.093)	(0.067)
Mother's education				
Secondary=1				1.606***
				(0.092)
Tertiary=1				3.162***
				(0.429)
Unknown=1				0.687**
				(0.111)
Head of household occupation				
Clerks and sales=1				0.738***
				(0.083)
Farming and Agriculture=1				0.572***
				(0.073)
Craftmen and operators=1				0.593***
				(0.063)
Unqualified worker=1				0.545***
				(0.060)
Unknown=1				0.711***
				(0.079)
Head of household years of schooling				1.102***
				(0.010)
Constant	0.739***	0.000***	0.000***	0.000***
	(0.048)	(0.000)	(0.000)	(0.000)
N	27631	27628	27628	27590
Pseudo R2	0.009	0.109	0.129	0.184
-2LL	2261679	2034261	1988296	1859637
BIC	2261740	2034374	1988449	1859883
AIC	2261691	2034283	1988326	1859685

* p<.10, **p<.05, *** p<.01

APPENDIX C (CHAPTER 5)

Table C.1. Relative risk ratios estimates for competing risk models, comparing risk of continuing/completion, relative to drop out. Single level models. Clustered standard errors reported

	Baseline RRR/se	Add background and debt RRR/se	Add demographic characteristics RRR/se	Add academic achievent & HEIs chars. RRR/se	Add expected income & interactions RRR/se
Outcome: Continuing registration					
Year 2= 1	1.309	1.205	1.311	1.368	1.443
	(0.140)	(0.160)	(0.172)	(0.183)	(0.258)
Year 3= 1	1.477	1.245	1.474	1.604	1.905
	(0.162)	(0.227)	(0.266)	(0.303)	(0.528)
Year 4= 1	1.723	1.321	1.709	1.951	2.818
	(0.186)	(0.310)	(0.398)	(0.472)	(0.973)
Year 5= 1	1.303	0.907	1.277	1.545	1.703
	(0.186)	(0.264)	(0.362)	(0.470)	(0.707)
Year 6= 1	1.389	0.889	1.372	1.759	1.689
	(0.217)	(0.340)	(0.520)	(0.700)	(0.762)
Partially CAE = 1		1.166	1.180	0.986	0.588
		(0.150)	(0.151)	(0.136)	(0.171)
Cumulated debt		1.182	1.176	1.066	1.080
		(0.161)	(0.154)	(0.139)	(0.230)
Quintile 2=1		1.149	1.163	1.175	1.100
		(0.122)	(0.126)	(0.123)	(0.119)
Quintile 3=1		1.199	1.189	1.173	1.095
		(0.132)	(0.131)	(0.128)	(0.124)
Quintile 4=1		1.437	1.408	1.322	1.184
		(0.191)	(0.184)	(0.166)	(0.164)
Quintile 5=1		1.460	1.410	1.231	1.138
		(0.172)	(0.162)	(0.137)	(0.139)
Public school =1		0.799	0.798	0.858	0.856
		(0.092)	(0.091)	(0.096)	(0.097)
Private mantained school=2		0.948	0.907	0.971	0.968
		(0.103)	(0.098)	(0.105)	(0.105)
Male=1			0.694	0.702	0.709
			(0.051)	(0.052)	(0.052)
Age			0.916	0.924	0.924
			(0.011)	(0.012)	(0.012)
PSU				1.261	1.280
				(0.075)	(0.076)
GPA				1.133	1.130
				(0.026)	(0.026)
Cruch Uni=1				1.526	1.459
				(0.260)	(0.253)
Complexity index				0.825	0.820
				(0.072)	(0.073)
Year=2 *Partially CAE funded=1					1.642
					(0.600)
Year=3 *Partially CAE funded=1					0.968
					(0.316)
Year=4 *Partially CAE funded=1					0.451
					(0.136)
Year=5 *Partially CAE funded=1					1.162
					(0.448)
Year=6 *Partially CAE funded=1					1.469
					(0.734)

	Baseline RRR/se	Add background and debt RRR/se	Add demographic characteristics RRR/se	Add academic achievent & HEIs chars. RRR/se	Add expected income & interactions RRR/se
Income Quintile=2 *Partially CAE funded=1					1.561
					(0.544)
Income Quintile=3 *Partially CAE funded=1					1.653
					(0.668)
Income Quintile=4 *Partially CAE funded=1					2.092
					(0.534)
Income Quintile=5 *Partially CAE funded=1					2.104
					(0.558)
Expected income					0.933
					(0.046)
c.zcaedebt#c.z~t					1.099
					(0.107)
Income Quintile=2 * St. Cum. Debt					0.853
					(0.115)
Income Quintile=3 * St. Cum. Debt					0.916
					(0.106)
Income Quintile=4 * St. Cum. Debt					0.915
					(0.111)
Income Quintile=5 * St. Cum. Debt					0.843
					(0.112)
Constant	21.560	21.739	147.972	96.534	88.476
	(1.690)	(4.354)	(45.890)	(30.027)	(32.485)
Outcome: Completed studies					
Year 2= 1	1.619	1.683	1.748	1.961	2.242
	(0.645)	(0.673)	(0.689)	(0.771)	(0.985)
Year 3= 1	2.398	2.576	2.784	3.506	4.287
	(0.779)	(0.978)	(1.050)	(1.357)	(2.160)
Year 4= 1	22.538	24.947	28.097	40.329	63.025
	(7.111)	(10.670)	(11.838)	(17.862)	(40.236)
Year 5= 1	195.925	224.729	264.207	443.165	575.413
	(58.451)	(99.084)	(113.358)	(204.774)	(359.699)
Year 6= 1	248.208	293.346	363.797	715.667	734.359
	(72.888)	(150.015)	(181.861)	(393.174)	(489.518)
Partially CAE = 1		0.580	0.605	0.391	1.035
		(0.120)	(0.123)	(0.085)	(0.708)
Cumulated debt		0.959	0.969	0.745	0.701
		(0.160)	(0.156)	(0.132)	(0.216)
Quintile 2=1		1.154	1.168	1.173	1.265
		(0.157)	(0.164)	(0.161)	(0.247)
Quintile 3=1		1.203	1.220	1.166	0.915
		(0.189)	(0.190)	(0.184)	(0.197)
Quintile 4=1		1.657	1.678	1.430	0.965
		(0.245)	(0.250)	(0.203)	(0.232)
Quintile 5=1		2.010	2.005	1.450	0.779
		(0.288)	(0.289)	(0.199)	(0.201)
Public school =1		0.758	0.744	0.899	0.889
		(0.106)	(0.105)	(0.120)	(0.121)
Private mantained school=2		0.919	0.870	1.036	1.027
		(0.119)	(0.113)	(0.132)	(0.132)
Male=1			0.555	0.572	0.605
			(0.056)	(0.059)	(0.061)
Age			0.951	0.970	0.970
			(0.015)	(0.016)	(0.016)
PSU				1.606	1.721
				(0.137)	(0.150)

	Baseline RRR/se	Add background and debt RRR/se	Add demographic characteristics RRR/se	Add academic achievent & HEIs chars. RRR/se	Add expected income & interactions RRR/se
GPA				1.315	1.343
				(0.058)	(0.063)
Cruch Uni=1				1.524	1.508
				(0.381)	(0.388)
Complexity index				0.830	0.758
				(0.092)	(0.093)
Year=2 *Partially CAE funded=1					0.924
					(0.834)
Year=3 *Partially CAE funded=1					0.518
					(0.402)
Year=4 *Partially CAE funded=1					0.153
					(0.107)
Year=5 *Partially CAE funded=1					0.159
					(0.107)
Year=6 *Partially CAE funded=1					0.240
					(0.182)
Income Quintile=2 *Partially CAE funded=1					0.777
					(0.421)
Income Quintile=3 *Partially CAE funded=1					1.362
					(0.746)
Income Quintile=4 *Partially CAE funded=1					4.213
					(1.774)
Income Quintile=5 *Partially CAE funded=1					4.686
					(2.131)
Expected income					0.675
					(0.058)
c.zcaedebt#c.z~t					1.127
					(0.124)
Income Quintile=2 * St. Cum. Debt					0.772
					(0.160)
Income Quintile=3 * St. Cum. Debt					1.055
					(0.200)
Income Quintile=4 * St. Cum. Debt					1.087
					(0.211)
Income Quintile=5 * St. Cum. Debt					1.215
					(0.253)
Constant	0.062	0.049	0.175	0.072	0.057
	(0.015)	(0.015)	(0.080)	(0.034)	(0.032)
N Obs	29892.000	29892.000	29892.000	29892.000	29892.000
AIC	20266.962	20168.490	20066.329	19904.526	19734.295
BIC	20366.626	20401.040	20332.100	20236.740	20315.670
ll	-10121.481	-10056.245	-10001.164	-9912.263	-9797.148

Table C.2. Relative risk ratio estimates for competing risk models, comparing risk of continuing/completion, relative to drop out. Random effects models. Clustered standard errors reported

	Baseline RRR/se	Add background and debt RRR/se	Add demographic characteristics RRR/se	Add academic achievent & HEIs chars. RRR/se	Add expected income & interactions RRR/se
Outcome: Continuing registration					
Year 2= 1	1.302	1.198	1.300	1.353	1.505
	(0.140)	(0.152)	(0.165)	(0.174)	(0.263)
Year 3= 1	1.460	1.229	1.450	1.569	2.073
	(0.160)	(0.222)	(0.266)	(0.298)	(0.566)
Year 4= 1	1.695	1.297	1.669	1.889	3.195
	(0.184)	(0.320)	(0.416)	(0.489)	(1.122)
Year 5= 1	1.276	0.885	1.240	1.482	1.985
	(0.185)	(0.257)	(0.361)	(0.460)	(0.824)
Year 6= 1	1.350	0.862	1.321	1.668	1.986
	(0.215)	(0.341)	(0.527)	(0.701)	(0.923)
Partially CAE = 1		1.183	1.189	1.007	0.595
		(0.159)	(0.159)	(0.142)	(0.175)
Cumulated debt		1.183	1.187	1.081	0.993
		(0.171)	(0.170)	(0.157)	(0.204)
Quintile 2=1		1.156	1.167	1.176	1.100
		(0.124)	(0.128)	(0.124)	(0.119)
Quintile 3=1		1.236	1.210	1.196	1.118
		(0.139)	(0.135)	(0.133)	(0.129)
Quintile 4=1		1.484	1.441	1.362	1.218
		(0.181)	(0.178)	(0.166)	(0.165)
Quintile 5=1		1.495	1.426	1.258	1.163
		(0.160)	(0.154)	(0.139)	(0.140)
Public school =1		0.764	0.776	0.826	0.824
		(0.085)	(0.086)	(0.092)	(0.092)
Private mantained school=2		0.937	0.906	0.962	0.958
		(0.101)	(0.097)	(0.105)	(0.105)
Male=1			0.723	0.732	0.734
			(0.054)	(0.054)	(0.054)
Age			0.915	0.923	0.923
			(0.011)	(0.012)	(0.012)
PSU				1.273	1.283
				(0.077)	(0.078)
GPA				1.131	1.128
				(0.027)	(0.027)
Cruch Uni=1				1.564	1.501
				(0.276)	(0.269)
Complexity index				0.829	0.832
				(0.073)	(0.075)
Year=2 *Partially CAE funded=1					1.631
					(0.597)
Year=3 *Partially CAE funded=1					0.945
					(0.309)
Year=4 *Partially CAE funded=1					0.433
					(0.132)
Year=5 *Partially CAE funded=1					1.109
					(0.435)
Year=6 *Partially CAE funded=1					1.414
					(0.724)
Income Quintile=2 *Partially CAE funded=1					1.572
					(0.550)
Income Quintile=3 *Partially CAE funded=1					1.649
					(0.675)

	Baseline RRR/se	Add background and debt RRR/se	Add demographic characteristics RRR/se	Add academic achievent & HEIs chars. RRR/se	Add expected income & interactions RRR/se
Income Quintile=4 *Partially CAE funded=1					2.028
					(0.515)
Income Quintile=5 *Partially CAE funded=1					2.034
					(0.545)
Expected income					0.948
					(0.054)
c.zcaedebt#c.z~t					1.110
					(0.113)
Income Quintile=2 * St. Cum. Debt					0.852
					(0.115)
Income Quintile=3 * St. Cum. Debt					0.915
					(0.105)
Income Quintile=4 * St. Cum. Debt					0.923
					(0.109)
Income Quintile=5 * St. Cum. Debt					0.850
					(0.110)
Constant	23.468	22.969	158.606	99.893	81.913
	(1.912)	(4.406)	(48.309)	(31.187)	(30.158)
Outcome: Completed studies					
Year 2= 1	1.605	1.418	1.495	1.583	1.743
	(0.643)	(0.566)	(0.591)	(0.629)	(0.766)
Year 3= 1	2.381	1.825	2.033	2.282	2.620
	(0.777)	(0.679)	(0.750)	(0.856)	(1.233)
Year 4= 1	23.315	15.305	18.064	22.029	33.322
	(7.489)	(5.959)	(7.026)	(8.837)	(18.843)
Year 5= 1	267.943	151.344	189.876	265.802	362.521
	(81.613)	(61.671)	(77.562)	(115.285)	(206.438)
Year 6= 1	438.329	213.387	286.791	464.540	546.349
	(132.723)	(105.365)	(141.816)	(246.716)	(331.856)
Partially CAE = 1		0.718	0.735	0.585	1.101
		(0.137)	(0.139)	(0.116)	(0.775)
Cumulated debt		1.338	1.354	1.143	1.245
		(0.229)	(0.229)	(0.207)	(0.332)
Quintile 2=1		1.115	1.126	1.124	1.309
		(0.162)	(0.169)	(0.162)	(0.266)
Quintile 3=1		1.263	1.261	1.202	1.059
		(0.200)	(0.199)	(0.191)	(0.240)
Quintile 4=1		1.730	1.718	1.578	1.435
		(0.250)	(0.254)	(0.232)	(0.313)
Quintile 5=1		1.801	1.747	1.464	1.214
		(0.248)	(0.242)	(0.209)	(0.295)
Public school =1		0.887	0.897	1.087	1.067
		(0.120)	(0.122)	(0.151)	(0.148)
Private mantained school=2		1.077	1.041	1.208	1.190
		(0.138)	(0.134)	(0.159)	(0.156)
Male=1			0.535	0.570	0.584
			(0.048)	(0.053)	(0.054)
Age			0.937	0.959	0.962
			(0.016)	(0.017)	(0.017)
PSU				1.981	2.052
				(0.159)	(0.163)
GPA				1.506	1.519
				(0.091)	(0.094)
Cruch Uni=1				0.975	0.990
				(0.246)	(0.246)
Complexity index				0.718	0.697
				(0.084)	(0.081)
Year=2 *Partially CAE funded=1					1.059

	Baseline RRR/se	Add background and debt RRR/se	Add demographic characteristics RRR/se	Add academic achievent & HEIs chars. RRR/se	Add expected income & interactions RRR/se
					(0.964)
Year=3 *Partially CAE funded=1					0.720
					(0.556)
Year=4 *Partially CAE funded=1					0.222
					(0.152)
Year=5 *Partially CAE funded=1					0.208
					(0.140)
Year=6 *Partially CAE funded=1					0.313
					(0.236)
Income Quintile=2 *Partially CAE funded=1					0.869
					(0.485)
Income Quintile=3 *Partially CAE funded=1					1.117
					(0.662)
Income Quintile=4 *Partially CAE funded=1					3.656
					(1.609)
Income Quintile=5 *Partially CAE funded=1					3.559
					(1.636)
Expected income					0.591
					(0.051)
Sq-cumulated debt					1.064
					(0.120)
Income Quintile=2 * St. Cum. Debt					0.731
					(0.157)
Income Quintile=3 * St. Cum. Debt					0.972
					(0.186)
Income Quintile=4 * St. Cum. Debt					0.857
					(0.166)
Income Quintile=5 * St. Cum. Debt					0.901
					(0.191)
Constant	0.051	0.048	0.229	0.095	0.060
	(0.014)	(0.015)	(0.112)	(0.048)	(0.037)
Random effects					
$\sigma^2_{\mu_0}$	1.230	1.203	1.149	1.139	1.128
	(0.089)	(0.081)	(0.068)	(0.066)	(0.071)
$\sigma^2_{\mu_2}$	2.787	2.841	2.701	3.010	2.522
	(0.437)	(0.454)	(0.423)	(0.510)	(0.406)
$\sigma_{\mu_0\mu_2}$	1.245	1.219	1.171	1.159	1.115
	(0.114)	(0.105)	(0.094)	(0.093)	(0.095)
N Obs	29892.000	29892.000	29892.000	29892.000	29892.000
AIC	19427.305	19334.845	19240.236	19004.187	18924.429
BIC	19551.885	19592.311	19530.923	19361.317	19530.719
ll	-9698.652	-9636.423	-9585.118	-9459.094	-9389.214

Table C.3. Relative risk ratios estimates for competing risk models, comparing risk of completion/dropout, relative to continuing

	Completion/ Continuing, single level	Completion/ Continuing, random effects
Year (*Wald* χ^2)	496.14***	551.79***
2	1.55	1.16
3	2.25	1.26
4	22.36***	10.43***
5	337.87***	182.66***
6	434.72***	275.06***
Debt structure		
Partially CAE funded=1	1.76	1.85
Cumulative debt	0.65	1.25
Sq_Cumulative Debt	1.03	0.96
Socieconomic background		
Income Quintile (*Wald* χ^2)	6.17	2.51
2	1.15	1.19
3	0.84	0.95
4	0.81	1.18
5	0.68	1.04
Secondary School (*Wald* χ^2)	0.58	9.08*
Public	1.04	1.30**
Maintained	1.06	1.24**
Demoraphic chars.		
Male=1	0.85*	0.80***
Age	1.05***	1.04**
Prior achievement		
PSU	1.34***	1.60***
GPA	1.19***	1.35***
University Chatacteristics		
Cruch University=1	1.03	0.66*
Complexity Index	0.93	0.84**
Year * Private debt (*Wald* χ^2)	26.19	28.09***
Year=2 * Partially CAE funded=1	0.56	0.65
Year=3 * Partially CAE funded=1	0.54	0.76
Year=4 * Partially CAE funded=1	0.34	0.51
Year=5 * Partially CAE funded=1	0.14**	0.19**
Year=6 * Partially CAE funded=1	0.16**	0.22*
Income Quintile *Private debt (*Wald* χ^2)	18.16***	13.99**
Income Quintile=2 * Partially CAE funded=1	0.50	0.55
Income Quintile=3 * Partially CAE funded=1	0.82	0.68
Income Quintile=4 * Partially CAE funded=1	2.01*	1.80
Income Quintile=5 * Partially CAE funded=1	2.23*	1.75
Expected income	0.72***	0.62***
Income quintile * St.Cum.Debt (*Wald* χ^2)	9.91*	5.55
Income Quintile=2 * St. Cum. Debt	0.91	0.86
Income Quintile=3 * St. Cum. Debt	1.15	1.06
Income Quintile=4 * St. Cum. Debt	1.19	0.93
Income Quintile=5 * St. Cum. Debt	1.44*	1.06
Intercept	0.00***	0.00***
Random effects		
	n/a	0.83***
	n/a	0.14
N	29892	29892
Goodness of fit statistics		
-2LL	19594	18778

AIC	19734	18924
BIC	20316	19531
Pseudo R^2	0.25	0.28

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Note: Wald χ^2 provided to test the null hypothesis that the effect of categorical variables equals to zero

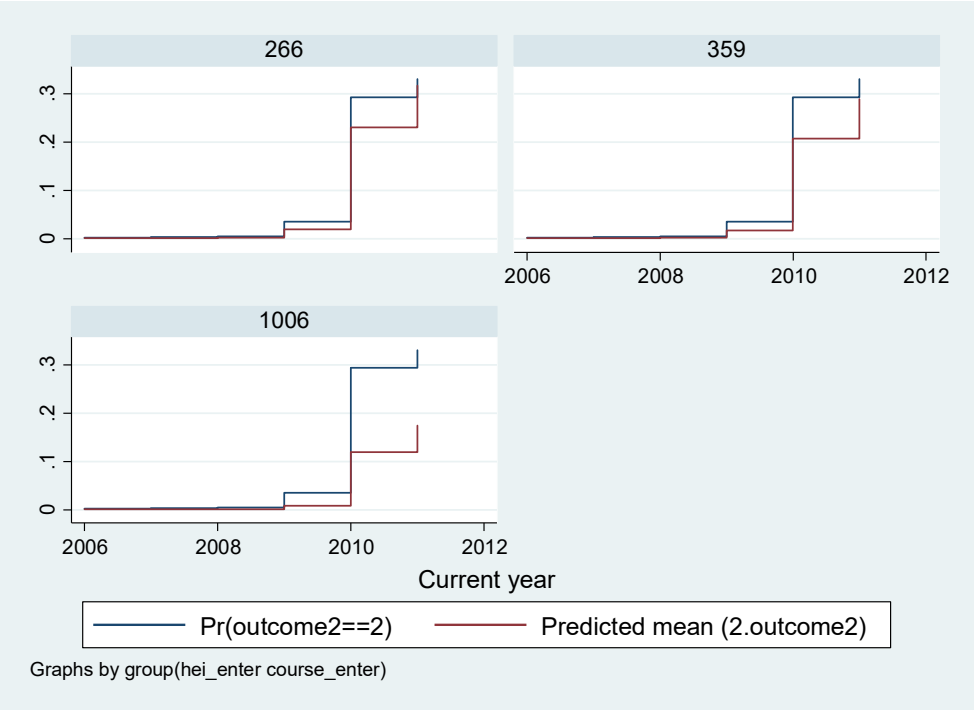
Table C.4. Predicted probabilities by type of university, income quintile and year (full single level model)

			CRUCH Universities						Private independent universities						Cumulated probabilities by year 6	
			Year						Year							
	Income quintile	CAE funded	1	2	3	4	5	6	1	2	3	4	5	6	Cruch Unis	Private Unis
Dropout		Partial	0.081	0.036	0.046	0.064	0.037	0.028	0.114	0.052	0.065	0.090	0.054	0.040	0.260	0.351
	1	Total	0.050	0.035	0.027	0.018	0.018	0.017	0.071	0.050	0.039	0.025	0.027	0.024	0.154	0.215
		Partial	0.049	0.022	0.027	0.039	0.024	0.018	0.070	0.031	0.039	0.055	0.034	0.026	0.166	0.231
	2	Total	0.046	0.032	0.024	0.016	0.016	0.014	0.065	0.046	0.035	0.023	0.023	0.021	0.139	0.196
		Partial	0.047	0.020	0.026	0.037	0.022	0.017	0.067	0.030	0.037	0.053	0.032	0.024	0.158	0.220
	3	Total	0.046	0.032	0.025	0.016	0.018	0.016	0.065	0.046	0.035	0.023	0.026	0.024	0.144	0.201
		Partial	0.035	0.015	0.019	0.027	0.014	0.010	0.050	0.022	0.028	0.038	0.021	0.015	0.114	0.161
	4	Total	0.042	0.030	0.023	0.015	0.017	0.015	0.061	0.043	0.033	0.022	0.024	0.022	0.134	0.189
		Partial	0.036	0.016	0.020	0.027	0.015	0.011	0.051	0.022	0.029	0.040	0.022	0.016	0.118	0.167
	5	Total	0.044	0.031	0.024	0.016	0.018	0.017	0.063	0.045	0.034	0.023	0.027	0.025	0.141	0.198
Continuing		Partial	0.915	0.960	0.950	0.910	0.819	0.767	0.883	0.945	0.930	0.885	0.809	0.762	0.395	0.317
	1	Total	0.948	0.962	0.968	0.937	0.576	0.518	0.927	0.947	0.957	0.931	0.579	0.522	0.147	0.098
		Partial	0.949	0.977	0.970	0.945	0.887	0.851	0.928	0.967	0.958	0.929	0.880	0.847	0.607	0.550
	2	Total	0.952	0.964	0.970	0.932	0.545	0.486	0.933	0.950	0.960	0.927	0.549	0.490	0.123	0.079
	3	Partial	0.951	0.977	0.971	0.944	0.872	0.830	0.931	0.968	0.960	0.929	0.866	0.827	0.579	0.526

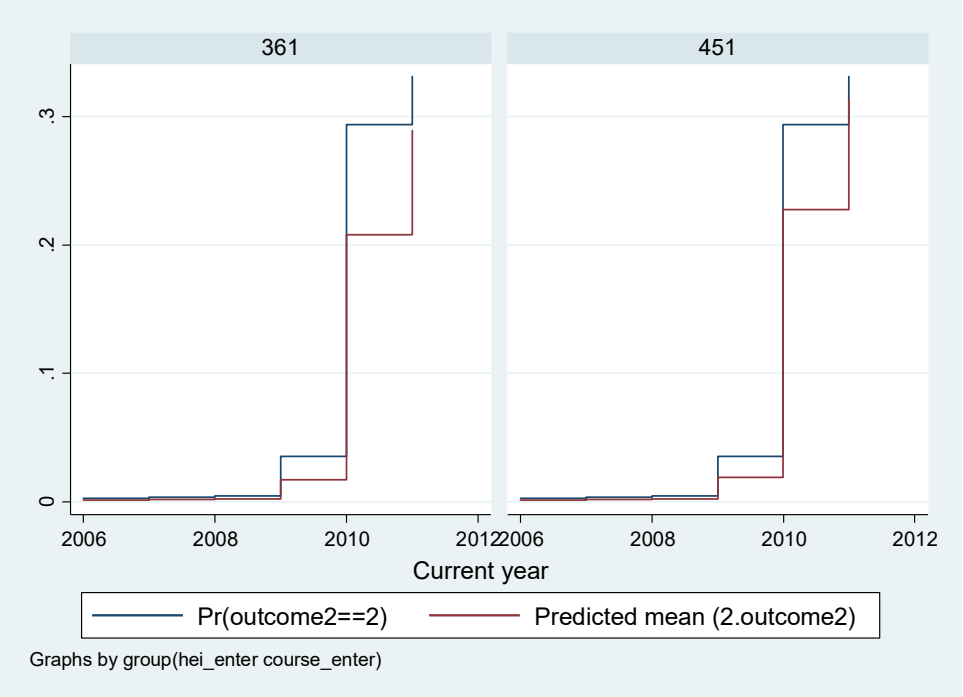
			CRUCH Universities						Private independent universities						Cumulated probabilities by year 6	
			Year						Year							
	Income quintile	CAE funded	1	2	3	4	5	6	1	2	3	4	5	6	Cruch Unis	Private Unis
		Total	0.953	0.965	0.971	0.945	0.617	0.561	0.933	0.951	0.961	0.939	0.619	0.564	0.206	0.161
		Partial	0.959	0.980	0.973	0.929	0.767	0.691	0.945	0.973	0.965	0.919	0.767	0.694	0.399	0.364
	4	Total	0.956	0.967	0.973	0.947	0.624	0.567	0.938	0.955	0.963	0.942	0.626	0.570	0.222	0.180
		Partial	0.959	0.979	0.973	0.931	0.778	0.705	0.943	0.973	0.965	0.921	0.778	0.708	0.417	0.380
	5	Total	0.954	0.967	0.973	0.952	0.660	0.606	0.936	0.953	0.963	0.947	0.661	0.608	0.265	0.222
Completed		Partial	*0.004*	*0.003*	0.004	0.027	0.144	0.205	*0.003*	*0.003*	0.004	0.025	0.138	0.197	0.345	0.333
	1	Total	0.002	0.003	0.005	0.046	0.405	0.465	0.002	0.003	0.005	0.044	0.395	0.454	0.700	0.687
		Partial	*0.002*	*0.002*	*0.003*	0.016	0.090	0.131	*0.002*	*0.002*	*0.002*	0.015	0.086	0.127	0.227	0.219
	2	Total	0.002	0.004	0.006	0.052	0.439	0.500	0.002	0.004	0.005	0.050	0.428	0.489	0.737	0.725
		Partial	*0.003*	*0.002*	0.003	0.019	0.106	0.154	*0.002*	*0.002*	0.003	0.018	0.102	0.148	0.264	0.255
	3	Total	0.002	0.003	0.004	0.039	0.365	0.423	0.002	0.003	0.004	0.037	0.355	0.412	0.651	0.638
		Partial	*0.006*	*0.005*	0.007	0.045	0.219	0.299	*0.006*	*0.005*	0.007	0.043	0.212	0.291	0.486	0.475
	4	Total	0.002	0.003	0.004	0.038	0.360	0.418	0.002	0.003	0.004	0.036	0.350	0.407	0.644	0.631
		Partial	*0.006*	*0.005*	0.007	0.041	0.207	0.284	*0.005*	*0.005*	0.007	0.040	0.200	0.276	0.465	0.453
	5	Total	0.001	0.002	0.003	0.032	0.321	0.377	0.001	0.002	0.003	0.031	0.312	0.367	0.594	0.581

Note: Non-significant probabilities in italics

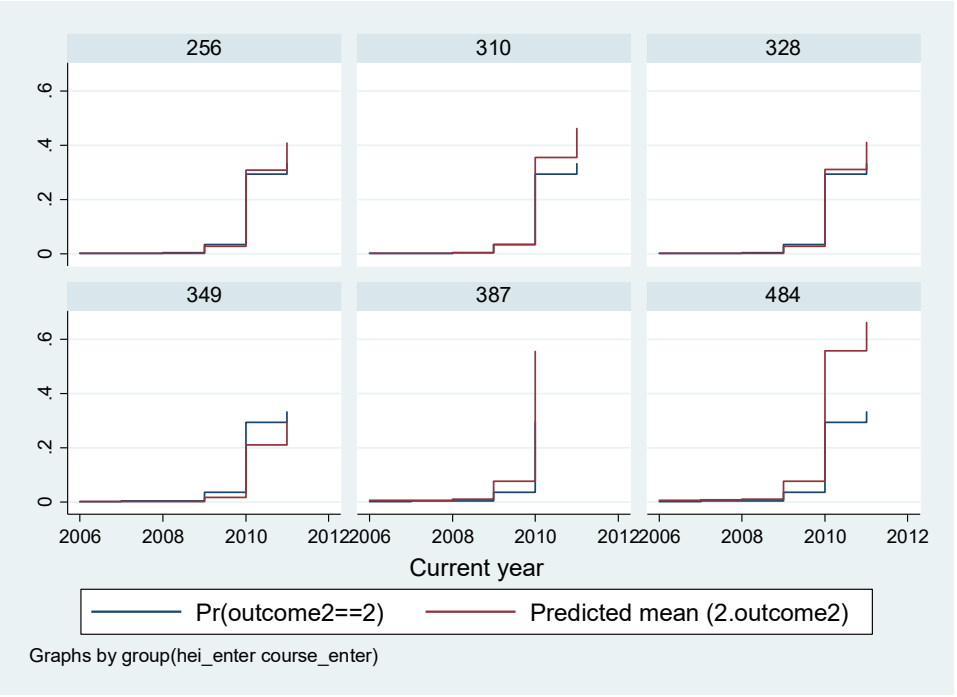
Graph C.1. Fitted completion hazard probabilities. Biochemistry programmes



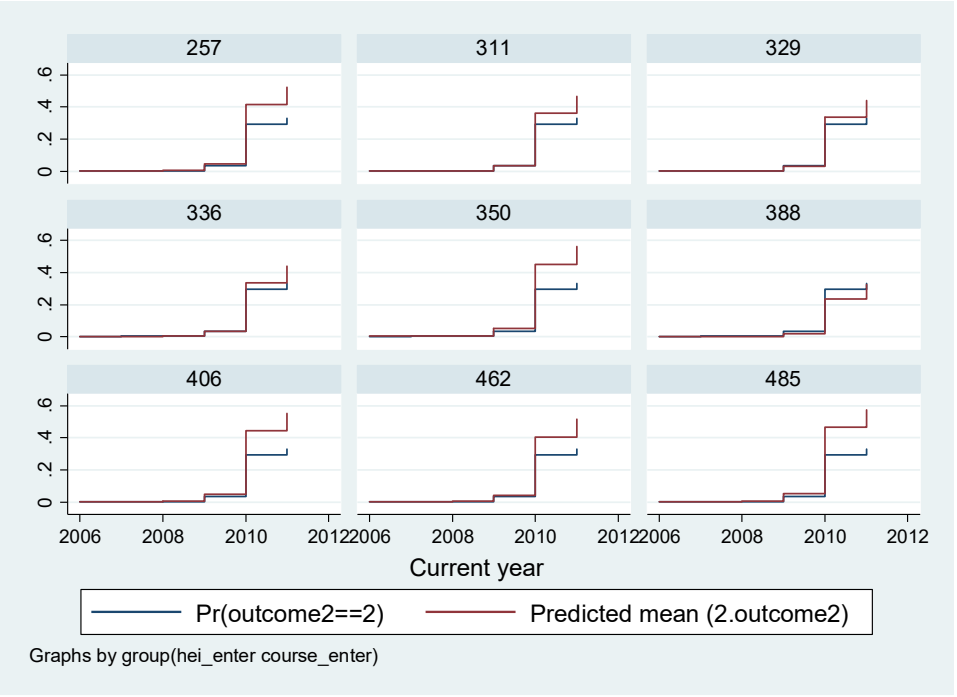
Graph C.2. Fitted completion hazard probabilities. Nursing programmes



Graph C.3. Fitted completion hazard probabilities. Journalism programmes



Graph C.4. Fitted completion hazard probabilities. Psicology programmes



Graph C.5. Fitted completion hazard probabilities. Social work programmes