Disentangling inequality of educational opportunities: The transition to higher education in Chile

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Humanities

2017

Alejandro Sevilla Encinas

School of Social Sciences

List of Contents

Word count: 65,365

List of Tables

List of Figures

Abstract

The University of Manchester
Alejandro Miguel Sevilla Encinas
For the degree of Doctor of Philosophy
Disentangling inequality of educational opportunities: The transition to higher education in Chile
30 September 2017

This thesis examines inequality of educational opportunities (IEO) in the transition to higher education. IEO measures the difference in higher education entry rates across social groups. The theoretical framework lays on Boudon's decomposition of IEO into primary and secondary effects of stratification. Furthermore, the theoretical propositions of *Maximally Maintained Inequality* (MMI) and *Effectively Maintained Inequality* (EMI) were also assessed to gain further understanding of IEO. The longitudinal data for the empirical analysis was created for a student cohort by linking administrative records of Chile's national student register, standardised tests and higher education enrolment. The student cohort was followed through the 12-years of compulsory education up to the transition to higher education, a year after completing secondary education.

The results from the empirical analysis showed that secondary effects were consistently predominant over primary effects, driving the overall IEO. On the other hand, controlling for school characteristics increased the relative importance of secondary effects. However, primary effects explained a large extent of IEO in the transition to traditional (most prestigious) universities, by the same token, in the transition to undergraduate programmes. Differences in parental education levels between secondary education completion and higher education transitions proved to be consistent with MMI. Likewise, the higher likelihood of less advantaged students to enrol in vocational colleges or vocational programmes, and the higher likelihood of advantaged students to enrol in traditional universities or undergraduate programmes, evidenced support for EMI.

The modelling setting was based on non-linear mediation modelling, accounting for sample-selection in the student cohort, two-level cross-classification between primary and secondary schools, and multinomial outcomes for type of institution and programme.

This thesis contributes to the educational attainment literature by finding evidence that, in emerging economies like Chile, educational inequality persists despite the sustained expansion of the educational system.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright Statement

Dedication

To my wonderful wife, Paulina, who supported and encouraged me through all these years. She has been my strength and inspiration to complete this thesis.

Acknowledgment

I would like to thank my supervisors Professor Yaojun Li, Professor Tarani Chandola and Professor Anthony Heath, who kept me on track and constantly persuaded me to reach improved versions of my research. I am also grateful to Dr Wendy Bottero, Dr Nick Shryane, and Dr Maria Pampaka, for their insightful comments on earlier versions of this thesis presented for internal review.

A special acknowledgement to CONICYT (*Comisión Nacional de Investigación Científica y Tecnológica*) for their financial support to conduct this doctoral research. Also, to the staff in *Centro de Estudios* (Ministry of Education), in particular to Roberto Schurch who assisted my request of data in higher education enrolment, being of chiefly importance to my research.

I would also like to thank staff and colleagues at CMIST and SoSS, particularly Marie Waite and Pip Walker for their patience and support, shown for all my enquiries.

I am very grateful to Vern for his proofreading, which went beyond the use of the language to motivate me when I needed this most.

Last but not the least, I thank my beloved parents for their endless encouragement and prayers that helped me to arrive at the end of this journey.

The Author

Alejandro Sevilla, prior to joining the Cathie Marsh Institute for Social Research (CMIST) at the University of Manchester, worked as a research assistant at the Centre for Advanced Research in Education (CIAE) at the University of Chile. He gained a BSc (Hons) in Economics from the University of San Simón, Bolivia, in 2004, and an MSc in Social Research Methods and Statistics from the University of Manchester in 2013. His research expertise is on national and international student assessments and schoolteachers' labour conditions. His current research interests are in the field of social stratification, with a particular focus on educational inequality and higher education.

Chapter 1: Introduction

1.1 What is equality of opportunities?

In this section, I attempt to define the term ‘equality of opportunities’ following the normative perspective of Swift (2005) and the social mobility perspective of Breen (2010a), with a particular emphasis on its relation to educational attainment and the purpose of this thesis.

It is a natural assumption that, in the majority of the cases, parents play, and have played, a pivotal role in the development of their children, having shaped the people they are as well as shaping the people they are yet to become. Parents, themselves, came with their own endowments, which varied in myriad ways, from genes through familial social background characteristics, and these endowments each held a relative position in society’s distribution of endowments. Parents, in turn and in the same manner, transferred these endowments to their children, including the inheritance of a relative position in society. Accordingly, some children are born advantaged, and some born less advantaged, with respect to other children, in terms of their parents’ position in society. Children’s chances to occupy a position in society are, therefore, determined, to a large extent, by their parents’ social position. Given the unfair consequences of this ‘accident of birth’, the concern naturally arises over how to level the field in order to help less advantaged children. Swift (2005) argued strongly for equality of opportunities as the removal of social barriers that discourage some people, amongst those with similar talents and aspirations that are not all affected by such barriers, to compete in similar conditions, regardless of their social origin.

Breen (2010a), on the empirical side, challenged the widespread claim that equality of opportunities can be measured in terms of the degree of social mobility in a society. He proposed the following example to illustrate his argument:

> Imagine that I had $1,000 and … I gave everyone a lottery ticket and then drew a ticket at random and gave the holder of that ticket the whole $1,000 … we would have equality of opportunity: since everyone had a lottery ticket, everyone had the same chance of success.
> But suppose we observed the outcomes of the lottery repeated many times and imagine that all the winners came from the front row of the audience. Then we would have better grounds for suspecting that opportunities were not, after all, distributed equally...
> This is closer to the situation we face when we look at mobility outcomes and observe, for example, that the lowest incomes and the worst jobs accrue most often to people whose parents also had low incomes and bad jobs. (Breen, 2010a, pp. 414, 424).

In terms of educational attainment Breen (2010a) argued that as the education system expands and participation in education increases in the population, the average educational attainment also increases, thereby the relationship between parent's and child's education declines. However, Breen insisted, the overall distribution of equality of opportunities remains unchanged, notwithstanding the overall increase in opportunities.

In this thesis the concern of equality of opportunities is centred in the empirical side, without losing sight of the definition in Swift (2005), and the understanding it advances concerning the contexts that shape children's opportunities in educational attainment. I measure inequality of educational opportunities as the difference between parental education groups on students' chances to progress into higher education.

In the next section, I briefly review the studies used as the theoretical framework in this thesis to operationalise inequality of educational opportunities empirically.

1.2 How do we measure inequality of educational opportunities?

The claim that the expansion of the education system will increase equality of educational opportunities can be tracked back to the studies that explain the process of the transition from a feudal society to an industrial society (Kerr *et al.*, 1960; Parsons, 1960; Treiman, 1970; Bell, 1973). The process of industrialism saw a restructuring of the labour force, whereby the demand for specialised skills was fulfilled through training and education. Thereafter, educational credentials became linked to social position. Under such analysis, it seems plausible that a general expansion of the education system, by increasing the number of places in education, providing a complementary decrease in the financial burden of education, and by providing universal access to compulsory education, will all result in an increase in educational opportunities. Therefore, the influence of parents' social position on their children social position will decrease as a consequence of enhanced childhood educational attainment (Breen, 2010a).

Boudon (1974), reluctant to accept that children hold a system of values according to their social groups that they attribute to education, proposed that inequality of educational opportunities (IEO) should, rather, be decomposed into primary and secondary effects of the process of stratification. Primary effects account for social background differences in academic performance, while the secondary effects account for social background differences in educational decisions, subject to academic performance. Boudon's model added rational and timeline factors to understand how IEO affects the process of class formation, by his definition of secondary effects, as children and their families make rational choices concerning the continuance of education, seeking resolution in what is the more appropriate path, given their interests. Breen and Goldthorpe (1997) later extended this concept to argue that children and their families act rationally to at least reach the social class of the parents, whilst avoiding downward social mobility. However, it was not until the work of Erikson and Jonsson (1996) that an empirical strategy was proposed to estimate the relative contribution of primary and secondary effects to the overall IEO (this work was subsequently refined, (Erikson *et al.*, 2005; Jackson *et al.*, 2007). Therefore, the measure of the overall IEO is the effect of social background on the educational transition, the point at which children and their families decide whether or not to continue in education.

Empirical evidence from industrialised countries that supports Boudon's IEO model has been growing in recent years (Jackson, 2013a). There has been little if any evidence, however, from emerging economies, as appropriate data are scarce in these countries, especially longitudinal data. Emerging economies, like Chile, experienced educational reforms and processes over the end of the last half of the century that reflects a history different from those of developed economies. As an illustration, the end of WWII is not considered a milestone to account for the expansion of the Chilean educational system, as it is in western countries. The expansion of the educational system was, instead, a result of the decentralisation of state services implemented in the 1980s during Pinochet's military dictatorship. Although this expansion of the educational system has increased higher education attainment in Chile, the securing of fair access to higher education is still a task that remains to be completed, in order to reduce inequality of opportunities. In this thesis, I seek to contribute to the discussion of educational inequality in Chile by providing new evidence with respect to the transition to higher education.

The next section positions the central importance of understanding the generation of inequality of educational opportunities within the context of recent events in Chile.

1.3 The influence of the Chilean student movement on the institutional context

The transition to higher education has become an important social selection mechanism in contemporary Chilean society. Despite a sustained, nationwide growth in higher education enrolment in recent decades, children in families of low social origin remain underrepresented in continuing to this educational level.

In 2011, a large number of school and university students in Chile took to the streets to protest against Chile's market-oriented educational system. They centred their demands on equal opportunities to access higher education, the restitution of education as a universal social right, and the direct participation of the state in the education provision[1]. The students mobilised, initially in their thousands, but, as the months progressed, they reached a peak in the hundreds of thousands by July. Notably, some journalists compared the 'Arab Spring', which occurred in the same year, to the student movement in Chile, which they, thus, termed the 'Chilean Winter', as it occurred during the winter season of the southern hemisphere (Bellei and Cabalin, 2013).

Increasing tuition fees and high interest rates in student loans triggered the students' protests. Students, and their families, had been expected to accept that this was a reasonable cost of the rapid expansion in the higher education system derived from the 1981 reform, which had led to an ever-heavier reliance on private institutions[2] (i.e., private universities and vocational colleges). The reform had restructured the role of the state from the provider role to a subsidiary role. Private institutions created after the reform lacked the regulation previously applied to traditional universities (i.e., universities created before the reform), and typically had weaker admissions policies than the academically selective admission policies that were enforced in traditional universities. Moreover, private institutions were

[1] They were many more demands, which can be seen in detail in Stromquist and Sanyal (2013).
[2] Only vocational colleges were allowed to operate as for-profit institutions, but latter it was discovered that a group of private universities were also operating as for-profit institutions.

allowed to compete alongside traditional universities for state funding to attract student enrolment through scholarships or student loans. Meanwhile, the funding to traditional universities was being gradually reduced, forcing them to charge tuition fees and to look for other sources of funding to cover their budgets. In this scenario, so favourable to private institutions themselves, higher education expansion was driven by their increasing participation in the education system (Stromquist and Sanyal, 2013; Bellei, Cabalin and Orellana, 2014).

The result of the negotiations between the student movement and the government was the creation of public agencies that regulate and monitor higher education institutions. The government also implemented full scholarships for the most vulnerable socioeconomic group, student loans were no longer intermediated by banks but by a public agency, and the interest rates were reduced. The government also investigated the illegal for-profit practices of various private universities. (Stromquist and Sanyal, 2013; Bellei, Cabalin and Orellana, 2014).

As increasing tuition fees and high interest rates in student loans were the reasons underpinning students' demands, it can be argued that the main concern in determining inequality of opportunities, in the access to higher education in Chile, is students' financial constraints. This is without doubt a relevant factor, in the institutional context of Chile, which contributes to create social inequality. However tempting, there is no absolute requirement to operationalise such financial constraints in order to examine IEO[3]. Indeed, in the strand of the literature of Boudon's IEO model, information concerning students' financial aids, such as scholarships or student loans, is not relevant to the decomposition of IEO into primary and secondary effects, which is the main theoretical framework of this

[3] Inequality of educational opportunities.

thesis. The use of parental education[4] to characterise students social background and its effect in determining IEO, is, by contrast, supported in the literature (Jackson, 2013a, p. 22).

1.4 Aims of the thesis

This thesis aims to examine social-background differences in Chilean students continuing to higher education, and to evaluate the applicability of Boudon's inequality of educational opportunities (IEO) model.

The specific aims are:

(i) To test the hypothesis that standardised tests (SIMCE) increase the contribution of secondary effects compared to grade point averages (GPA) (both measures of academic performance are available for the student cohort). Complementary to the comparison of both measures, to test the hypothesis of anticipatory decisions in GPA levels compared to standardised tests applied long before the transition to higher education (Jackson *et al.*, 2007; Erikson and Rudolphi, 2010).

(ii) To compare the contribution of secondary effects in the overall IEO of the transition to higher education in Chile with other countries.

(iii) To account for sample-selection in the student cohort, as only 64% completed the 12-year period of compulsory education.

(iv) To assess *Maximally Maintained Inequality* (MMI) (Raftery and Hout, 1993; Hout, Raftery and Bell, 1993) between secondary education completion and the transition to higher education, since secondary education is a universal level of education in Chile.

(v) To investigate the effect of the stratification in the education system on the contribution of primary and secondary effects to the overall IEO.

[4] Only parental education was available in the data for the student cohort analysed in this thesis.

(vi) To assess *Effectively Maintained Inequality* (EMI) (Lucas, 2001) by investigating the change in the likelihood of higher education enrolment between higher-tier and lower-tier institutions.

1.4.1 Thesis structure

Chapter 2 provides a review of all the relevant studies to support the research questions and hypotheses of the thesis. Two parallel strands in the literature are identified to explain how inequalities of educational opportunities are generated. The first related to the Mare educational transition model (Mare, 1980, 1981) and derived theories such as MMI and EMI, while the second related to Boudon's model of inequality of educational opportunities (Boudon, 1974) and the recent cross-country empirical applications.

Chapter 3 describes the data and gives a brief overview of the methods used in the following three analytical chapters.

Chapter 4 provides a description of the Chilean educational system with a particular emphasis on higher education expansion, since the 1980s educational reforms. The chapter concludes examining trends in higher education attainment in Chile and a selection of European countries.

Chapter 5 derives primary and secondary effects of the transition to higher education in Boudon's IEO model, by selecting the most appropriate measure of academic performance for the student cohort, between standardised tests and grade point averages.

Chapter 6 accounts for sample-selection in the student cohort, in terms of the previous educational transition, which is the completion of secondary education. Then there is a comparison of the effects of parental education between the transitions, in order to evaluate *Maximally Maintained Inequality* (MMI).

Chapter 7 examines the effect of the stratification in the education system in changing the relative contribution of primary and secondary effects from those derived in previous chapters. Finally, *Effectively Maintained Inequality* (EMI) is assessed, by testing whether or not parental education is associated with differential higher education enrolment (higher-tier versus lower-tier institutions).

Chapter 8 discusses the results from the analysis and presents the conclusions. Future areas of research are also suggested.

Chapter 2: Literature Review

2.1 Introduction

Sociological studies in the field of social stratification and mobility have been increasingly interested in the association of class origins and class destinations. In this field, the association of class and educational attainment has been chiefly to understand the process of class formation. This chapter aims to provide the theoretical and empirical grounds to support the research questions motivating this thesis.

This chapter is organised in the following way: section 2.1 briefly reviews social class theory relating to educational attainment. Sections 2.3 and 2.4 are two parallel but complementary strands of the literature on educational inequalities, which are the basis of the research questions in this thesis. The first strand is related to the Mare educational transition model (Mare, 1980, 1981) and derived hypotheses, while the second strand is related to Boudon's model of inequality of educational opportunities (Boudon, 1974). Sections 2.5, 2.6 and 2.7 compare the effect of social background on educational attainment in Chile and other countries, as well as describe Chile's education system and its expansion. Finally, the chapter concludes stating the research questions of this thesis.

2.2 Social class theory and educational attainment

Goldthorpe (2007a) differentiates two main strands in class theory, which are the Marxist theory and liberal theory. Class formation in Marxist theory is based on the historical social conflict between the groups that own the means of production and the groups that are enforced to participate in the market by selling their labour-power. Elster (1985, 1986) stated that in market economies, the basis of class is determined by individuals' *endowment-necessitated behaviour*, conditional to their constraints in the endowment structure of society. The liberal theory, in response to its Marxist counterpart, established that class formation experiences social change resulting from the development of industrialism, in which the inheritance of the social class of the father decreases its influence on the social class of his offspring. Sorokin (1927) related this social change to the transition from a rigid feudal society to an industrial society, in which the restructuration of the labour force creates an increasing demand for trained workers. In this regard, education constitutes the vehicle of social mobility so individuals can fulfil the requirements of specialised job positions. The theoretical framework of this thesis, follows the liberal theory proposition to examine the change in inequality of educational opportunities in light of the expansion of the educational system, in post-industrial societies[5].

Kerr *et al.* (1960) highlighted that constant advances in science and technology require an education system which provides training for professions created by industrial development under what they called the 'logic of industrialism'. Parsons (1960) also agreed that the content of formal education is bound to societies' technological advancement through industrialism. Moreover, he stated that educated individuals from less advantaged social backgrounds would tend to seek better-remunerated occupations than those of their

[5] In contrast, the Marxist theory in class formation does not directly account for education as the mean to change the working-class paradigm.

parents, while Treiman (1970) went further in his assertions, providing empirical evidence to support the claim that industrialisation is correlated with an increase of the educational opportunities. He recognised that the effect of education in the process of inheritance of social status would vary according to the level of educational provision in the society. The advancements of industrialism were associated with making the educational attainment less dependent on social origins and status attainment more dependent on education. Asserting that individuals themselves are the only ones responsible for their failure or success.

Bell (1973) argued that in the post-industrial society, higher education is the basis of social position, where school achievement and job competition ultimately positions a person in society. Bell acknowledge that wealthy families can provide cultural advantages to their children for schooling; for example, giving them access elite schools. He proposed that post-industrial societies are in essence meritocratic; where differences in social status and income are the consequence of differences in educational credentials. There is another important point in Bell's arguments about a meritocratic society. He states that the admission to higher education should be free of any type of discriminations regarding sex, race or religion, in order to provide equality of opportunities to all members of society. Moreover, the education system should assure that all social groups in the society have equal access to higher education, which as he sees it is the only path to higher occupational positions.

Goldthorpe and Marshall (1992) criticised this liberal position on the absence of evidence for a meritocratic society in industrialised countries. In contrast, the association of social background and educational attainment seemed to be *persistent* as Shavit and Blossfeld (1993) demonstrated in their cross-country study, despite of an increasing expansion of the education systems in industrial countries through the course of the 20th century. However, more recently, new evidence suggests a reduction in educational inequality in industrialised countries (Ganzeboom and Luijkx, 2004; Breen and Jonsson, 2005; Breen *et al.*, 2009; Breen, 2010b).

Erikson and Goldthorpe (1992) found no evidence for the assertions of liberal theory on social mobility as an expected consequence of the process of industrialism. The group of European countries analysed showed either increased social mobility, convergent rates of absolute mobility, or gradually equal relative rates of social mobility. Interestingly, they found that countries with a lesser degree of industrialisation such as Hungary, Ireland and Poland, evidenced rates of upward mobility, contrary to the expectations of liberal theory, which states that upward mobility would be more predominant in countries with a higher degree of industrialisation.

With regard to educational attainment, Erikson and Goldthorpe (1992) reasserted the mediating role of education in the social status attainment process. They assumed that actors have a general understanding of the opportunities they might have in their social context. With this information, they rationally pursue their aspirations of status acquirement. It is the social context, which will determine the degree of influence of social origins on education attainment, and educational attainment on status acquirement[6]. I will come back to this point in the review of the Breen and Goldthorpe (1997) model.

[6] In particular, they compare two Eastern European countries (Hungary and Poland) and two Western European countries (England and France).

2.3 The Mare educational transition model

The educational transition model proposed by Mare (1980) was influenced by the status attainment model of Blau and Duncan (1967), in which they introduced a new approach for investigating social mobility. In the same way, Sewell's and colleagues' (in the University of Wisconsin) subsequent studies on educational attainment and Hauser and Featherman (1976)[7] intriguing findings —for liberal theorists— on the stable effect of social background on educational attainment, inspired Mare's work.

Blau and Duncan (1967) developed a conceptual model capable of empirical applications. Their attainment model used an index of occupational status as their outcome variable. The advantage of having a continuous variable instead of a categorical variable, traditionally used in mobility tables (Glass, 1954, cited in Payne, Ford and Robertson, 1977), was the use of path analysis with linear regression. In the basic model, social origins of the individuals are measured by father's education and occupational status. These social origins then predicted the educational attainment of their sons, followed by their first occupational status, ultimately predicting their later occupational status. For men aged 20–64 in the United States, educational attainment was consistently the greater influence when determining later occupational status.

Sewell, Haller and Portes (1969); Sewell, Haller and Ohlendorf (1970) and Sewell and Hauser (1972) developed a model[8] of the status attainment process with a particular emphasis on social-psychological factors, which could account for the effect of social background on educational attainment. The social-psychological factors were measured by individuals' perceptions of the influence of significant others (e.g. parents, teachers and

[7] Their study used Occupational Changes in a Generation (OCG) survey data (1962 and 1973) for men between 20–65 years-old in the United States.
[8] Based on the basic model of Blau and Duncan (1967).

peers), educational and occupational aspirations. They gathered data from Wisconsin high-school seniors in 1957, and then followed them for seven to ten years. In 1965, information from government offices about their parents' occupation, income, and the students' annual earnings over the same time-period were linked to the data. The *Wisconsin model* demonstrated the significant predictive power of educational and occupational attainment, while it was less predictive of income attainment. The model showed the great influence of social background on educational attainment, regardless of the combinations of social background measures (e.g. parental income, education or occupation). Social-psychological factors were found to mediate the effect of social background on educational and occupational attainment, but only to a certain extent, as there is a persistent influence of social background after controlling for these factors and others such as mental ability and academic performance.

Mare (1980), knowing the aforementioned studies, proposed viewing the effect of social background on educational attainment as a sequence of individual decisions. He sought to understand how dependent educational decisions are on social background. He argued that the effect of social background on educational attainment decreases over the schooling process. As Sewell and his colleagues demonstrated earlier, he acknowledged the mediating effect of social-psychological factors, mental ability and academic performance in the influence of social background on educational attainment.

It is noteworthy that Mare's (*ibid.*) first model used to support his arguments employed a system of two regressions. The first was a linear regression for academic ability and the second a non-linear model of the decision as to whether to continue at school or not. These two models are similar to the mediation model applied in Chapter 6 of this thesis. However, he finally estimated a series of binary logistic models of school continuation decisions over different educational levels, discarding the linear model owing to the lack of a mental ability measure in his main dataset.

The data used in Mare (*ibid.*) came from the 1973 OCG survey. Moreover, he used the 1964 veterans survey of the United States military because it could be linked with a mental ability measure, which was the Armed Forces Qualifying Test (AFQT). As a mental ability measure could not be linked to the OCG survey. The sample in the OCG data were men between 20–65 years old and the veterans' data were men between 18–34 years old. In the OCG data, the social background variables such as parental education, family income and parental occupation were asked retrospectively, when the respondent was 16 years old.

Mare (*ibid.*) found evidence supporting his initial argument that the effect of social background showed a declining pattern through educational levels (i.e. from elementary education to post-college attendance). He interpreted this declining pattern as the result of selective attrition across educational levels. Moreover, he added that differences in social-psychological factors could be more relevant than social background in determining educational inequalities at higher levels of educational attainment, while differences in socioeconomic resources could be more relevant to precollege years.

Later, Mare (1981) criticised the common interpretation of inequality of educational opportunities, which was related to the rate of participation of a given birth cohort at different educational levels. For instance, if the rate of participation in upper secondary education is close to the 100%, it can be inferred that inequality of educational opportunities have declined at this level. However, this interpretation disregards how social background differences influence individuals' decisions to continue further in the education system. Therefore, he concludes that the effect of social background in a logistic model of sequential educational transitions constitute a more 'pure' measure of inequality of educational opportunities than the linear models of years of schooling upon social background had tended to evidence. The work of Mare was very influential at that time, as other studies replicated the declining pattern of social background across educational levels using logistic educational transition models (Shavit and Blossfeld, 1993).

2.3.1 Proposed theories of the social background on educational transition association

Maximally Maintained Inequality

Raftery and Hout (1993) and Hout, Raftery and Bell (1993) argued that once the higher class have satisfied their demand of higher levels of education, middle and lower classes can increase their educational opportunities at universal levels of education. In other words, background inequalities decrease at bottom educational levels as these levels become compulsory in younger cohorts. They hypothesised that *Maximally Maintained Inequality* (MMI) can explain the stable pattern across cohorts in transitions rates at higher educational levels and the stable effect of social background on these educational transitions, as long as the influence of increasing enrolment does not modify this pattern[9]

Raftery and Hout (*ibid.*) sought to explain their results by rational-choice assumptions. They argued that students and their families evaluate the costs and benefits of whether to remain in the education system or not. Families with parents holding higher educational levels would tend to value the benefit of education at a higher level, assuming the implied costs, thereby their children are more likely to continue in the education system. On the other hand, the costs associated to education would generally be higher for families from lower social background, and the future benefits unknown. As parents from low social

[9] In this thesis, I test for MMI across secondary education completion and the transition to higher education, following Lucas (2001).

background had not reached higher levels of parental education, they lack the own experience to guide their children on their educational decisions.

Raftery and Hout (*ibid.*) were not the first or the last to propose that educational inequalities can be explained by families' evaluation of costs and benefits. The first in doing so was Boudon (1974). Inspired by Boudon's study, Breen and Goldthorpe (1997) developed a mathematical model embedded in the rational choice framework, which is presented below.

Relative Risk Aversion

The Breen and Goldthorpe (1997) study sought to explain increasing participation rates and the associated effect of social background on educational attainment. They based their model on Boudon (1974) distinction of primary and secondary effects in producing inequality of educational opportunities. Primary effects are social background differences in academic ability[10]. Secondary effects are social background differences in educational decisions given the students' academic ability. These decisions in light of the model are understood to be rational. Families and their children evaluate the costs and benefits associated to each educational transition relative to their social position.

An important assumption of the model is that families' evaluations are unaffected by any cultural influence, related to the norms or beliefs that they may have about their children's education. Moreover, they assume that there is class structure defined by the contractual status in the labour market and their relationship with the production units in the economy, following a hierarchical order. Lastly, they assume a diversified education system that provides different types of education.

[10] Academic ability and academic performance are synonymous in Breen and Goldthorpe (1997).

The proposed model through three mechanisms explains class differentials in educational attainment. The first is aspirations, through which families seek to ensure their children reach at least the social class of the family (risk aversion). The second is academic ability, through which students fulfil the requirements to continue in the next level of education. Students' academic ability is associated to students own expectations of completion of the subsequent levels of education. The final mechanism are the families' resources required for education costs, which might constraint students' chances to continue in education. The Breen and Goldthorpe's model can explain stable differences of social class in educational attainment, despite growing participation rates in education, through the third mechanism. The result in the model is an increase in the proportion of service and working class children, as higher educational credentials are highly demanded in the labour market. The overall pattern of inequality will remain relatively unchanged, as the likelihood to continue in education will increase in a similar magnitude for each class.

The *Maximally Maintained Inequality* (MMI) theory in the Breen and Goldthorpe model will hold when the costs of continuing in education are less than the resources available for the service class. Additional reductions in the costs of education will have no further influence in the proportion of service class children that continue in education; rather it will increase the proportion of working class children. In terms of the model, it is not necessary that all children in the service class continue in education for MMI to hold. It is sufficient that those children in the service class, who recognise that is in their interest to continue in education, will do so. Moreover, class groups' response to reductions to the cost of education will be different, which will lead to unequal continuation rates in education.

Effectively Maintained Inequality

Lucas (2001) argued that two strands in the literature converge to explain students' progression through the education system. These are the track mobility literature, which focus on the stratification of the education system, and the educational transitions literature, which focus on sequential schooling decisions (Mare, 1980). In the light of both strands of the literature, Lucas proposed the theory of *Effectively Maintained Inequality* (EMI). This theory states that the higher social group will seek to secure some degree of advantage over other social groups. The advantage can be quantitative where, for a non-universal level of education, the participation rates of the higher social group (at that level of education) will be larger than in other social groups. Alternatively, the advantage of the higher social group can be qualitative where, for a universal level of education, the higher social group will have access to advantageous types of education (e.g. elite educational institutions) which has repercussions in their educational attainment.

Lucas' (*ibid.*) empirical analysis focused in one student cohort, as in this thesis, using data of sophomore students (grade 10) from the longitudinal survey High School and Beyond (HS&B) in the United States. He used three waves of these data (1980, 1982 and 1986), following the student cohort up to their transition to college. Moreover, these data included curricular tracks, besides social background and academic performance measures. He modelled these curricular tracks, in an ordered probit model, as different educational transitions.

Lucas (*ibid.*) sought to assess empirically MMI alongside EMI. He emphasised that MMI is useful for explaining the pattern of social background across educational transitions, besides its common use to explain variations across birth cohorts. He suggested that the observed larger effect of social background on the transition to college, compared with the lesser effect of social background on high school completion, is consistent with MMI theory. Lastly, he found evidence for EMI as the advantaged social group exhibited a larger predicted probability of continuation to the college preparation track than the less advantaged social group. Similarly, the predicted probabilities of dropout or no-course were larger for the less advantaged group.

The following subsections addresses criticisms to the Mare model and its empirical application to Chilean birth cohorts to assess MMI and EMI.

2.3.2 Criticisms to the Mare educational transition model

Cameron and Heckman (1998) criticised the declining pattern of social background on educational transitions found in the Mare model (Mare 1980, 1981) and other studies (Shavit and Blossfeld, 1993). They argued that comparisons in logistic models is not appropriate as the coefficients are estimated using different scales. The observed pattern of declining coefficients in social background is associated to the functional form of the logistic model, rather than a decreasing effect of social background on educational attainment.

The comparison of coefficients across or between logit models is not similar to the comparisons of coefficients in linear models. In linear models, the variance of the dependent variable is fixed and is the result of adding the variance of the estimated model and the residual. The variance of the estimated model and the residual can change, as explanatory variables are included into the model. In contrast, in logistic models, the variance of the dependent variable depends only of the variance of the estimated model, which can also change as predictors are subsequently added into the model. The residual variance is constant ($\pi^2/3$) in logistic models. Therefore, as the scaling is different in every logistic model, comparisons between coefficients will not reflect the true underlying pattern of the effect of social background across transitions.

As time-varying predictors in direct relation to specific transitions were not included in the logit models, only the error term is accounting for social background differences across educational transitions. Cameron and Heckman (*ibid.*) emphasised that educational transition models do not consider the effect of unobserved heterogeneity influencing educational transitions, derived from sample-selection or, to use their term, *dynamic selection bias*.

2.3.3 The educational transitions model in the Chilean context

Torche (2005) examined changes in educational attainment by social background, following the 1980s educational reform, which significantly expanded the educational system in Chile. She used data from the 2001 Chilean Social Mobility Survey (CMS), which sampled the male head of households aged between 24–69. She modelled four educational transitions: the completion of primary education, entry to secondary education, completion of secondary education and entry to higher education. The sample size was 3,544 participants, spread across seven birth cohorts. Measures of social background considered fathers' and mothers' schooling years and fathers' occupational status. Moreover, type of school funding was also included as a predictor in the models, to distinguish between privately-funded, voucher and municipal schools.

Torche's (*ibid.*) modelling strategy followed the Mare model (Mare 1980, 1981) by estimating logit models for each educational transition and birth cohort. Additionally, interaction effects between the cohorts and social background variables were considered. In the second step, dummy variables for the type of school were added into the models. The results from the transition models evidenced a stable tendency of the effect of social background on educational attainment across cohorts. She emphasised the growing influence of father's education in the transition to secondary education for the two youngest cohorts, contrary to the older cohorts. She interprets this result as evidence contrary to the MMI theory, arguing that advantaged social groups reached saturation in secondary education completion for the younger cohorts. Finally, she found qualitative differences in educational attainment by type of school funding across cohorts. Individuals that attended privately-funded schools were more likely to attain higher levels of education, compared to those attending voucher or public schools. She asserts that the distinction of type of school funding increases inequality in educational attainment, rather than mediating its effect on social background. However, she did not evaluate the EMI theory in her findings, comparing predicted probabilities of the social groups at this level.

In the following section, I review Boudon's IEO model, to continue with the second strand in the literature of inequality of educational opportunities for the purpose of this chapter.

2.4 Boudon's inequality of educational opportunities (IEO) model

The model of inequality of educational opportunities (IEO)[11], developed by Boudon (1974), has been highly influential in understanding the process of social class formation. Boudon attempted to explain the origins of IEO by challenging the assumption that a decrease in IEO will lead to a decrease in the inequality of social opportunities (ISO). Boudon emphasised two theories to frame his analysis of IEO.

First, the *value theory* (Hyman, 1953) stated that children were more likely to remain in the social circumstances they were born into because they tended to place a different value on the returns from a higher education. Boudon challenged this theory by adding a third variable to the relationship between social background and the decision to continue at various educational levels, namely *academic performance*. He demonstrated cases that diverged from *value theory*, thus questioning its validity. Second, in the 1960s the term 'cultural theory' was largely used to explain differences between the social environments children experienced at school and at home. Children from less-advantaged backgrounds may have felt less prepared to face the academic demands of school than their peers from middle or upper-class backgrounds.

Boudon (1974) decomposed IEO into primary and secondary effects, where primary effects seek to isolate the effect of social background on academic performance, and secondary effects to isolate the effect of social background on educational decisions subject

[11] Contemporary to Boudon's work, there were other studies that aimed to explain educational inequalities questioning whether the educational system was designed to reproduce social inequalities. Bourdieu and Passeron (1977) accordingly identified the social sphere in which the privileged classes endure, consolidating their cultural capital, which further enables them to negotiate and even dictate the cultural climate of the schools. However, Bourdieu's conceptual propositions are difficult to operationalise in quantitative studies and have provoked scepticism among scholars (Goldthorpe, 2007b, 2007c). In this thesis, due to lack of other measures of cultural capital, social capital and economic capital in the data at hand I concentrate on a general measure of social background following Boudon's IEO model.

to performance. Boudon’s work, therefore, aimed to explain inequalities in educational attainment, as a process of class formation, following the same line of the studies reviewed previously. Boudon’s conclusions on his IEO model can be summarised, as:

(i) Educational inequalities could only be removed in the absence of social stratification and differentiated education systems. However, the tendency is the opposite, despite the increasing developments of industrialism in western societies.

(ii) Cross-sectional data might seem to support the importance of cultural differences in determining educational inequalities. Cultural differences in Boudon’s IEO model are the primary effects. In contrast, longitudinal data will evidence that primary effects become less influential than secondary effects over time, as the effect of secondary effects grows.

(iii) The proposed model seems to provide a better account of how educational inequalities unfold, as the model assumes individuals act rationally pursuing what is best for their interests. Furthermore, cultural theory was accounted into the model, alternatively to Hyman’s (1953) *value theory*. *Value theory* is unsuccessful in explaining divergent cases, as individuals do not always behave according to their system of values associated to their social class. Moreover, Boudon criticised the frequent use of the proportion of explained variance to assess a theory. He argued that a small proportion of explained variance might be due to the theory being false, rather than assuming that other factors may explain additional proportions of the variance in the dependent variable, as in the case of the *value theory*.

(iv) The social composition of the student population can change at a faster rate than the change in educational inequality, because of increasing participation rates.

(v) The premise that in industrial societies, technological advances and economic change increase the demand of specialised training and the demand of education grows accordingly. Then it follows from the model that individuals would tend to demand more education in subsequent time-period. Thereby, the aggregate demand for education might be or might not be related to changes in the economic and social structure.

I now turn to the empirical operationalisation of Boudon's IEO model, developed in recent studies.

2.4.1 Empirical decomposition of primary and secondary effects

The methodological breakthrough of Boudon's IEO model came with the work of Erikson and Jonsson (1996, pp. 65-93), Erikson *et al.* (2005) and Jackson *et al.* (2007), who used innovative statistical methods to estimate both primary and secondary effects simultaneously. Their strategy consists of separating the distribution of academic performance and the transition probabilities of educational choices by family class and then, using a counterfactual analysis, evaluating the relative contribution of primary and secondary effects in the educational transition in question. The counterfactual analysis consists of a numerical integration that combines the academic performance and transition probabilities of students from different social classes. This allows the researchers to investigate the change in the proportion of students of a given class that make the educational transition as if they had the same academic performance and/or transition probabilities as members of another class.

Recently in the literature the method devised by Erikson *et al.* (2005) have been extended. For instance, Buis (2010) generalised the method by allowing the distribution of academic performance (primary effects) to take any distributional form, instead of assuming that performance comes from a normal distribution. Moreover, Buis allowed to derive bootstrap standard errors for both primary and secondary effects. Karlson, Holm and Breen (2010) and Karlson and Holm (2011) proposed a non-linear mediation model, in which academic performance (primary effects) mediates the effect of social background on educational transitions (secondary effects) of interest. This method performs a linear decomposition of direct, indirect and total effects for the inclusion of a predictor in non-linear models (i.e. logit and probit models) by overcoming the scale identification issue emphasised by Cameron and Heckman (1998).

Empirical studies based on Boudon's IEO model have proven to be robust to different measures of family background (e.g. social class or parental education) (Jackson, 2013a). There is, however, a difficulty on the sensitiveness that different measures of academic performance might generate on the relative contribution of primary and secondary effects to the overall IEO. Erikson *et al.* (2005), Jackson *et al.* (2007), and Erikson and Rudolphi (2010) hypothesised that students might anticipate their academic performance given their expectations of future educational attainment. These studies suggest testing the influence of anticipatory decisions by comparing the change in primary and secondary effects attributable to measures of academic performance taken close and taken long before the transition.

Jackson *et al.* (2007) and Erikson and Rudolphi (2010) found that measures academic performance taken close to the transition produce smaller secondary effects than primary effects. While secondary effects are larger than primary effects when the measure of academic performance was taken long before the transition.

The resulting effects in a recent cross-country study (Jackson, 2013a) exhibited a larger influence of secondary effects for the transition to higher education compared to the transition to upper secondary education, in the majority of the compared countries. Chapter 5 includes a more detailed discussion comparing Chile's secondary effects and the countries found in Jackson (2013a).

2.4.2 Criticisms on Boudon's IEO model

Hauser (1976) review on Boudon (1974) criticised Boudon's study on several grounds, but only criticisms referring to the IEO model are considered. Hauser argued that he failed to understand Boudon's rationale to reject the propositions of *value theory*. He believed Boudon's exposition of *value theory* was not fully developed and complemented with the empirical findings of other studies. Moreover, Boudon's model seemed to be simplistic, as social background —in effect— has a direct effect on individuals' social position and not just through education.

Hauser doubted Boudon's conclusions, as they were not fully supported empirically, as neither were his model assumptions. He emphasised that Boudon's data analysis lacked a proper interpretation, as a detailed review of the reported tables revealed some inconsistencies in Boudon's interpretations. Moreover, Boudon's simulation model was questioned as being unrealistic in its parameters, reflecting an *ideal* western society and assuming constant continuation rates across transitions. Boudon (1976) responded to Hauser (1976) criticisms. However, Boudon's (*ibid.*) response adds very little to the discussion of what was already stated in Boudon (1974).

On the other hand, Halsey, Heath and Ridge (1980) evidenced that in the British education system, class differentials are attenuated over time. They argued that it might a result of fewer branching points compared with the large number of branching points showed in Boudon's model, which reassert secondary effects over time. Therefore, they reasoned, this difference might be the result of differences between *sponsored* and *contested* systems of mobility through education[12]. In *contested* systems, a greater relative contribution of secondary effects is more likely than in *sponsored* systems. Certainly, Boudon's model was assumed in a *contested* education system.

The resulting primary and secondary effects in Halsey, Heath and Ridge (1980), for England and Wales, did not support Boudon's claim of the greater importance of secondary effects in determining IEO. They observed major class differences between schools (greater primary effects) than within schools (lesser secondary effects). This result can be explained by the organisational features of the British educational system as follows: (i) the British system of secondary education is *sponsored*, (ii) it has fewer branching points, (iii), the tendency of survival rates[13] to converge over schooling, and (iv) survival and attendance rates differ among types of secondary schools. To conclude, Halsey, Heath and Ridge critique on Boudon's IEO model was the lack of anticipation on how the relative importance of primary and secondary effects might change, in education systems with different organisational features, such as those in *sponsored* systems.

Focusing on the British minority ethnic population, Waters *et al.* (2013) proposed that secondary effects have an opposite influence than in Boudon's thesis of the greater importance of secondary effects in explaining class differences. Based on Heath and Brinbaum (2007, cited in Waters *et al.*, 2013, p. 133), they argued that a minority ethnic student tends to remain in education if they anticipate discrimination in the labour market for lower educational credentials. Moreover, ethnic minorities are more likely to have higher

[12] Turner (1960) argued that the *sponsored* mobility involves individuals' selection by the elite, giving them elite education necessary to place them into positions of power. While in *contest* mobility, although there is no selection, contestants constantly compete through schooling; the top achievers earn elite status.
[13] Survival rates refer to the proportion of students that remain in education.

aspirations in educational attainment. Their results suggested that secondary effects did not influence ethnic minorities likelihood of secondary education completion.

Finally, Sullivan *et al.* (2014) identified four major difficulties in Boudon's model, which are:

(i) it should not be assumed that students have an unvarying cognitive ability, as several studies in the literature tend to assume, ability should be treated as variable over time;

(ii) the interpretation of primary effects cannot be seen as determined by only cultural differences, as Boudon argued, cultural factors might also explain secondary effects;

(iii) the interpretation of secondary effects cannot be seen solely in relation to economic differences, economic factors can also explain primary effects;

(iv) the assumption of branching points in the education system does not clearly explain the accumulation of educational credentials in adults, and it ignores previous years of education such as primary education years;

(v) students' expectations of educational attainment are not explicitly accounted for in Boudon's model, which can also influence students' performance at school.

Regarding the argument in (ii), Bulle (2009) responded to similar criticisms made earlier in the literature. She argued that Boudon did not formulate an *ex-ante* hypothesis explicitly separating cultural and economic differences in determining primary and secondary effects. However, it seems that in the literature primary effects have been traditionally related to cultural differences, and secondary effects to economic differences among children. She stated that inferences from Boudon's model are not compromised in the set social background factors affecting the overall educational inequality.

Sullivan *et al.* (2014) analyses of the 1970 British Cohort Study emphasised that secondary effects were determined to a large extent by parental education, which can be considered a cultural factor. This is in contrast to the non-significant effect of social class in secondary effects, where social class is considered as an economic factor. They concluded that Boudon's proposition that secondary effects are mainly accounted for by economic factors was not supported in their study. However, Jackson (2013b) found no significant differences between social class and parental education or even a cross-classification of both measures to derive primary and secondary effects. Similarly to Sullivan *et al.* (2014), Jackson (2013b) based her analyses on the 1970 British Cohort Study (BCS), with links to the 1958 National Child Development Study (NCDS) and the 1985 Youth Cohort Study (YCS).

At the time this chapter was written, no published study estimated Boudon's primary and secondary effects using data from Chile. However, a longitudinal study of Chilean students investigating their educational careers was found. It has a different focus than Torche (2005) study and will be discussed in the next subsection.

2.4.3 Longitudinal study of educational careers in Chile

Schiefelbein and Farrell (1982) collected a sample of Chilean students attending their last grade of primary education (grade 8) in 1970, and then following them until the first years of university (1977). They found that social background has a significant effect on the likelihood of completing secondary education in the four years of its duration. However, their results showed that national standardised tests, administered in grade 8, seemed to have an even greater predictive effect than social background differences (Schiefelbein and

Farrell, 1982, p. 93). In other words, children that performed well in grade 8 were likely to succeed in completing school, irrespective of their social background.

Similarly, higher achievers in grade 8 tests were more likely to take university entrance exams, despite being from low social backgrounds. Finally, the transition to university for low social background students was largely explained by their performance in university entrance exams, where the effect of social background was less pronounced for this group (Schiefelbein and Farrell, 1982, pp. 104–110).

The results from Schiefelbein and Farrell (1982) can be related to Boudon's IEO model, as the influence of academic performance in shaping students educational careers. Moreover, Schiefelbein and Farrell (1982) findings are comparable to the findings in Chapter 5 of this thesis, as it is striking that in both cohorts, though separated by 39 years, the effect of national standardised tests in grade 8, despite social background, is still influencing the transition to higher education.

What follows is a review of international studies in educational attainment in which data from Chile was included.

2.5 International comparative studies in educational attainment

Pfeffer (2008) sought to investigate the effect of institutional arrangements in educational systems influencing educational mobility patterns in 20 countries, among them Chile. Based on Kerckhoff (1995) he argued that institutional features in the education system determine individuals' opportunities of educational attainment. Moreover, he added that the degree of stratification in the education system affects the role of parental education in educational attainment. That is to say, the higher degree of stratification, and the higher the effect of parental education on educational attainment. He used the International Adult Literacy Survey (IALS), collected between 1994 and1998, including individuals aged 26–65 years old. The countries in the sample are Canada, Germany, Ireland, the Netherlands, Poland, Sweden, Switzerland, the United States, Australia, Belgium, Great Britain, Northern Ireland, New Zealand, Chile, the Czech Republic, Denmark, Finland, Hungary, Italy, Norway, and Slovenia.

The results from log-linear and log-multiplicative models showed that in the majority of the countries the association between parental education and educational attainment has remained constant across age groups. Chile is among this group of countries. The divergent results were found in the Czech Republic, Hungary, Norway and Finland. However, countries differ significantly in the extent of educational mobility. The ranking of educational mobility presented in the study classified countries as Finland, Northern Ireland, New Zealand and Denmark, with the highest educational mobility. While educational mobility is the lowest in Slovenia, Germany, Belgium and Switzerland. Chile and Ireland have average educational mobility for the countries considered in the study.

To investigate the effect of the stratification of the education system, Pfeffer (*ibid.*) classified education systems in three groups: weakly stratified, highly stratified and very highly stratified. In weakly stratified systems, students are mobile between tracks and the access to post-secondary education is independent their choice of track. In highly stratified systems, schools differ in their curriculum orientation as they prepare students to enter post-secondary education or to enter the labour market, and students are less mobile between schools. In very highly stratified systems, further to the characteristics on highly stratified systems, schools select students at an early age and mobility between school types is restricted. Moreover, the degree of standardisation of the education system across the country is considered as being standardised or not (i.e. two groups of classification).

Chile is classified within the group of highly stratified and standardised education systems. Other countries in this group are the Czech Republic, Hungary, Italy and Poland. In one extreme of the classification, Germany and Switzerland have very highly stratified and standardised education systems, while in the other Canada, Great Britain and the United States have weakly stratified systems with low standardisation. Pffefer (*ibid.*) found evidence to support his initial hypothesis that the effect of parental education tends to be larger in more highly stratified education systems.

On the other hand, Hout (2006), sought to assess *Maximally Maintained Inequality* (MMI) and *Effectively Maintained Inequality* (EMI) using a cross-country dataset. The data comes from the Social Inequality module of the International Social Survey Programme (ISSP), collected between 1999 and 2001. The participating countries were Ireland, Norway, Sweden, France, West Germany, East Germany, Spain, Portugal, Cyprus, Hungary, the Czech Republic, Slovakia, Slovenia, Poland, Russia, Latvia, Australia, New Zealand, the United States, Canada, Brazil, Chile, the Philippines, Israel, and Japan. The sample was limited to those 25 years old and over.

Levels of education were modelled using ordered logit regression. The levels of education were complete primary, incomplete secondary, complete secondary, incomplete post-secondary, and complete post-secondary education. The social background measures were father and mother educational attainment and number of books at home. Moreover, a control for gender, interaction effects between social background and country and interactions across birth cohorts were added to the model. Ireland was selected as the reference category in the set of dummy variables of countries.

Hout (*ibid.*) showed that only in market economies (in contrast to socialist economies[14]) MMI seemed to be supported. The effect of social background on educational attainment seem to be larger in West Germany, Portugal and Spain compared to Ireland, despite this group of countries (and Chile) having lower enrolment rates in post-secondary education. The effects of social background in Chile are not significantly different from those in Ireland[15]. On the other hand, the rest of market economies had a smaller effect of social background on educational attainment, but larger post-secondary enrolment rates compared to Ireland. In other words, the effect of social background in post-secondary education attainment in Chile was ranked in the middle of the distribution of the compared countries.

The main result of the study is that the effect of social background, from the coefficients of the model, decreases across cohorts as post-secondary enrolment increases. Similarly, this pattern was replicated for the comparison of social background and GDP per capita. Therefore, Hout (*ibid.*) concludes that these results are consistent with EMI, but not with MMI, since the decrease of the influence of social background should have been seen after saturation was reached.

[14] The socialist countries are the Czech Republic, East Germany, Hungary, Latvia, Slovakia, Slovenia, Poland and Russia.
[15] Surprisingly, Pfeffer (2008) also ranked Chile alongside Ireland, but in terms of educational mobility.

2.6 Summary

Two strands in the literature were identified that sought to explain social inequalities in educational attainment. This thesis relies on the strand of the decomposition of the primary and secondary effects of inequality of educational opportunities proposed by Boudon (1974) and their subsequent empirical developments. However, the hypotheses derived from the educational transitions strand of the literature (Mare 1980, 1981); namely *Maximally Maintained Inequality* (MMI) (Raftery and Hout, 1993; Hout, Raftery and Bell, 1993) and *Effectively Maintained Inequality* (EMI) (Lucas, 2001) provide alternative explanations worth to explore in the analyses. These two hypothesis (i.e., MMI and EMI) as well as the criticisms to the Mare model are not ignored in the studies using Boudon's IEO model (Jackson, 2013a). Hence, both strands of the literature can be complementary for the aim of this thesis, which is to examine social-background differences in Chilean students continuing to higher education.

This chapter concludes presenting the research questions of this thesis in the following section.

2.7 Research questions

An important feature of the reviewed studies is that they argue that the continuous expansion of the educational system weakens the association of social background on educational attainment. It was shown that the Chilean educational system expanded rapidly since the reforms in the 1980s. Accordingly, the trend in higher education attainment has been steadily increasing over birth cohorts. However, studies on educational attainment in

Chile suffer from inadequate data[16] and ignored limitations of their empirical analysis[17] (Torche, 2005). Further, they are outdated[18] or do not address any of the proposed theories to explain educational inequality considered in this thesis (i.e. the Mare model and the subsequent theories or Boudon's IEO model) (Schiefelbein and Farrell, 1980). Therefore, this thesis seeks to fill these gaps in the literature by examining social-background differences in Chilean students continuing to higher education, and to evaluate the applicability of Boudon's inequality of educational opportunities (IEO) model.

The two main research questions in this thesis are:

1. To what extent are social inequalities associated with students' chances of progression into higher education?
2. To what extent does the Chilean educational system increase or decrease the effect of social background on the transition to higher education?

The first research question is addressed in Chapters 5 and 6, while the second research question is addressed in Chapter 7. Both research questions are addressed by decomposing primary and secondary effects in Boudon's IEO model over the three analytical chapters. In Chapter 5, the average primary and secondary effects are derived for the student cohort using the *ad hoc* methods developed in the literature of Boudon's IEO model (Erikson *et al.*, 2005; Buis, 2010; Karlson, Holm and Breen, 2010). In Chapter 6, the critique to the Mare transition model in relation to sample-selection is accounted for in the modelling strategy to test the validity of the effects found in Chapter 5. Moreover, in Chapter 6, *Maximally Maintained Inequality* (MMI) is assessed across the different transitions of

[16] Limitations of the data in Torche (2005) are: (i) women are not part of the sample; (ii) there is a lack of academic performance or cognitive ability measures, and (iii) the data is cross-sectional.
[17] Torche (2005) ignored all the criticisms in Cameron and Heckman (1998) to the Mare model, which was her modelling strategy.
[18] The student cohort on Schiefelbein and Farrell (1980) attended grade 8 in 1970 and was followed until 1977.

secondary education completion and the transition to higher education. It is noteworthy to emphasise that a strict assessment of MMI cannot be accomplished, as only one student cohort was available at the time the analyses were conducted. I follow Lucas (2001), who assessed MMI across educational transitions for one student cohort. Finally, in Chapter 7, *Effectively Maintained Inequality* (EMI) is assessed when the transition to higher education is disentangled by type of higher education institution and programme.

The specific research questions for each analytical chapter are presented below:

Chapter 5

- Are SIMCE tests in grade 8 more appropriate than GPA levels to identify social background differences in academic performance? Are anticipatory decisions underestimating secondary effects in GPA levels?

  This research question is twofold. First, in Boudon's rationale students can only decide whether to continue in the education system, by knowing beforehand their individual academic performance. In Chile, students acknowledge their individual GPA levels, but they ignore their individual SIMCE scores because the Ministry of Education uses SIMCE scores for educational policies[19]. The influence of academic performance to continue to higher education depends on the quality of the measure of academic performance. I reason both GPA levels and SIMCE tests assess students' performance according to the national curriculum. However, GPA levels can differ in their calculation across schools as assessment practices depend on schools' internal policies as well as differences in curriculum tracks (i.e. general or

[19] If individual test scores would be known, schools could refuse to enrol students with lower performance in SIMCE tests, as schools receive economic incentives from public programmes, to reward high-achiever schools. For more details on these programmes see Mizala and Urquiola (2013).

vocational). SIMCE tests, on the other hand, avoid all the aforementioned difficulties, as they are a standardised measure. Moreover, scholars and policy-makers in Chile rely on SIMCE tests to assess school differences in academic performance, (subsection 2.5.1). Therefore, I hypothesise that SIMCE tests will capture a larger share of social background differences in academic performance than GPA levels, meaning larger primary effects.

Second, differences in secondary effects between GPA levels and SIMCE tests could be the result of anticipatory decisions (Erikson *et al.*, 2005; Jackson *et al.*, 2007; Erikson and Rudolphi, 2010). The meaning of anticipatory decisions is that students made decisions about their educational aspirations in advance; then, they accordingly increase or decrease their efforts on academic performance to achieve them. The opposite of Boudon's rationale of the operation of primary and secondary effects, in which academic performance precedes educational decisions. As a result, secondary effects are underestimated (Jackson *et al.*, 2007; Erikson and Rudolphi, 2010). I hypothesise secondary effects in GPA levels will have an influence similar to secondary effects in SIMCE tests in the overall IEO. Therefore, the anticipatory-decisions hypothesis will be unsupported.

- Are secondary effects more important (>50%) than primary effects in the transition to higher education?

  Boudon (1974) argued that secondary effects are larger for latter transitions, due to the higher cost associated to the access of higher education and the benefits are unknown to students and families from less advantaged social backgrounds. Moreover, Breen and Goldthorpe (1997) argued that the overall pattern of social inequality remains relatively unchanged, despite increasing participation rates in education. In the same line, I argue that social background differences determine the transition to higher education in Chile, as higher education is mainly financed through tuition fees (section 2.6) . Moreover, in cross-country comparisons of higher education transitions (Jackson and Jonsson, 2013), secondary effects exhibited a larger share than primary effects for the majority of countries. Therefore, I hypothesise secondary effects will be larger (>50%) than primary effects in the transition to higher education.

- What is the relative position of Chilean students' secondary effects in comparison with the countries studied in Jackson and Jonsson (2013), in the transition to university?

  Chile falls in the middle of the distribution in two key international comparative studies which rank the effect of social background on educational attainment (Hout, 2006) and the level of educational mobility (Pfeffer, 2008). Both studies included a majority of European countries. I hypothesise that Chile's ranking of secondary effects will lay in the middle of the distribution of the Jackson and Jonsson (2013) cross-country classification of secondary effects. Among the compared countries is Germany, which has a highly stratified and selective educational system, while England falls at the other end of the distribution, having a low stratified educational system (Pfeffer, 2008; Jackson and Jonsson, 2013) .

Chapter 6

- Is sample-selection producing a downward bias in the effect of social background on the transition to higher education?

Educational transition models using the Mare model tend to exhibit a declining pattern of social background at higher levels of education. Cameron and Heckman (1998) argued that it is owing to the effect of unobserved heterogeneity, which derives in sample-selection. The empirical strategy proposed to test for sample-selection in Boudon's model has not been applied before[20] in the literature, owing to the large attrition in the student cohort, just 64% completed secondary education. I hypothesise that the effect of social background on the transition to higher education is biased downward; thereby secondary effects will tend to be larger than those found in Chapter 5.

- Do secondary effects in the completion of secondary education and higher education transition have opposite contributions to the overall IEO? If so, is this result consistent with *Maximally Maintained Inequality* (MMI)?

Lucas (2001) found evidence consistent with MMI, as the effect of social background was larger in the transition to higher education compared with the smaller effect of social background in high school completion. Moreover, Jackson (2013b) interpreted the larger contribution of secondary effects in the transition to higher education compared to the smaller contribution of secondary effects in the transition to A-levels as consistent with MMI theory.

In Chile, Torche (2005) argued that her results contradict MMI because of the growing effect of father's education in the transition to secondary education for the youngest cohorts. Hout (2006) argued that his findings are not consistent with a strict

[20] At least at the time this empirical strategy was applied.

version of MMI, in reference to the group of compared countries but without any particular reference to Chile. However, both studies assessed MMI across birth cohorts but not across transitions.

Breen and Goldthorpe (1997) argued that MMI holds in relation to the costs of higher education for the higher class. I showed in Figure 2.4 that households in the last two years (2013–2014) still account over 50% in the distribution of higher education expenditure, implying greater financial constraints for lower class families. In contrast, in the schooling system municipal schools are free of charge, which increases lower class participation at this level of education. Therefore, I hypothesise that social background will be found to matter to a larger extent for the transition to higher education than it will for secondary education completion, which is consistent with MMI across transitions.

Chapter 7

- Does the stratification of the schools[21] increase social background differences for the transition to higher education?

Pfeffer (2008) classified the Chilean school system as highly stratified owing to the existing curriculum tracks (i.e. general and vocational) and the low student mobility across schools. However, other school characteristics were not considered such as (i) type of school funding, (ii) location area, and (iii) the large number of primary or secondary education only schools, which causes that about 64% of students change schools at the end of grade 8 (Lara, Mizala and Repetto, 2010). I hypothesise that accounting for school characteristics increase the importance of secondary effects in the transition to higher education.

- Does the distinction of the transition to higher education by type of institution and programme change the contribution of primary and secondary effects?

Rudolphi (2013) argued that individuals of higher social background tend to select the most prestigious educational tracks to preserve their advantage. These prestigious courses generally have higher entrance requirements, thereby implying stronger primary effects. In Chile, traditional universities are the most prestigious in the country. These institutions are highly selective in terms of academic performance, meaning that for these institutions primary effects are likely to be higher than for private universities and vocational colleges. Since primary and secondary effects add

[21] I refer to the stratification of the schools to differences in type of school funding, curriculum tracks and location area among schools for the period of compulsory education. The educational system, on the other hand, encompasses schools and higher education institutions.

to the total effect of educational inequality, secondary effects will be proportionally smaller in traditional universities. Similarly, primary effects by type of programme will have a greater contribution in undergraduate programmes than in vocational programmes. I hypothesise that the contribution of primary effects will increase for the transition to traditional universities and undergraduate programmes, as admission requirements rely on students' academic performance. In contrast, as high academic performance is not required in vocational colleges or programmes, secondary effects will tend to increase even more for these transitions.

- Are students from a lower social background more likely to enrol in vocational colleges and students from a higher social background more likely to enrol in traditional (prestigious) universities? If so, is this result consistent with *Effectively Maintained Inequality* (EMI)?

Torche (2005) found evidence of qualitative inequality in the organisational features of the educational system, by considering type of school funding. However, she did not properly assessed EMI by comparing predicted probabilities of advantaged and less advantaged social groups to access the different types of school. I hypothesise, that students from lower social backgrounds are more likely to be enrolled in vocational colleges, while students from higher social backgrounds are more likely to be enrolled in traditional (prestigious) universities. If this is the case, there is consistent evidence to support EMI.

Chapter 3: Data and methods

This chapter outlines all data sources required to create the longitudinal data to track the student cohort examined in this thesis, followed by the methodology applied for each analytical chapter.

The section that follows describes the three main data sources to create the longitudinal dataset for the student cohort examined in this thesis, these are: (i) national student register, (ii) national standardised tests, and (iii) university entrance exams. Moreover, other complementary data sources are referred, such as the European Social Survey (ESS) and Chile’s household cross-section surveys (CASEN). The subsequent sections describe the process of data linkage, the student cohort trajectories over the 12-years period of compulsory education. Finally, this chapter ends with a brief overview of the methodology employed in the subsequent chapters, which aim to answer the research questions proposed in Chapter 2.

3.1 Data

The major challenge for the analysis was to create a longitudinal dataset suitable to investigate a student cohort throughout the period of compulsory education in Chile’s educational system. Unlike developed countries, birth cohort studies are non-existent in Chile. The alternative relies upon the use of administrative data sources such as the national student register (*Registro de Estudiantes de Chile*, RECH) renamed in 2008 to (*Sistema Información General de Estudiantes*, SIGE), which collects data of all children in the country enrolled in compulsory education. The national student register began in 2002, meaning that one cohort can be tracked throughout the 12 years of compulsory schooling.

With additional data, their higher education trajectories can also be tracked. This makes the national student register the ideal available baseline data for this longitudinal study.

3.1.1 National student register (RECH-SIGE)

The Chilean Ministry of Education collects student data from all the schools of the country every year in order to keep a record of students' enrolment and academic performance data, along with information about the schools themselves. The enrolment data gathers basic demographic information about students such as gender, date of birth and municipality of residence. Moreover, it collects information related to the school such as grade, educational level, type of school funding (i.e. municipal, voucher, private), school location area (i.e. urban or rural), among others. Every school in the country sends this information to the Ministry of Education until the 30th April of every year, since the school year in Chile begins in March and ends in December, as it is common in south hemisphere countries. The academic performance dataset contains similar information as the enrolment dataset, adding students' grade point averages (GPAs), attendance over the year (e.g. 0–100%) and their academic promotion status at the end of the school year (i.e. promotion, retention, dropout).

The baseline dataset for the analysis is the academic performance data (AP) because this study aim is to look at students' transitions at the end of every school year. However, the information from the enrolment data (EN) allowed tracking students that might have dropped out over the year; therefore, would not have been included in the AP data at the end of the school year.

3.1.2 National standardised tests (SIMCE)

Information about students' social background can be linked to the student register from parents' questionnaires collected in national standardised tests (*Sistema de Medición* *de la Calidad de la Educación*, SIMCE). SIMCE tests were designed to monitor schools with the purpose to evaluate students' knowledge of the national curriculum regarding core subjects[22] such as Spanish and mathematics. SIMCE tests are applied to all students in the country enrolled in the grade selected for assessment in a particular year.

The selected student cohort took SIMCE tests in grades 4 and 8 of primary education that is 2005 and 2009 respectively. The variables of interest for the student cohort are highest parental education (SIMCE grade 4) and the average of scores in Spanish and maths in both grades of SIMCE tests' applications.

3.1.3 University entrance exams (PSU) and higher education enrolment (SIES)

Students attending their last grade of secondary education (grade 12) can voluntarily register to take university entrance examinations (*Prueba de Selección Universitaria*, PSU). In other words, PSU exams are not compulsory for all secondary education graduates. All university admission's criteria in Chile require PSU exams, in contrast to vocational colleges, which just require a certificate of secondary education completion.

Similarly, to SIMCE tests, PSU exams assess students in their knowledge of core subjects such as Spanish and mathematics. Other subjects such as chemistry, physics, history and biology are required for admission to specific undergraduate programmes (e.g. medicine, engineering, geology among others). For the aims of this thesis, I only consider the average score of Spanish and mathematics. The student cohort took PSU exams at the end of grade 12 (2013) in order to attend a higher education institution by 2014.

The data of higher education enrolment comes from the administrative datasets collected by SIES (*Servicio de Información de Educación Superior*) a department specialised

[22] Other subjects' assessment depends on the guidelines of the Ministry of Education for a particular grade, these can be sciences and/or social studies.

in higher education information within the Ministry of Education. The variables of interest in these data were higher educational enrolment by type of institution and programme.

3.1.4 European Social Survey (ESS) and Chile's Household Survey (CASEN)

The European Social Survey (ESS) and Chile's household survey (*Encuesta de Caracterización Socioeconómica Nacional*, CASEN) were used as supporting datasets, to compare Chile's trends in higher educational attainment alongside European countries. Seven rounds of the ESS from 2002–2014 were grouped to conform the five cohorts to be analysed (Table 3.1). The selected countries were all those that had data available for the period. The data used for Chile come from two cross-sectional household surveys (CASEN) collected in 2003 and 2013. The ESS cohorts weighting was adjusted by design effects to correct for different probabilities of sample selection in each round, following Ganninger (2007). For Chile's surveys, it was not necessary to adjust for design effects, since the first 4 cohorts are self-contained in CASEN 2003 and only the younger cohort comes from CASEN 2013.

Table 3.1 Cohorts in ESS Rounds 2002 to 2014 and Chile's CASEN 2003 and 2013

Country	Cohort					Total
	1933 to 1942	1943 to 1952	1953 to 1962	1963 to 1972	1973 to 1982	
Belgium	1,247	1,848	2,175	2,156	1,127	8,553
Chile*	9,417	12,614	17,898	20,347	26,200	86,476
Denmark	1,249	1,971	1,870	1,917	885	7,892
Finland	1,669	2,679	2,355	2,106	1,136	9,945
France	1,475	2,205	2,183	2,287	1,314	9,464
Germany	2,747	3,160	3,799	3,641	1,501	14,848
Hungary	1,346	1,858	1,937	1,768	1,238	8,147
Ireland	1,680	2,449	2,367	2,668	1,954	11,118
Netherlands	1,617	2,396	2,455	2,611	1,125	10,204
Norway	1,017	1,916	2,087	2,257	1,064	8,341
Poland	1,154	1,736	2,239	1,763	1,122	8,014
Portugal	2,176	2,297	1,990	2,017	1,253	9,733
Slovenia	1,128	1,474	1,625	1,537	758	6,522
Spain	1,399	1,689	2,088	2,513	1,587	9,276
Sweden	1,460	2,270	1,907	2,143	1,098	8,878
Switzerland	1,421	1,916	2,141	2,419	1,026	8,923
United Kingdom	1,984	2,551	2,435	2,776	1,433	11,179

Notes: Chile's cohorts range from 1934 to 1983.
Source: Data retrieved from ESS (2002; 2004; 2006; 2008; 2010; 2012; 2014) and CASEN (2003; 2013).

3.1.5 Chile's Household Survey (CASEN) 2011

This particular year of Chile's household survey (CASEN 2011) was linked to the student cohort data to create instruments to control for sample-selection. Despite Chapter 6 goes in detail about sample-selection, students' trajectories through the period of compulsory education reveal a periodic reduction of the cohort's original sample, especially for the year 2011. Students' trajectories can be seen in section 3.3.

Cameron and Heckman (1998) suggested that there is an associated downward bias in the effect of social background at latter transitions. To address this, I control for sample-selection in the transition to higher education. To identify the sample-selection specification, I create three instrumental variables, which are labour force participation rate for the age-group 15–19, separated for boys and girls, and the percentage of teen maternity. These instruments are proportions at the regional level using data from CASEN (2011) and linked to the student cohort data using area codes of the region of students' residency. These variables were among the most frequent reasons for dropping out of school in 14–17 year-olds, for the lower quintiles of income reported in CASEN (2011) (Ministerio de Desarrollo Social, 2013, p. 141).

3.2 Data linkage

Official statistics from the Ministry of Education reported that 272,955 students were enrolled in grade 1 in 2002 (Ministerio de Educación, 2003). Common to international educational research using developing countries' official statistics, there is a lack of validations implemented by the Ministry of Education to evaluate the consistency of the figures reported compared to previous years. This single available figure also includes those students that are repeating grade one and those that will drop out before the end of the year. I have obtained a more reserved and consistent estimation of the cohort of interest linking enrolment data (ED) with data from the academic performance data (APD), collected at the end of the school year, and validating the data using key markers. The steps applied over the cohort's compulsory education (2002–2013) are detailed in Table 3.2. The initial estimation of enrolment shows a reduction in the official statistics by almost 12,000 students. This reduced figure is consistent in terms of students' identification numbers (ID), sex and date of birth over the period.

Table 3.2 Student cohort data linkage

Validation steps	Students	Percentage
The student cohort in grade 1 (2002), linking the student register data (ED and APD) over the period 2002–2013. Every student in the cohort was validated in *ID number*, *sex* and *date of birth* over the period.	260,993	100
The student cohort born in 1995–1996, who by 2002 were 6–7 years-old. The required age to start compulsory education.	254,393	97.5
The student cohort validated through their promotion status over the period. For instance, if the students were promoted in 2002 to grade 2, they should be found in the student register enrolled in grade 2 in 2003.	233,905	89.6
The student cohort who remained in the educational levels for children and adolescents over the period. That means that they did not enrolled in special education or adult education.	210,340	80.6
The student cohort validated in their corresponding grade in both the register and standardised tests (SIMCE). For instance, students with information in grade 8 SIMCE tests should be found in the same grade in the register.	210,168	80.5

The following steps of data validation in Table 3.2 sought to make the student cohort as homogenous as possible in order to make valid comparisons. One of the criteria was to restrict the sample to those born in 1995–1996, who in 2002, reached the required age to enter grade one (6–7). Another important criterion of selection was that students remained in regular education over the period, which is the educational experience of typically developing children. There were a group of students in the cohort that started education in grade one, but in the following years were enrolled in courses of special education, oriented to children with special educational needs such as learning disabilities. Similarly, another group of students in the cohort in secondary education were enrolled in adult education. The adult education pathway offers a simplified curriculum to advance two regular school years in one calendar year. Students who were retained in previous grades or decided to participate in the labour market are the most likely to be enrolled in adult education.

3.3 Students' trajectories over the period of compulsory education (2002–2013)

This subsection sheds light on the cohort's promotion, dropout and retention rates relative to each of the 12 grades of compulsory education, in relation to the final sample size of the student cohort found in the data linkage section.

To illustrate, the promotion rate of grade *n+1* is the result of dividing the number of students enrolled in grade *n+1* over the total 210,168 students in the cohort. A grade *n* can take the values from 1–12, which correspond to the period 2002–2013. The retention rate in grade *n+1* was found by dividing the number of students held back in grade *n* over the total student cohort. Similarly, the dropout rate in grade *n+1* was found by dividing the number of students not enrolled in school in year *n+1* over the total student cohort. The three rates add to 100% each year.

Figure 3.1: Grade promotion, retention and dropout rates over 12-years of compulsory education

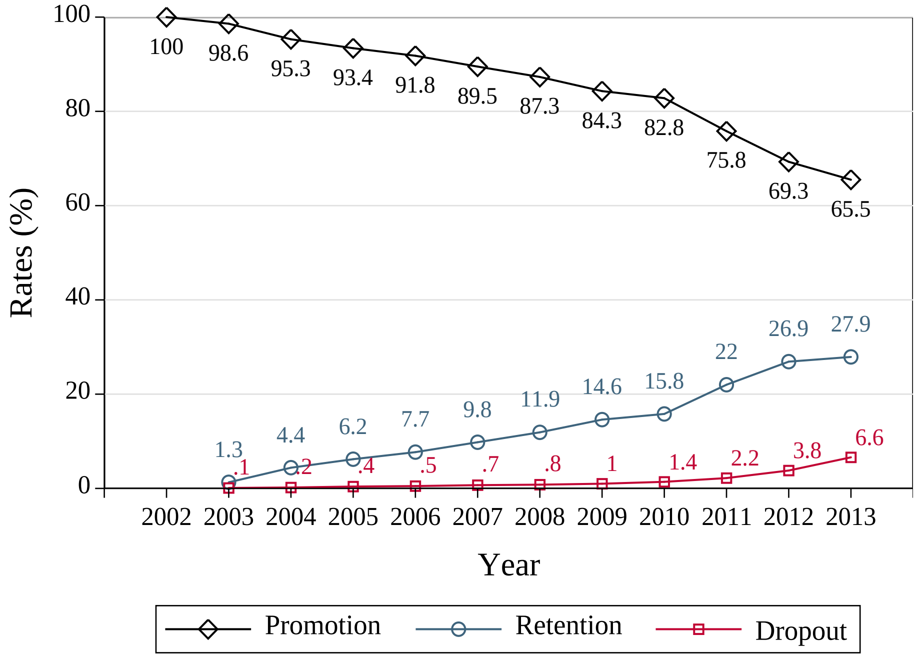


Figure 3.1, shows there was a steadily decreasing trend in promotion rates in the student cohort over the first 9-years, from 2002 to 2010. When the cohort was in grade 9 (when students were between 14 to 15 years-old), the promotion rate decreased, on average, by 2.3% annually, over this 2002 to 2010 period. Conversely, the retention and dropout rates increased, on average, by the same proportion. Over the next two years (2011 and 2012), the promotion rate decreases sharply by 6.7% annually, on average.

The sharp decrease in promotion rate, evident during the years 2011 and 2012, might have been, at least in part, due to the legal working age in Chile, which is 15 years. It could also have been a consequence of the student protests held in 2011, which interrupted the school year for several months. In Chapter 6, I will investigate the possible influence of external factors in the sample-selection of the cohort such as the rate of teens (15–19) participating in the labour force and the percentage of teen maternity for the same age group, both measured at the regional level.

The key statistic for the aim of this thesis is the successful completion of secondary education[23], which is the necessary condition for students to attend higher education. The ultimate completion rate of secondary education was 64% for the student cohort, which is a little less than the promotion rate of 65.5% in 2013, owing to the fact that this rate was calculated with respect to students' enrolment in grade 11, the year before.

A comparison with a similar group of European countries than those selected from the ESS survey (Table 3.1), can put Chile's cohort completion rate into perspective. In Table 3.3, according to OECD (2012)[24], successful completion of secondary or upper secondary education was over 70% in Ireland, Poland, Slovenia, Sweden and Finland. A second group of countries had a similar rate to Chile (over 60%). This group included Belgium, Hungary,

[23] Successful completion of upper secondary education can also be understood as successful completion of school. Both concepts are used interchangeably.

[24] Cohorts used in the analysis range from 1999 and 2007 as the first year of entry to upper secondary. Chile's cohort entered upper secondary in 2012.

the UK and the Netherlands. Finally, the rest of the countries had below 60% completion of secondary or upper secondary education, which included Denmark, France, Spain and Norway. Nevertheless, OECD (2012) also reports rates of completion up to two years after the theoretical duration of secondary or upper secondary programmes. Consequently, it may be possible for the European countries described in the second and third group to reach higher successful completion rates than the rates described in this particular Chilean cohort.

Table 3.3 Completion of secondary or upper secondary programmes in European countries (%)

Country	Method	Year used for new entrants Duration of programme (G: general, V: vocational)	N = theoretical duration	Total	Boys	Girls
Belgium (Fl.)	True cohort	2004-05	within N	69	62	77
		4 years G & V	2 years after N	85	82	89
Denmark	True cohort	2002-03	within N	59	54	64
		3-4 years G & 2-5 years V	2 years after N	73	69	76
Finland	True cohort	2004	within N	70	68	72
		3 years G & V	2 years after N	80	78	83
France	Longitudinal sample survey	1999-2005	within N	59	54	64
		3 years G & 2 years V	2 years after N	82	78	85
Hungary	Proxy cohort data	2006-07	within N	68	64	72
		4 years G & V	2 years after N	*m*	*m*	*m*
Ireland	True cohort	2004	within N	87	84	90
		2-3 years G & V	2 years after N	*m*	*m*	*m*
Netherlands	True cohort	2007	within N	61	56	67
		2-3 years G & 2-4 years V	2 years after N	78	75	82
Norway	True cohort	2004	within N	57	48	66
		3 years G & 4 years V	2 years after N	72	68	76
Poland	True cohort	2006-07	within N	80	75	84
		3 years G & 3-4 years V	2 years after N	*m*	*m*	*m*
Slovenia	Proxy cohort data	2007	within N	76	72	81
		4 years G & 3-4 years V	2 years after N	*m*	*m*	*m*
Spain	Proxy cohort data	2006-2007	within N	57	53	61
		2 years G&V	2 years after N	82	80	84
Sweden	True cohort	2006	within N	72	70	74
		3 years G&V	2 years after N	79	77	81
United Kingdom	True cohort	2006	within N	61	56	67
		2 years	2 years after N	80	76	85

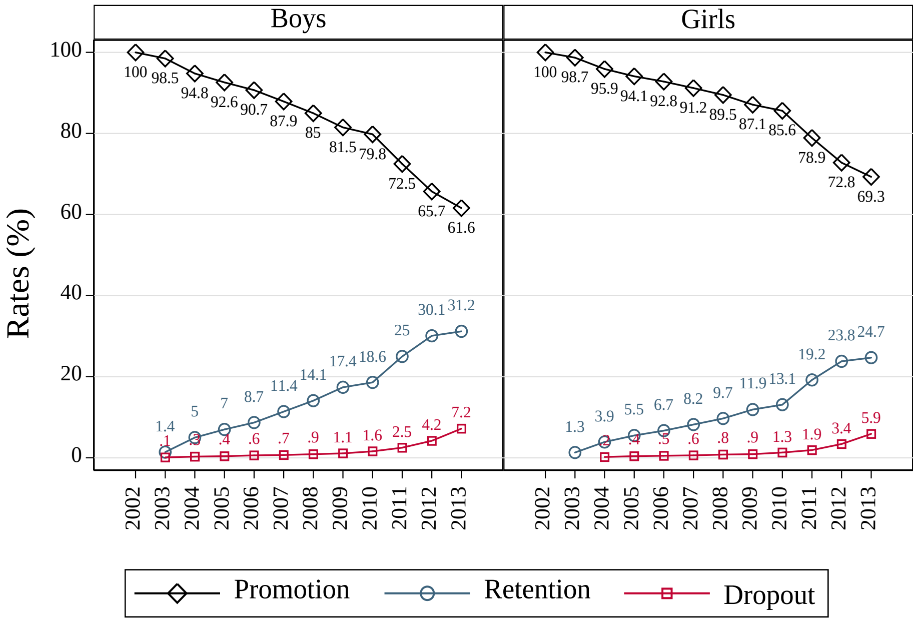
Source: OECD (2012, p. 57).
Note: This table is a shorter version of Table A2.5. Successful completion of upper secondary programmes by gender.
m is missing data.

Proportion tests for the rates presented in Figure 3.3 indicate that annual changes in promotion and retention rates were significantly different ($p<0.05$) over the entire period, whereas dropout rates were only significantly different ($p<0.05$) during the last four years (corresponding to the four years of secondary education in Chile's educational system).

3.3.1 Students' trajectories by gender

In Figure 3.2, the trend of promotion rate by gender indicates that a larger proportion of girls were promoted than boys. For both girls and boys, there is a decline in the promotion rate in the last three years of secondary education, but it is more marked for boys. The tendency is the opposite for retention and dropout rates. Girls have a lower proportion of retention and dropout rates, whereas boys have a larger proportion of the same rates.

Figure 3.2: Student cohort promotion, retention and dropout rates by gender



A difference in proportions test between boys and girls, confirms that differences in promotion rates since grade 3 (2004) up to grade 12 (2013) are statistically significant ($p<0.05$). The completion rate of secondary education for boys is 60% and for girls 68%. Likewise, completion rates in OECD European countries (OECD, 2012) were larger for girls.

3.4 Analytical approach

The methods in each of the analytical chapters are based on regression modelling. There are two recurrent dependent variables in the regression models, these are (i) transition to higher education and (ii) academic performance. The transition to higher education is generally[25] measured as a binary variable to whether or not the student was enrolled in higher education a year after completing secondary education. While academic performance is a standardised measure with mean zero and standard deviation of one for GPA levels, SIMCE average scores of maths and Spanish, and PSU average scores of maths and Spanish as well.

The recurrent independent variable over all analytical chapters is parental education, which is measured in four educational levels, these are (i) university degree, (ii) vocational qualifications, (iii) secondary education and (iv) primary education or none formal education. However, variables such as gender, school funding and school location area are used as controls in some models.

It follows a brief description of the methods implemented in the analytical chapters.

3.4.1 Methods in Chapter 5

This chapter's aim is to measure primary and secondary effects in Boudon's inequality of educational opportunities (IEO) model (Boudon, 1974) for the transition to higher education. In most recent studies, different methods have been proposed to estimate primary and secondary effects. However, the Erikson *et al.* (2005) method has gained popularity in the literature and has been preferred for cross-country comparisons (Jackson, 2013a). This method estimates separated logit models for each social group, in this case parental educational levels, where the transition to higher education is regressed upon

[25] The exception is the multinomial mediation model in Chapter 7.

academic performance[26]. Then, through numerical integration over the assumed normal distribution of academic performance, predicted and counterfactual probabilities of transition are derived for different combinations of academic performance and the log-odds of transition —from logit models— for each social group. The specific aspects of this method can be seen in Chapter's 5 methodological appendix.

The relative contribution of primary effects and secondary effects are calculated from counterfactual odds ratios of various combinations of performance and the log-odds of transition to higher education, by groups of parental education.

3.4.2 Methods in Chapter 6

This chapter's aim is to test whether sample-selection might be creating a downward bias in the effect of parental education upon the transition to higher education. The regression model is a non-linear mediation model (MacKinnon, 2008; Hayes and Preacher, 2014) for the transition to higher education. The mediating effect of academic performance is captured by a separated linear regression of SIMCE scores. Both models control for parental educational levels. Primary effects are the relative proportion of the indirect effect of SIMCE scores upon the transition to higher education over the total effect of parental education upon the transition. Secondary effects are the relative proportion of the direct effects of parental education —controlling for SIMCE scores— upon the transition over the total effect.

The resulting primary and secondary effects from the non-linear mediation model are similar to the effects found in Chapter 5 using the Erikson *et al.* (2005) method. The advantage of the mediation model over Erikson *et al.* (2005) method is that can also account

[26] The academic performance measure is the average of Spanish and mathematics scores in national standardised tests (SIMCE) or the average of grade point averages of the four years of secondary education. A detailed explanation of the selection of both measures of academic performance can be seen in Chapter 5.

for sample-selection within the structural equation modelling framework (Skrondal and Rabe-Hesketh, 2004; Muthén, Muthén and Asparouhov, 2016).

Missingness in parental education and SIMCE scores is addressed by a two-level multiple imputation model following multiple imputation routines implemented in Mplus (Muthén and Muthén, 1998–2017), addressed in Asparouhov and Muthén (2010), and Enders (2015).

3.4.3 Methods in Chapter 7

This chapter's aim is to investigate the effect of the stratification of the school system in the transition to higher education. Furthermore, it differentiates the transition to higher education by type of institution and programme.

The first regression builds on the non-linear mediation model applied in Chapter 6, by accounting for a two-level cross-classified design of the data. The cross-classified design was applied because 65.1% of the student cohort changed schools between grade 8 and grade 12, which are the grades, associated with SIMCE tests and the secondary education completion. Students change schools in that proportion because there is a great number of primary-only schools and secondary-only schools in Chile's school system. For more details see Lara, Mizala and Repetto (2010) and Racynski (2011).

The second group of models are multinomial mediation models, which seek to disentangle the transition to higher education by type of institution and programme. Therefore, primary and secondary effects are differentiated by these variables. Unfortunately, multinomial outcomes cannot be estimated in cross-classified models in Mplus, which was used to estimate the models in this chapter.

Chapter 4: The Chilean educational system

4.1 Introduction

After the period of Chile's independence from Spain (1810–1818), a series of educational policies were implemented, which aimed to increase population's participation in education. According to Schiefelbein and Farrell (1980), Chile was the first country in Latin America to create a public education system. Years later, in 1860, free primary education was established for all children up until age 11. In 1920, primary education was made compulsory. Presidential decrees at the end of the 19th century allowed women to access public secondary and higher education, resulting in increasing female participation at these educational levels. In 1970, the proportion of women enrolled in secondary education was over 50% of the total enrolment (Barrera 1976, cited in Shiefelbein and Farrell, 1980).

The educational reforms in direct relation to the student cohort studied in this thesis took place during the military government headed by General Augusto Pinochet, who over-took the democratically-elected socialist government of Salvador Allende, starting a period of 17 years of dictatorship (1973–1990). One of the first measures implemented by the military government was the decentralization of the central government services by the division of the territory into regions, provinces and municipalities (Parry, 1997). Having prepared the field for more decentralisation measures, new educational reforms were applied in the 1980s. These reforms, as before, also aimed to increase participation by instigating a major structural change in the country's school system: restructuring the administration and funding of state and private-subsidized schools.

Before the 1980s reforms, the school system was separated into three types of schools: state schools, publicly subsidised but privately-owned schools, and private schools. Private schools remained free of any compulsory structural change.

Three features of the reforms salient to this thesis are: (i) delegating state administrated schools to municipalities; (ii) introducing a voucher system of finance to replace subsidies, vouchers depend on the number of students attending these schools; and (iii) teachers ceased to be public employees, their contracts were transferred to private administration in voucher and private schools, or municipalities.

Some researchers (Bellei and Cabalin, 2013; Mizala and Torche, 2012) stated that the 1980s reforms were ideologically inspired by Friedman (1962) propositions of privatisation schemes, encouraging school competition (between municipal and voucher schools) in order to boost families' selection set of available schools.

The current education system comprises of eight years of primary education (grades 1–8) and four years of secondary education (grades 9–12). Secondary education has been compulsory since 2003. Moreover, two main curriculum tracks, general and vocational, exist in the last two grades of secondary education (grades 11 and 12). The general track prepares students for university education, while the vocational track trains them for entry into the labour market. However, students can still access university education after completing the vocational track, the higher education system does not restrict the enrolment of students graduated from the vocational track. Both academic and vocational tracks can be offered at the same school, although there are also specialised schools that offer only one the aforementioned tracks. Schools can be comprehensive, including both primary and secondary education, or just primary or secondary schools.

Higher education offers two types of programmes: (i) undergraduate programmes of 4 to 6 years leading to a degree, offered in traditional[27] and private universities; and (ii) advanced vocational qualifications of 2 to 4 years, offered in vocational colleges such as

[27] Traditional universities are those state and private funded before the 1981 reform. More details will be seen in section 4.2.

professional institutes (PI) and technical training centres (TTC). Table 4.1 summarises the educational levels in the Chilean educational system.

Table 4.1 Educational levels in the Chilean educational system

Education	Level	Grade (From)	Grade (To)	Age (From)	Age (To)	Years	Notes
Primary	Primary Education	1	8	6/7	13/14	8	After the first 2 years of secondary, students continue to one of the following tracks.
Secondary	Common Curriculum	9	10	14/15	15/16	2	
	General Curriculum	11	12	16/17	17/18	2	Secondary ed. graduate.
	Vocational Curriculum	11	12	16/17	17/18	2	Secondary ed. graduate.
Tertiary	Advanced Vocational Qualifications					2–4	Technical training centres (TTCs), 2 years. Professional institutes (PIs), 4 years.
	First Level					4–6	Degree[28] (undergraduate programmes of 4 to 6 years)
	Second Level					2	Master's degree
	Third Level					3–5	Doctoral degree

It follows a revision of the most important studies in Chile centred on the 1980s reforms, as well as other following policy measures to gain a glance of the effects of the reforms in the educational system.

[28] In Chile, university degree or bachelor degree is called *Licenciatura*.

4.2 The effects of the 1980s reforms

In 1993, after the dictatorship period, the voucher system was modified, entitling schools to charge families additional fees in order to increase schools' monetary resources. The co-payment system was created, allowing voucher school owners to legally establish their schools as *for-profit* institutions. However, not all voucher schools are for-profit or fee-charging institutions: such is the case in religious schools.

Over time and as a direct consequence of the voucher funding system, a great proportion of students migrated from public schools (now municipality schools) to private-voucher schools. In 1981, public school enrolment was at 80%, which was reduced to 38% by 2012. Voucher schools, by contrast, increased their enrolment from 12% in 1981 to 55% by 2012 (Mizala and Torche, 2012).

The change in the enrolment rate by type of school led to an increase in socioeconomic segregation for families unable to afford the costs of voucher schools, even though paradoxically voucher schools were still publicly subsidized (Hsieh and Urquiola, 2003). Moreover, it is interesting to note that voucher schools are overrepresented in urban areas, in contrast to rural areas. Apparently, these schools seem to attract middle-class families (Mizala and Romaguera, 2000).

Mizala and Romaguera (2000) found no consistent differences in students' performance in national standardised tests (*Sistema de Medición de la Calidad de la Educación*, SIMCE) between voucher and municipal schools when sufficient control variables were added into the model. Similarly, Tokman (2002) showed that municipal schools were more efficient in terms of performance in standardised tests than voucher schools, in particular for students from less advantaged backgrounds. However, other studies have shown that students attending voucher schools do have significantly higher performance in standardised tests—although in small magnitude—compared to students' performance in municipal schools (Sapelli and Vial, 2005; Anand, Mizala and Repetto, 2009).

Other studies attempted to identify the impact of competition between voucher and municipal schools measured by students' performance in SIMCE tests. Hsieh and Urquiola (2006) found that voucher schools benefited from students of socioeconomic advantaged families, while municipal schools attracted a larger proportion of students from less advantaged families. They demonstrated that the effect of competition between schools was negligible in educational outcomes such as performance in SIMCE tests, grade retention rates and years of schooling. In contrast, Gallego (2006, 2013) showed that greater levels of competition between schools significantly increased educational outcomes.

Critics to the school voucher system have seen the introduction of vouchers, as a mechanism of sorting students, generating socioeconomic segregation between schools and, hence, increasing inequality (Torche, 2005; Hsieh and Urquiola, 2006; Mizala and Torche, 2012; Elacqua, 2012). Other researchers have argued that vouchers should have a higher value for socioeconomically less advantaged students (Neal, 2002; Gallego and Sapelli, 2007; Epple and Romano, 2008).

In 2008, a new-targeted voucher system looked to have a direct effect on socioeconomically less advantaged students (40% of the poorest student population) in municipality and voucher schools. The studies of Mizala and Torche (2012); Neilsen (2013); Villarroel (2014); Correa, Parro and Reyes (2014), among others have sought to measure the impact of this new measure in standardised test scores and other educational outcomes. Importantly, for the purpose of this thesis, the new-targeted voucher system did not benefit the selected student cohort.

4.3 Higher education expansion in Chile

The educational reform in higher education, implemented in 1981, is a milestone on the restructuring of the education system in Chile. New private universities and vocational colleges were granted access into the higher education system, as the system before the reform had become inefficient due to its dependency of the central government (Fried and Abuhadba, 1991). Moreover, the total enrolment in higher education had plummeted by 30% between 1975 and 1980 (Bernasconi and Rojas, 2003, p. 21). Therefore, the military government, at that time, aimed to diversify the system to increase student enrolment.

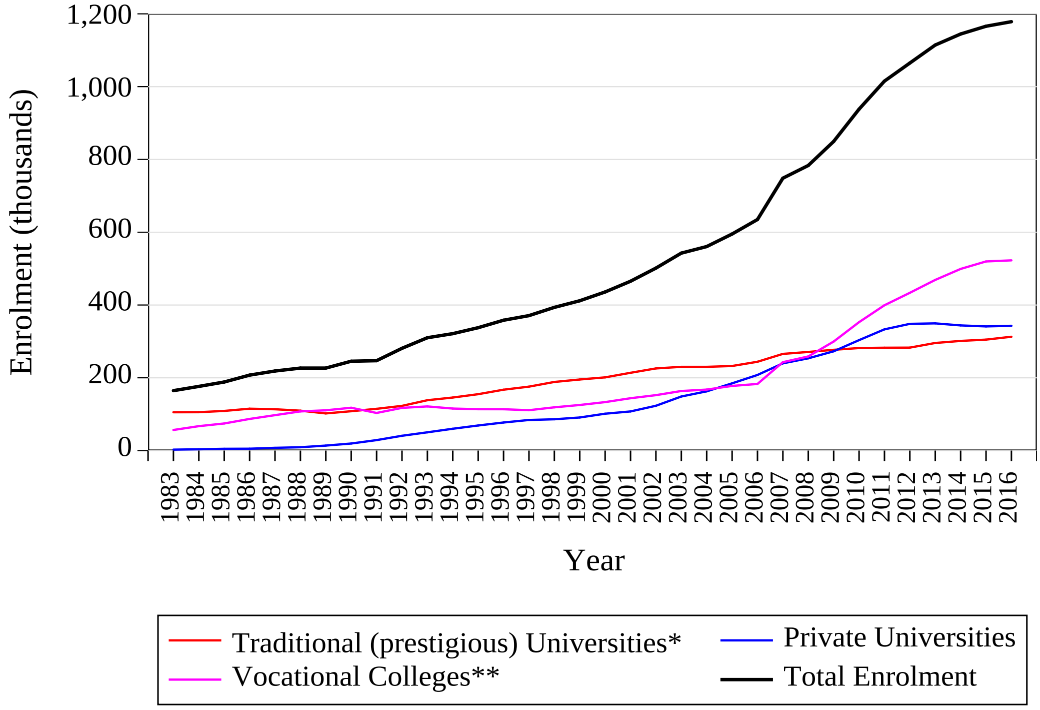
The entry of new private higher education institutions created two additional tiers of institutions in the higher education system. The first-tier institutions are the traditional universities (*Consejo de Rectores de las Universidades Chilenas*, CRUCH universities), state and private universities founded before the 1981 reform. The second-tier institutions are private universities founded after the 1981 reform; and the third-tier institutions are vocational colleges such as professional institutes (PI) and technical training centres (TTC) (*Institutos Profesionales* and *Centros de Formación Técnica*) that offer advanced vocational qualifications. The growth in enrolment after the 1981 reform is presented in Figure 2.2 by the three tiers of higher education institutions.

The black line in Figure 4.1 depicts the total enrolment for the period of 1983 to 2016[29]. Since private higher-education institutions (i.e. private universities and vocational colleges) entered into the education system, enrolment started to grow. In the 1980s, the growth rate was modest[30], 38% (i.e., [(enrolment(1989)-

[29] This is the period of data available in SIES (2016a), Ministry of education.
[30] Bernasconi and Rojas (2003, p. 23) argue that the fact that the Chilean Home Office required authorisation to approve private higher education institutions might have impacted in the expansion of higher education enrolment up to 1988.

enrolment(1983))/enrolment(1983)]×100)[31]. In the 1990s, higher education enrolment was steadily growing at 68% rate. In the 2000s enrolment reached its peak at 95% rate. Apparently, there has been a deceleration in the growth rate since 2010, compared with the previous decade; the growth rate has been 26% so far. The most significant expansion of the higher education enrolment occurred in the 2000s, in which student enrolment was nearly double (95%).

Figure 4.1: Total enrolment by higher education institutions 1983–2016



Notes: * In Chile known as *Consejo de Rectores de las Universidades Chilenas* (CRUCH).
** In Chile known as Professional Institutes and Technical Formation Centres (IP & CFT).
Data retrieved from SIES (2016), Ministry of Education[32].

The enrolment growth in traditional universities has increased steadily over the period. It tended to be above the enrolment of the other institutions, although at the end of the 1980s and the beginning of 1990s was at the same level of the enrolment in vocational colleges. Since 2009, however, the enrolment in vocational colleges and private universities

[31] ((226,703-164,665) /164,665)*100 = 38%;((411,650-245,561) /245,561)*100 = 68%; ((849,340-435,884) /435,884)*100 = 95%.

[32] It is possible that one student can be enrolled in more than one programme in the same year.

has surpassed the enrolment in traditional universities. In particular, the enrolment in vocational colleges is noteworthy higher. According to Arum, Gamoran and Shavit (2007), institutional differentiation of higher enrolment rates is expected in market systems of higher education. Another distinctive feature in this type of system is that first-tier institutions become more selective by raising their admission criteria. In contrast, lower tier institutions might be engaged on client-seeking activities, since these institutions are mainly funded by tuition fees. These features of a market system of higher education match the features found in the Chilean higher education system. Moreover, their findings suggest that the reliance on private funding to expand the education system is beneficial to a certain extent. The privatisation of the system will increase social inequality, as tuition fees will be not affordable for all.

Higher education funding

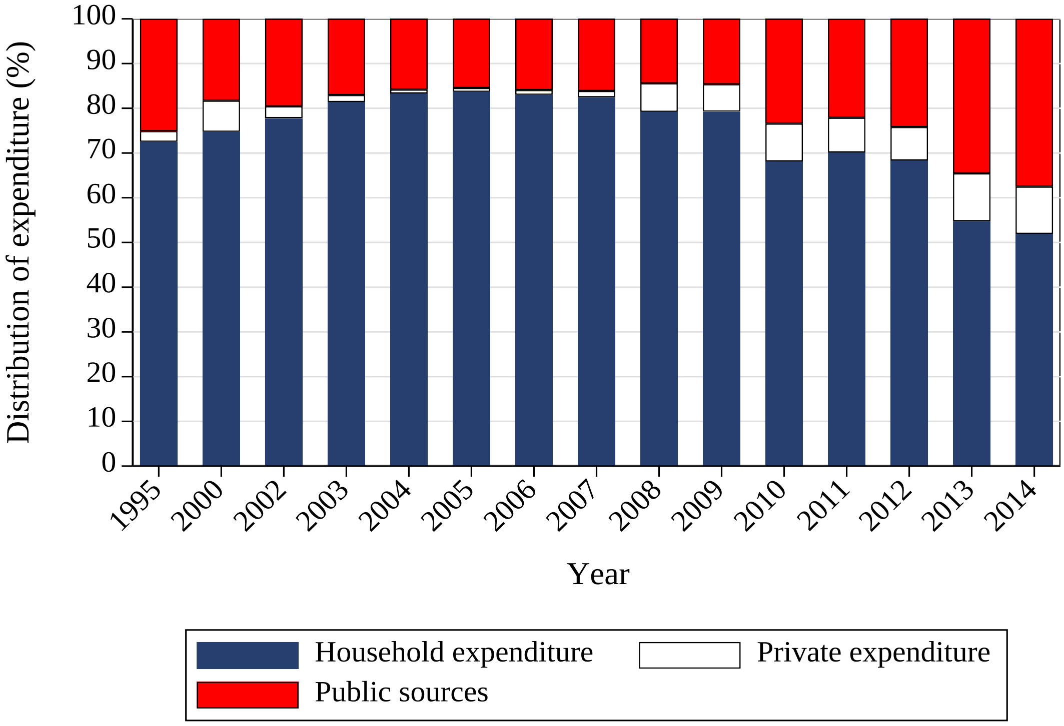
Traditional universities[33] were the only existing universities before the 1981 education reform. These universities used to be funded almost entirety (Letelier, 2014) by the central government. Nowadays, public funding is estimated to constitute up to 17% of each traditional university annual budget (Herrera and Muñoz, 2013). This is despite the fact that, on average between 2011 and 2015, over a third (34%) of the total public funding for higher education[34] is spent on these institutions (Contraloría General de la República, 2012, 2013, 2014, 2015, 2016).

[33] Traditional universities are called the universities that were created before Pinochet's military government. In this group, there is a separation between state traditional universities and private traditional universities. There are 16 state traditional universities and 9 private traditional universities. The private traditional universities are catholic universities, whereas the state universities are secular. Traditional universities are considered more prestigious than private universities created after the 1980s reform.

[34] The public funds that private institutions created after the 1981 reform are less than 40% in average for the period 2011 to 2015. Traditional universities (CRUCH), state and private, receive the larger share of public funds, 61.4% in average for the same period.

The degree of marketisation in higher education can be operationalised as the share of private expenditure on higher education institutions (Arum, Gamoran and Shavit, 2007, p. 22). The OECD (2003–2016) data on higher education expenditure by public and private sources in Chile, for the period 2000 to 2014, exhibits the larger proportion of funding as coming from private sources (Figure 4.2). Particularly, the share of households' expenditure on higher education through tuition fees was above 70% between 2000 and 2009. It was between 2010 and 2012 that funding streams began to grow in public and private institutions. The most significant growth came from public sources, which had an average rate of 16.4% in the 2000s, increasing to 23.2% between 2010 and 2012. Finally, the data from the 2013 and 2014 showed a major increase in public sources: 34.6% and 37.5% respectively. Private institutions expenditure was also highly increased to an average of 10.6% in the aforementioned years. The result was that household expenditure experienced a sharp decrease from over 70% in the 2000s to 54.8% in 2013 and 52% in 2014.

Figure 4.2: Distribution of public and private expenditure on higher education institutions, 1995–2014



Note: Data retrieved from OECD (2003–2016).

A possible explanation of the decrease in household expenditure in the last few years (2012–2014), and the subsequent boost of public expenditure in higher education might be the result of the negotiations between the student movement and the government, to tackle inequality of access to higher education.

The Chilean government issued new regulations to forbid for-profit private institutions of higher education. It also reduced interest rates on university loans and increased scholarships and the number of students that benefited from them. For more details of the policies implemented after the student movement in 2011, see Bellei and Cabalin (2013); Stromquist and Sanyal (2013) and Bellei, Cabalin and Orellana (2014).

OECD indicators in education as well as the OECD and World Bank (2009) report of tertiary education in Chile, provided the evidence to sustain students' claims of the unequal access to higher education in Chile resulting from the market-oriented higher education reform of 1981. In 2011, the OECD reports used data available up to 2008, as shown in Figure 4.2, which was reported as household expenditure in higher education in the 2000s (80.7% in average). Although, the household expenditure was already showing signs of decrease in 2011, the data was unknown to students, apparently, as well as by the government officials who handled students' demands.

Returns to higher education attainment

Hastings, Neilson and Zimmerman (2013) sought to investigate the returns to higher education in Chile. They used administrative data from 1985 to 2011 of 30 cohorts of applications to traditional universities (CRUCH) linked to administrative records of labour market income. Their findings showed larger positive returns to selective undergraduate programmes in health, law, social science, engineering and business. In contrast, non-selective undergraduate programmes showed small or negative returns in arts, humanities and education. However, their study did not include all higher education institutions (e.g. private universities or vocational colleges).

The OECD and World Bank (2009) reported differences in salaries that are up to four-times between degrees and secondary education attainment, using data from the national household survey (CASEN 2006). In terms of the returns to higher education attainment, González-Velosa *et al.* (2015) found that, on average, degrees have a rate of return of 62% over individuals with only secondary education attainment. While advanced vocational qualifications have, on average, a null rate of return over the same group. González-Velosa *et al.* (*ibid.*) used data from the Ministry of education on tuition fees and programmes' duration, as well as on employment and salaries up to four years after graduation. The estimated return rates are the average of the net present value of projected earnings of individuals between 24 and 65 years-old, graduated from higher education programmes, compared to the projected earnings of individuals between 18 and 65 years-old who attained secondary education.

González-Velosa *et al.* (*ibid.*) found a large dispersion in the rates of returns over the variety of higher education programmes, offered by universities and vocational colleges. The rate of return on vocational programmes for the 90th percentile reached 52% and for the 10th percentile was -41%. In contrast, the return rate on university degrees was 167%, for the 90th percentile, and -15% for the 10th percentile. It is noteworthy, that the proportion of higher education graduates with positive returns, at 78%, is substantially larger for university degrees. Whereas the proportion of higher education graduates, holding vocational qualifications with positive returns is 49%. These results suggest that there is a higher financial risk in acquiring an advanced vocational qualification in contrast to pursuing a university degree.

Remarks on higher education expansion

As it was seen in Figure 4.2, in the last few years (from 2009 onwards) the growth of higher education enrolment was driven by the enrolment in vocational colleges. However, returns associated to advanced vocational qualifications seem to be null or even negative. As individuals from lower social backgrounds are more likely to attain vocational qualifications, what can be inferred from these results? The work of Brint and Karabel (1989) sheds some light on this matter for the case of junior or community colleges in the United States. Their thesis is that community colleges divert students' opportunities from universities and consequently divert them from job positions of higher order in the labour market hierarchy. On the positive side, Brint and Karabel (*ibid.*) argued that students who are not able to attend a university might see community colleges as an alternative to access higher education. Moreover, they criticised that community colleges (at the time of their study) lacked any transfer opportunity to university degrees, as it is the case in Chile's current vocational colleges (OECD and World Bank, 2009, p. 44).

Walters (2000) argued that it is important to understand the distribution of individuals across the social strata in the country in order to increase the opportunities of educational attainment. In Chile, it seems that higher education expansion acknowledges the social stratification in the country but, as Boudon (1974) argued, social inequalities are a consequence of society rather than the education system. Nevertheless, it is still to be seen how inequality of educational opportunities unfold for the student cohort studied in this thesis.

In the following subsection, I compare higher educational attainment between Chile and a selection of European countries, according to their data availability on the European Social Survey (ESS).

4.4 Higher education attainment across European countries and Chile

In order to gain a further understanding in trends of higher education attainment in Chile, and its relative position compared to a selection of European countries[35], five birth cohorts in the age range 30–69, following Breen *et al.* (2009), were examined using 7 rounds of the European Social Survey (2002; 2004; 2006; 2008; 2010; 2012; 2014) and Chile's cross-section household surveys CASEN (2003; 2013). The birth cohorts were grouped as: (1) 1933–1942, (2) 1934–1952, (3) 1953–1962, (4) 1963–1972 and (5) 1973–1982.

The definition of higher education graduates was adjusted to the one employed by the highest authority of higher education in Chile (*Consejo Nacional de Educación*, CNED), which is to have completed at least a first university degree or advanced vocational qualifications (technical or professional), conditional on secondary education completion. The International Standard Classification of Education (ISCED 1997) used in the ESS and in Chile to allocate national education programmes, classifies university degrees as level 5A and advanced vocational qualifications (technical or professional) as level 5B. In the ESS the ISCED 5B also defines advanced vocational qualifications.

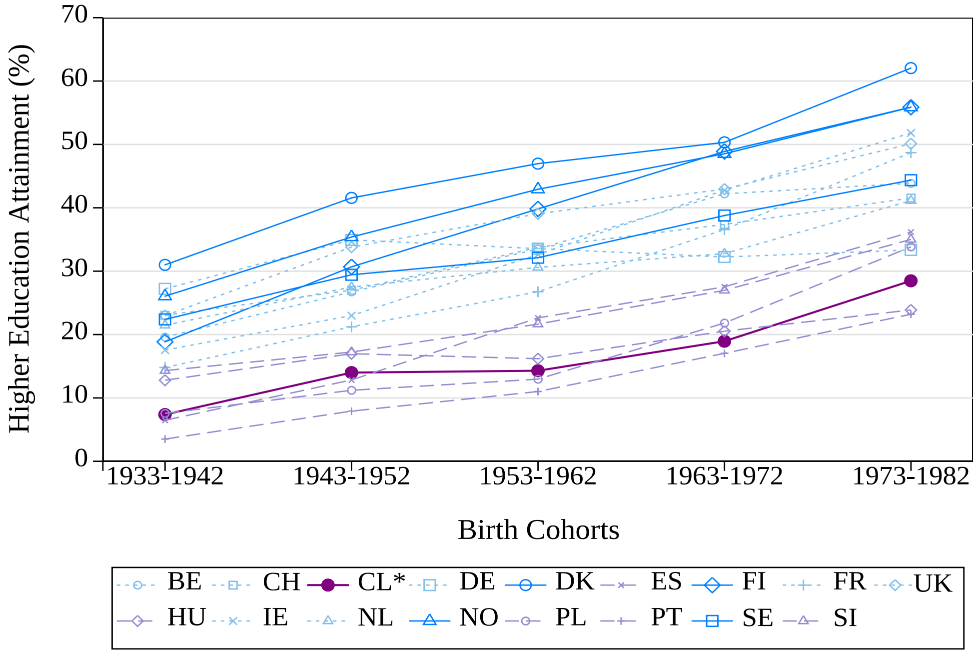
Trends in higher educational attainment (Figure 4.3), highlight three groups of countries. The first group are Scandinavian countries that tend to have higher proportions of higher education graduates in contrast to the rest. These are depicted in blue solid lines. Danish cohorts are in the top of the trend in this group while Sweden is at the bottom.

[35] It was not possible to access cohort data for other Latin American countries, as it was not publicly available at the time this comparative analysis was done. The only international data available were cross-country surveys such as *Latinobarómetro*. However, the sample sizes in *Latinobarómetro* are too small (around 1,200 observations per country) to compare birth cohorts. Moreover, design effects or weights that serve the purpose to combine different rounds of the *Latinobarómetro* survey were not available. For these reasons, trends in educational attainment in Chile were compared to European countries with information available for the period 2002–2014.

The second group includes other western European countries, with the exception of the Iberian Peninsula. This group (dashed blue lines in Figure 4.3) show a mix of tendencies among cohorts. Nevertheless, the UK comes on top of the group with a larger proportion of higher education graduates, alongside the Republic of Ireland in the two younger cohorts. France was the country with a smaller number of graduates at this level in the first three cohorts, while Germany was at the bottom in this group in the last two cohorts.

The third group includes Spain and Slovenia, with a larger proportion of higher education graduates in the last three cohorts. Portugal is the country that is at the bottom of this group from all selected countries in all cohorts. Chile's higher education graduates are just above Portugal, and the younger cohort of Chile is above Hungary's graduates too.

Figure 4.3: Higher Education Attainment across European countries and Chile



Notes: *Chile's cohorts range from 1934 to 1983.
BE=Belgium, CH=Switzerland, CL=Chile, DE=Germany, DK=Denmark, ES=Spain, FI=Finland, FR=France, UK=United Kingdom, HU=Hungary, IE=Ireland, NL=the Netherlands, NO=Norway, PL=Poland, PT=Portugal, SE=Sweden, SI=Slovenia.
Data retrieved from ESS (2002; 2004; 2006; 2008; 2010; 2012; 2014) and CASEN (2003; 2013).

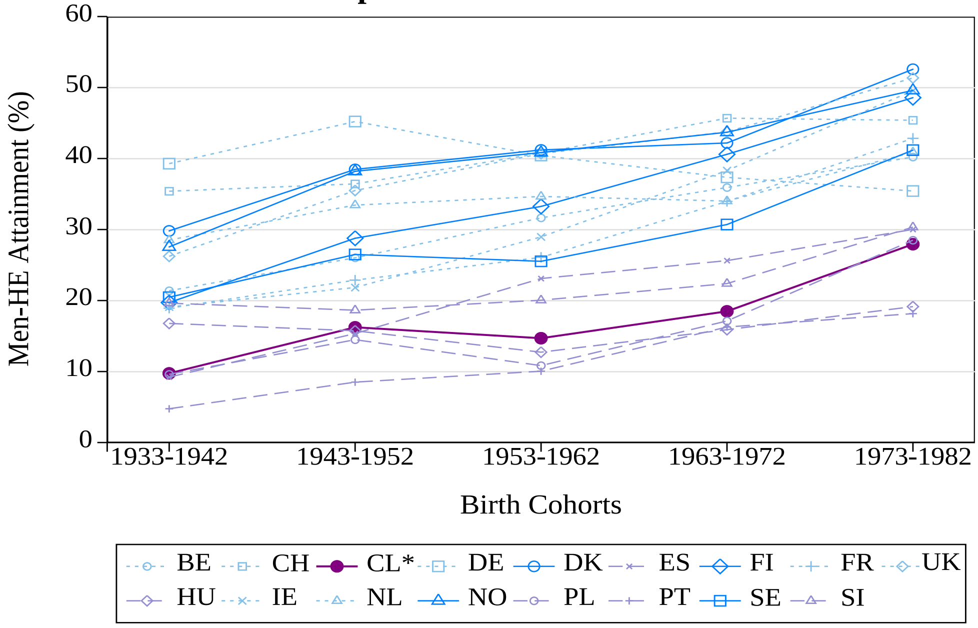
The majority of the analysed countries experienced the largest growth rate of graduates in the period relating to the 1933–1942 and 1943–1952 cohorts, which might be related to the expansion of the education system that European countries experienced after World War II. Between these cohorts (1933–1942 and 1943–1952), Chile seemed to have been following a similar trend equal to a 90% growth rate.

Chile’s second largest growth rate, 50%, was between the two youngest cohorts (1963–1972 and 1973–1982). Poland (55%) had the highest growth rate between the same cohorts. This may have been the result of the expansion of the education system due to the educational reforms in Chile, during the 1980s.

Unfortunately, the data in household surveys in Chile (CASEN) lacked parents’ social background information to develop a more in-depth analysis of the effect of parental education on the cohorts’ educational attainment. However, this will be address in the subsequent chapters of this thesis.

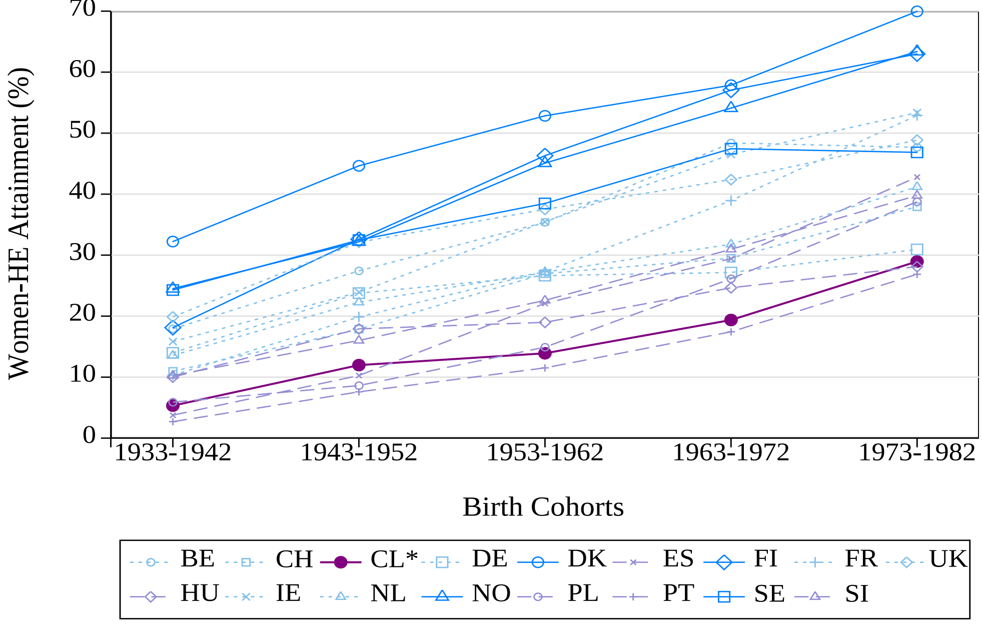
Figure 4.4 and Figure 4.5 display differences in higher educational attainment between men and women. In the majority of European countries, there is a larger proportion of women, who attained higher education. In Chile, this tendency is more noticeable in the two younger cohorts, but not as marked as in Scandinavian countries.

Figure 4.4: Higher Education Attainment in Men across European countries and Chile



Notes: *Chile's cohorts range from 1934 to 1983.
BE=Belgium, CH=Switzerland, CL=Chile, DE=Germany, DK=Denmark, ES=Spain, FI=Finland, FR=France, UK=United Kingdom, HU=Hungary, IE=Ireland, NL=the Netherlands, NO=Norway, PL=Poland, PT=Portugal, SE=Sweden, SI=Slovenia.
Data retrieved from ESS (2002; 2004; 2006; 2008; 2010; 2012; 2014) and CASEN (2003; 2013).

Figure 4.5: Higher Education Attainment in Women across European countries and Chile



Notes: *Chile's cohorts range from 1934 to 1983.
BE=Belgium, CH=Switzerland, CL=Chile, DE=Germany, DK=Denmark, ES=Spain, FI=Finland, FR=France, UK=United Kingdom, HU=Hungary, IE=Ireland, NL=the Netherlands, NO=Norway, PL=Poland, PT=Portugal, SE=Sweden, SI=Slovenia.
Data retrieved from ESS (2002; 2004; 2006; 2008; 2010; 2012; 2014) and CASEN (2003; 2013).

Chapter 5: Primary and secondary effects in the transition to higher education

5.1 Introduction

The aim of this chapter is to investigate how differences in parental social background influence students' chances to make the transition to higher education. The theoretical framework of the analysis rests on the Boudon (1974) model of inequality of educational opportunities (IEO). Boudon's approach conceives IEO as a process of individual decisions whereby students, acknowledging their academic performance, repeatedly face situations of choice in their progression through the school system. His proposal was a decomposition of IEO into *primary effects,* and *secondary effects*. Primary effects explain differences in academic performance between social groups, whereas secondary effects account for differences in individuals' decisions about whether to continue in education, conditional on their academic performance, between social groups.

The empirical decomposition of IEO into primary and secondary effects requires the careful selection of measures of social background and academic performance in relation to the measures used in the literature. It was stated in Chapter 2 that social class and parental education have proven to be compatible measures of social background, as no significant differences between them have been evidenced when both are available (Jackson, 2013a, p. 22). Therefore, the selection of parental education as the measure of social background for the student cohort is reasonable, as social class is not available in the data.

In terms of the academic performance measure, Boudon's rationale states that students make educational decisions after knowing their academic performance. Along the same lines, Jackson (2013a), in her cross-country study, preferred academic performance measures the results of which the students were informed, such as grade point averages (GPA), rather than standardised tests, the results of which the students were not informed. Jackson and Jonsson (2013) also offer caution with the use of standardised tests, as they may derive artificially larger secondary effects, giving an inflated account for secondary effects, in the overall IEO. (See also Jackson *et al.*, 2007; Erikson and Rudolphi, 2010). Furthermore, it could be the case that the academic performance close to the educational transition is contaminated by anticipatory decisions (Erikson *et al.*, 2005; Jackson *et al.*, 2007; Erikson and Rudolphi, 2010). The meaning of anticipatory decisions is that students make decisions about their educational aspirations well in advance of school-leaving age, thereafter, through the remainder of their compulsory education, they accordingly increase or decrease their efforts on academic performance to achieve these aspirations —this is the opposite of Boudon's rationale, whereby secondary effects are underestimated (Jackson *et al.*, 2007; Erikson and Rudolphi, 2010).

The data available for the student cohort includes both GPA levels and national standardised tests (SIMCE). Students know their GPA levels, but are unaware of their SIMCE scores. The difficulty with GPA levels is that they suffer from differences in their calculation across schools as assessment practices depend on schools' internal policies as well as differences in curriculum tracks (i.e., general or vocational). SIMCE tests, on the other hand, avoid all the aforementioned difficulties, as they are a standardised measure. Moreover, scholars and policy-makers in Chile rely on SIMCE tests to assess school differences in academic performance as studies in the outcomes of the 1980s reforms have focused (as seen in Chapter 4). It is also questionable that students are completely unaware of their individual SIMCE scores, as the Ministry of Education makes available to schools SIMCE practice tests[36] in which teachers inform students about their individual progress in the test.

In this chapter, I test whether SIMCE tests or GPA levels are more appropriate to make comparisons of academic performance across social groups. Furthermore, I derive primary and secondary effects in both measures, to test if the SIMCE test accounts for a larger share of the overall IEO in secondary effects than GPA levels would. I also test if anticipatory decisions operate in connection with GPA levels, which are closer to the transition to higher education than SIMCE tests.

Boudon (1974) argued that secondary effects are larger for latter transitions, e.g., the transition to higher education, because they are associated with higher economic costs than the costs at lower levels of education, and the benefits are unknown to students and families from less advantaged social backgrounds. Moreover, Breen and Goldthorpe (1997) argued that the overall pattern of social inequality remains relatively unchanged, despite increasing participation rates in education. As it was discussed in Chapter 4 (Figure 4.2), between 1995 and 2012 families financed over 60% of the higher education expenditure of Chile through tuition fees. Despite this share being offset by increments of public expenditure, in recent years, it has only been reduced to 50%. This ongoing heavy financial burden on families has been added to by rising tuition fees and by rising interest rates in higher education loans (triggering the students' protests in 2011), and is congruent with an enduring contextual setting of market-oriented reforms[37] reliant on private institutions' participation to expand the education system (Bellei, Cabalin and Orellana, 2014). Moreover, there is recent international evidence that secondary effects in the transition to higher education exhibited

[36] The simulation tests are available in the website of the Ministry of Education.
[37] The reform of 1981 was discussed in Chapter 4.

a larger share than primary effects, conditional on the degree of stratification and selectivity of the education system (Jackson and Jonsson, 2013).

Data from older birth cohorts[38] positioned Chile in the middle of a group of compared countries (mainly European) which ranked the effect of social background on educational attainment (Hout, 2006) and the level of educational mobility (Pfeffer, 2008). Jackson and Jonsson (2013) published a comparative classification of secondary effects in countries with data available for the transition to university, which ranked the Netherlands, Italy and Germany with the largest secondary effects, while at the opposite end of the ranking is the US, England and France –Chile was absent from their study. In this chapter, the resulting secondary effects in Chile will be compared with the aforementioned group of countries.

A brief description follows of the process of admission to higher education in 2014, the year the student cohort made the transition to higher education after completing secondary education. The next section proposes the research questions and hypothesis of the chapter. Then, follows the description of the data and variables. The methodological approach is discussed before presenting the descriptive and modelling results. The chapter ends with the discussion and conclusions.

[38] Birth cohorts from the 1980s and older.

Box 5.1 Transition to higher education in Chile – Admission 2014

Students, conditional on completing secondary education, from a general or vocational secondary programme (grades 9 to 12), can be enrolled in any higher education institution. Higher education institutions in Chile are both universities and vocational colleges. Universities offer degrees as well as technical and professional qualifications. Vocational colleges can only offer technical and professional qualifications.

There is a group of 33 universities that share a common policy of admissions. From this group, 25 are traditional universities that receive public funding. These 25 universities were public universities prior to the 1981 reform and did not then charge student fees (as seen in Chapter 4). The remaining eight are entirely private universities. This group of universities publishes the number of vacancies every year and the minimum academic performance requirements for admission to each offered programme.

There are three academic performance requirements to apply to these universities. Firstly, a minimum average score (i.e., 450 out of 850, in Spanish and mathematics) in entrance exams (*Prueba de Selección Universitaria*, PSU). Secondly, the average of grade point averages (GPA) over the 4 years of secondary education (grades 9 to 12). Thirdly and lastly, students' ranking relative to their peers in GPA, within schools. Each performance requirement receives a different weight according to the programme of application. PSU scores tend to weigh at least 50%, GPA between 10% to 30%, and the remaining weight for students' ranking relative to their school peers.

There is a higher demand in students' applications for programmes such as engineering and medicine rather than social sciences programmes, as they lead to highly paid jobs, which are respected in society (Stromquist and Sanyal, 2013). However, tuition fees in engineering and medicine is up to 8 times that of social sciences programmes, according to Zibechi (2012).

Private universities that do not share the common admission process also request the three aforementioned academic performance requirements. However, these private universities do not assign weights by programme for each performance requirement. Vocational colleges' admission policy only requires secondary education GPA. PSU exams are not necessary in applying to these institutions, although students' ranking might be required for scholarship applications.

5.2 Research Questions

The general research question of this chapter and Chapter 6 is:

> To what extent are social inequalities associated with students' chances of progression into higher education?

The specific research questions in this chapter are:

- Are SIMCE tests in grade 8 more appropriate than GPA levels to identify social background differences in academic performance? Are anticipatory decisions underestimating secondary effects in GPA levels?

  I hypothesise that SIMCE tests will capture a larger share of social background differences in academic performance than GPA levels, meaning larger primary effects. In relation to the adjunct question, I hypothesise secondary effects in GPA levels will have an influence similar to secondary effects in SIMCE tests in the overall IEO. Therefore, the anticipatory-decisions hypothesis will be unsupported.
- Are secondary effects more important (>50%) than primary effects in the transition to higher education?

  I hypothesise secondary effects will be larger (>50%) than primary effects in the transition to higher education.
- What is the relative position of Chilean students' secondary effects in comparison with the countries studied in Jackson and Jonsson (2013), in the transition to university?

  I hypothesise that Chile's ranking of secondary effects will lay in the middle of the distribution of the Jackson and Jonsson (2013) cross-country classification of secondary effects.

5.3 Data and Variables

The student cohort analysed in this thesis was selected from administrative records of the national student register linked to national standardised tests (SIMCE) and enrolment data in higher education, as explained in Chapter 3. The student cohort was born in 1995 to 1996; they were enrolled in compulsory education in 2002 at age 6 to 7. There were 210,168 children in first grade of primary education. However, only 134,595 (64%) graduated from secondary education in 2013, after completing secondary education.

Variables

The dependent variable is the transition to higher education. It takes the value of one for the students that transitioned to higher education, and zero for the students that did not enrolled in higher education the year after completing secondary education, 2014, at age 18–19.

Social background variables were linked to the students' register from parents' questionnaires in SIMCE tests, at age 9–10. These questionnaires did not include specific socioeconomic variables necessary to construct the Erikson–Goldthorpe–Portocarero class scheme. However, parental education was available for the student cohort. The highest parental education attainment was computed using the dominance approach between the father's and the mother's educational attainment to represent the educational level of the student's family. According to Erikson (1984) we assume that higher qualifications dominate the lower, meaning the labour market hierarchy influences also the family The resulting parental education variable has four-categories, of which the highest educational level is university degree, followed by advanced vocational qualifications, secondary education, and primary education or non-formal education, which is the lowest educational level.

The academic performance measures available for the student cohort, as previously stated, are grade point averages (GPA) in secondary education and national standardised tests (SIMCE) applied in grade 4 and grade 8 of primary education. GPA levels are the average of the four-years of secondary education and are known to the students, while SIMCE scores are unknown to students.

According to the admissions policy in 2014, GPA levels were weighted up to 30%[39] in the global score for application, lower than PSU exams for the requirements of traditional (prestigious) universities (see Box 5.1). PSU scores[40], in contrast, are weighted to at least 50%, and students' rankings within their schools are weighted between 10% to 40%. The difficulty in using PSU exams is that they were limited only to students applying to traditional (prestigious) universities. However, SIMCE tests in grade 8 were available for all students in the cohort, since those tests are mandatory. SIMCE scores in grade 8 are standardised tests that are highly correlated with PSU scores, the Pearson correlation was 0.78 ($p<.001$)[41].

The average of SIMCE test scores in Spanish and mathematics assessed in grade 8 is standardised through the sample that completed secondary education in 2013, with a mean of zero and a standard deviation of one. Similarly, GPA levels were standardised with a mean of zero and a standard deviation of one. The reason for averaging students' scores in Spanish and maths is because these subjects are consistently assessed in SIMCE tests in grade 4 and grade 8, other assessed subjects are not comparable between these evaluations. Moreover, university entrance requirements consider the average of PSU scores in Spanish

[39] The 30% weight was considered in two universities out of the 33 universities, which share a common admissions policy. The majority of this group of universities weigh GPA less than 20% of the overall score of application. See Larroucau (2014, p. 6) for more details of the weights assigned by university.
[40] PSU scores considered in the weighting are the average of Spanish and mathematics exams.
[41] Pearson correlation between std. GPA and std. PSU scores is 0.60 ($p<0.001$).
Pearson correlation between std. GPA and std. SIMCE grade 8 scores is 0.56 ($p<0.001$).
Pearson correlation between std. SIMCE grade 4 scores and std. PSU scores is 0.69 ($p<0.001$).
Pearson correlation between std. SIMCE grade 8 scores and std. SIMCE grade 4 scores is 0.76 ($p<0.001$).

and mathematics, which happens to be highly correlated to the average SIMCE sores in the same subjects for the cohort. Therefore, the average of Spanish and mathematics constitutes a good measure of the students' academic performance, as required to transition to higher education.

Other control variables, such as gender, were included in the descriptive analysis but left from the modelling as the methodological approach of Erikson *et al.* (2005) to derive primary and secondary effects does not allow for control variables. However, the exclusion of gender does not alter the size and significance of the primary and secondary effects presented in this chapter, which will be seen in Chapter 6.

5.4 Descriptive Analysis

The original student cohort of 210,168 students was reduced to 134,595 students (64%)[42], who completed secondary education in 2013, thereby, they were in the position to be enrolled in higher education from the following year (2014). Missingness in parental education reduced the student cohort to 121,088 students from the 134,595 who completed secondary education. Missingness[43] in SIMCE tests (grade 8) reduced the student cohort to 114,585 students.

Table 5.1 displays the distributions of students by parental education, gender, average by quartile of SIMCE scores (grade 8) in Spanish and mathematics, and the average of GPA in secondary education[44] by quartile as well. Quartiles were used in academic performance measures to show the distribution of students that accessed higher education, because this cannot be seen by taking the mean of a standardised measure with mean zero

[42] In this chapter, I simply ignore the attrition in the student cohort. However, in Chapter 6 I account for sample-selection.
[43] Missingness in both parental education and SIMCE scores in grade 8 is accounted for in Chapter 6 through multiple imputation techniques. In this chapter, it is simply ignored.
[44] In the rest of the chapter, it will be referred as GPA.

and standard deviation of one. Moreover, quartiles can be matched to the four levels of parental education for cross-tabulations used in the following subsection.

There is a larger proportion of students with parents from lower educational levels, these are the 70% of the student cohort. While the remaining 30% of the students come from families whose parents have at least advanced vocational qualifications. Differences between boys and girls are marked, being the girls the larger proportion, greater than 50%, than the boys. The difference by gender in the cohort might be related to girls to be more likely to complete compulsory education in the 12-years period, which was seen in Figure 3.2 of Chapter 3. The average of SIMCE scores and GPA by quartiles display a descending gradient in the cohort, from the highest to the lowest quartile. The average performance in first two highest quartiles surpasses the mean zero of the standardised measure of SIMCE and GPA, while the rest of the two quartiles have an average performance below the mean of the student cohort.

Table 5.1 Student cohort by parental education and SIMCE scores

	N	Percentage	
Highest parental education			
Degree	19,429	16.1	
Advanced vocational qualifications	16,664	13.8	
Secondary education	52,538	43.4	
Primary education or none	32,457	26.8	
Girls	65,186	53.8	
Boys	55,902	46.2	
Quartiles of SIMCE average scores in grade 8 [†]	**N**	**Mean**	**Standard Deviation**
Q(high)	28,646	1.27	0.41
Q(3)	28,646	0.36	0.20
Q(2)	28,646	-0.32	0.21
Q(low)	28,646	-1.31	0.46
Quartiles of Secondary ed. average GPA levels			
Q(high)	30,272	1.36	0.47
Q(3)	30,272	0.28	0.24
Q(2)	30,272	-0.44	0.19
Q(low)	30,272	-1.21	0.32
Total	**121,088**	**0.0**	**1.0**

Note: [†] There are 114,585 students with information in SIMCE scores in grade 8.

5.4.1 Academic performance interaction within parental education and educational transition association

Boudon (1974) argued that differences in academic performance by social background are only a first group of effects that determine childhood inequality of educational opportunity. He asserted that there is a second group of effects: the decisions that children and their families take, given their academic performance, to choose what educational pathway to follow. An empirical analysis, using Boudon’s model, begins by separating the aforementioned effects by tabulating the proportion of students following the next educational pathway, in terms of social background and net of academic performance.

Table 5.2 cross tabulates the proportion of students that make the transition to higher education given their parental education and academic performance. The first academic performance measure to be analysed, was the average of GPA in secondary education.

Table 5.2 Rates of transition to higher education by parental education and GPA (cell %)

	GPA (quartiles)				
Parental education	**Q(Low)**	**Q(2)**	**Q(3)**	**Q(High)**	**All**
Degree	59.8	68.5	74.2	86.9	78.0
Advanced Vocational Q.	52.5	61.9	69.2	82.2	67.6
Secondary Education	41.5	53.2	61.9	77.1	57.5
Primary or none	28.1	40.4	50.3	71.4	43.6
All educational levels	**39.0**	**52.3**	**62.1**	**79.9**	**58.5**

Note: Std. average GPA of the four years of secondary education. Cochran-Mantel-Haenszel test of linear association between parental education and transition, stratified by performance ($p < .001$).

Students' transition rates to higher education rose as levels of parental education and students' academic performance rose. A Cochran-Mantel-Haenszel test confirms that there was a linear association between parental education and the transition to higher education, which is stratified by GPA (Tang, He and Tu, 2012, pp. 110–111). These results support Boudon (1974, pp. 25–28) early findings in other countries, and different educational transitions.

Odds-ratios of successful student transition to higher education, as functions of GPA and of parental educational attainment are presented in Table 5.3. There is a descending gradient from the left to the right in each row. In the first row, the highest and the lowest parental educational levels are compared. Students, whose parents had a university education, were 3.8 times more likely to make the transition to higher education as those whose parents only had primary or no education, within the year after completing school. As GPA quartiles reaches the highest quartile (from left to right in Table 5.3), the odds-ratios steadily decrease, to 2.7.

The odds ratios of children whose parents have advanced vocational qualifications and secondary education are lower compared to the odds ratios for children with highly educated parents (second and third row in Table 5.3), although there is a similar gradient from the left to the right in the GPA distribution. The effect of parental education on transition to higher education is modified by children's GPA—the parental education gap in the odds of transitioning into higher education—is much smaller for children with high GPA compared to children with low GPA.

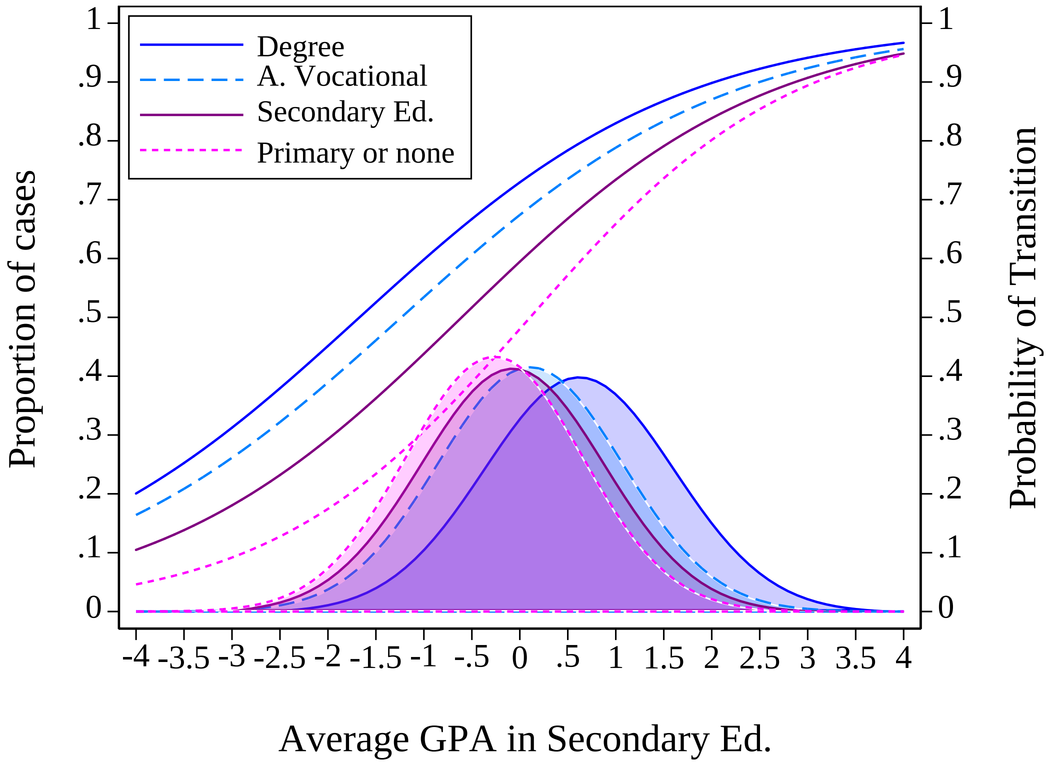
Table 5.3 Odds-ratios of the transition to higher education by parental education and quartiles of GPA

	Q(Low)	**Q(2)**	**Q(3)**	**Q(High)**
Degree/Low	3.8	3.2	2.8	2.7
	[3.4;4.2]	[2.9;3.5]	[2.6;3.1]	[2.5;2.9]
Vocational/Low	2.8	2.4	2.2	1.9
	[2.6;3.1]	[2.2;2.6]	[2.1;2.4]	[1.7;2.0]
Secondary/Low	1.8	1.7	1.6	1.4
	[1.7;1.9]	[1.6;1.8]	[1.5;1.7]	[1.3;1.5]

Note: Low refers to primary ed. or none formal education. CI at 95%, in square brackets.

Following the Erikson *et al.* (2005) method, the transition probabilities and average GPA in secondary education were drawn in Figure 5.1. Probabilities of transition were estimated using separated logistic regression models for each level of parental education and controlling for GPA. GPA assume a normal distribution.

Figure 5.1: Transition to higher education probability curves and GPA distributions by parental education



Note: The Data range in normalised GPA goes from -3.0 SD to +2.88 SD.

In Figure 5.1, the transition advantage that students with higher educated parents (university degree) enjoyed over students, whose parents had lower levels of education can clearly be seen. The darker, solid line represents this top group for parental education, and its curve reaches higher probabilities of transition as the average GPA rises. As the GPA rises, however, the distances from this top group's curve to all other curves get progressively narrower. The GPA distribution, for the same group, also shows an advantage over lower levels of parental education.

Students, whose parents had advanced vocational qualifications, had a very similar distribution of GPA to students whose parents had secondary education. In contrast, there was a clear advantage for the former, 'advanced vocational qualifications' group, over the latter, 'secondary education' group, in terms of transition probabilities.

Finally, those with low parental education lagged far behind in transition probabilities to higher education, compared with the rest. In terms of GPA, however, the distance between this group and the group of students who had parents with secondary education appears to be closer. Notwithstanding this, the largest proportion of students in the lower half of the GPA distribution (i.e. below the mean zero) belonged to students with low parental education.

Standardised test scores (SIMCE)

I turn now to analysing the second selected measure of academic performance, the standardised test scores. These SIMCE tests, as they are called in Chile, were collected as the average score of each student for Spanish and Mathematics in grade 8 (i.e., at age 13 to 14), for the purpose of standardisation.

The pattern in Table 5.4 closely resembles the pattern of Table 5.2. The transition rates increase from the left to the right, as the group of higher achievers in SIMCE test increases. Moreover, the transition rates increase from the bottom level of parental education to the top. The linear association test confirms that there was a significant association ($p<0.001$) among the three variables.

Table 5.4 Rates of transition to higher education by parental education and SIMCE grade 8 (cell %)

	SIMCE-Grade 8 (quartiles)				
Parental education	**Q(Low)**	**Q(2)**	**Q(3)**	**Q(High)**	**All**
Degree	58.8	66.2	74.5	85.2	78.2
Advanced Vocational Q.	51.4	61.0	69.0	79.9	67.8
Secondary Education	43.7	52.7	62.1	75.4	57.8
Primary or none	33.0	42.7	53.1	68.9	43.8
All educational levels	40.6	51.9	63.2	78.9	58.8[45]

Note: Std. Average (Spanish & Maths) scores in SIMCE G8. Cochran-Mantel-Haenszel test of linear association between parental education and transition, stratified by performance, $p < .001$.

The odds-ratios for students to transition to higher education, as a function of test score and parental education level, were found to have been noticeably lower for the SIMCE test (Table 5.5) than they were for the GPA (Table 5.3). The SIMCE test also gave a narrower range of odds-ratios than the GPA test. While it is important to note that the GPA do not reflect an average of the same number of subjects[46] as the SIMCE, this discrepancy may also have been caused by variability in school practices, in terms of standards of assessment and teaching ability. In contrast, the SIMCE test reflects a nationally standardised measure, which normalised schools' practices and differences.

Table 5.5 Odds-ratios of the transition to higher education by parental education and quartiles of SIMCE grade 8 scores

	Q(Low)	**Q(2)**	**Q(3)**	**Q(High)**
Degree/Low	2.91	2.63	2.58	2.61
	[2.6;3.3]	[2.4;2.9]	[2.4;2.8]	[2.4;2.9]
Vocational/Low	2.15	2.10	1.97	1.79
	[2.0;2.3]	[1.9;2.3]	[1.8;2.1]	[1.6;2.0]
Secondary/Low	1.58	1.49	1.45	1.38
	[1.5;1.7]	[1.4;1.6]	[1.4;1.5]	[1.3;1.5]

Note: Low refers to primary ed. or none formal education. CI at 95%, in square brackets.

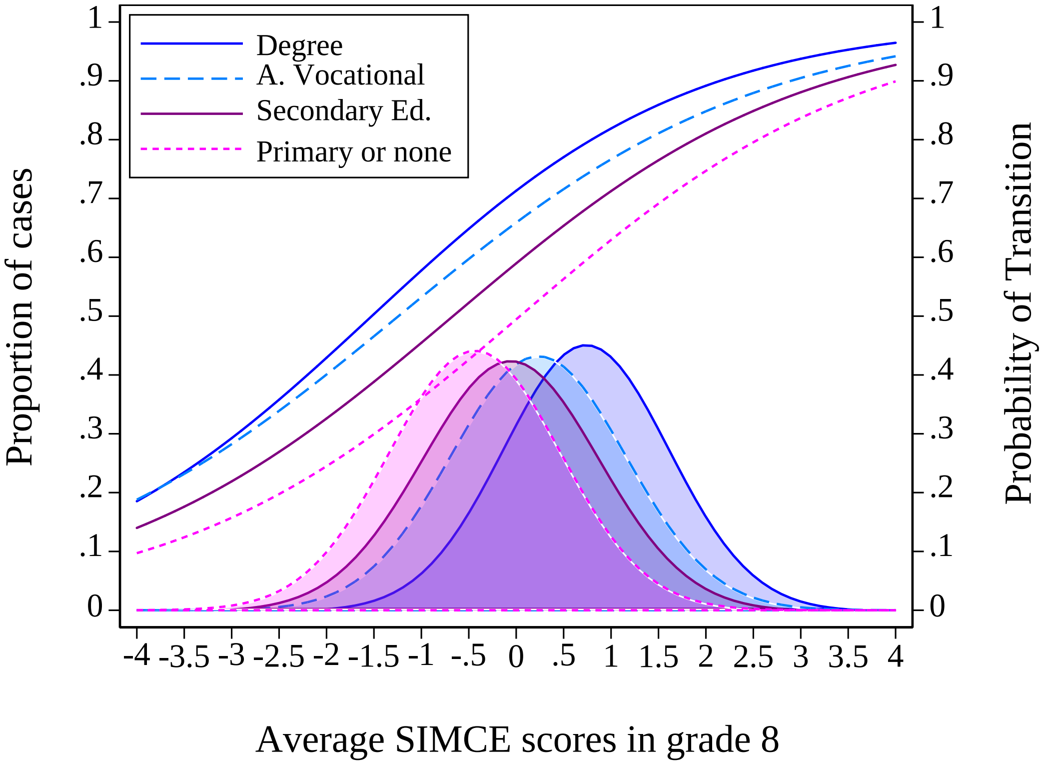
[45] Differences in the last column rates compared to Table 5.2, total rates of transition by parental education levels, are due to missingness in SIMCE grade 8 scores. GPA have a lower percentage of missingness in the data.
[46] This was highlighted on the documentation provided along with students' registry data due to differences in schools' assessment policies.

In Figure 5.2, transition probabilities and academic performance were plotted using average SIMCE scores. The curves and distributions for parental education levels closely follow the pattern seen in Figure 5.1 for the average of GPA. Students whose parents had attained higher education were clearly more likely to make the transition to higher education than their peers in the three groups whose parents had lower levels of education, and they also made up the right-hand side of the distribution of SIMCE scores.

The major differences between these two figures, however, are in the shape of the curves. The different transition curves by parental education levels in Figure 5.2 are less distant to each other, below the mean (zero) of SIMCE scores, than they are in Figure 5.1. Students with less educated parents, and those who were also in the lower half of the GPA distribution, paid a particularly high penalty in the probability of transition to higher education. For these students, GPA below the mean were associated with a clearly observable widening of the distance, in the probability of transition, from students with more educated parents. SIMCE scores, on the other hand, appear to have generated parallel distances, and more uniformity, in the trajectories of the transition probabilities, by parental education levels.

The shape of the distribution curves in academic performance, suggest that there was a greater variability in GPA (Figure 5.1) than in SIMCE scores (Figure 5.2). This result could evidence differences in school assessment practices that influence the variability seen in GPA.

Figure 5.2: Transition to higher education probability curves and SIMCE grade 8 distributions by parental education



Note: The Data range in normalised GPA goes from -3.5 SD to +2.90 SD.

In the following sections, the empirical strategy used to derive Boudon's primary and secondary effects is described.

5.5 Primary and secondary effects

In this final section of the analysis, primary and secondary effects are derived using the most frequently applied method in the literature.

The method was envisaged by Erikson and Jonsson (1996), and, in turn, this was further refined by Erikson *et al.* (2005)[47]. The first step of this method, was applied in the figures of transition probabilities and academic performance distributions in the previous section. The next step of the method is the numerical integration of the estimates from logit models with the normalised distribution of academic performance, which provides a

[47] There are, also, a few improvements in the methodology found in Jackson *et al.* (2007).

predicted probability of transition for each social background group. Then, counterfactual odds ratios are calculated using the latter predicted probability of transition, over the predicted probability of a counterfactual analysis. The Erikson *et al.* (2005) method then applies counterfactual analysis to the various combinations of integrating the distributions of academic performance of social group *j* with the probability of transition of social group *k*, or vice versa.

Finally, the total effect[48] is the product of the counterfactual odds ratios of social group *j*[49], with the academic performance distribution of group *k*; and the counterfactual odds ratios of social group k, with the transition probabilities of group *j*. The relative contribution to the first term to the total effect is the primary effect. Conversely, the relative contribution of the second term to the total effect is the secondary effect.

The equations of the method, and the subsequent derivation of primary and secondary effects, can be seen in Chapter's 5 methodological appendix in section 5A.2.

5.5.1 Primary and secondary effects by the Erikson *et al.* (2005) method

In this section I investigate which of the two academic performance measures, GPA or SIMCE test scores, is more appropriate to capture greater differences among parental education attainment groups, within Boudon's model framework.

Primary and secondary effects by GPA

The first step in the Erikson *et al.* (2005) approach is the estimation of separate logit models for each social background group. The social background measure in this study is

[48] The total effect can alternatively be derived as the sum of the logarithm of the odds ratios, i.e. the log-odds. See equations (7) and (8) in the chapter's methodological appendix.
[49] Jackson *et al.* (2007) call it the first method, because it is also possible to derive the total effect by taking first the group *k* and then the counterfactual of group *j*. See equations (5) and (6) in the chapter's methodological appendix.

highest parental education attainment, which was separated into four groups. Table 5.6 shows the results from the four logit models of the transition to higher education based upon the standardised average of GPA over the four years of secondary education, according to the Chilean educational system. The estimates of the coefficients are expressed in odds ratios rather than coefficients (log-odds), in order to simplify their interpretation. Specifically, the odds ratios expressed are the increase in the odds of transition to higher education estimated for one standard deviation increase in GPA.

In the first column of Table 5.6, the odds for students in the lowest parental education group (primary or none) are over 2 times for a corresponding increase of one standard deviation in standardised GPA. It seems that there was a greater reward of academic performance (over 20%) on entry to higher education for the lowest group of parental education, than there was for students in the other parental educational levels.

On the other hand, the average pattern of the odds ratios of making the transition in the intercepts of Table 5.6, when the standardised GPA were set to the mean of the sample (i.e., mean zero on the GPA slopes), showed an increasing gradient effect from the lowest to the highest group of parental education. For the two higher groups of parental education the odds were more pronounced than the odds of the lower groups. This result suggests that students of parents with higher education attainment still have greater odds to be enrolled in higher education programmes, regardless of their GPA. While the two groups at the bottom of parental education have smaller odds of attending higher education. What stands out is that the lowest group is the only group that is 8% less likely to make the transition to higher education, when their GPA are at the level of the sample mean.

Table 5.6 Logit models of the transition to higher education upon GPA, separated by parental education

Transition to Higher Education	Primary or none	Upper Secondary	Advanced Vocational Q.	Degree
	Coefficients expressed in odds ratios			
Std. GPA	2.09	1.88	1.80	1.81
	(.03)	(.02)	(.03)	(.03)
Intercept	0.92	1.47	2.07	2.70
	(.01)	(.01)	(.04)	(.05)
Pseudo R-square	0.074	0.058	0.050	0.054
N	32,457	52,538	16,664	19,429

Note: odds ratios are the estimated increase in the odds of transition to higher education estimated for one standard deviation increase in GPA. Standard errors in parentheses. All odds ratios are significant at $p<.001$ level.

Table 5.7 displays the distribution of standardised GPA by parental education attainment. The mean in GPA gradually increase from lower to higher levels of parental education.

Table 5.7 Standardised GPA by parental education

Std. GPA in Sec. Ed.	N	Mean	Std. Dev.	Min.	Max.
Primary or none	32,457	-0.26	0.92	-3.01	2.77
Secondary Education	52,538	-0.07	0.96	-2.95	2.82
Advanced Vocational Q.	16,664	0.11	0.96	-2.64	2.82
Degree	19,429	0.62	1.00	-2.44	2.82

The second step, following the Erikson *et al.* (2005) approach, was to calculate the counterfactual combinations of all parental education groups. Thus, coefficients from the logit models (Table 5.6) and the GPA means and standard deviations (Table 5.7) were combined in a numerical integration, under an assumption of a normal distribution for GPA.

The diagonal of transition probabilities in Table 5.8 show the predicted transition probabilities associated with each of the four parental education groups. Off-diagonal transition probabilities show how much the transition probability changed, for the reference group, by combining the average performance in GPA (Table 5.7) and/or log-odds of transition probabilities (from the models in Table 5.6) from other parental education groups. These probabilities are the counterfactual transition probabilities.

Table 5.8 Predicted and counterfactual transition probabilities between parental education groups (GPA)

If they had the log-odds of… / If they had the avg. GPA of…	Primary ed. or none	Secondary education	Advanced Vocational Q.	Degree
Primary ed. or none	0.44	0.55	0.63	0.68
Secondary education	0.47	0.58	0.65	0.70
Advanced Vocational Q.	0.49	0.60	0.68	0.72
Degree	0.57	0.67	0.74	0.78

The datum, 0.44, in the first row and column of Table 5.8, represents the transition probability found for the low group of parental education. For this low group, the counterfactual transition probability steadily increases when average GPA from the other groups of parental education are combined with the estimated log-odds of this group (that is to say, the counterfactual transition probabilities from rows 2 to 4, within column 1). The change in the transition probability for the same low group, also steadily increases after taking into account the estimated log-odds from higher parental education groups (that is to say, the counterfactual transition probabilities from columns 2 to 4, within row 1).

The changes, for this low group, in the counterfactual transition probabilities were larger when considering the estimated log-odds from the other parental education groups (up to 0.68/0.44 = 1.5 times) in comparison to the changes amongst the GPA (up to 0.57/0.44 = 1.3 times) from the other parental education groups. This may signal that the effects of whether to continue or not in the educational attainment pathway, which are associated with social background, had a greater influence on the transition to higher education than academic performance.

The overall pattern in the changes, by average GPA (rows) and log-odds of transition probabilities (columns), shows that a consistently increasing gradient of transition probability favoured the higher social background groups.

In Table 5.9, predicted probabilities from Table 5.8 were transformed into odds ratios. For instance, the predicted probability for the high parental education group (0.78) was divided by its complement in probability. Therefore, the predicted probability expressed in odds ratios was 0.78/(1-0.78) = 3.50.

Table 5.9 Odds ratios of predicted and counterfactual transition probabilities between parental education groups

If they had the log-odds of… / **If they had the avg. GPA of…**	Primary ed. or none	Secondary education	Advanced Vocational Q.	Degree
Primary ed. or none	0.78	1.22	1.68	2.10
Secondary education	0.87	1.36	1.87	2.35
Advanced Vocational Q.	0.97	1.52	2.09	2.62
Degree	1.30	2.03	2.79	3.50

The effect of changes in academic performance, the effect of changes in transition probabilities, and the total effect were then, in turn, calculated as odds ratios comparing the reference group of low parental education with each of the other three parental education groups (see Table 5.10).

The first column of Table 5.10 shows that the overall odds ratio of attending higher education for students with parents holding a degree is 4.49 times larger than for students in the low group (the total effect). The total effect may be expressed as the ratio of the odds ratios reported in the extremes of the diagonal of Table 5.9 for both groups, i.e. 3.50/0.78=4.49.

Table 5.10 Odds ratios between parental education groups (GPA)[50]

	Total Effect	Different performance, similar transition probability	Similar performance, different transition probability
Degree/Low	4.49	1.67	2.69
Vocational/Low	2.69	1.25	2.15
Secondary/Low	1.75	1.12	1.56

Note: Relative to low parental education (primary ed. or none).

The odds ratio between the high and low groups in parental education is reduced to 1.67 (top of second column of Table 5.10), when dividing the counterfactual odds ratio of the low group log-odds at the GPA of the high group (first column and fourth row of Table 5.9), over the predicted odds ratio of the low group (first column and row of Table 5.9). This ratio of odds ratios is 1.30/0.78=1.67.

Similarly, the odds ratio between the high and low groups in parental education is reduced to 2.69 (top of third column of Table 5.10), when dividing the predicted odds ratio of the high group (fourth column and row of Table 5.9) over the counterfactual odds ratio of the high group GPA low group at the log-odds of the low group (first column and fourth row of Table 5.9). That is, the ratio of odds ratios 3.50/1.30=2.69.

The procedure continued by following analogous calculations for the groups of parents with vocational qualifications and secondary education.

The last step of the Erikson *et al.* (2005) method was to derive the relative contributions of the primary effects and the secondary effects to the total effect, for each parental educational group comparison. Accordingly, the natural logarithms of the odds

[50] Bootstrapped standard errors for the odds ratios presented in this table were included in Table 5A.1 in the chapter's appendix, by using Buis (2010) generalisation method.

ratios from Table 5.10 were used to calculate the proportions of these log-odds ratios, from the second and third columns, over the log total effects from the first column. The primary effects were derived from the resulting second column's relative contributions to the total effects. The third column's relative contributions were the secondary effects, according to Boudon (1974) model. The resulting proportions are expressed as percentages in Table 5.11 and add up to 100%.

Table 5.11 Relative primary and secondary effects between parental education groups (GPA)

	Primary Effects (%)	Secondary Effects (%)
Degree/Low	34.1	65.9
Vocational/Low	22.4	77.6
Secondary/Low	20.2	79.8

Note: Relative to low parental education (primary ed. or none).

Table 5.11 shows that secondary effects are greater (>50%) than primary effects for the transition to higher education, when compared to the low group of parental education.

Between the high and low group, the secondary effects were 65.9% of the total effect, which means that inequality of educational opportunities is mainly driven by the choices that students and their families in the high group can make, in contrast to the choices of the low group. This result is compelling. Albeit GPA were equalised through the counterfactual odds ratios for these two groups of parental education (third column of Table 5.10), the effect of different GPA only accounts for 34.1% of the total effect in both groups.

Secondary effects in the other groups of parental education (secondary education and vocational qualifications) are larger in contrast with the low group, 77.6% and 79.8% respectively.

Primary and Secondary Effects by SIMCE Grade 8 scores

In the following, I emphasise the differences between GPA and SIMCE results on the decomposition of primary and secondary effects using the Erikson *et al.* (2005) approach.

The major difference[51], in the results of the logit models, between the GPA and the SIMCE test scores is that the odds ratios associated with GPA tended to decrease, from the low to the high group of parental education. As mentioned above, this may have been the result of the greater relative importance of the GPA, for groups of lower parental education, in increasing the odds of making the transition to higher education.

In contrast to the pattern found for GPA, an increase in SIMCE score preferentially enhanced the odds ratio, for enrolling in higher education, for the student group whose parents had attained higher education compared to all other groups (Table 5.12). While there was a modest increasing gradient from vocational to low parental education groups, the odds ratios of all three of these groups clearly fell below that of the high group.

[51] Another important difference between SIMCE scores and GPA is the sample size. The GPA sample size is larger, 121,088, whereas the SIMCE simple size is 114,585. However, re-estimating the analysis of GPA by restricting the sample size to be equal to the sample size used in SIMCE scores, confirmed that sample size does not explain the difference between SIMCE scores and GPA results

Table 5.12 Logit models of the transition to higher education upon SIMCE scores, separated by parental education

Transition to Higher Education	Primary or none	Upper Secondary	Advanced Vocational Q.	Degree
	Coefficients expressed in odds ratios			
SIMCE in grade 8	1.74	1.72	1.70	1.82
	(.02)	(.02)	(.03)	(.04)
Intercept	0.98	1.44	1.93	2.48
	(.01)	(.01)	(.03)	(.05)
Pseudo R-square	0.042	0.043	0.040	0.046
N	30,218	49,693	15,967	18,707

Note: odds ratios are the estimated increase in the odds of transition to higher education estimated for one standard deviation increase in SIMCE level. Standard errors in parentheses. All odds ratios are significant at p<.001 level.

The predicted and counterfactual transition probabilities, as well as transition odds ratios for SIMCE scores, are not presented here, although they can all be seen in the chapter's appendix. This is because the patterns by parental education groups are similar to the ones presented for GPA (Table 5.8 and Table 5.9).

Table 5.13 illustrates the total effects by comparing the three higher parental education groups against the low parental group and the results of counterfactual odds ratios by combinations of performance and transition probabilities.

The results were strikingly similar to those for GPA (Table 5.10). The odds ratio for attending higher education, in the high group, was over 4 times than that for the low group of parental education. Whilst the odds ratio with SIMCE scores was only slightly higher than for GPA (4.57 versus 4.49) in the high/low comparison, what is particularly noteworthy was that the total effects for the vocational/low comparison, and the total effects for the secondary/low comparison were each identical in the SIMCE and GPA analyses. The likenesses in the total effects of the middle groups of parental education follow from the similarity in predicted and counterfactual transition probabilities (see Tables 5A.3 and 5A.4 in the appendix).

The differences between the GPA and the SIMCE scores analyses lie in the primary and secondary effects (second and third columns of Table 5.10 and Table 5.13). In terms of different performance but similar transition probability (second columns), the odds ratios of the transitions were larger for SIMCE scores compared to GPA, each were approximately 110% of the values found in the GPA analysis. In terms of similar performance but different transition probability (third columns), the SIMCE odds ratios of the transition were each approximately 90% the corresponding value of that for the GPA.

Table 5.13 Odds ratios between parental education groups (SIMCE grade 8)

	Total Effect	Different performance, similar transition probability	Similar performance, different transition probability
Degree/Low	4.57	1.84	2.48
Vocational/Low	2.69	1.41	1.90
Secondary/Low	1.75	1.22	1.44

Note: Relative to low parental education (primary ed. or none)[52].

Comparing Table 5.11 (GPA) with Table5.14 (SIMCE) shows there was a systematic inflation of primary effects, across the three group comparisons, in the SIMCE analysis compared to the GPA analysis. In keeping with this, there was a systematic inflation in secondary effects for the GPA analysis, compared to the SIMCE analysis. Between the high and the low groups of parental education, SIMCE primary effects were 6.2% higher compared to the primary effects derived from GPA. SIMCE secondary effects were 6.2%

[52] Bootstrapped standard errors for the odds ratios presented in this table were included in Table 5A.1 in the chapter's appendix, by using Buis (2010) generalisation method.

lower in parallel. Parallel differences, like this, of very similar magnitude, were evident in all comparisons of effects (primary and secondary) generated by using GPA with those generated by using the standardised SIMCE tests. For the second group comparison (vocational/low) the difference in primary effect was +12.7% using SIMCE and for secondary effect it was -12.7% (compared to the GPA analysis). For the third (secondary/low) comparison, the GPA analysis had a primary effect that was 14.8% lower than that for the SIMCE analysis whilst it was 14.8% higher for secondary effects.

Table 5.14 Relative primary and secondary effects between parental education groups (SIMCE in grade 8)

	Primary Effects	Secondary Effects
Degree/Low	40.3	59.7
Vocational/Low	35.1	64.9
Secondary/Low	35.0	65.0

Note: Relative to low parental education (primary or none).

5.5.2 Secondary effects from Erikson *et al.* (2005), Karlson, Holm and Breen (2010), and Buis (2010) methods

In this subsection, I compare the relative contribution of secondary effects to the overall inequality of educational opportunities (IEO) from alternative methods to Erikson *et al.* (2005).

The Karlson, Holm and Breen (2010) method decomposes the indirect effect of academic performance over the total effect of social background on the transition higher education, which is the primary effect of Boudon's model. The direct effect accounts for the effect of social background on the transition, controlling for academic performance, which is the secondary effect of Boudon's model. Despite the KHB method is similar to a mediation model, the models involved in the estimation of primary and secondary effects are different to the models applied in a mediation model under structural equation models (SEM) (MacKinnon, 2008)[53]. The mediation models applied in Chapters 6 and 7 will be estimated under the SEM framework. The models derived from the KHB method can be seen in Tables 5A.7 to 5A.10 of Chapter's 5 appendix, and the explanation of the method is in Chapter's 5 methodological appendix.

The Buis (2010) method relaxes the assumption of normal distribution of the academic performance to find the predicted and counterfactual transition probabilities, by numerically integrating over the empirical distribution of academic performance instead of the normal distribution as in Erikson *et al.* (2005) method. Moreover, bootstrapped standard

[53] The main differences are that the KHB method estimates a 'reduced' model, which is a non-linear model for the transition accounting for the residuals of the linear model of the academic performance (see equation 2 in section 5A.2.2 of the methodological appendix). Then, the indirect effect of the academic performance on the transition is calculated as a difference between the coefficients of the 'reduced' model and the 'full' model (see equations 3 and 8 of the methodological appendix). The mediation model under SEM does not estimate a 'reduced' model, but a 'full' model, using KHB terminology, which is the non-linear model of the transition controlling for academic performance and parental education. The indirect effect of academic performance is calculated as the product of the coefficients of the linear model of academic performance and the coefficients of the non-linear model of the transition (see section 6A.2.1 in the methodological appendix of Chapter 6).

errors can be computed for all estimates. The results derived from Buis (2010) method can be seen in Tables 5A.1, 5A.2, 5A.5 and 5A.6 in the chapter's appendix, and the explanation of the method is in the methodological appendix.

What stands out in Table 5.15 is that, ignoring trivial differences in decimal places, the three methods found the same relative sizes of secondary effects, as both primary and secondary effects add to 100%, primary effects also have similar relative sizes. It can be inferred from these results that is safe to assume a normal distribution in GPA levels and SIMCE scores, as the Erikson *et al.* (2005) method does, since the Buis (2010) method yields the same results by not assuming a normal distribution of both measures of academic performance. Similarly, Karlson and Holm (2011) in a Monte Carlo study found comparable results between the three methods, assuming a normal distribution in academic performance and even in the presence of a strong confounding. Karlson and Holm (2011) demonstrated that the strengths of the KHM method over the Erikson *et al.* (2005) and Buis (2010) methods can be seen in the presence of a continuous social background variable and confounding variables that might affect the decomposition of primary and secondary effects. However, a categorical social background variable and no-confounding variables were considered in the present comparison, that is the reason of finding similar results between the three methods.

Table 5.15 Relative secondary effects for the transition to higher education between parental education groups

	GPA			**SIMCE G8**		
	Erikson et al.	**KHB**	**Buis**	**Erikson et al.**	**KHB**	**Buis**
Degree/Low	65.9	65.2	65.2	59.7	59.7	59.4
Vocational/Low	77.6	77.4	77.4	64.9	64.8	64.7
Secondary/Low	79.8	79.7	79.7	65.0	64.9	64.8

Note: Relative to low parental education (primary ed. or none).

Secondary effects seem to be the dominant effect (>50%) for the transition to higher education, when comparing students with parents holding low formal education against students with parents holding higher educational credentials. However, secondary effects between the low and high parental education groups are smaller to the secondary effects found in the comparison of secondary education and vocational qualifications, regardless of the measure of academic performance (GPA or SIMCE).

The dissimilarity of impact between GPA and SIMCE tests on secondary effects suggests that SIMCE tests are more capable of capturing an important share of educational inequality than are GPA; thereby, yielding larger primary effects which decreased the impact of secondary effects. A possible explanation could be the standardised nature of SIMCE tests, which make them able to capture differences in academic performance than GPA are capable.

5.6 Secondary effects cross-country comparison

Jackson and Jonsson (2013) study compares secondary effects on the transition to university degree found in European countries and the USA. The majority of the countries used birth cohorts from the early 1990s, whereas the Chilean cohort was from the mid-1990s, and the Italian and Dutch cohorts were from the 1980s. Secondary effects compare the transition chances of the most advantaged and most less advantaged groups, in social class or parental education accordingly to data availability[54].

There are important distinctions in the way that measures of academic performance were used in each country. These distinctions need to be appreciated in order to effectively

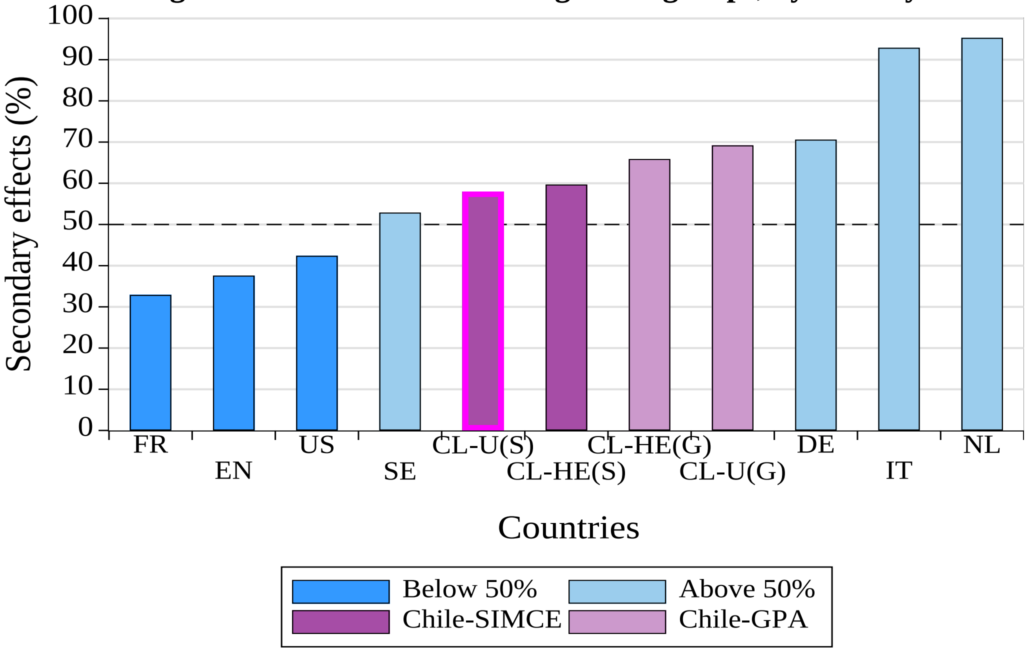
[54] The average of social class and parental education was considered when both were available.

compare differences in secondary effects, when using GPA and standardised tests (SIMCE), in the manner that this difference was investigated for the Chilean student cohort.

Teachers in Sweden and Germany assigned grade point averages, while in France, England and Italy came from final examinations regulated by a national authority. Thus, GPA for the latter group of countries tended to be more homogenous than in the former group. The Netherlands used the average of language and arithmetic in standardised tests taken at age 12. The United States used a standard combined score of mathematics scores at age 17 or 18, and reading scores at age 15 or 16, as reading tests at age 17 or 18 were not available. It is important to emphasise that the US and the Netherlands were the only countries that used standardised tests. Jackson and Jonsson (2013) argued that secondary effects will account for a larger proportion of IEO when standardised tests are used instead of GPA, which could explain the large secondary effects observed in the Netherlands.

Jackson and Jonsson (2013) focused on the transition to university, whereas the Chilean cohort analysed in the transition to higher education, including universities and vocational colleges. For comparability purposes, primary and secondary effects were recalculated for the Chilean cohort, accounting only for the transition to university. In Figure 5.3, the differences in secondary effects, between the transition to higher education (CL-HE(S)) and the transition to university (CL-U(S)), is only 1.8% when SIMCE scores are the measure of academic performance. Conversely, when GPA levels are the measure of academic performance, the difference in secondary effects is 3.3% larger for the transition to university (CL-U(G)) than the transition to higher education (CL-HE(G)).

Figure 5.3: Secondary effects on the transition to university between the highest and lowest social-background groups, by country



Notes: CL-U(S)=Chile transition to university using SIMCE (preferred),
CL-HE(S)=Chile transition to higher education using SIMCE,
CL-HE(G)=Chile transition to higher education using GPA,
CL-HE(G)=Chile transition to university using GPA,
DE=Germany, EN=England, FR=France, IT=Italy,
NL=the Netherlands, SE=Sweden, US=the United States.
The secondary effects in the figure were retrieved from 'Why does inequality of educational opportunity vary across countries?' Jackson and Jonsson (2013, p. 321).

Figure 5.3 shows that France, England and the US are the group of countries where secondary effects explain a smaller proportion (<50%) of the overall IEO in the transition to university. This result implies that social background differences in academic performance (primary effects) are the dominant effect for these countries. On the other hand, there is a group of countries where secondary effects are the dominant effects (>50%) which are Sweden, Germany, Italy, the Netherlands and Chile. In this group of countries, the overall inequality in the transition to university was mainly influenced by differences in students' choices, conditional on their academic performance.

Jackson and Jonsson (2013, p. 310) classified countries' educational systems in terms of stratification and selectivity, at the secondary or upper secondary level. Noteworthy, England and the United States, which have lower secondary effects (<50%), were classified in the group of low stratification and low selectivity. In contrast, the Netherlands and Germany, which have higher secondary effects (>50%), were classified in the group of high stratification and high selectivity because children are selected at age 10–11. Sweden's classification seemed to match the magnitude of its secondary effects as well, classified in the intermediate group for both stratification and selectivity.

France was classified in the intermediate group of stratification, but with high selectivity, although it exhibited the lowest secondary effects. Italy, was in the intermediate group for both stratification and selectivity, although its secondary effects are among the highest (>50%). Jackson and Jonsson (2013, p. 315) made a note of awareness regarding the low quality of the measure of academic performance in Italy, which may, indeed, explain the discrepancy with the classification of the Italian system.

Chile's educational system could fit in Jackson and Jonsson's classification in the group of countries with high stratification, as there are differences in curriculum tracks (i.e. general and vocational) and the quality or status of schools. Pfeffer (2008) also classified Chile's school system as highly stratified, owing to differences in curriculum tracks in the last two years of secondary education. Schools differences in Chile are related to the type of school funding (i.e. municipal, voucher and private), which reflects families' ability to pay voucher or private schools as municipal schools are free of charge. Furthermore, the three types of school has been evidenced to be associated with student's performance in standardised tests, which subsequently influences educational attainment (Rosas and Santa Cruz, 2013). In terms of selectivity, schools can overtly select[55] children according to previous academic performance, student's behaviour, family income, or other selection

[55] Since 2016, the ministry of education has issued a law against selectivity practices on schools. However, this law had no effect on the student cohort considered in this thesis.

mechanisms (Bellei and Cabalin, 2013). However, selectivity is not determined for all students from an early age, as seen in Germany and the Netherlands. Therefore, it seems that Chile can fit in the intermediate group of selectivity along with Italy and Sweden, following Jackson and Jonsson's classification.

The relative contribution of secondary effects for the Chilean cohort in the transition to university using SIMCE scores—CL-U(S) in Figure 5.3—is between Sweden and Germany, although closer to Sweden. It seems that the rationale of the classification of the Chilean school as highly stratified and with intermediate selective, in the same way conforms in-between the classification of the highly stratified and selective German system and the intermediate level of stratification and selectivity of the Swedish system.

5.7 Discussion

The specific research questions of the chapter are addressed in this section, while the general research question is addressed in the conclusion.

Are SIMCE tests in grade 8 more appropriate than GPA levels to identify social background differences in academic performance? Are anticipatory decisions underestimating secondary effects in GPA levels?

Differences between the two measures of academic performance became more apparent, after applying the Erikson *et al.* (2005) method. The resulting primary effects accounted for a larger proportion in SIMCE scores than they did in GPA levels over levels of parental education. This result was consistent with alternative methods developed in the literature (Karlson, Holm and Breen, 2010; Buis, 2010). As both primary and secondary effects add to 100% of the total inequality of educational opportunities (IEO), thereby secondary effects using SIMCE scores accounted for a smaller proportion than GPA levels.

Jackson and Jonsson (2013) strongly recommended the use of grades, whenever they are available, over standardised tests, to appropriately apply Boudon's IEO model. Boudon (1974) assumes that students and their families choose whether to continue in the educational system conditional on their knowledge of the student academic performance. However, in the introduction it was stated that individual scores in SIMCE tests are not entirely unknown, since students practice with simulated tests, and they learn of their scores from the practice tests. Therefore, use of the SIMCE test does not contradict Boudon's rationale.

Jackson *et al.* (2007) and Erikson and Rudolphi (2010) show that secondary effects tend to be larger when using standardised tests than when using grades. In the student cohort the contribution from secondary effects was, to the contrary, smaller in SIMCE tests when compared to GPA levels. However, neither study (Jackson *et al.*, 2007; Erikson and Rudolphi, 2010) analysed the transition to higher education. The former study, investigated the transition to A-levels in England and Wales, and in the latter, the transition was to upper secondary education in Sweden.

The Jackson *et al.* (2007) and Erikson and Rudolphi (2010) studies, demonstrated that standardised tests lack the association with the educational transition, compared to grades. However, I argue in favour of SIMCE scores over GPA levels, as I consider SIMCE scores the more relevant measure in capturing inequality of academic performance, from the overall inequality of opportunities in the transition to higher education, through primary effects. Moreover, in the introduction I argued that GPA levels suffer from not being comparable across schools, as differing assessment policies exist amongst schools owing to differences in school funding or curriculum tracks.

If there was evidence of anticipatory decisions, secondary effects in GPA levels, given they are closer to the transition than SIMCE tests, would have been smaller than primary effects. Therefore, secondary effects would have been underestimated (Jackson *et al.*, 2007; Erikson and Rudolphi, 2010). However, the opposite was evidenced, as secondary effects in GPA levels are the dominant effect (>50%), which is akin to the dominance of secondary effects in SIMCE tests. It seems that there is no support for the anticipatory-decisions hypothesis, in this student cohort, in regard to GPA levels.

Are secondary effects more important (>50%) than primary effects in the transition to higher education?

Secondary effects contributed to the overall inequality in the transition to higher education to a larger extent than did primary effects (>50%). As was stated in section 4.5, primary and secondary effects were derived by comparing the higher groups of parental education (i.e., from degree to secondary education) to the low group (i.e., primary ed. or no formal education). On the analysis, inequality of opportunities in the transition to higher education was found to be at least 60% influenced by differences in parental education and up to 40% influenced by differences in standardised (SIMCE) tests.

Chile's access to higher education has, then, remained closely linked to family background despite frequent efforts by policy-makers to improve less advantaged students' chances of entering higher education. In effect, various initiatives of financial aid[56] were implemented since the student protests in 2011. Prior to 2011, higher education funding was strongly linked to students' performance on university entrance exams[57] (PSU) and overall academic performance on higher education programmes. It is not surprising that students with higher academic performance tend to come from private schools and from the very selective municipal schools[58].

Parents with higher education tend to enrol their children in higher performance schools according to average tests scores (SIMCE and PSU) for the school, although there is evidence that parents also look for schools that reflect their social status, and those close to their residential area (Mizala and Torche, 2012). Parents' behaviour, thus, increases socioeconomic segregation among schools (Elacqua, 2012). This reflects how the educational system bridges schools and higher education institutions following the neoliberal reforms of the 1980s. Students completing secondary education in higher performance schools have a greater chance of entering higher education than students from

[56] The main government initiatives after the student movement in 2011, were to decrease the interest rate in university loans subsidised by the State from 6% to 2%. Moreover, new programmes of scholarships were introduced by high academic performance in less advantaged students (Stromquist and Sanyal, 2013). Moreover, from 2016, a new funding programme '*Gratuidad*' was implemented by allowing the 50% most less advantaged students in socio-economic background, to be exempt from any fees for the total duration of the university degree programme in which they were enrolled. By 2017, the programme has been extended to vocational colleges certified by the ministry of education (www.gratuidad.cl).

[57] Albeit secondarily, academic ranking relative to their peers, within the school, has also been considered in the last few years, as part of the admission requirements requested for higher education institutions. To compensate for the academic disadvantage, students coming from non-deprived schools may have to reach higher scores in PSU exams.

[58] This group of municipal schools select students based on their past academic performance and in some cases admission exams. These schools are no more than 30 and are located only in the capital city, Santiago (Allende González and Valenzuela, 2016).

average schools or those at the bottom of the scores distribution in standardised tests. The consequence is an ever-widening gap in educational opportunity amongst students planning to attend higher education programmes.

In Chapter 6, the effects of schools were introduced to Boudon's model to investigate their influence on primary and secondary effects.

What is the relative position of Chilean students' secondary effects in comparison with the countries studied in Jackson and Jonsson (2013), in the transition to university?

Jackson and Jonsson (2013) cross-country classification of the relative contribution of secondary effects to the overall IEO, concentrated on the transition to university due to comparability purposes, as themselves asserted. For the Chilean student cohort secondary effects between the highest and lowest parental education groups were not as different, between the transition to higher education[59] and the transition to university. The difference was 1.8% larger for the transition to higher education compared to the transition to university, when SIMCE scores were the measure of academic performance.

The secondary effects of the Chilean student cohort was over the 50% benchmark, meaning that secondary effects drove the overall IEO as in Sweden, Germany, Italy and the Netherlands. The preferred measure of secondary effects using SIMCE tests and restricting the transition to university to be able to compare with the rest of the countries, was between Germany and Sweden. In terms of the stratification of the school system due to differences in type of school funding and curriculum track, following Jackson and Jonsson (2013) and Pfeffer (2008) I argue that Chile has as a highly stratified system. On the other hand, I classified the selectivity of the system as intermediate, because not all students are selected

[59] As a reminder, the transition to higher education includes vocational colleges besides universities.

from an early age as in Germany and the Netherlands. Both the classification of the Chilean system and the relative contribution of secondary effects seem to agree.

5.8 Conclusion

The inequality of educational opportunities, for the transition to higher education in Chile, was found to be driven by social inequalities acting on student choices concerning whether or not to continue to this educational level, conditional on their academic performance. These are the secondary effects in Boudon's IEO model. Differences in performance can be better accounted for using standardised tests (SIMCE), rather than the average of GPA levels, in the four-years of secondary education. The caveat with the SIMCE test is that individual scores were unknown to students, although, students typically had knowledge of their individual scores in practice SIMCE tests.

In the Chilean educational system, SIMCE test results are substantially important to monitor schools because of reward programmes[60] that incentivise the improvement of school performance or give extra benefits to schools in the top group of the average SIMCE scores. Moreover, schools with higher SIMCE scores attract greater student enrolment. In this regard, the Ministry of Education allows schools to train their students, through practice tests to improve their average SIMCE scores.

In Chapter 6, I address sample-selection in the student cohort as only 64% completed secondary education at the end of the compulsory education period. Moreover, I examine social inequalities in secondary education completion to assess *Maximally Maintained Inequality* (MMI) across transitions.

[60] For a detail description of the reward programmes see Mizala and Torche (2012) and Mizala and Urquiola (2013).

In Chapter 7, the influence of Chile’s education system is examined at the school level and at higher education level, using Boudon’s IEO model and assessing *Effectively Maintained Inequality* (EMI) by type institution and programme of enrolment.

Chapter 6: Inequality of educational opportunities across transitions

6.1 Introduction

In the previous chapter, I argued that SIMCE tests were more appropriate than GPA levels to measure academic performance in the student cohort, as SIMCE tests avoid comparability issues across schools. The larger relative contribution of primary effects of SIMCE compared to GPA levels evidenced larger social background differences in academic performance. This result does not contradict Boudon's rationale of students making informed educational decisions subject to their academic performance, as the Ministry of Education makes available to schools SIMCE practice tests in which teachers inform students about their individual scores. Moreover, I found no support for the anticipatory decisions hypothesis in the student cohort; since secondary effects in GPA levels were significantly larger than primary effects, as was found with SIMCE tests.

However, one key limitation in Chapter 5 findings is that the student cohort was reduced to 64% by the time they completed secondary education, the precise point in time of a possible transition to higher education. Students in the lowest parental education group had the largest attrition in the cohort (38%), while the highest parental education group was found to have had the lowest attrition (16.3%)[61]. This finding is not surprising, because it would be expected that a larger proportion of students from less-advantaged backgrounds would not complete the 12-year period of schooling. In the literature of educational transition models, the attrition of the sample is a recurrent issue that has proven to downwardly bias

[61] The secondary education and vocational qualifications groups of parental education had attritions of 28.3% and 24.3%, respectively.

the pattern of social background across transitions. Cameron and Heckman (1998) argued that unobserved heterogeneity influences educational transitions, producing continuous sample-selection. They asserted that this selection bias might explain the declining pattern of social background across transitions in the Mare model (Mare, 1980, 1981). While the literature of empirical applications of Boudon (1974) acknowledges the issues of unobserved heterogeneity and potential sample-selection across transition, these problems have gone unresolved (Neugebauer *et al.*, 2013; Rudolphi, 2013; Jackson and Jonsson, 2013).

The main parts of this chapter are: First, to test if sample-selection in the student cohort downwardly biases the effect of parental education in the transition to higher education, in comparison with the results found in Chapter 5. The specific measurement to account for sample-selection is completion of secondary education. Second, to examine the pattern of secondary effects between the completion of secondary education and the transition to higher education, in order to assess *Maximally Maintained Inequality* (MMI) across transitions. In Boudon's inequality of educational opportunities model, secondary effects are associated with the effect of social background on educational transitions, controlling for academic performance. If sample-selection is diminishing the effect of social background across transitions, secondary effects accounting for sample-selection would be larger than those found in Chapter 5. Therefore, the secondary effects not contaminated by sample-selection are unbiased.

The modelling setting in this chapter is a mediation model, in which the association of parental education and transition to higher education is mediated by the academic performance in SIMCE tests in grade 8. The mediation model is similar to the Karlson, Holm and Breen (2010) method, and the resulting primary and secondary effects resemble those found using the Erikson *et al.* (2005) approach, and will be shown in the mediation modelling section. Then, I address missingness in parental education levels and SIMCE scores in grade

8 as a sensitivity test of primary and secondary effects. The sample selection specification is then introduced into the mediation modelling-framework, as a parallel mediation model of secondary education completion where SIMCE in grade 4 is the mediator. Within this setting of two parallel mediation models, primary and secondary effects can simultaneously be derived for both transitions. The residuals of higher education transition and completion of secondary education are allowed to be correlated, to test if sample-selection holds between transitions. More details of the modelling setting will be discussed in the empirical strategy section and in the methodological appendix.

The sample-selection strategy is based on Lucas, Fucella and Berends (2011), which gives due consideration to the following: (i) instruments in the selection model, (ii) standardised coefficients in relation to the dependent variable (y) to overcome the scale identification issue in non-linear models, and (iii) previous measures of academic performance. The selected instruments influencing secondary education completion are the rate of labour force participation[62] by gender, and percentage of teen maternity[63], both for the age-group 15–19. These indicators were computed as proportions at the regional level[64] in the cross-sectional household survey (CASEN, 2011), which was described in Chapter 3. The rationale to select these regional indicators as instruments is because official statistics showed that among the reasons for dropping out of school in 14–17 year olds, economic and maternity/paternity were higher in the two-lowest quintiles of income (Ministerio de

[62] The rate of labour force participation is measured as the ratio of the population aged 15–19 employed or looking for work over the total population aged 15–19, at the time the survey was conducted (2011).
[63] The percentage of teen maternity is not related to the rates generally reported in demographic indicators. In this case it is just the ratio of the female population aged 15–19 who reported to have at least one child over the total female population aged 15–19, at the time the survey was conducted (2011). The question in the questionnaire was directed to females, which is the reason this proportion was computed only for girls.
[64] Indicators at the municipal level were not possible to use because not all municipalities in the country were sampled in CASEN (2011). Thus, these indicators generated additional missing data when they were linked to the student cohort for the same year. Moreover, an attempt to estimate the model with the aforementioned indicators did not converge in Mplus (Muthén and Muthén, 1998–2017).

Desarrollo Social, 2013)[65]. I will come back in more detail to this point in the descriptive analysis. In (ii), the purpose of comparing y-standardised coefficients was to identify unbiased changes in the effect of parental education across transitions. I also derive average partial effects (APEs) for the coefficients of parental education to validate the findings of y-standardised coefficients, and to compare coefficients across non-linear models on the probability scale (Karlson, Holm and Breen, 2012). Finally, in (iii) I account for SIMCE tests in grade 4, when the cohort aged 9–10, as the measure of academic performance previous to the SIMCE tests in grade 8, used for the transition to higher education.

The next section presents the specific research questions and hypotheses for the chapter, following the description of the variables and modelling strategy. The descriptive analysis precedes the modelling results. Finally, this chapter ends with the discussion of the findings and conclusions.

[65] I will discuss the rationale showing the referred official figures in the descriptive analysis.

6.2 Research Questions

As a reminder of Chapter 2 (section 2.7), the general research question of this chapter and Chapter 5 is:

> To what extent are social inequalities associated with students' chances of progression into higher education?

Then, the specific research questions in this chapter are:

- Is sample-selection producing a downward bias in the effect of social background on the transition to higher education?

  I hypothesise that the effect of social background on the transition to higher education is biased downward; thereby secondary effects will tend to be larger than those found in Chapter 5.
- Do secondary effects in the completion of secondary education and higher education transition have opposite contributions to the overall IEO? If so, is this result consistent with *Maximally Maintained Inequality* (MMI)?

  I hypothesise that social background will be found to matter to a larger extent for the transition to higher education than it will for secondary education completion, which is consistent with MMI across transitions.

6.3 Data and Variables

The data come from the national student register linked to national standardised tests (SIMCE[66] grades 4 and 8) for the same student cohort analysed in Chapter 5. The cohort was born in 1995 to 1996, starting compulsory school in 2002 at age 6 to 7. The original sample size is 210,168 children. Of these, 134,595 (64%) graduated from secondary education by

[66] *Sistema de Medición de la Calidad de la Educación* (SIMCE).

2013, after 12 years of compulsory education. Social background information was collected in grade 4 through SIMCE tests, when students were aged 9–10, by which time the student cohort was reduced to 196,231 students (93.4%). Therefore, the sample-selection specification will account for the cohort from grade 4, as students who were not attending grade 4 (i.e., due to dropout or grade repetition) lacked the social background information collected in SIMCE tests in grade 4.

Dependent variables

The main dependent variable is binary, capturing whether or not students enrolled in higher education in 2014, after completing secondary education in 2013. The second dependent variable is secondary education completion, which is also binary, capturing whether or not students completed secondary education in 2013.

Independent variables

Parental education and gender are used in both models of secondary education completion and transition to higher education models.

Parental education remains unchanged from the previous chapter, as a four-category variable, of which the highest educational level is university degree, followed by advanced vocational qualifications, secondary education, and primary education or none.

Gender is a binary variable that takes the value of one for girls and zero for boys.

National standardised tests (SIMCE) assessed in grade 8 are only considered for the transition to higher education, but not for secondary education completion. The reason is that missingness in SIMCE scores in grade 8 is related to the attrition in the cohort. From the 196,231 students found in grade 4, 90.3% progress to grade 8 after four years. The remaining students (9.7%) dropped out school or they were repeating a lower grade at that time. Conversely, grade 4 SIMCE tests were not considered in the transition to higher education because: (i) they exhibit a lower correlation (0.69) with university entrance exams (PSU) than SIMCE grade 8 (0.78), although, as I stated in Chapter 5, the PSU would have been the preferred measure of academic performance if it had been compulsory for all higher education institutions; (ii) primary and secondary effects using SIMCE grade 4 scores show no significant difference between parental education groups, while primary effects for the higher group of parental education is higher in SIMCE grade 8; and (iii) grade 8 in the timeline is closer to the transition to higher education, i.e., after completing grade 12, rather than grade 4.

SIMCE scores in both grade 4 and 8 are the average of Spanish and Mathematics scores, then standardised with a mean zero and a standard deviation of one for the 196,231 students with information available in both tests.

Instruments

As a reminder, in 2011, the student cohort was made up of grade 10 students who were 15–16 years old. At that time the promotion rate sharply decreased from 2010 compared to the decreasing rate exhibited in previous years (see Figure 3.1 in Chapter 3). It is for this reason that instruments were selected for the year 2011.

It was stated in the introduction that three indicators were included as instruments in the specification for secondary education completion; this is the selection model. These indicators are the labour force participation rate for the age-group 15–19, separated for boys and girls, and the percentage of teen maternity, only for girls. These three indicators were computed for each of the 15 regions in Chile using data from the cross-sectional household survey (CASEN, 2011). I will discuss the rationale of selecting these indicators in the following descriptive analysis.

6.4 Descriptive Analysis

The sample size used in this section matches the sample size in the modelling estimation under the structural equation modelling (SEM) framework of the mediation model[67], which is the modelling setting in this chapter. The sample size in the independent variables, such as parental education, gender and the instruments is 170,856 students, thereby it is the sample size of the model accounting for sample selection. The sample size for the transition to higher education is 121,088 students.

Table 6.1 displays a comparison of the proportion of students who completed secondary education (2013) and the proportion who transitioned to higher education, following graduation (2014). All groups of parental education show a significant reduction between transitions. The largest reduction (38%) was in the lowest-group of parental education, as was stated in the introduction. Gender differences are meaningful in secondary education completion, a larger proportion of girls progress swiftly through the period of compulsory education. However, gender differences are not significant in the transition to higher education. Performance differences in SIMCE scores between transitions, show a similar pattern in parental education groups, but a larger reduction in the two lowest quartiles of SIMCE scores in both grades, which was at least 50%. SIMCE in grade 8 seems to capture a larger porportion of students who completed both transitions than the proportions seen from SIMCE grade 4. Therefore, there seems to be ample evidence of sample-selection between secondary education completion and the transition to higher education.

[67] The main analysis were conducted in Mplus (Muthén and Muthén, 1998–2017), which estimates models conditional on the observations in the independent variables, thereby the sample size of the models correspond to the observations in parental education, gender and the instruments. SIMCE scores in grades 4 and 8 and the transitions are dependent variables in the models, as well as being treated as mediators in the models, having 167,920 and 147,412 observations each, respectively.

Table 6.1 Parental education and SIMCE scores by transition

	Secondary education completion (%)	**N**	**Percentage**	**Transition to HE (%)**	**N**	**Percentage**
Highest Parental Ed. Attainment						
Degree	83.7	23,206	13.6	78.0	19,429	16.0
Advanced Vocational Q.	75.7	22,023	12.9	67.6	16,664	13.8
Upper Secondary	71.7	73,292	42.9	57.5	52,538	43.4
Primary or none	62.0	52,335	30.6	43.6	32,457	26.8
Girls	74.3	87,768	51.4	58.2	65,186	53.8
Boys	67.3	83,088	48.6	58.8	55,902	46.2
SIMCE Average score in grade 4 (†)						
Q(high)	87.3	41,980	25.0	74.8	36,658	30.7
Q(3)	78.6	41,980	25.0	60.9	32,982	27.7
Q(2)	68.0	41,980	25.0	49.5	28,550	23.9
Q(low)	50.1	41,980	25.0	39.7	21,033	17.6
SIMCE Average score in grade 8 (‡)						
Q(high)	90.8	36,853	25.0	77.6	33,459	29.2
Q(3)	83.5	36,853	25.0	61.0	30,773	26.9
Q(2)	74.7	36,853	25.0	49.6	27,527	24.0
Q(low)	61.9	36,853	25.0	39.5	22,826	19.9
Total		**170,856**	**100.0**		**121,088**	**100.0**

Note: Proportion tests between secondary ed. completion (%) and the transition to higher education (%), for all variables are significantly different (p<.001).
(†) There are 167,920 students with information in SIMCE scores in grade 4 for the first transition, and 119,223 students for the second transition.
(‡) There are 147,412 students with information in SIMCE scores in grade 8 for the first transition, and 114,585 students for the second transition.

Table 6.2 summarises the proportion of youngsters participating in the labour market and the proportion of the teenager mothers. Boys are more likely to be involved in the labour force than girls, and teenager maternity reached an average of 10% in all regions of Chile.

Table 6.2 Instruments

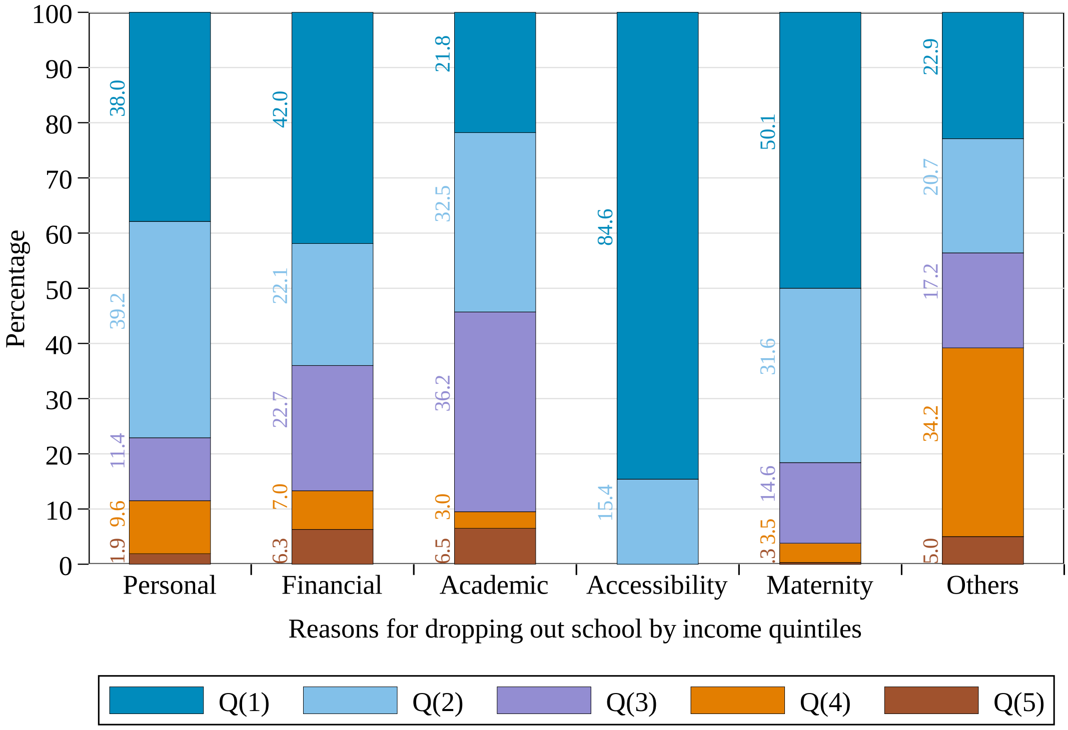
Regional indicators	N	Mean	Standard Deviation
Labour force participation (15–19, boys)	170,856	0.21	0.04
Labour force participation (15–19, girls)	170,856	0.12	0.02
Teen maternity (15–19)	170,856	0.10	0.02

Instruments

The selection of the aforementioned instruments is based on the social policy report (Ministerio de Desarrollo Social, 2013, p. 141), which used the household survey (CASEN, 2011) for its question related to reasons for dropping out of school in 14–17 year-olds. The reasons that could be approximated by instruments included: (i) financial (14.4%), (ii) pregnancy, maternity or paternity (12.2%), (iii) academic performance (8.6%), and (iv) accessibility to schools (0.4%). The largest proportions of dropping out reasons were non-measurable for the purpose of creating instruments, these were: personal (25.5%), and others (38.9%), these relate to stated alternatives given for this question in the survey (e.g., stays at home, disability, family issues, lack of motivation). Financial reasons were related to school costs or being in the labour market (i.e., employed or looking for a job). Accessibility to schools was related to the inexistence of schools in the area of residence or difficulties in transportation to access a school. Academic performance reasons were related to low academic performance or to having been expelled from school.

The distribution of the aforementioned reasons for dropping out of school by income quantiles is shown in Figure 6.1. The two-lowest quantiles, Q(1) and Q(2), have the largest proportions for financial and maternity reasons, which can be approximated through the instruments described above. Accessibility to schools was not considered because only 0.4% indicated this as a reason for dropping out of school. Academic performance, on the other hand, is considered in the models through SIMCE tests.

Figure 6.1: Reasons dropping out of school by income quintiles (14–17) CASEN (2011)



Note: The indicators in the figure were retrieved from 'Informe Política Social 2013' (Ministerio de Desarrollo Social, 2013, p. 142) using data from CASEN (2011).

It may be argued that the selected instruments of labour force participation rate for the age-group 15–19 could have been chosen to be more specific to a sector, or sectors, in the economy, as in Lucas, Fucella and Berends (2011), who selected as instruments the percentage of workers in manufacturing and agriculture in the regions of students' residence. Regrettably in CASEN (2011) the sample size of the age-group who is engaged in the labour market is small (2,924) compared to the sample size of the age-group not engaged in the labour market (15,085). This sample size was reduced even further, by 29.3%, due to missingness in the economic sector variable, which substantially decreased the sample by regions. Therefore, the proportion of teens engaged in the labour market, further subdivided by the finer distinctions of economic sectors, would have been insignificant in the model.

Another important feature of instruments is that they should be related to the selection model because they are used for identification of this specification, in this case secondary education completion, but not the subsequent outcome, the transition to higher education. However, instruments that accomplish this desired property are difficult to find because it depends of factors beyond the researcher's efforts, such as the quality and availability of the data (Holm and Jæger, 2011). In this regard, I tested the effect of the instruments in the transition to higher education model, comparing the sign and significance found for secondary education completion (the selection model), following Lucas, Fucella and Berends (2011).

Boys' participation in the labour force was significant ($p<0.001$) in both transitions, where it decreased the likelihood of secondary education completion but increases the likelihood of the transition to higher education. Girls participation in the labour force is significant ($p<0.05$) in the first transition and in the second transition ($p<0.001$), where it decreased the likelihood of both transitions. Teen maternity increased significantly ($p<0.001$) the likelihood of secondary education completion, but decreased the likelihood of higher education transition ($p>0.05$). The three instruments are jointly significant ($p<0.001$) in both transitions. I consider the opposite effects of boys' participation and teen maternity[68] in the transitions to be sufficient, for the purposes of this chapter, to argue their validity. Girls' participation in the labour force is included, anyway, in the secondary completion model, as its effect follows boys' participation.

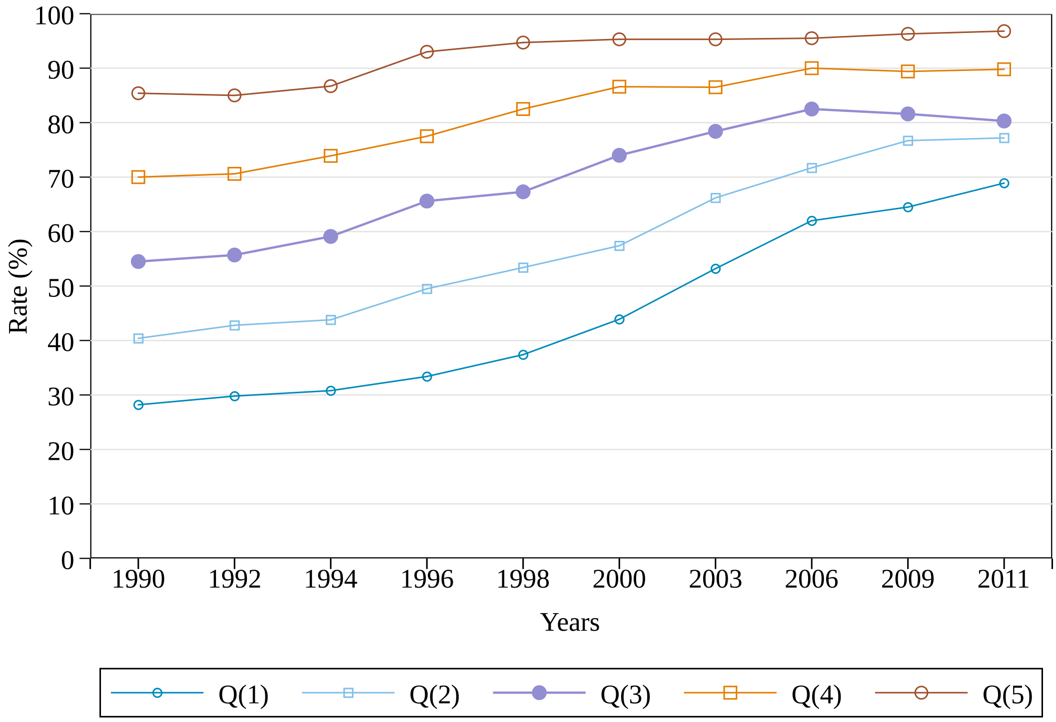
[68] Teen maternity has more support as a valid instrument in view of their significance in the transition to higher education ($p>0.05$) at 95% of confidence level.

Is the demand for secondary education satisfied in the advantaged group?

To close this section, I present evidence to argue that the advantaged group might have been close to *saturation* in secondary education completion. Thereby, it might indicate that MMI is consistent in the educational context of Chile.

In Figure 6.2, the trends of secondary education completion in the highest income quantile, Q(5), surpassed the 90% benchmark since 1996. The fourth quintile, Q(4), has a completion rate above 80% since 1998. While, the lowest quintile, Q(1), reached a completion rate above 50% since 2003 and it was close to 70% by 2011. The distance between the lowest and highest income quantile between 1990 and 2011 was reduced from 57.2% to 27.9%.

Figure 6.2: Secondary education completion rates (20–24) by income quintiles

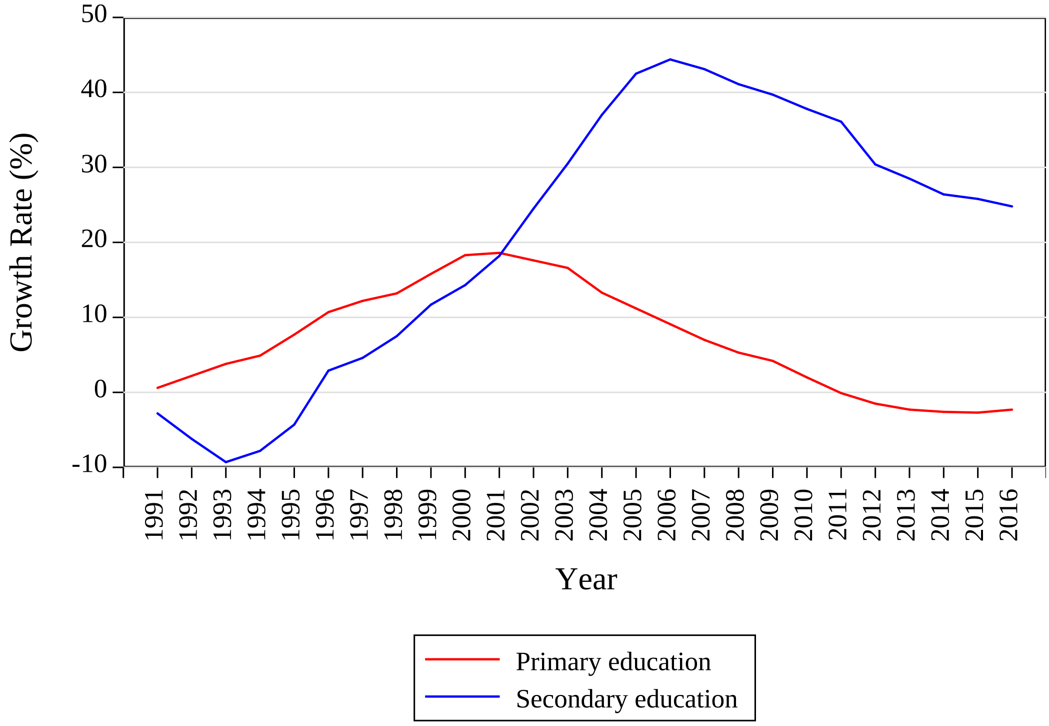


Note: The indicators in the figure were retrieved from 'Informe Política Social 2013' (Ministerio de Desarrollo Social, 2013, p. 140) using data from CASEN (1990–2011).

The 90% rate of secondary education completion in the advantaged group of income resembles the benchmark used in Haim and Shavit (2013) to measure saturation in the MMI theory. They examined trends of educational expansion and its influence on inequality of educational opportunities on 24 European countries. They defined a 90% benchmark of enrolment in upper secondary and tertiary education to measure saturation for children whose fathers had attained upper secondary and tertiary education and they attained the same levels themselves. However, the information presented in the report of Ministerio de Desarrollo Social (2013) lacked other measures of social background to follow the method of Haim and Shavit (2013). Figure 6.2 attempts to show an apparent saturation of the highest quantile group in secondary education completion, resulting in a reduction of educational inequalities between advantaged and less advantaged groups.

On the other hand, it can be argued that income tends to be volatile and dependant of the economic structure of the considered period. In contrast, other measures of social background, such as parental education or social class, show greater stability. Moreover, the age group in Figure 6.2 is older than the expected age at which students complete secondary education (17–18), which increases the proportion of completion over time. It is also important to notice that secondary education was made compulsory in Chile since 2003, but the reforms conducted in the 1980s aimed to expand the educational system at all its levels. In effect, the enrolment in primary and secondary education grew substantially since 1990 (Figure 6.3), although only secondary education continued in the same trend up to mid-2000s, despite the growth rate has decreased is still over 20% than the enrolment in 1990. Conversely, primary education enrolment has decreased since mid-2000s, reaching slightly lower levels (-2%) of enrolment than in 1990.

Figure 6.3: Primary and secondary education growth rate of enrolment since 1990 to 2016



Note: Data retrieved from Ministerio de Educación (2010, 2017).

A possible explanation to decreasing rates of enrolment in both educational levels might be related to the change in the population composition, which is not addressed in this thesis but can be the focus of future research. However, the focus of this chapter is to examine MMI theory by testing the change in the coefficients of parental educational levels between transitions (Lucas, 2001), and by the change in secondary effects from Boudon's IEO model, following Jackson (2013b).

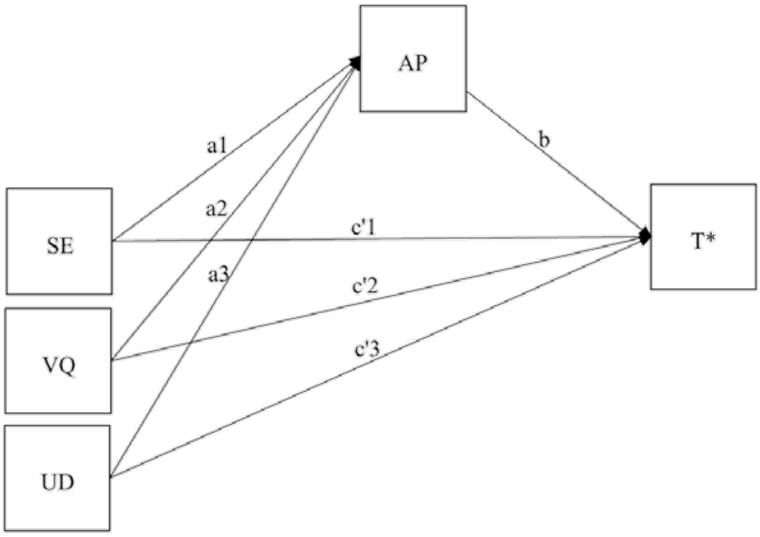
6.5 Modelling Strategy

The modelling strategy in this chapter has three stages: (i) estimation of a mediation model for the transition to higher education and compare primary and secondary effects with the methods applied in Chapter 5; (ii) missingness in parental education and SIMCE scores in grade 8 being addressed through multiple imputation in the mediation model; and (iii) secondary education completion and the transition to higher education being modelled as parallel mediation models, accounting for sample-selection in the student cohort in the secondary education completion specification. See Chapter's 6 methodological appendix in section 6A.2 for more technical details of the modelling strategy.

In the first stage, I propose to empirically estimate Boudon (1974) IEO decomposition in primary effects and secondary effects using a mediation model[69]. For the reader not familiar with mediation models, Figure 6.4 provides an illustration. The left-hand-side illustrates parental education levels (i.e., secondary education, vocational qualifications and university degree), the variable in the middle, AP, is the academic performance in SIMCE tests playing the role of the mediator between the parental education levels and the transition to higher education (T*), on the right-hand-side. The equations describing the mediation model and the derivation of indirect, direct and total effects can be seen in section 6A.2.1 of the methodological appendix.

[69] Karlson, Holm and Breen (2010) and Karlson and Holm (2011) followed a similar methodological approach.

Figure 6.4: Mediation model diagram of the transition to higher education



Primary effects resemble the indirect effects of the mediation model, capturing the influence of the mediator (academic performance) in the independent variable-outcome association (parental education-transition to higher education). These are the product of the *a* coefficients with the *b* coefficient in Figure 6.4. Secondary effects, in contrast, resemble the direct effects, accounting for the effect of parental education in the transition while controlling for the mediator (academic performance), and these are the *c'* coefficients in Figure 6.4. Both indirect and direct effects sum to 1 or, alternatively, to 100%, which is the total effect of parental education on the transition to higher education. It is important to emphasise that the resulting primary and secondary effects found in the mediation model are *relative effects* against the lowest-level of parental education (primary education or no-formal education), as parental education is a categorical variable.

Complementary to the mediation analysis, the method of multiple imputation[70] addressed missingness on parental education groups (10%) and SIMCE average scores in

[70] I only consider multiple imputation in the variables of the transition model to higher education because the secondary education completion model is used as the sample-selection model in the following stage of the empirical strategy.

grade 8 (5.8%)[71]. Under the assumption of 'missing at random' (MAR), Bartlett (2013) suggested the application of a Chi-square test between the variables with complete information and those with missingness. However, the Chi-square test only allows the testing of a pair of variables at once. Therefore, two separated logistic regressions were, instead, applied to test for missing data on parental education and SIMCE scores, recoded as 1=missing and 0=not missing, upon the rest of variables used in the mediation model (i.e., gender, and transition to higher education). The resulting coefficients in both logit regressions were significant ($p<0.01$)[72] suggesting that is not plausible that the missing data in parental education and SIMCE scores were 'missing completely at random' (MCAR). Enders (2015), however, remarked that testing mechanisms are not sufficient to conclude whether the missingness is 'missing at random' (MAR) or 'not missing at random' (NMAR). Moreover, the methods dealing with NMAR are similar to those applied to MAR, in terms of efficiency.

Muthén, Muthén and Asparouhov (2016) argued that the proportion of non-missing data in all variables of the model allow the identification of the model parameters to estimate for the set of variables in the model. The lowest coverage of the data is 85.1% when SIMCE scores (grade 8) and parental education groups are observed together[73]. This percentage of coverage fits Muthén, Muthén and Asparouhov (2016) recommendation of the 80–90% range for estimates to be trustworthy. Moreover, a higher coverage in the data means less dependence on the assumption and the methods to address missing data. Therefore, it is safe to assume that the missing data in parental education and SIMCE scores (grade 8) are MAR.

[71] The percentage of missingness is relative to the 134,595 students completing secondary education after the 12-years period.
[72] The coefficient for girls was significant at $p<0.05$ in the logistic regression of SIMCE scores as the dependent binary variable.
[73] The coverage of the data in other variables in the model is no less than 90%.

The two variables with missing data come from the SIMCE tests. These tests were designed to evaluate the academic performance of the schools. In this regard, multiple imputation should consider the multilevel structure of the missingness of both variables. I follow the multiple imputation routines implemented in Mplus (Muthén and Muthén, 1998–2017) for a two-level multilevel model, described in Asparouhov and Muthén (2010), and Enders (2015). The number of imputed datasets were 100[74] for both parental education and SIMCE scores. A technical description of Mplus multilevel imputation routines was developed in section 6A.2.3 of the methodological appendix.

6.6 Mediation modelling

The estimates of a mediation model consist of two separated regression models (Table 6.3). The mediator model is the left-hand-side panel of the table, which is a linear regression of the standardised average SIMCE scores of mathematics and Spanish, collected in grade 8 when the cohort was 13–14 years old. The predictors in the linear model are dummy variables from parental education levels, such as secondary education, vocational qualifications and university degree, and gender (girl). The reference category in parental education is the lowest level, i.e., primary education or none. The right-hand-side panel of Table 6.3 is the probit model for the transition, which includes the parental education levels, gender and SIMCE scores in grade 8. Both models follow the mediation model diagram seen in Figure 6.4.

The coefficients of the probit model are y-standardised[75] to allow comparisons between probit models in the same scale, the coefficients of the probit model are divided by

[74] It is common in studies with sample sizes less than 10,000 observations to select around 20 imputed datasets. The cohort is over 100,000 observations. I argue that 100 imputed datasets are the ideal number for that sample size and for the time needed to compute the imputations.
[75] The linear model coefficients are exactly the same as y-standardised coefficients.

the estimated standard deviation of the dependent variable, which is the sum of the estimated variance of the model and the theoretical residual variance of the probit (1), (Lucas, Fucella and Berends, 2011; Karlson, Holm and Breen, 2012).

The coefficients from SIMCE and the probit for the transitions suggest that there was a gradient effect of parental education levels, which was higher for students whose parents held a university degree and gradually reduced to the group of students whose parents attained secondary education. There is a negative and significant effect on girls in SIMCE scores, although of low impact. Gender is not significant in the transition to higher education. This latter result agrees with the proportion of boys and girls transitioning to higher education (Table 6.1). Moreover, SIMCE scores have a positive and significant influence in the transition.

Table 6.3 Mediation model of the transition to higher education, mediated by SIMCE in grade 8

	SIMCE scores (Linear)	Transition to HE (Probit)
	y-standardised coefficients	
Parental Education		
Primary ed. or none (reference)		
Degree	1.18	0.53
	(.01)	(.01)
Advanced Vocational Q.	0.67	0.38
	(.01)	(.01)
Secondary Education	0.38	0.22
	(.01)	(.01)
SIMCE in grade 8 (Spanish+Maths)	-	0.31
		(.00)
Girl	-0.03	0.01
	(.01)	(.01)
Intercept	-0.44	-0.02
	(.01)	(.01)
McKelvey and Zavoina's R-square		0.17
		(.00)
R-square	0.15	
	(.00)	
N	121,088	

Note: Standard errors in parentheses. All coefficients are significant (p<0.05), except girl (p>0.05) in the transition model.

The magnitude of the effects in parental education levels in the transition are not relevant at this point. However, I will return to them in detail in the next section for the comparison with the effects of the mediation model accounting for sample-selection.

Finally, I derive primary and secondary effects from indirect and direct effects of the mediation model, respectively. Primary effects are the proportion of the indirect effects over the total effect. Similarly, secondary effects are the proportion of direct effects over the total effect. These effects are relative to the reference group of parental education (primary education or none). I compare these effects alongside the effects found using the methods of

Erikson *et al.* (2005), and Karlson, Holm and Breen (2010), in Chapter 5[76], and the resulting effects derived from 100 imputed datasets accounting for missingness in parental education and SIMCE scores in grade 8 (Table 6.4).

The primary and secondary effects in the mediation model were found to be fairly similar to the effects found using the Erikson *et al.* (2005) and Karlson, Holm and Breen (2010) methods. These results confirm that secondary effects are the dominant effect (>50%) in the overall inequality of the transition to higher education. This is particularly evident in the effect of the choices of the students and their families to continue further in the educational system -subject to their academic performance- in accounting for greater inequality among social groups, in terms of parental education.

The primary and secondary effects from multiple imputation models in 100 datasets also show significantly larger secondary effects (>50%) explaining IEO, although the primary effects seem to be slightly smaller, and consequently secondary effects slightly larger, than the effects found in the mediation model which did not address sample selection (Table 6.4). Their confidence intervals at the 95% level overlap, thereby these differences are not significant. Moreover, the overall pattern of primary and secondary effects in the first two panels of the table is reflected in the multiple imputation models, notwithstanding the modestly enhanced importance of secondary effects for the three groups of parental education therein. Therefore, it can be stated that missingness in parental education and SIMCE scores in grade 8 did not significantly influence the resulting primary and secondary effects found in Chapter 5.

[76] As a reminder, the effects in both methods, and in the Buis (2010) approach, were all very similar.

Table 6.4 Primary and Secondary effects of the Transition to Higher Education by different methods (SIMCE grade 8)

Parental Education	Erikson *et al.*/KHB[77]		Mediation Model		Mediation model (*m*=100 datasets)	
	Primary Effects (%)	Secondary Effects (%)	Primary Effects (%)	Secondary Effects (%)	Primary Effects (%)	Secondary Effects (%)
Degree	40.3	59.7	40.8	59.2	39.3	60.7
			[39.3, 42.2]	[57.8, 60.7]	[37.9, 40.6]	[59.4, 62.1]
Advanced Vocational Qualifications	35.2	64.8	35.1	64.9	34.0	66.0
			[33.4, 36.7]	[63.3, 66.6]	[32.4, 35.6]	[64.4, 67.6]
Secondary Education	35.1	64.9	34.8	65.2	32.5	67.5
			[32.8, 36.9]	[63.1, 67.2]	[30.6, 34.5]	[65.5, 69.4]

Note: Confidence intervals at 95% in square brackets. Relative to primary education or none.

In the next section the second stage of the modelling strategy follows, which is the mediation model accounting for sample-selection.

6.7 Sample selection

Table 6.5 displays the transition estimates for the completion to secondary education (selection model) alongside the transition to higher education, expressed in y-standardised coefficients to allow comparisons in the same scale across transitions (Lucas, Fucella and Berends, 2011; Karlson, Holm and Breen, 2012). The linear models of SIMCE scores in grades 4 and 8, which are also part of the mediation model setting similar to the mediation model presented in Table 6.3, are detailed only in the appendix (Table 6A.1), allowing attention to be focused on the transition models here.

[77] KHB stands for Karlson, Holm and Breen (2010) methodology to derive primary and secondary effects.

At the end of Table 6.5 can be seen that the correlation among residuals (ρ) between secondary education completion and the transition to higher education is significant ($p<0.001$), which suggests that selection bias affects the transition to higher education. The negative sign of the correlation (-0.42) indicate that jointly the unobserved factors that increase the likelihood of secondary education completion, decrease the likelihood of the transition to higher education. As it was stated in section 6.4, two out of the three instruments yielded opposite effects between transitions (i.e. rates in boys' participation in the labour force and teen maternity), thereby the negative sign of the correlation between transitions seems to be reasonable. Furthermore, in the literature other studies support the negative correlation between school completion and higher education entry (Riphahn and Schieferdecker, 2012; Lucas, Fucella and Berends, 2011), which were found in Germany and the United States respectively.

The labour participation in boys and girls reduces the likelihood of secondary education completion, while teen maternity increases it (Table 6.5). This latter result might seem to be counterintuitive, as it is expected that maternity might be inversely related to school continuation. However, it can be related to gender differences, as girls[78] in the cohort are more likely to complete secondary education than boys (Table 6.1), which can also be seen in the significant effect of girls in completing secondary education. Moreover, Santos (2009) also found that girls are less likely to dropout school than boys in Chile.

I turn, now, to the effect of parental education between transitions to answer the research questions as well as to support or discard the proposed hypotheses.

[78] This tendency is not only a characteristic in the student cohort, but it has been a trend since 1990, as can be seen in Figure 6A.1 in the appendix. However, an older age group than the student cohort was considered (20–24).

The effect of parental education levels once more exhibits a decreasing gradient, in both transitions, from the degree attainment group to the secondary education group. The most important comparison in parental education levels is, however, between transitions, as parental education exerts a stronger effect in the transition to higher education than completion of secondary education. These differences are also significant ($p<0.001$) in Wald tests between transitions at the same educational levels. Therefore, parental education effects are significantly larger for the transition to higher education than secondary education completion, which might signal a larger importance of social background at higher educational transitions.

On the other hand, the y-standardised coefficients of the transition to higher education in the mediation model that does not address sample-selection (Table 6.3) showed larger effects on parental education levels than those found in the model accounting for sample-selection in secondary education completion (Table 6.5). These differences are significant ($p<0.05$) as their confidence intervals by educational levels do not overlap. Nevertheless, the validity of y-standardised coefficients as a reliable measure for comparison across non-linear models has been questioned (Karlson, Holm and Breen, 2012). Average partial effects (APE), alternatively, measure the average change in the population probability of making the transition by the difference between the three parental educational levels (i.e., degree, vocational and secondary education) in relation to the lowest group (primary ed. or none), while other variables in the model are held constant (Holm and Jæger, 2011).

Table 6.5 Probit models for secondary ed. completion and transition to higher education[79]

	Secondary ed. completion	**Transition to HE**
	y-standardised coefficients	
Parental Education		
Primary ed. or none (reference)		
Degree	0.15	0.38
	(.01)	(.01)
Advanced Vocational Q.	0.07	0.27
	(.01)	(.01)
Secondary Education	0.06	0.15
	(.01)	(.01)
SIMCE in grade 4 (Spanish+Maths)	0.27	-
	(.00)	
SIMCE in grade 8 (Spanish+Maths)	-	0.21
		(.00)
Girl	0.15	-0.02
	(.00)	(.01)
Instruments (regional % in 2011)		
Labour force participation (15–19, boys)	-0.26	-
	(.07)	
Labour force participation (15–19, girls) (+)	-0.25	-
	(.11)	
Teen maternity (15–19)	0.84	-
	(.13)	
Intercept	0.25	0.13
	(.02)	(.01)
McKelvey and Zavoina's R-square	0.59	0.43
	(.00)	(.02)
Residuals correlation (ρ)	-0.422	
	(.01)	
N	170,856	

Note: Standard errors in parentheses.
All coefficients are significant (p<.001), (+) (p<.05).

Table 6.6 shows average partial effects (APE) for secondary education completion, transition to higher education conditional on secondary education completion, and the

[79] The regular coefficients in both probit models match the coefficients found using the *heckprobit* command in Stata, with a slight variation in decimal places.

transition to higher education without accounting for sample-selection. I consider the conditional probability on secondary education completion, given the correlation between the residuals of both transitions is significantly different from zero ($p<0.001$). Therefore, the probabilities of both transitions depend on each other, which is precisely the case in practice as students can only access to higher education after completing secondary education.

The APEs exhibit changes in probability between each of the three levels of parental education in reference to the lowest level. As before, for every transition there is evidence of a decreasing gradient form the highest educational attainment. The most compelling finding is the likeness in the APEs between all models of higher education transition, whether they account for or do not account for sample selection. The APEs between these models are not significantly different (at $p<0.05$) as their confidence intervals overlap. The aforementioned result contradicts the differences between these models on y-standardised coefficients, whereas the APEs in secondary education completion are significantly smaller, similarly to the result found on y-standardised coefficients.

Table 6.6 Average partial effects of parental educational levels across transitions

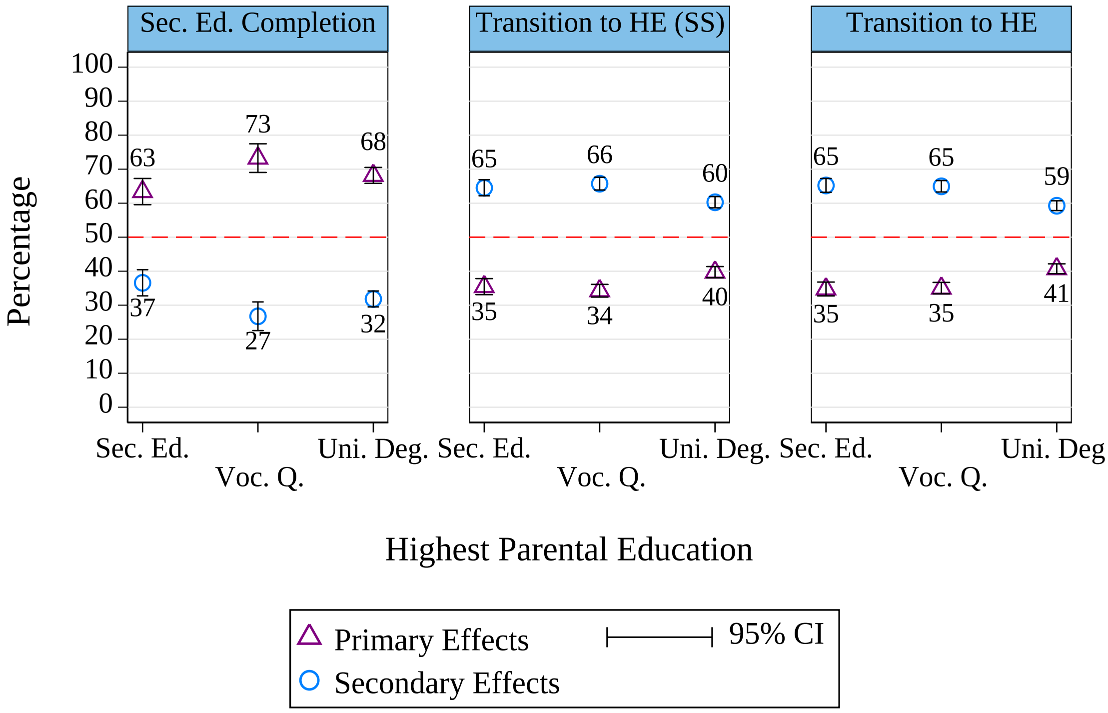
	Average Partial Effects (APEs)[80]		
	Secondary ed. completion	**Transition to HE | Sec. ed. c.**	**Transition to HE**
Parental Education			
Degree	0.08	0.21	0.21
	[.070, .085]	[.197, .216]	[.199, .218]
Advanced Vocational Q.	0.04	0.15	0.15
	[.029, .044]	[.138, .157]	[.145, .163]
Secondary Education	0.03	0.08	0.09
	[.027, .037]	[.077, .091]	[.083, .097]

Note: Confidence intervals at 95% in square brackets. Relative to primary ed. or none.

[80] APEs here are estimates from using the *margins* command in Stata; APEs estimated in Mplus were alike.

To conclude this section, the question of whether sample-selection influences the effect of parental education on the transition to higher education is further investigated. To this end, I derived primary and secondary effects from the sample-selection model under the mediation modelling setting, and I compared them with the effects of the mediation model ignoring sample-selection (Figure 6.5). As a reminder, primary and secondary effects were derived using the unstandardised coefficients of the models, following the explanation given in section 6.5.

Figure 6.5: Primary and secondary effects across transitions



Note: Primary and secondary effects are relative to low parental education (primary or none). SS=Sample-selection accounted in the transition to HE.

Figure 6.5 illustrates the patterns found in APEs between secondary education and higher education transitions, as well as between transitions only to higher education, accounting for or not accounting for sample selection. Thus, it appears that not accounting for sample-selection in the transition to higher education does not bias the effect of parental education. Moreover, the general trend of significantly larger secondary effects, over the 50% benchmark, was unaltered. While, the effects for secondary education completion were completely opposite to the effects of the transition to higher education, primary effects were the dominant effects (>50%) over secondary effects in the overall inequality of educational opportunities.

6.8 Discussion

The general question is addressed as the conclusion of the chapter, while the specific research questions are addressed in this section.

The first research question in this chapter sought to test the influence of sample-selection in producing a downward bias in the effect of social background on the transition to higher education. I hypothesised that the effect of social background is downward biased on the transition to higher education, resulting in smaller secondary effects.

The results from APEs analysis, concerning primary and secondary effects, showed evidence contrary to my hypothesis; there were no significant differences between the models accounting or not for sample-selection. However, the comparison between y-standardised coefficients between the aforementioned models revealed that parental education coefficients are larger by not accounting for sample-selection, which might indicate the presence of upward bias instead. I argue that y-standardised coefficients are only rescaled coefficients, by the estimated variance of the dependent variable. These measures ignore the effect of the correlation between the residuals of both transitions, which was used in the computation of conditional probabilities in APE. It is noteworthy that conditional probabilities had to be considered in this case, as students' eligibility to make the transition to higher education depends on them having completed secondary education. Similarly, Jackson and Jonsson (2013), considered conditional transition rates in the transition to university education. Further support against y-standardised coefficients was found in the relative contribution of primary and secondary effects to the overall IEO, reasserting the findings from APEs of non-significant differences in parental education levels by accounting for sample-selection (sec. ed. completion) in the transition to higher education.

The significant correlation of residuals (ρ) in both transitions and smaller[81] coefficients in parental education in the transition to higher education, suggested that the relative contributions of primary and secondary effects might have changed. In the mediation model framework, the indirect effect of SIMCE scores in the transition are the primary effects, while the direct effect of parental education in the transition, controlling for SIMCE scores, are the secondary effects. The change in the coefficients involved in the computation of primary and secondary effects, in the model that accounts for sample-selection, were smaller in a similar proportion than the coefficients in the model that ignores sample-selection. Thereby, in terms of the model this explains why there were not significant differences in primary and secondary effects, by accounting or not for sample-selection (secondary education completion). Future research, employing new instruments or methods may confirm or refute the validity of these findings.

The second research question addressed whether or not secondary education completion and the transition to higher education show opposite primary and secondary effects, and, if they do, whether or not such a result is consistent with MMI theory.

The change in the probability of transition to higher education by parental education levels (under APEs analysis) was larger than the probability of transition for secondary education completion. This pattern is consistent with MMI theory (Lucas, 2001). Moreover, primary effects were over the 50% benchmark for the secondary education completion. In contrast, it was consistently found that secondary effects were over the 50% benchmark for the transition to higher education. Jackson (2013b) also interpreted the larger contribution

[81] In relation to the parental education coefficients of the transition model without accounting for sample-selection (secondary ed. completion).

of secondary effects in the transition to higher education compared to the larger contribution of primary effects in the transition to A-levels as consistent with MMI theory.

Torche (2005) argued that there was not enough evidence to support MMI theory in Chile, in her findings across birth cohorts. I refrain from claiming the opposite, acknowledging that comparisons across birth cohorts are also necessary to fully assess MMI, rather than making comparisons across educational transitions alone. However, the trends in secondary education completion by income quantiles, seen in the descriptive analysis, seem to indicate that birth cohort comparison might lend additional support to MMI. Figure 6.2 shows a stable pattern of secondary education completion over 90% for the highest quantile of income since 1996.

In the context of the Chilean educational system, in which families financed in a larger extent the costs of higher education (Figure 4.2 in Chapter 4), less advantaged families have scarce opportunities to access to higher education. Following Breen and Goldthorpe (1997) model, it can be argued that the advantaged groups might have satisfied their demands placed on lower levels of education, i.e., secondary education. Then, it follows that MMI theory to hold.

Finally, other robustness checks to the mediation model validated the results presented in this chapter. These include multiple imputation[82], the moderating effect of academic performance (not significant, $p>0.05$), and bootstrapped standard errors on 1,000 replications, and, indeed, none of these changed the interpretation that the predominant contribution to the overall inequality in the transition to higher education was from secondary effects.

[82] Missingness in parental education and SIMCE scores was 10% of the 134,595 students, of the total cohort, that completed secondary education.

In Chapter 7, I seek to shed some light on the role of schools in the overall educational inequality for the students in the cohort transitioning to higher education, a year after completing compulsory education.

6.9 Conclusion

The possibility was considered, that a hidden pattern in secondary education completion had a systematic impact on the transition to higher education, specifically in terms of the relative contribution of primary and secondary effects to the inequality of educational opportunities. Such sample-selection was found not to have significantly modified the relative contribution of primary and secondary effects to the overall inequality of educational opportunities; students' chances of progression were still, notwithstanding this analysis, predominantly influenced by social background differences.

Although it may be argued that the findings of this chapter are still to be more rigorously validated, this result conforms with the contextual setting of social inequality in Chile. To wit, the expansion and universalism of the country's primary and secondary education proving to have reduced the overall effect of social background in secondary completion, along with advantaged social groups typically having their demands of education at these levels satisfied (as was also shown in the descriptive analysis of rates of secondary education completion by income quintiles). Therefore, these findings are consistent with MMI theory.

Chapter 7: The influence of the Chilean educational system on IEO

7.1 Introduction

The results from Chapters 5 and 6 provided an overall measure of inequality of educational opportunity (IEO) in the transition to higher education, for the student cohort studied in this thesis. Yet, these results offer little understanding of how, in the view of critics in the student movement of 2011, the Chilean educational system itself contributed to widening educational inequality. The student movement emphasised that the market-orientation of the educational system was the source of pronounced social differences across schools. In Chapter 5, the school system in Chile was classified as highly stratified and selective following Jackson and Jonsson (2013, p. 310) classification, owing to the quality and status of schools by type of school funding (i.e., municipal, voucher and private). Furthermore, it was argued that school's selection[83] by family income tended to segregate students by their socio-economic background and subsequently to determine students' chances to advance in the educational system (Villalobos and Valenzuela, 2012; Wormald *et al.*, 2012; Bellei, 2013; Valenzuela, Bellei and de los Ríos, 2014). Following Jackson and Jonsson (2013) comparative analysis it is expected that secondary effects will increase, in relation to the stratification of the schools. Therefore, it is this chapter aim to explicitly model the effects of the school system to test the change in primary and secondary effects.

To investigate the effect of the educational system in Boudon's model, school differences in grade 8 (the mediator, SIMCE scores) and school differences in grade 12 (the

[83] That was before 2016, when the Ministry of Education introduced a new measure to regulate students' selection within schools.

dependent variable, transition to higher education) were considered, as 65% of the students in the cohort change schools between grades 8 and 12. Students change schools in that proportion because there is a great number of primary-only schools and secondary-only schools in Chile's school system (Lara, Mizala and Repetto, 2010; Racynski, 2011) Furthermore, the transition to higher education was specified, in terms of higher education institutions (i.e., vocational colleges, private universities, traditional universities) and programme of enrolment (i.e., vocational or undergraduate).

The results found in Chapter 6 evidenced that there was non-significant difference in primary and secondary effects after controlling for missingness in parental education groups, and in SIMCE scores. Similarly, it seems to be enough evidence to support the claim that the student cohort was not influenced by sample selection, over the 12-years of compulsory education. Therefore, the modelling strategy in this chapter does not account either for missingness or sample-selection to model the stratification of the educational system, in terms of school characteristics and type of higher education institutions and programmes.

The following sections present the research questions to be addressed in this chapter, as well as the data and methods used to answer the proposed research questions. The descriptive analysis focuses on differences in the student cohort by school characteristics. The first part of the modelling stage, a cross-classified mediation modelling assesses the primary and secondary effects, accounting for school effects, in the transition to higher education. Then, a multinomial mediation model examines the primary and secondary effects by higher education institutions and programmes. At the last stage, the chapter ends with a discussion of the findings and conclusions.

7.2 Research Questions

As a reminder of Chapter 2 (section 2.7), the general research question of this chapter is:

> To what extent does the Chilean educational system increase or decrease the effect of social background on the transition to higher education?

The specific research questions in this chapter are:

- Does the stratification of the schools[84] increase social background differences for the transition to higher education?

  I hypothesise that accounting for school characteristics increase the importance of secondary effects in the transition to higher education.

- Does the distinction of the transition to higher education by type of institution and programme change the contribution of primary and secondary effects?

  I hypothesise that the contribution of primary effects will increase for the transition to traditional universities and undergraduate programmes, as admission requirements rely on students' academic performance. In contrast, as high academic performance is not required in vocational colleges or programmes, secondary effects will tend to increase even more for these transitions.

- Are students from a lower social background more likely to enrol in vocational colleges and students from a higher social background more likely to enrol in traditional (prestigious) universities? If so, is this result consistent with *Effectively Maintained Inequality* (EMI)?

- I hypothesise, that students from lower social backgrounds are more likely to be enrolled in vocational colleges, while students from higher social backgrounds are

[84] I refer to the stratification of the schools to differences in type of school funding, curriculum tracks and location area among schools for the period of compulsory education. The educational system, on the other hand, encompasses both schools and higher education institutions.

more likely to be enrolled in traditional (prestigious) universities. If this is the case, there is consistent evidence to support EMI.

The description of the data and the variables follows below.

7.3 Data and Variables

The data refer to the same student cohort examined in Chapters 5 and 6. In this chapter, all analyses concern the 134,595 students that completed secondary education after the period of compulsory education.

Variables

In the first stage of the analysis, the response variable reflects the overall transition to higher education. In the second stage, the response variable is a multinomial variable specifying the type of higher education institution and type of higher education programme.

Parental educational level is the social background variable, divided into four educational levels, i.e., degree, vocational qualifications, secondary education, and primary education or none.

The variables for school characteristics are school funding in the last grade of primary education (grade 8), as well as in the last grade of secondary education (grade 12), and ‘general’, ‘vocational’, or ‘both’ curriculum track schools in secondary education (grade 12 only).

School funding is a three-category variable, which differentiates municipal, voucher and private schools.

In terms of the curriculum track, two separated variants were considered in the models. The first variant accounts for students’ choice of track, from which they completed secondary education, as a binary variable at the student-level (1=general track, 0=vocational track). The second variant is a three-category variable at the school-level (1=general track schools, 2=general and vocational schools, 0=vocational track schools).

Lastly, gender is a binary variable (1=girl, 0=boy), used as a control in the model. The section below presents the descriptive analysis.

7.4 Descriptive Analysis

The descriptive analysis in this chapter was separated into two subsections. The first one focuses on school characteristics and their incidence on the student cohort social background, performance in SIMCE tests and the transition to higher education. The second one seeks to disentangle differences in higher educational institutions as well as type of higher education programme.

7.4.1 School characteristics

To begin with, I assume that the direct influence of school characteristics in Boudon's model were the schools that the cohort attended in grade 8 (last year of primary education), and grade 12 (last year of secondary education). The schools in grade 8 were the schools in which SIMCE tests were collected, which is assumed to have a direct influence on the mediator of the model. Similarly, students completed their secondary education in grade 12, then decided whether to continue further into higher education the following year; this was assumed to have a direct influence on the outcome of the model.

Table 7.1 School characteristics in grades 8 and 12

School characteristics	Primary schools (grade 8) (col %)	Secondary schools (grade 12) (col %)
School Funding		
Municipal	51.3	27.3
Voucher	42.0	59.1
Private	6.7	13.7
Location area		
Rural	26.8	5.4
Urban	73.2	94.6
Curriculum track school[85]		
Vocational	-	22.7
General and Vocational	-	12.2
General	-	65.2
Total schools	**5,735**	**2,684**

It is apparent from Table 7.1 that schools in grade 12 (2,684) were less than half of the schools found in grade 8 (5,735), which concurs with the Racynski (2011) study of the reduced supply of secondary schools and its effect on students' educational opportunities. By school funding, it can be seen that the majority of the cohort attended municipal schools (51.3%) in grade 8, while the majority attended voucher schools (59.1%) in grade 12. Schools tended to be located in urban areas (over 70%) in both grades, although in grade 12, rural schools were significantly smaller in proportion. Similarly, general track schools were the largest proportion, 65.2%, followed by vocational schools (22.7%) and schools that offer both tracks (12.2%).

To conclude the description of schools, the average number of students per school in grade 8 was 24 students, and the average number of students in grade 12 per school was 50 students. Students that changed school between grade 8 and grade 12 were 65.1% of the 134,595 students that graduated from high school.

[85] At the student level, the curriculum track only refers to students graduated from the vocational or general track.

Students' differences by school characteristics

The following tables examine how students were allocated by the school characteristics, seen above, for the proportion of students that transitioned to higher education, SIMCE scores, and parental educational groups.

In Table 7.2, the proportion of students that make the transition to higher education, by the school characteristics at their last year of secondary education (grade 12), can be seen below. Students completing secondary education by type of school funding show a gradient effect from municipal to private schools, where students from private schools transitioned to higher education in the largest observed proportion (83.5%). The larger proportion of students enrolled in a higher education institution, a year after completing secondary, were from schools located in urban areas (58.6%) rather than rural areas, and students who chose the general curriculum track here in a majority of 70.5% over the rest who had chosen the vocational track.

A comparison of the proportion of students that transitioned to higher education by school characteristics in grade 12 and grade 8 (Table 7.2 and Table 7A.1 in the appendix) yield significant differences ($p<0.001$) by location area and type of school funding, with the exception of private schools ($p>0.05$). The trend is similar across both grades.

Table 7.2 Proportion of students that transitioned to HE by school characteristics in grade 12

School characteristics	Transition to HE (%)	Total
School Funding		
Municipal	49.7	42,029
Voucher	58.2	79,031
Private	83.5	13,535
Location area		
Rural	41.4	3,644
Urban	58.6	130,951
Curriculum track		
Vocational	38.8	52,495
General	70.5	82,100
Total students	**58.1**	**135,495**

The mediator, SIMCE scores, also exhibits a gradient pattern by school funding in grade 12 (Table 7.3). The average students' scores in schools located in urban areas resemble the standardised scores for the cohort (i.e., mean zero and standard deviation of one), while the mean scores from schools in rural areas are below the mean (-0.5). The average performance in SIMCE tests is comparatively higher for students that graduated from the general track (0.29), in contrast to the average performance for students in the vocational track (-0.46).

Table 7.3 Students' SIMCE scores by school characteristics in grade 12

School characteristics	N	Mean	Standard Deviation
School Funding			
Municipal	38,129	-0.29	0.98
Voucher	75,821	-0.01	0.94
Private	12,884	0.90	0.82
Location area			
Rural	3,421	-0.50	0.93
Urban	123,413	0.01	1.00
Curriculum track			
Vocational	48,759	-0.46	0.88
General	78,075	0.29	0.96
Total students	**126,834**[86]	**0.00**	**1.00**

The average performance in SIMCE tests by school characteristics in grade 8 can be seen in the appendix (Table 7A.2). The patterns of average scores by school funding and school location are similar in both grades, although significantly different ($p<0.001$).

Table 7.4 compares the proportion of students by their parents' highest educational attainment and school characteristics in grade 12.

Table 7.4 Proportion of students by parental education and school characteristics in grade 12

Parental ed. (col %) / School grade 12	Primary education or none	Secondary education	Vocational Q.	Degree	Total
School Funding					
Municipal	49.0	32.1	19.2	9.6	31.2
Voucher	50.8	65.8	70.5	45.7	59.2
Private	0.2	2.1	10.4	44.8	9.6
Location area					
Rural	5.2	2.1	1.2	0.6	2.6
Urban	94.8	97.9	98.9	99.4	97.4
Curriculum track					
Vocational	65.5	41.5	18.5	4.2	38.8
General	34.5	58.5	81.5	95.8	61.2
Total students	**32,457**	**52,538**	**16,664**	**19,429**	**121,088**[86]

[86] The total number of students reported in Table 7.3, Table 7.4, and Table 7.5 does not add to the 135,494 who completed secondary education in the compulsory period of 12-years. This is due to missingness in SIMCE scores and parental educational levels.

In the first row of Table 7.4, the proportion of students who attended municipal schools decreases by levels of parental educational attainment (from left to right). An increase in the proportion of students of higher levels of parental educational attainment was found among students that attended private schools. Students that attended voucher schools also followed an increasing pattern through parental educational levels, but only up to advanced vocational qualifications. Then, for parents who had attained a university degree, the share of students in voucher schools (45.7%) was similar to the share in private schools (44.8%).

Over 90% of the students that attended schools in urban areas were from higher groups of parental education than the lowest (primary ed. or none). In contrast, among those students that attended schools in rural areas (5.2%), the largest proportion had parents whose educational level was primary education or none.

By curriculum track, the proportion of students tended to increase through parental educational levels in the vocational track. In contrast to this, students in the general track increased in proportion, as parental educational levels were higher.

Table 7.5 Proportion of students by parental education and school characteristics in grade 8

Parental ed. (col %) / School grade 8	**Primary education or none**	**Secondary education**	**Vocational Q.**	**Degree**	**Total**
School Funding					
Municipal	67.4	42.0	23.4	10.9	41.3
Voucher	32.5	56.3	67.1	45.9	49.7
Private	0.1	1.7	9.6	43.2	9.0
Location area					
Rural	21.4	6.2	3.0	2.2	9.2
Urban	78.6	93.8	97.0	97.8	90.8
Total students	**32,457**	**52,538**	**16,664**	**19,429**	**121,088**[87]

[87] The total number of students reported in Table 7.3, Table 7.4, and Table 7.5 does not add to the 135,494 who completed secondary education in the compulsory period of 12-years. This is due to missingness in SIMCE scores and parental educational levels.

In grade 8 (Table 7.5), the differences in proportions of students between municipal and voucher schools were more noticeable than in grade 12. Nonetheless, the proportions of students in private schools by parental education remained similar to the proportions seen in grade 12. There was a significantly larger proportion of the student cohort in rural schools in grade 8 than that found in grade 12. This was particularly so for students whose parents had primary education or none (21.4%).

Overall, these results suggest that school characteristics in grade 12 and 8 would have a notable impact when we assess the data using Boudon's model, since significant differences in parental educational levels, academic performance, and the transition to higher education were found. Above all other school characteristics, differences between types of school funding seem to be noteworthy.

The next subsection explores performance in SIMCE tests and parental educational levels differences, by type of higher educational institution and higher education programme.

7.4.2 Higher education institutions and type of programme

In the Chilean higher educational system there are two types of organisational features associated with the student cohort data. These are *higher education institutions* and *higher education programmes*, and can be examined to understand their effect in producing inequality of educational opportunities.

The type of higher education institution, which students decided to attend after completing secondary education, were vocational colleges, private universities, and traditional universities, as described in Chapters 1 and 4. The types of higher education programmes offered in these institutions are advanced vocational programmes and undergraduate programmes. Vocational colleges can only offer advanced vocational programmes. While private and traditional universities, to a large extent, offer undergraduate programmes, they can also offer advanced vocational programmes.

A comparison of access to higher education institutions and programme, by performance in SIMCE tests and parental educational attainment (Boudon's model), was made in the following tables.

In Table 7.6, the first row (no-transition) and second to the last row (transition) add to the total students (126,834) who had information of their individual SIMCE scores. Students by higher education institution add to the students who made the transition (74,182). Similarly, students by type of higher education programme add to the students who made the transition to higher education (74,182).

Table 7.6 Students' SIMCE scores by type of transition

Type of transition	N	Mean	Standard Deviation
No-transition to higher education	52,652	-0.36	0.93
Higher education institution			
Vocational college	25,507	-0.40	0.84
Private university	19,808	0.34	0.83
Traditional university	28,867	0.78	0.81
Type of HE programme			
Advanced vocational programme	28,664	-0.37	0.85
Undergraduate programme	45,518	0.65	0.83
Transition to higher education	74,182	0.25	0.97
Total students	**126,834**	**0.00**	**1.00**

The gradient difference in SIMCE average scores is noticeable by higher education organisational features (Table 7.6). Students who enrolled in vocational colleges had an average score below the mean zero (-0.40) in SIMCE tests, while the average score of students enrolled in universities was above the mean of the student cohort. Students enrolled in traditional universities had a noteworthy advantage (0.78), in the average of SIMCE scores, over students enrolled in private universities (0.34). A possible explanation of the aforementioned difference in performance may be related to the requirements of admission in traditional universities (Box 5.1 in Chapter 5), which give greater weights to students' academic performance[88], compared to private universities.

By type of higher education programme, the differences in SIMCE scores are also noticeable. Again, as stated above, it seems that universities admission requirements for undergraduate programmes were related to students' performance in SIMCE scores, for whom the average was well above the mean at 0.65. While the average SIMCE score for students enrolled in advanced vocational programmes was well below the mean (-0.37), as it was for the performance of students enrolled in vocational colleges (-0.40).

Finally, higher education organisational features by parental education and the type of transition (i.e., institution and programme) are shown in Table 7.7. Similar to Tables 7.4 and Table 7.5 in the previous subsection, the proportion of students that did not transition to higher education or enrol in vocational colleges decreased for higher levels of parental educational attainment (from left to right in Table 7.7). In contrast, the proportion of students enrolled in universities increased for higher levels of parental educational attainment.

The same tendencies discovered for levels of parental educational attainment were reproduced for type of higher education programme (Table 7.7). There was a larger proportion of students enrolled in undergraduate programmes, whose parents attained higher educational levels. Conversely, there is a larger proportion of students enrolled in vocational programmes, who had parents with lower educational levels.

[88] The university entrance exam (PSU) was demonstrated to be highly correlated to SIMCE scores (see Chapters 5 and 6).

Table 7.7 Proportion of students by parental education and type of transition

Parental education (col %) / Type of transition	Primary education or none	Secondary education	Vocational Q.	Degree	Total
No-transition to higher education	56.4	42.5	32.4	22.0	41.5
Higher education institution					
Vocational college	26.5	22.9	16.1	6.3	20.3
Private university	6.8	13.7	21.8	28.1	15.3
Traditional university	10.3	21.0	29.7	43.6	22.9
Type of HE programme					
Advanced vocational programme	29.1	25.8	18.9	7.5	22.8
Undergraduate programme	14.5	31.7	48.7	70.4	35.7
Transition to higher education	43.6	57.5	67.6	78.0	58.5
Total students	**32,457**	**52,538**	**16,664**	**19,429**	**121,088**

The following section describes the steps of the multilevel modelling strategy.

7.5 Multilevel modelling Strategy

Cross-classified mediation modelling was used to account for school characteristics simultaneously (Snijders and Bosker, 2012) in grade 8, in direct relation to the mediator (SIMCE scores), and in grade 12, from which schools students graduate and transitioned or not into higher education. The first level of the model was student and the second level was divided between schools in grade 8 and schools in grade 12. See Chapter's 7 methodological appendix in section 7A.2 for more technical details of the modelling strategy.

The cross-classified models were estimated using a Bayesian structural equation model (BSEM). The traditional frequentist approach in statistics treat the model parameters as constants, and by large-sample theory it is assumed that the parameter estimates follow a normal distribution. In contrast, the Bayesian analysis treats the parameters of the model as variables, yielding the whole distribution of the parameters referred as the posterior distribution, which is not assumed to be normal and does not rely on large-sample theory (Asparouhov and Muthén, 2012, p. 314). Furthermore, Bayesian analysis refers to *prior* to the parameters distribution, which can be informative or non-informative. The informative priors can be used from previous studies or any knowledge of the parameter distribution (e.g. pilot studies), this is commonly used in typical Bayesian analyses. However, non-informative priors can also be used in the analysis, which tends to follow a uniform or a normal distribution with a large variance. The large variance purpose is to account for the uncertainty in the parameter estimate (Asparouhov and Muthén, 2012, p. 315).

In this chapter, I employ non-informative priors as there is no published study using data from Chile with a similar focus on the transition to higher education, from where to derive the parameters' priors. Therefore, I rely on Mplus' (Muthén and Muthén, 1998–2017) routines to estimate the cross-classified mediation model under BSEM with non-informative priors. Furthermore, it should be acknowledged that cross-classified mediation models cannot be empirically estimated using the frequentist approach.

The first step was to estimate a separated variance model for the continuous SIMCE scores (mediator) and the binary transition to higher education (response variable) separately. Subsequently, a variance cross-classified model was estimated and compared with the previous models, using intra-class correlations (ICC).

The second step was to set up the cross-classified mediation models, which are random intercept models[89] accounting for school characteristics build upon the variables used in the mediation model from Chapter 6 (i.e. parental education, SIMCE scores and the transition to higher education). The model fit was evaluated using the *posterior predictive checking* providing the posterior predictive p-value (PPP), which measures the distance between the data and the model (Asparouhov and Muthén, 2017)[90]. A model with excellent fit should have a PPP value close to 0.5 (Gelman, 2007; Asparouhov and Muthén, 2012). However, a PPP value greater than 0.05 also reflects a good fit of the model. Likewise, 95% confidence intervals for the difference between the observed and the replicated Chi-square values, associated with the PPP value should be symmetrical and centred on zero to suggest an excellent model fit. I use both PPP and 95% confidence intervals to evaluate model fit in the cross-classified mediation models.

[89] Random intercept models were the models estimated in this chapter. Random slope models were not computationally feasible to estimate due to the small variability of parental educational levels within schools, which was seen in Table 7.4 and Table 7.5 from the previous section. There seems to have been a degree of social segregation within schools, as the literature suggested. However, the random slope models can be investigated in future research.

[90] Cross-classified models were estimated using Bayesian Markov Chain Monte Carlo (MCMC) algorithm with 10,000 iterations as recommended in Snijders and Bosker (2012).

In the third and last step, I derive primary and secondary effects for each the cross-classified models that proved to have a good fit of the data. Finally, I compare the aforementioned effects with the effects found on Chapter 6[91] which did not account for school characteristics and the multilevel design of the data, to assess the change in the effect of the stratification of the schools in determining inequality of educational opportunities in the transition to higher education.

[91] The multiple imputation applied in Chapter 6, to address missingness in parental educational levels and SIMCE scores, is not available to models estimated using a Bayesian estimator in Mplus as the Bayesian estimator derives the posterior distribution of the parameters over the knows and unknowns of the data. However, primary and secondary effects estimated without multiple imputations are not different from those estimated using imputed datasets, as evidenced in Chapter 6. Therefore, all the estimated models had 121,088 students out of the 135,495 students that completed secondary education, in the 12-years period of compulsory education.

7.5.1 Two-level variance model

The first step in the modelling strategy was addressed in this subsection by comparing the variance (null) of two separated two-level multilevel models for the mediator (SIMCE scores) and the outcome (transition to HE), alongside a variance cross-classified mediation model.

Table 7.8 Two-level variance models for the mediator, outcome, and cross-classified mediation model

	SIMCE scores	**Transition to HE**	**Cross-classified mediation model**	
			SIMCE scores	**Transition to HE**
Fixed effects				
Intercept	-0.19 (.01)	0.26 (.01)	-0.12 (.01)	0.22 (.01)
Random effects				
Student-level variance	0.67 (.00)	-	0.60 (.00)	-
School grade 8-level variance	0.29 (.01)	-	0.10 (.00)	0.03 (.00)
School grade 12-level variance	-	0.33 (.01)	0.23 (.01)	0.29 (.01)
Posterior predictive p-value (PPP)	0.508	0.473	0.000	
Difference (Observed-Replicated) Chi-square values	[-7.5, 8.3]	[-7.8, 8.8]	[1402.9, 1613.2]	
Deviance (DIC)	313113	-	-	
Intra-class correlation (ICC)	0.31 (.01)	0.25 (.01)	-	-
ICC-same school G8, different G12	-	-	0.11 (.00)	0.02 (.00)
ICC-different school G8, same G12	-	-	0.24 (.01)	0.22 (.01)
ICC-same school G8, same G12	-	-	0.36 (.01)	0.24 (.01)

Note: All estimates are significant (p<.001). Posterior S.D. in parentheses.
Confidence intervals at 95% level in square brackets.

The first column of Table 7.8 shows the linear model of SIMCE scores accounting for schools in grade 8. Similarly, in the second column, the probit model for the transition to higher education accounts for the schools in grade 12. The third and fourth columns account for schools in both grades in the cross-classified mediation model. To evaluate the goodness of fit of the models posterior predictive p-values (PPP) were used, which can be seen in bottom panel. The deviance information criterion (DIC) is only available for models with continuous dependent variables, in Mplus.

The two-separate variance two-level models for SIMCE scores and the transition indicate a good fit of the models (PPP>0.05). In contrast, the variance cross-classified model fit was poor (PPP<0.05). Likewise, the 95% confidence intervals of the difference of Chi-square values in relation to PPP, for the two separate two-level models, were close to being symmetrical around zero. Confidence intervals for the PPP of the cross-classified model were wide and not around zero. However, the most relevant feature of the variance models is the variance decomposition, by the levels of interest, rather than the overall fit of the model, which will be referred to in the next subsection.

Intra-class correlations (ICC) measure the correlation in the dependent variables of the models within a given school. For the variance model of SIMCE scores (first column in Table 7.8) the ICC is 0.31, which indicates a high variation of SIMCE scores at the school-level in grade 8. The ICC for the variance model of the transition to higher education (second column in Table 7.8) is 0.25[92], which also indicates a high variation of the propensity to make the transition at the school-level in grade 12. There are three different ICC for cross-classified models as can be seen in the bottom panel of the last two columns.

[92] As non-linear models such as probit and logit models have a fixed variance at the student-level, the ICC was calculated assuming a variance of 1 for the probit model of the transition to HE (Muthén and Muthén, 2011).

The first ICC for the cross-classified model examines the correlation in SIMCE scores and the transition, for students who attended the same schools in grade 8, but went to different schools in grade 12. The ICC for SIMCE scores was moderate (0.11), while the propensity of transition was low (0.02), which suggests a lower homogeneity of the students in both variables that shared the same school in grade 8 but completed secondary education in a different school (grade 12). The second ICC examines the correlation in SIMCE scores and the transition, for students who attended different schools in grade 8, but went to the same schools in grade 12. The ICC for SIMCE scores was high (0.24), and for the propensity to transition as well (0.22), which indicates a greater homogeneity in both variables for students that came from different schools in grade 8 but went to the same schools in grade 12. Finally, the third ICC examines the correlation in SIMCE scores and the transition, for students who attended the same schools in grade 8 and grade 12. The ICC for these SIMCE scores was the highest (0.36), the propensity to transition was also higher but of a lower magnitude (0.24), which is evidence that students who attended the same schools in both grades were the most homogeneous in their performance in SIMCE, as well in their propensity to transition to higher education.

The intercepts in SIMCE scores are the average performance in standardised SIMCE scores, which tend to be higher when the cross-classified structure of the school grades was taken into account (-0.12) compared to the two-level model (-0.19). In contrast, the average z-scores in the probit models was lower in the cross-classified model (0.22) compared to the two-level model of the transition (0.26). This difference is marginal in terms of probabilities, 0.59 and 0.60 respectively, in the average probability to make the transition to higher education.

The following subsection describes steps 2 to 4 of the model strategy for the cross-classified mediation model, which were designed to control for school characteristics.

7.5.2 Cross-classified Mediation Model

In the first part of this subsection, the remaining steps of the modelling strategy (steps 2 to 4), followed in order to select the best fitting cross-classified mediation model, are described. In the second part of the subsection, the selected model will be compared against the mediation model derived in Chapter 6, along with a comparison of primary and secondary effects in both models.

The baseline model (model 1) is the mediation model from Chapter 6, which accounts for a cross-classified two-level design of schools and students in grades 8 (mediator), and grade 12 (transition to higher education). The subsequent model (model 2) controlled further for school funding in grades 8 and 12. Then model 3 added a control for school location area in schools in the aforementioned grades. Lastly, model 4 additionally controlled for school curriculum track at the student-level (model 4a) and school-level in grade 12 (model 4b). The estimates of models 1-3 are displayed in Table 7.9 and those on models 4a and 4b in Table 7.10.

The goodness of fit of each model was evaluated using the posterior predictive p-value (PPP). Models 1 to 3 and 4b each evidenced a good fit (PPP>0.05), which can be seen in the bottom panel of Table 7.9 for models 1 to 3 and in Table 7.10 for model 4b. Similarly, their confidence intervals of the difference of Chi-square values in relation to PPP were all nearly symmetrical and around zero. In contrast, model 4a had a poor fit (PPP<0.05) and wide confidence intervals not centred on zero (Table 7.10). The poor fit of model 4a was enough justification to discard this model.

Table 7.9 Cross-classified mediation models

	Model 1		Model 2		Model 3	
	SIMCE scores	**T. to HE**	**SIMCE scores**	**T. to HE**	**SIMCE scores**	**T. to HE**
Fixed effects (Student-level)						
Intercept	-0.21	0.10	-0.54	-0.18	-0.54	-0.14
	(.01)	(.02)	(.02)	(.02)	(.02)	(.02)
Girl	-0.05	-0.01	-0.05	-0.01	-0.05	-0.01
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
SIMCE G8 scores (std.)	-	0.27	-	0.27	-	0.27
		(.01)		(.01)		(.01)
Parental Education						
Primary or none (reference)						
Degree	0.37	0.23	0.32	0.19	0.32	0.18
	(.01)	(.02)	(.01)	(.02)	(.01)	(.02)
Advanced Vocational Q.	0.17	0.20	0.15	0.18	0.15	0.17
	(.01)	(.02)	(.01)	(.01)	(.01)	(.02)
Secondary Education	0.13	0.13	0.11	0.12	0.12	0.11
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Fixed effects (School grade 8-level)						
School Funding						
Municipal (reference)						
Voucher	-	-	0.15	0.12	0.17	0.11
			(.01)	(.01)	(.01)	(.01)
Private	-	-	0.45	0.27	0.47	0.27
			(.03)	(.04)	(.03)	(.04)
Location area						
Urban (reference)						
Rural	-	-	-	-	0.10	-0.05
					(.02)	(.02)
Fixed effects (School grade 12-level)						
School Funding						
Municipal (reference)						
Voucher	-	-	0.25	0.22	0.24	0.21
			(.02)	(.02)	(.02)	(.02)
Private	-	-	0.64	0.70	0.63	0.67
			(.03)	(.05)	(.03)	(.05)
Location area						
Urban (reference)						
Rural	-	-	-	-	-0.35	-0.40
					(.04)	(.05)

Table 7.9 Cross-classified mediation models (continued)

	Model 1		Model 2		Model 3	
	SIMCE scores	**T. to HE**	**SIMCE scores**	**T. to HE**	**SIMCE scores**	**T. to HE**
Random effects						
Student-level variance	0.60	-	0.60	-	0.60	-
	(.00)		(.00)		(.00)	
School grade 8-level variance	0.10	0.03	0.08	0.02	0.08	0.02
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
School grade 12-level variance	0.19	0.25	0.13	0.19	0.12	0.19
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
Posterior predictive p-value (PPP)	0.494		0.479		0.466	
Difference (Obs.-Rep.) Chi-square values	[-18.8, 19.3]		[-21.7, 22.4]		[-22.2, 24.2]	
ICC-same school G8, different G12	0.11	0.02	0.10	0.02	0.10	0.02
	(.00)	(.00)	(.00)	(.00)	(.00)	(.00)
ICC-different school G8, same G12	0.21	0.20	0.16	0.16	0.15	0.15
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
ICC-same school G8, same G12	0.32	0.22	0.26	0.18	0.26	0.17
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
N	**121,088**		**121,088**		**121,088**	

Note: All estimates are significant (p<.001) except 'Girl' in all transition specifications (p>.1). Posterior S.D. in parentheses. Confidence intervals at 95% level in square brackets.

All coefficients were significant (p<0.001) in models 1 to 3 and 4b (Table 7.9 and Table 7.10), except for the coefficient of the dummy variable for girls (1=girl, 0=boy) in the transition specifications. The effects of parental educational levels were larger in model 1 in both specifications of SIMCE G8 scores and the transition than in model 2, 3 and 4b, since model 1 does not control for school characteristics. The coefficients of parental educational levels in SIMCE G8 scores are very similar from each other in models 2, 3 and 4b. However, there is a pattern of diminishing effect in parental education in the aforementioned models in the transition specification. School characteristics in grade 8 also show a similar effect on both specifications in models 2, 3 and 4b, whereas school characteristics in grade 12 are alike in models 2 and 3, the effect of school's curriculum track in model 4b decreases the effect of school funding and rural schools.

The between-school variance in both grades in models 1 to 4b (Table 7.9 and Table 7.10) compared to the variance component model (Table 7.8), decreased significantly by adding individual and school predictors. The largest reduction in between-school variance was in grade 12, in model 4b, which reduced the variance in 59.3% (-0.593 = (0.09-0.23)/0.23) [93] for the SIMCE scores specification, and 63.3% for the transition specification. Therefore, school characteristics in grade 12 such as curriculum track, school funding and schools' location area (urban/rural) explain 59.3% and 63.3% of the school variance in SIMCE scores and the propensity of transition respectively. On the other hand, the school variance in grade 8 remained similar from models 2 to 4b, which are the models that include school-level predictors.

The intra-class correlations (ICC) compared to the variance component model (Table 7.8) also tend to decrease as individual and school related predictors are added to the models. Particularly, for students nested in a common school in grade 12 but different school in grade 8, and for students that shared the same school in both grades (bottom panels in Table 7.9 and Table 7.10). The largest reduction in ICC compared to the variance component model (right panel of Table 7.8) with the rest of the models can be seen in model 4b. For students nested in the same school in grade 12 but different school in grade 8, the ICC is reduced by 57.1% (-0.571 = (0.09-0.22)/0.22) in the transition model, and by 51.2% in the SIMCE scores model. Similarly, for students in the same school in both grades, the ICC is reduced by 52.7% (-0.527 = (0.12-0.24)/0.24) in the transition model, and by 36% in the SIMCE scores model.

[93] The ICCs used in the calculation of the percentage of the reduction had three decimal places.

Table 7.10 Cross-classified mediation models

	Model 4a		Model 4b	
	SIMCE scores	**Transition to HE**	**SIMCE scores**	**Transition to HE**
Fixed effects (Student-level)				
Intercept	-0.55	-0.38	-0.74	-0.51
	(.02)	(.02)	(.02)	(.02)
Girl	-0.05	-0.02	-0.05	-0.01
	(.01)	(.01)	(.01)	(.01)
SIMCE G8 scores (std.)	-	0.26	-	0.27
		(.01)		(.01)
Parental Education				
Primary or none (reference)				
Degree	0.33	0.13	0.32	0.15
	(.01)	(.02)	(.01)	(.02)
Advanced Vocational Q.	0.15	0.13	0.14	0.15
	(.01)	(.02)	(.01)	(.01)
Secondary Education	0.12	0.09	0.11	0.10
	(.01)	(.01)	(.01)	(.01)
Curriculum track grade 12				
Vocational (reference)				
General	-	0.50	-	-
		(.02)		
Fixed effects (School grade 8-level)				
School Funding				
Municipal (reference)				
Voucher	0.17	0.09	0.16	0.09
	(.01)	(.01)	(.01)	(.01)
Private	0.47	0.26	0.46	0.25
	(.03)	(.04)	(.03)	(.04)
Location area				
Urban (reference)				
Rural	0.10	-0.04	0.10	-0.05
	(.02)	(.02)	(.01)	(.02)

Table 7.10 Cross-classified mediation models (continued)

Fixed effects *(School grade 12-level)*				
School Funding				
Municipal (reference)				
Voucher	0.24	0.15	0.15	0.11
	(.02)	(.02)	(.02)	(.02)
Private	0.63	0.46	0.42	0.38
	(.03)	(.04)	(.03)	(.04)
Location area				
Urban (reference)				
Rural	-0.35	-0.26	-0.24	-0.19
	(.04)	(.04)	(.04)	(.04)
Curriculum track				
Vocational (reference)				
General and Vocational	-	-	0.11	0.37
			(.02)	(.03)
General	-	-	0.42	0.69
			(.02)	(.02)
Random effects				
Student-level variance	0.60	-	0.60	-
	(.00)		(.00)	
School grade 8-level variance	0.08	0.02	0.08	0.02
	(.00)	(.00)	(.00)	(.00)
School grade 12-level variance	0.12	0.11	0.09	0.11
	(.01)	(.01)	(.00)	(.01)
Posterior predictive p-value (PPP)	0.000		0.477	
Difference (Obs.-Rep.) Chi-square values	[219.4, 352.8]		[-22.2, 27.1]	
ICC-same school G8, different G12	0.10	0.02	0.11	0.02
	(.00)	(.00)	(.00)	(.00)
ICC-different school G8, same G12	0.15	0.10	0.12	0.09
	(.01)	(.00)	(.00)	(.00)
ICC-same school G8, same G12	0.26	0.12	0.23	0.12
	(.01)	(.00)	(.00)	(.00)
N	**121,088**		**121,088**	

Note: All estimates are significant (p<.001) except 'Girl' in the transition specification of model 4B (p>.1).

Posterior S.D. in parentheses. Confidence intervals at 95% level in square brackets.

The aforementioned reductions on between-school variance and also in intra-class correlations suggest that model 4b should be the preferred model as it explains a larger percentage of school variance in SIMCE scores as well as the propensity of transition. However, following a strict statistical evaluation of model fit in terms of PPP value and 95% confidence intervals of Chi-square values (discussed in section 7.5), model 1 shows to be the best fitting model. As model 1 has the PPP value closest to 0.5 and also has the 95% confidence intervals nearly symmetrical around zero (Table 7.9). Therefore, model 1 should be the selected in statistical terms. Nonetheless, before to reach a final decision of the preferred model an inspection of primary and secondary effects will be seen in Table 7.11.

Table 7.11 Primary and Secondary effects

Parental Education	Model 1		Model 2	
	Primary Effects (%)	**Secondary Effects (%)**	**Primary Effects (%)**	**Secondary Effects (%)**
Degree	30.3	69.7	31.9	68.1
	[27.1, 34.0]	[66.0, 72.9]	[28.2, 36.6]	[63.4, 71.8]
Advanced Vocational Q.	18.7	81.3	18.0	82.0
	[16.0, 21.8]	[78.2, 84.0]	[15.3, 21.4]	[78.6, 84.7]
Secondary Education	21.0	79.0	20.8	79.2
	[18.0, 24.7]	[75.3, 82.0]	[17.7, 24.6]	[75.4, 82.3]
Parental Education	**Model 3**		**Model 4b**	
	Primary Effects (%)	**Secondary Effects (%)**	**Primary Effects (%)**	**Secondary Effects (%)**
Degree	33.2	66.8	36.1	63.9
	[29.2, 38.3]	[61.7, 70.8]	[31.3, 42.1]	[57.9, 68.7]
Advanced Vocational Q.	19.1	80.9	20.6	79.4
	[16.2, 22.9]	[77.1, 83.8]	[17.2, 24.9]	[75.1, 82.8]
Secondary Education	22.4	77.6	23.6	76.4
	[19.1, 27.1]	[72.9, 80.9]	[19.8, 28.5]	[71.5, 80.2]

Note: Primary and secondary effects relative to primary education or none.
Confidence intervals at 95% level in square brackets.

Table 7.11 displays the four best fitting models to this point. It shows primary and secondary effects are very similar for the comparisons of each of the educational levels, i.e., secondary to degree (against the low group, primary or none, as reference). A closer inspection of the corresponding confidence intervals shows that all of them overlap; thereby there is not a significant difference ($p<0.05$) of primary and secondary effects between models 1 to 4b.

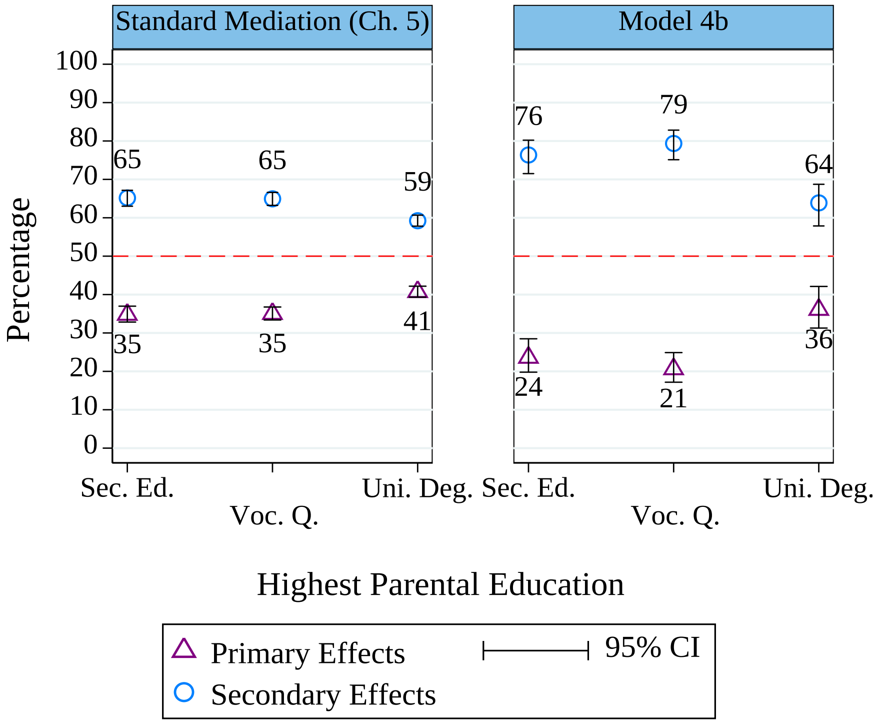
It is, apparent from Table 7.9 and Table 7.10 that the PPP values, decreased in models 2 to 4b compared to model 1. However, all PPP values are over 0.46 which seems to be reasonable to state that models 2 to 4b have a very good fit as those values are still closer to the 0.5 benchmark of excellent fit. Likewise, the 95% confidence intervals in Chi-square values tend to be increasingly wider for models 2 to 4b compared to model 1, but there is not a criterion of model selection based on small differences of how symmetrical the confidence intervals should be to prefer one model over another. In this regard, it goes in line with the aim of this chapter to explain a larger share of school variance and still show a very good fit of the data. The model that serves best the purpose to gain a better understanding of the effects of the stratification of the schools in the Chilean educational system is model 4b, thereby it is the preferred model.

Primary and secondary effects comparison

Figure 7.1 provides a comparison of primary and secondary effects between preferred cross-classified model 4b from the previous analysis, and the standard mediation model applied in Chapter 6 which does not account for the nested design of the data or schools' characteristics. The overall pattern of larger importance of secondary effects in Boudon's model (>50%) remains unchanged in both models. It can be seen that primary effects are smaller in model 4b compared to the model from Chapter 6, and secondary effects are larger in model 4b. The 95% confidence intervals do not overlap for the comparisons of secondary education and vocational qualifications with the low level of parental education, meaning that the differences between both models are significant ($p<0.05$). However, there is not a significant difference ($p<0.05$) in the comparison between the low and high (degree) parental education levels, despite the primary effect observed in model 4b (36%) is smaller compared to the model from Chapter 6 (41%).

Notice that the 95% confidence intervals of primary and secondary effects between the low and high parental educational level in models 1 to 3 (Table 7.11) do not overlap with those from the standard mediation model (Figure 7.1), thereby all primary effects are significantly smaller and all the secondary effects are significantly larger in models 1 to 3. The difference of the aforementioned models and model 4b is the curriculum track variables in grade 12, which seems to compensate for the effects of the other school characteristics (i.e. school funding and schools' location area) that might increase the social background differences between the low and high levels of parental education.

Figure 7.1: Primary and Secondary Effects from the standard mediation model and the preferred cross-classified mediation model 4b



Note: Primary and secondary effects relative to primary ed. or none.

The increasing effect of social background differences by accounting for school characteristics also agrees with Jackson and Jonsson (2013) analysis. They stated that the stratification of the school system will increase the relative importance of secondary effects to the overall inequality of educational opportunities (IEO), which is precisely what the cross-classified models have demonstrated

The next section of the empirical analysis seeks to understand differences in the transition to higher education by type of institution and programme.

7.6 Multinomial Mediation Model

The multinomial mediation model examines changes in primary and secondary effects by type of higher education institution and programme, to be compared with the average effects of the transition to higher education found in Chapter 6.

It was not computational feasible to estimate the cross-classified mediation model, from the previous section, with a multinomial dependent variable, since Bayesian estimation in Mplus does not support this type of dependent variable. Instead, two separated logistic[94] multinomial mediation models were estimated at a single-level; one by type of higher education institution, and the other by type of higher education programme. The results from both of these models can be seen in the appendix (Tables 7A.4 and 7A.5); the focus in this section is, however, the comparison of primary and secondary effects.

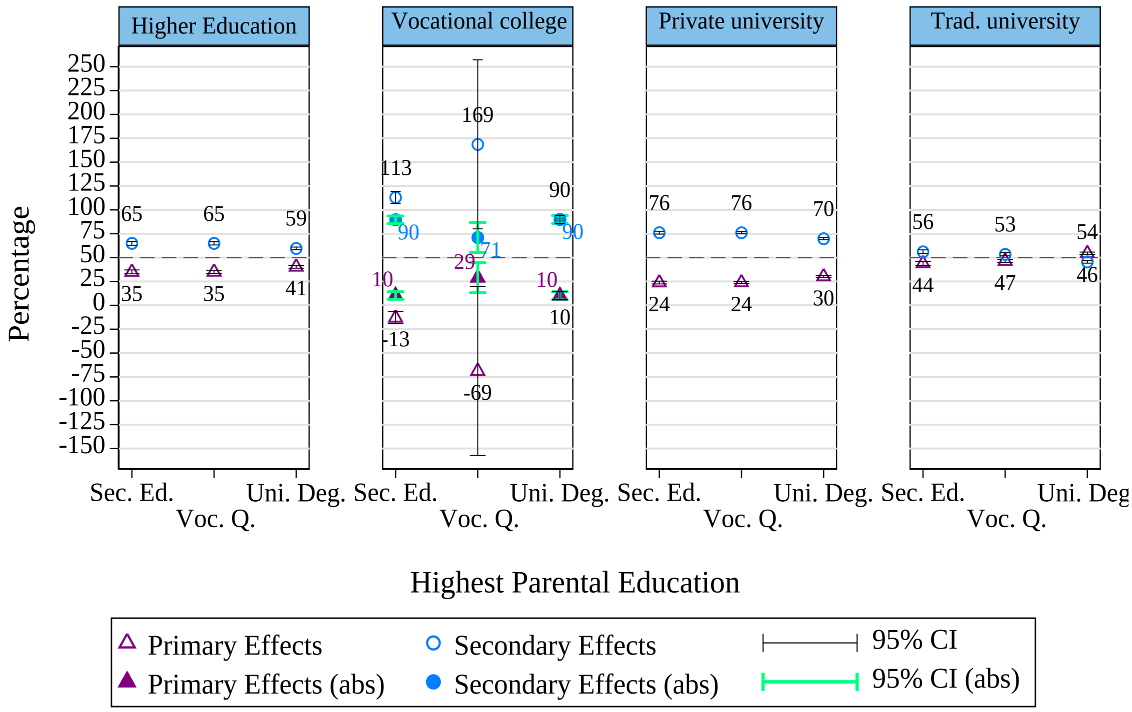
The panels in Figure 7.2 compares primary and secondary effects, when the transition to higher education is treated as a binary variable, the left panel, and when the transition differentiates type of higher education institution, in the remaining panels. The labels of the effects were included due to the long range of the y-axis (-150 to 250), otherwise distinguishing effect magnitudes would be difficult, particularly in the right panel (transition to traditional university).

Before proceeding to examine primary and secondary effects, it is important to specify the proportions enrolled by institution of students included in the estimation of the multinomial model. The proportion of students enrolled in vocational colleges was 20.3% (24,528), 15.3% (18,483) in private universities, and 22.9% (27,778) in traditional universities. These three groups add to 58.5%, i.e., the 70,789 total students enrolled in a

[94] The transition to higher education was estimated using a probit model in Chapter 6. To test if primary and secondary were unchanged by the functional form of the non-linear model, a logit model was estimated for the binary dependent variable of the transition (Table 7A.5 in the appendix), yielding primary and secondary effects of the same magnitude which are reported in the first panel of Figure 7.2.

higher education institution, after completing secondary education. The remaining students who did not make the transition to higher education were 41.5% (50,299). Similarly, the proportion of students enrolled by type of higher education programme was 22.8% (27,619) in vocational programmes, and 35.7% (43,170) in undergraduate programmes.

Figure 7.2: Primary and secondary effects by type of higher education institution



Note: Primary and secondary effects relative to primary ed. or none.
(abs) computed from absolute values of the coefficients in the model.

In Figure 7.2, the black primary effects for students enrolled in vocational colleges are below zero, thereby not significant ($p<0.05$). These primary effects are the product of the negative coefficient of SIMCE scores in the multinomial model, for the transition to vocational college (Table 7A.4 in the appendix). The negative effect of SIMCE scores in the students enrolled in vocational colleges reflects a performance below the mean of the total student cohort (zero). Furthermore, primary effects below zero are also associated with larger secondary effects (over 100%). Therefore, it seems that the transition to vocational colleges might be a case of an *inconsistent* mediation model (MacKinnon, Krull, and Lockwood, 2000 cited in MacKinnon, 2008, p. 83).

The proposed solution in the case of inconsistent mediation is to take the absolute values of the coefficients before computing the proportion mediated (Alwin and Hauser, 1975, cited in MacKinnon, 2008, p. 83). In Chapter 6, primary effects were defined as the proportion mediated by SIMCE scores over the total effect of parental education upon the transition to higher education. Secondary effects, on the other hand, were defined as the proportion of the direct effects of parental education over the total effect.

The blue effects in the second panel of Figure 7.2 reflect primary and secondary effects computed from absolute values. The comparison of parental educational levels of secondary education and vocational qualifications to the low parental educational level (primary or none) seems to be successfully corrected with absolute values, as the effects are in the expected range (0 to 100%), and have narrower confidence intervals. The comparison between university degree and primary or none remains unchanged, as the primary effect was already above zero, applying absolute values yields the same result.

Primary effects for parental educational levels of secondary education and university degree were smaller ($p<0.05$) than the related primary effects in the transition to higher education (first panel of Figure 7.2). Accordingly, the associated secondary effects are larger ($p<0.05$) than those in the transition to higher education, while the effects in the parental educational level of vocational qualifications are not different from those in the transition to higher education, as confidence intervals overlap.

In the third panel of Figure 7.2, the effects for the transition to private universities resembles the pattern of dominant secondary effects (over 50%) seen in the transition to higher education (first panel). However, the secondary effects of the transition to private universities are significantly larger ($p<0.05$), and, correspondingly, primary effects are significantly smaller ($p<0.05$).

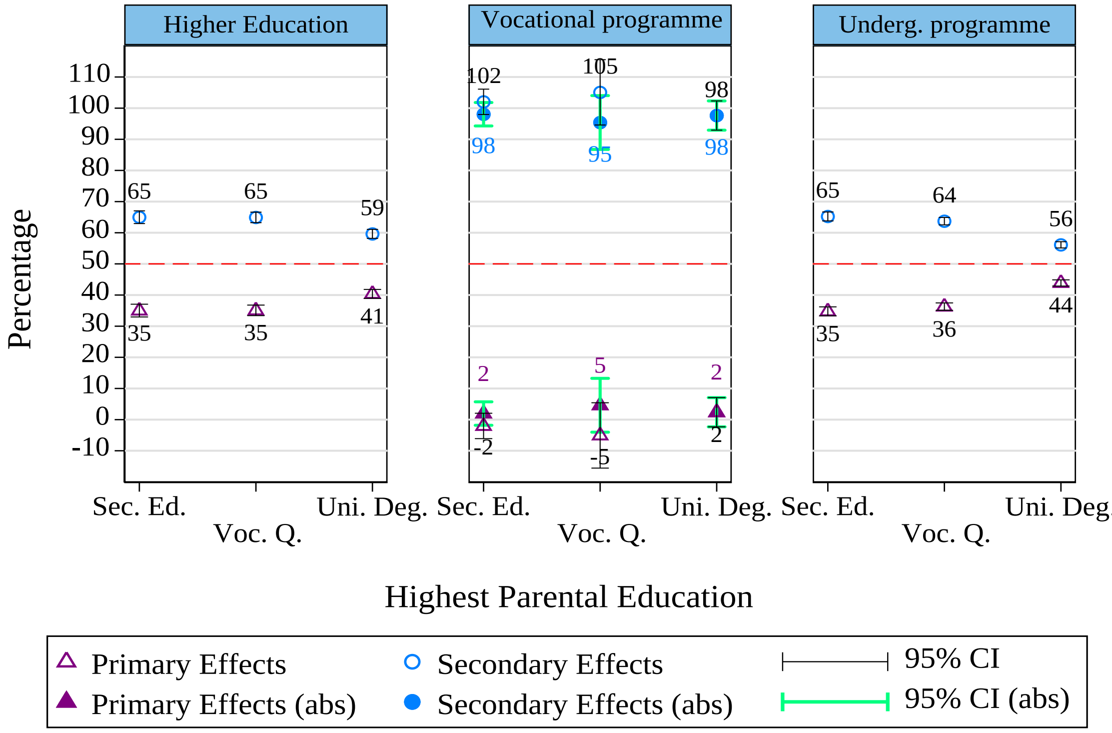
In the last panel of Figure 7.2, the gap between primary and secondary effects in the transition to traditional universities seems to be closing. In effect, the comparison between the high group (university degree) and the low group (primary or none) of parental educational levels shows that the primary effect is the dominant effect (54%). The latter result suggests that academic performance has a greater influence on accounting for inequality of educational opportunities (IEO), between high and low groups of parental education.

In the last part of this section, the transition to higher education by type of programme of enrolment, i.e., vocational and undergraduate programmes, was examined in Figure 7.3.

The second panel of Figure 7.3, the transition to vocational programmes, reasserts the trends found in the transition to vocational colleges. Primary effects, in black, are below zero for the comparison of the first two parental education levels (i.e., secondary and vocational). Moreover, the primary effect for the comparison between the high and low levels of parental education is not significantly different from zero, as its confidence intervals contain zero. The corresponding secondary effects, in black, are larger (over 95%) and have wider confidence intervals. Thus, there is ample evidence that the mediation model for the transition to vocational programmes is inconsistent.

Primary and secondary effects for vocational programme were, therefore, recalculated by their absolute values (blue features, second panel, Figure 7.3), yielding effects in the expected range (0 to 100%) but with their confidence intervals surpassing this range. Therefore, primary and secondary effects are not significant ($p<0.05$), which is also related to the non-significant coefficient of SIMCE scores ($p>0.05$) found in the transition to vocational programmes. (See Table 7A.5 in the appendix).

Figure 7.3: Primary and secondary effects by type of higher education programme



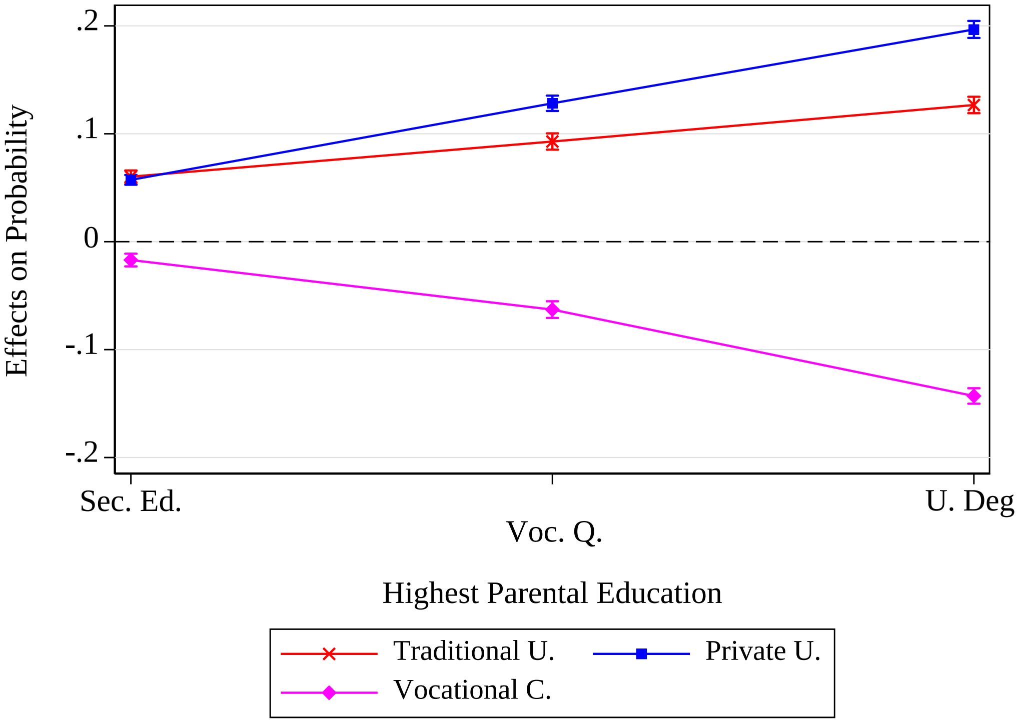
Note: Primary and secondary effects relative to primary ed. or none.
(abs) computed from absolute values of the coefficients in the model.

The effects for the transition to undergraduate programmes (right panel in Figure 7.3) are closer to the average effects for the transition to higher education (left panel). However, similar to the transition to traditional universities, the distance between primary and secondary effects seem to be closing, particularly between the comparison of high and low levels of parental education, although the secondary effect is still the dominant effect ($p<0.05$) in explaining IEO.

To assess *Effectively Maintained Inequality* (EMI) (Lucas, 2001) the average partial effects were computed by parental education groups. It can be seen from both Figure 7.4 and Figure 7.5 that the average change in probability increases as higher the parental education group is for the enrolment in universities as well as for undergraduate programmes. In contrast, the average change in probability decreases for higher groups of parental education. This results follows Lucas (2001) comparison in predicted probabilities to assess EMI.

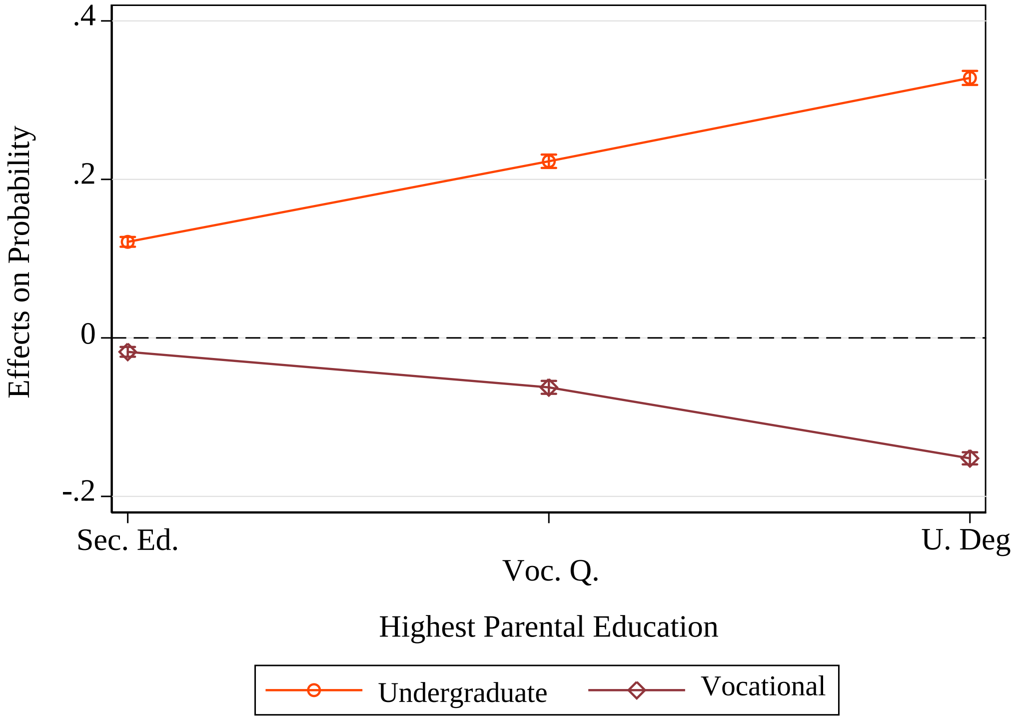
Therefore, there it seems that there is evidence of EMI for the transition to higher education by type of institution and programme.

Figure 7.4: Average Partial Effects with 95% CIs by HE institutions



Note: Relative to primary education or none.

Figure 7.5: Average Partial Effects with 95% CIs by HE programmes



Note: Relative to primary education or none.

A discussion of the findings follows which accounts for the effects of the educational system, and thus addresses the proposed research questions in this chapter.

7.7 Discussion

Prior studies in Boudon's inequality of educational opportunities (IEO) model have considered a country-specific organisation of educational systems as branching points, such as whether or not students transitioned to upper secondary, vocational secondary, A-level, among others (Ress and Azzolini, 2014; Jackson, 2013a; Erikson and Rudolphi, 2010; Jackson *et al.*, 2007; Halsey, Heath and Ridge, 1980). However, the direct influence of schools in a multilevel model setting has not been addressed, to my knowledge of published studies at the time this chapter was written, despite multilevel mediation models being the subject of considerable discussion in the literature (Krull and MacKinnon, 2001; Bauer, Preacher and Gil, 2006; MacKinnon, 2008; Preacher, Zyphur and Zhang, 2010; Preacher, Zhang and Zyphur, 2011; Tofighi, West and MacKinnon, 2013; Tofighi and Thoemmes, 2014; MacKinnon and Valente, 2014). Therefore, primary and secondary effects accounting for school effects, at the end of primary and secondary education, imply a relevant and timely contribution to the literature.

Primary and secondary effects estimated from a multinomial mediation model, on the other hand, is not a novelty in the literature (Sullivan *et al.*, 2014; Holm and Jæger, 2013) of Boudon's model. Also they can be derived using the Karlson, Holm and Breen (2010) method, which was applied in Chapter 5. In this regard, the results from the multinomial model can also contribute to the shedding of some light on decomposing IEO, in scenarios of more than two educational pathways.

The first research question was the comparison of primary and secondary effects from the standard mediation model, derived in Chapter 6, with the cross-classified mediation model seen in this chapter. It was found that secondary effects are significantly ($p<0.05$) larger in the cross-classified mediation model than in the standard mediation model. Therefore, conditional on the schools that students attended in grades 8 and 12, social differences in parental educational levels in the transition to higher education, subject to SIMCE performance, account for at least the 64%[95] of Boudon's IEO model. This result agrees with the Jackson and Jonsson (2013) interpretation of the effect of the stratification of the school system, which they refer to as the vertical differentiation of schools. They argued that secondary effects are expected to increase in relation to the degree of stratification found in the schools. Secondary effects, by definition, are related to the educational choices students and their families make, thereby students whose parents have higher educational attainment will seek to choose the most advantageous school characteristics for their transition to higher education. It is notable, then, that the predominance of secondary effects in the overall IEO was, indeed, enhanced by the influence of school characteristics and the nested design of the data.

It seems that the larger influence of secondary effects in the overall IEO, support the claims of the student movement in 2011, that students' opportunity to access higher education was restricted due to the stratification of the schools, because most socio-economically less advantaged students attended only municipal schools, due to financial constraints. Similarly, their only alternative, in most cases, was to choose the vocational secondary track, as they cannot afford to delay their entrance to the labour market (OECD and World Bank, 2009; Zibechi, 2012; Bellei and Cabalin, 2013; Stromquist and Sanyal,

[95] In the comparison of the low and high parental educational levels of model 4b in Figure 7.1. The comparison of secondary education and vocational qualifications with the low parental education group, secondary effects are over 75%.

2013; Bellei, Cabalin and Orellana, 2014). It is precisely those students coming from municipal schools and/or the vocational track, in the student cohort studied in this thesis, who transitioned to higher education in a lower proportion, performed below the mean in SIMCE tests, and had parents whom tended not to have attained higher educational levels.

A note of caution is due here since the interpretation of the results of the above research question is related to one student cohort, and cannot be extended to conclude that the entire Chilean educational system influences socio-economic less advantaged students. These results simply add new evidence to other related studies.

The second research question examined differences in primary and secondary effects, by type of higher education institution and programme. The transition to vocational colleges and programmes[96] was particularly noteworthy, because primary and secondary effects were divergent from the effects of the transition to higher education. Despite this, the effects seemed to be corrected by establishing absolute values for the transition to vocational colleges. This was not the case in the transition to vocational programmes, due to the non-significant effect of SIMCE scores in the transition specification.

A possible explanation for the disparate effects for the transition to vocational programmes is that the admissions requirements did not consider students' performance in university entrance exams (PSU)[97]. The requirement to access these programmes is to have completed secondary education, (see Box 5.1 in Chapter 5 for more details). Therefore, primary and secondary effects in Boudon's model cannot account for IEO in this particular transition.

[96] Let us recall that universities can also offer vocational programmes.
[97] Academic performance in university entrance exams is the PSU. However, as it was seen in Chapter 5 and 6, SIMCE scores are highly correlated to PSU scores and they yield similar primary and secondary effects, after accounting for sample selection.

The pattern of predominant secondary effects explaining IEO was reversed in the transition by higher education institutions and undergraduate programmes. The only exception was found in the transition to traditional universities, between the high and low groups of parental education, in which primary effects were over 50%. Moreover, the gap between primary and secondary effects was significantly reduced in the transition to traditional universities. This result may also be explained by the fact that traditional universities weigh PSU scores at around 50% of the academic performance requirements of applications (as seen in Box 5.1, in Chapter 5).

The last research question in this chapter sought to explore an alternative interpretation to the effect of social background on the transition to higher education institutions (HEIs) and programmes. Tentatively, these results might suggest that *Effectively Maintained Inequality* (EMI) was in operation[98], as qualitative inequality in HEIs and programmes was evidenced (Lucas, 2001). Torche (2005) also stated the presence of EMI for older cohorts[99] (1936 to 1976) in Chile, by examining differences in educational attainment by type of school funding (i.e. municipal, voucher and private).

[98] Following Lucas (2001, p. 1674), it seems that maximally maintained inequality (MMI) was also in operation for the student cohort, as the effect of parental educational levels were significantly larger in the transition to higher education, than they were to the completion of secondary education. Therefore, yielding significantly (>50%) larger primary effects, in Boudon's model, for secondary completion, in contrast to the smaller primary effects (<50%) found for the transition to higher education. Torche (2005) challenged the MMI theory for Chile.

[99] Only male heads of households were sampled.

7.8 Conclusion

The aim of this chapter was to examine the effect of the Chilean educational system in explaining inequality of educational opportunities, using Boudon's decomposition of primary and secondary effects. The influence of the educational system was separated in two key areas. The first area addressed the school system, and the second area the higher education system. Notwithstanding, it was not possible to combine both, the findings are informative of how educational inequalities unfold in the Chilean education system.

The main finding in the school system is that social differences among schools seem to determine students' chances to continue further in the educational system. In the higher education system, the same tendency was observed but with a lesser influence of secondary effects on the transition to undergraduate programmes and private universities. However, the transition to traditional universities showed a more balanced pattern between primary and secondary effects. Even primary effects turned out to explain a greater proportion of IEO between students of high and low backgrounds of parental education. In contrast, the transition to vocational programmes might suggest that the overall IEO was explained by students' social background differences, regardless of their academic performance.

The findings of this chapter seem to support the claims of many studies (mentioned in the introduction) that the stratification of the schools produces social segregation. It also contributes to the understanding of differences among the types of higher education enrolment, not frequently studied with Boudon's IEO model framework. These findings add to the evidence, from the studies of the Chilean education system that suggests that policy-makers should continue to seek to improve fairness and equality of access to all levels of the educational system, particularly for students from less advantaged backgrounds.

Chapter 8: Conclusion

8.1 Introduction

This thesis aimed to examine social-background differences in Chilean students continuing to higher education, and to evaluate the applicability of Boudon’s inequality of educational opportunities (IEO) model (Boudon, 1974). It has also assessed *Maximally Maintained Inequality* (MMI) (Raftery and Hout, 1993; Hout, Raftery and Bell, 1993) and *Effectively Maintained Inequality* (EMI) (Lucas, 2001) to explain the association of social background in the transition to higher education. This was done by following a student cohort from grade one of primary education up to secondary education completion, in the 12-year period of compulsory education, and their subsequent enrolment in higher education. The decomposition of IEO into primary and secondary effects was not estimated in Chile before the time this thesis was written. Therefore, this thesis fills important gaps of knowledge in social inequalities in educational attainment for the country.

One important finding in this thesis is the larger contribution of secondary effects in producing educational inequalities in the transition to higher education. As a reminder, secondary effects are social background differences in educational decisions conditional on academic performance, while primary effects are social background differences in academic performance. The magnitude of secondary effects increased by accounting for the stratification of the schools, by school characteristics, and the transition to vocational colleges/programmes or private universities. In contrast, the magnitude of secondary effects decreased, as a result of the increment on primary effects, in the transition to traditional universities. Smaller secondary effects for the completion to secondary education compared with the larger secondary effects for the transition to higher education is consistent with MMI theory, following Lucas (2001) interpretation of MMI across transitions. On the other hand, other findings are consistent with EMI theory; these are the higher likelihood of students from lower background of parental education having enrolled in vocational colleges/programmes, and the higher likelihood of students with higher background of parental education having enrolled in universities or undergraduate programmes.

The next section discusses the *Contribution to Knowledge* of this thesis, followed by the *Summary of Findings*, *Strengths and Limitations*, *Future Research* and the chapter concludes with the *Implications for Policy*.

8.2 Contribution to Knowledge

This thesis contributes to the understanding of social inequalities in educational attainment in the context of emerging economies such as Chile. The educational system in Chile, as in other countries, has experienced a significant expansion that has increased the average educational attainment in the population. Notwithstanding this general expansion, students' chances to continue in education remains unequally distributed, as evidenced by the secondary effects found, using Boudon's IEO model, that account for the larger extent of the overall inequality in the transition to higher education. This result agrees with the tendency of the larger relative importance of secondary effects found amongst western countries for the transition to university (Jackson, 2013a). It also agrees with the institutional context in Chile where the expansion of the system has been driven by the participation of private institutions, which fall under weaker statutory regulation than traditional universities, but, nonetheless, receive state funding based on student enrolment (Stromquist and Sanyal, 2013; Bellei, Cabalin and Orellana, 2014).

This thesis contests the recommendation in the literature about preferring grades to standardised tests to properly estimate Boudon's IEO model, owing to the strict interpretation of secondary effects that require that students make informed educational decisions (Jackson and Jonsson, 2013). Grades, however, suffer from not being a comparable measure of academic performance across schools, since schools have different internal policies of assessment. Moreover, standardised test results are not entirely unknown to students as the Ministry of Education itself promotes the widespread use of practice tests, for which students routinely learn of their individual scores. Jackson *et al.* (2007) and Erikson and Rudolphi (2010) emphasised that secondary effects will be overestimated when using standardised tests, but this was not supported in my findings; in contrast, the secondary effects were larger when using grades than when using standardised tests. In the same studies, it was argued that anticipatory decisions might cause an underestimation of secondary effects when using academic performance measures that were recorded close to the transition of interest. In both grades and standardised tests, the secondary effects contributed to the larger share (>50%) of the overall IEO, thereby there was no evidence to support the anticipatory decisions thesis.

Finally, based on the parallel strand in the literature of educational transition models, this thesis was able to account for sample-selection techniques to control for the attrition in the student cohort that occurred during the period of compulsory education. Moreover, the proposed hypotheses, in this strand of the literature, that aimed to explain the effect of social background in the process of educational attainment, namely *Maximally Maintained Inequality* (MMI) and *Effectively Maintained Inequality* (EMI), accordingly accounted for differences in social background between secondary education completion and the transition to higher education (MMI) by type of institution and programme (EMI), in this thesis. The mediation modelling setting, as an alternative to the *ad hoc* methods developed in the literature to empirically derive primary and secondary effects, allowed the implicit modelling of the effect of the stratification of the schools, as well as controlling for sample-selection. These latter empirical strategies have not previously been explored in the literature.

8.3 Summary of Findings

The first aim of the thesis was to find a suitable measure of academic performance to derive primary and secondary effects of the inequality of educational opportunities (IEO) in the transition to higher education. The literature (Jackson *et al.*, 2007; Erikson and Rudolphi, 2010) suggests a preference for grade point averages over standardised tests, when both measures are available. However, these recommendations might be influenced by the design and the purpose of standardised tests in certain countries (i.e., England and Sweden). In Chile, national standardised tests (SIMCE) are an instrument of public policy which function as reward programmes to incentive schools to increase their average performance in SIMCE tests (Mizala and Torche, 2012; Mizala and Urquiola, 2013). Consequently, primary effects using SIMCE tests were, indeed, capable of capturing larger differences between parental education groups than were GPA levels, notwithstanding that overall inequality is driven by secondary effects in both measures, as was seen in Chapter 5.

There was no-evidence of anticipatory decisions in the student cohort, when the average of the four years of secondary education[100] in GPA levels was employed as the measure of academic performance. The aforementioned studies in England and Sweden found that academic performance measures close to the educational transition of interest underestimated secondary effects. As referred to above, secondary effects in GPA levels were above the 50% benchmark, implying that they account for the larger proportion of the overall IEO. Similarly, secondary effects derived using SIMCE tests collected in grade 8,

[100] That is from grade 9 to grade 12.

before the four years of secondary education, were above the 50% benchmark of the overall IEO.

The second aim of the thesis was to arrive at an acceptable educational classification for Chile. To this end, I followed the Jackson and Jonsson (2013) classification of educational systems at the secondary and upper secondary level, in terms of stratification (vertical differentiation) and selectivity (choice). I classified, the Chilean school system[101] as highly stratified and of an intermediate level of selectivity. The stratification of the Chilean educational system is higher due to school differences in the type of school funding (i.e., municipal, voucher and private), reflecting differences in family budgets and whether they afford a voucher school or a private school. Municipal schools[102], on the other hand, are free, but they tend to be of low status and low performance in standardised tests (SIMCE, PSU). At the time the student cohort attended compulsory education, schools were allowed to select students, however, the selectivity was not a general measure imposed from an early age, as in Germany and the Netherlands, which were classified as highly selective systems. The Chilean cohort's secondary effects were, thus, ranked between Germany and Sweden, which belong to the group of countries with predominant secondary effects (>50%). The classification of the German system (i.e., highly stratified and selective) and the Swedish system (i.e., intermediate level of stratification and selectivity), respectively agree with the classification of the Chilean system.

The third aim was to investigate the effect of selection in the group of students in the cohort that completed secondary education in the 12-years period of compulsory education. From the description of the student cohort in Chapter 3, it can be seen that only the 64% of

[101] I refer to the school system as schools that offer primary and secondary education, whereas the Jackson and Jonsson (2013) classification is for secondary or upper secondary level. I refer to the education system when higher education is also considered along with primary and secondary education.
[102] At the time the student cohort attended compulsory education, there was a group of no more than 30 municipal schools of high status and high performance in standardised tests. For more details see Allende González and Valenzuela (2016).

the students that started grade 1 in 2002 finished school in 2013. Grade repetition and dropout rates gradually increased as promotion rates gradually decreased for the period. To address sample-selection in the student cohort, secondary education completion was included as a separate model in the mediation-modelling framework proposed in Chapter 6. To identify the sample-selection specification, which is secondary education completion, three instruments were linked to the region of students' residency in the cohort. The proportion of boys and girls in the age group 15–19 who participate in the labour force was considered along with the proportion of girls that became mother for the same age group. These measures were calculated from Chile's household survey (CASEN, 2011), which is the year the promotion rate sharply decreased in the cohort. Based on a government report (Ministerio de Desarrollo Social, 2013), financial and maternity reasons were among the most frequent reasons given for dropping out of school, and can be approximated using the CASEN survey.

The parental education coefficients were smaller in the transition to higher education when sample-selection was accounted for, compared to the parental education coefficients in the specification that ignored sample-selection. However, primary and secondary effects derived from indirect and direct effects of the mediation model, respectively, were not significantly different between the aforementioned specifications. Furthermore, secondary completion average partial effects (APEs) that were conditional on sample-selection were not significantly different from the APEs in the transition specification that did not account for sample-selection.

Further robustness checks of the model included controlling for missingness in parental education and SIMCE scores in grade 8, potential moderating effects (i.e., interaction terms), and bootstrapped standard errors for the indirect effect of SIMCE scores; none of these significantly altered the contribution of primary and secondary effects to the overall IEO found in Chapter 6. Therefore, based on these findings, the student cohort's chances of progressing to higher education were driven, to a large extent, by secondary effects. In other words, differences in parental education groups, despite performance in SIMCE tests, mattered most for student's access to higher education institution a year after finishing school.

Conversely, social background differences in the performance of SIMCE tests accounted for a larger contribution of the overall IEO to complete secondary education, that is, larger primary effects (>50%) than secondary effects. Completion of secondary education in Chile seems nearly to be saturated for the socioeconomically advantaged groups, as the results from the model and trends by income quartiles from household surveys have shown (Figure 6.2). Moreover, the enrolment in secondary education seems to have been decreasing since 2007 (Figure 6.3), in contrast to the sustained growth in higher education enrolment (Figure 4.1). Therefore, it seems to be enough evidence to support *Maximally Maintained Inequality* (MMI) (Raftery and Hout, 1993; Hout, Raftery and Bell, 1993) between secondary education completion and higher education transition for the student cohort.

The last two aims of this thesis were pursued in Chapter 7, in relation to the influence of the Chilean education system on increasing or decreasing the effect of social background on the transition to higher education. The effect of the education system was separated in two parts. First, the stratification of the schools was examined in terms of type of school funding (i.e., municipal, voucher and private), location area (i.e., urban or rural), and curriculum track (i.e., general and vocational). Second, the transition to higher education distinguished between institutions (i.e., traditional universities, private universities and vocational colleges) and programmes (i.e., vocational and undergraduate academic).

The stratification of the schools increased the relative importance of secondary effects by approximately 10% for the transition to higher education (when using the low groups of parental education as the reference group). This result agrees with the Jackson and Jonsson (2013) interpretation of the effect of the stratification in the system (which they refer to as the vertical differentiation of school systems). They argued that secondary effects are expected to increase in relation to the degree of stratification of the system. Secondary effects, by definition, are related to the educational choices students and their families make, thereby students whose parents have higher educational attainment will seek to choose the most advantageous school characteristics for their transition to higher education. It is notable, then, that the predominance of secondary effects in the overall IEO was, indeed, enhanced by the influence of the stratification in the Chilean schools.

The distinctions of the transition to higher education by type of institution and by type of programme resulted in an increment of the relative importance of secondary effects in vocational colleges and in vocational programmes, while primary effects increased or even surpassed the 50% benchmark for students enrolling in traditional universities. Rudolphi (2013), following Breen and Goldthorpe (1997), noticed that primary effects will tend to increase in prestigious higher education pathways, since advantaged social groups seek to preserve their advantage by selecting them. Indeed, the admission requirements for traditional (prestigious) universities mostly rely on students' academic performance in university entrance exams (PSU) (see Box 5.1, in Chapter 5), which are highly correlated with the SIMCE tests used in the analyses of this thesis. Conversely, secondary effects explain over 90% of the overall IEO in students accessing vocational programmes and at least 70% of vocational college enrolment. The latter result, again, is a consequence of the admission requirements to those institutions, since only the certificate of secondary education completion is required.

To assess *Effectively Maintained Inequality* (EMI) (Lucas, 2001) for the transition to higher education institutions and programmes, I calculated average partial effects by parental education groups. The average change in probability of enrolment in universities (traditional or private) increases with increasing parental education background, and similarly when enrolment to programmes were compared, in undergraduate programmes. In contrast, the average change in probability of enrolment in vocational colleges decreases with increasing levels of parental education background, and similarly with enrolment in vocational programmes. However, the average change in probability of enrolment in private universities is higher than the enrolment in traditional universities for parents holding higher education credentials. The reason of this difference seems to be related, once more, to the effect of the importance of academic performance in the admission requirements of traditional universities, which would account for an enhanced proportion of primary effects, or even the larger overall proportion of IEO being attributable to primary effects. Lucas (2001) argued that advantaged social groups seek to position themselves in high-status types of education *within* elite educational institutions, which alludes to the horizontal dimension of educational inequality and considers differences in educational pathways (in contrast to the vertical dimension, i.e., stratification of institutions). These results suggest that EMI might be in operation in the transition to higher education institutions and programmes.

8.4 Strengths and Limitations

This thesis benefited from the availability of administrative data from the national student register, which was close to having census data of the population in the age group of compulsory education. The large sample size of the cohort seems to have contributed to non-significant differences when robustness checks were conducted, such as addressing missingness through multiple imputation methods. Moreover, the structure of the data accounted for the school system in Chile, which allowed me to model the effect of the stratification of schools and their influence in Boudon's IEO model. Connecting this data structure with information about students' destinations in higher education enrolment, also enabled me to disentangle differences of the relative contribution of primary and secondary effects to the overall IEO.

The structural equation modelling framework of the mediation model was an appropriate method to assess Boudon's IEO model, as an alternative to the *ad hoc* methods developed in the literature (Erikson *et al.*, 2005). Furthermore, it allowed me to investigate the effect of sample-selection in the student cohort and the stratification of the schools and higher education institutions. Previous studies acknowledged the influence of sample-selection and the education system in creating inequalities of educational attainment, but these were not empirically tested using Boudon's decomposition of primary and secondary effects (Jackson, 2013a).

First, the major limitation of this thesis was not to be able to compare across other student cohorts, in order to evaluate trends of social inequality in educational attainment. This is important to understand how educational inequalities might have been influenced by the expansion of the education system, which Chile and other nations have experienced in the 20th Century.

Second, the estimation of random slopes in the cross-classified model of Chapter 7 was not computationally feasible, given the reduced variability of parental education groups within schools, since parental education is a categorical variable. This issue might be addressed in future research by aggregating the proportion of parental education groups within schools as school-level covariates. While it seems likely that random slope models would recover a greater magnitude of secondary effects than those found by the random intercept models, the interpretation that the stratification of the schools increases the relative importance of secondary effects in the overall IEO will be unaffected.

Finally, modelling school characteristics and higher education destinations in one model was not computationally feasible either, as it was not possible for the restriction of multinomial outcomes in cross-classified models in Mplus (Muthén and Muthén, 1998–2017). It would be of interest for policy implications to understand the connection of school and higher education institutions in determining educational inequalities. This limitation must also be left for future research.

8.5 Future Research

This thesis was based on Boudon's IEO model (Boudon, 1974), in which the mediating mechanism between social background and the transition to higher education was the academic performance. Other mediating mechanisms can also be explored that are insightful for the context of the Chilean education system. One could be students' financial aid through scholarships and student loans. In Chapters 1 and 4, it was argued that higher tuition fees and higher interest's rates in university loans triggered students' protests in 2011. Since then, the government has increased its efforts to secure fair access to higher education, in particular for socioeconomically less advantaged students (Stromquist and Sanyal, 2013; Bellei, Cabalin and Orellana, 2014). Furthermore, linking the data of the student cohort of this thesis with information of students' financial aid, in order to compare it with subsequent student cohorts, could be of potential use in the evaluation of policies implemented since 2011.

Students' trajectories through the 12-years of compulsory education could, alternatively, be explored from the life course approach. The Pallas (2003) review of various studies supports the explanation that initial inequalities shape the individual life course. He emphasised the hypothesis of the *Matthew effect* (Merton 1968, cited in Pallas, 2003, p. 174), which states that those that are rich are likely to get richer, and those who are poor are likely to get poorer. In terms of educational inequalities, the initial location of students in the educational system will be influential to their subsequent educational attainment. In addition to the *Matthew effect*, other similar hypotheses, in this strand of the literature, have been proposed, such as the *cumulative dis/advantage* hypothesis. Guidici and Pallas (2014) tested the *cumulative dis/advantage* hypothesis for the transition out of high school, where students from advantaged social backgrounds are more likely to attend prestigious post-secondary institutions, in contrast to their peers from less advantaged social backgrounds. This hypothesis, also assumes that initial inequalities are determinant to social positions later in the life course.

Following the line of reasoning of the life course literature, Valenzuela B. *et al.* (2013) have shown that early age academic performance in Chilean students is an important determinant of students' performance in secondary education. In the case of the student cohort in this thesis the high correlation of SIMCE scores and university entrance exams (PSU) suggested that the hypotheses of the life course literature ought, indeed, be investigated.

8.6 Implications for Policy

This thesis showed that secondary effects are the prevailing force in generating inequalities for the transition to higher education. However, for the transition to traditional universities the gap between primary and secondary effects were narrowed, even reversed in the comparison of low and high groups of parental education (i.e., primary effects>50%). The caveat is that students who attended private or voucher schools are more likely to transition to traditional universities, as they are also more likely to reach the academic requirements[103] to access those institutions, than their peers who attended socioeconomically less advantaged schools. These findings add to the current discussion of alternative mechanisms to access traditional universities and the eight private universities, which follow the same admission policy[104].

The aforementioned scenario of less opportunities for students attending socioeconomically less advantaged schools has not gone unnoticed by scholars and policy-makers in Chile (Rosas and Santa Cruz, 2013). In response to this awareness, two chief policy ideas have been proposed and implemented. One concerns the alteration of the academic admission requirements of traditional universities[105], where students' ranking relative to their peers in GPA levels is also included. This was designed to allow high-performance students from socioeconomically less advantaged schools (with low averages in university entrance exams, the PSU, and low GPA levels) the ability to access these universities. The studies of Larroucau, Ríos and Mizala (2013) and Larroucau (2014) found that the effect of this 'ranking' has actually been effective in increasing student access to these universities. This measure has been applied since 2012[106], which implies that the

[103] See Box 5.1 in Chapter 5 for more details of the admission process for the year after students in the cohort completed secondary education.
[104] *Ibid.*
[105] *Ibid.*
[106] The weight of the 'ranking' in the overall score of application to university was increased gradually since 2012.

students in the cohort benefited from it. The effects of the ranking might explain the decrease in the relative importance of secondary effects for the transition to traditional universities[107].

The other measure of alternative access to traditional universities, currently underway but still at an early phase, is one of the first education policies applied under the current government. The programme PACE (*Programa de Acompañamiento y Acceso Efectivo*) prepares students from less advantaged backgrounds during the last two years of secondary education, to achieve the academic requirements to access traditional universities. The number of students benefitted from this programme has continuously grown since its implementation in 2015.

These government efforts to tackle fairness of access to higher education are to be assessed in the following years.

[107] The eight private universities that follow the same admissions policy as traditional universities were not distinguished among the group of private universities reported in Chapter 7.

References

Allende González, C. and Valenzuela, J. P. (2016). 'Efectividad de los liceos públicos de excelencia en Chile', Universidad de Chile, Centro de Investigación Avanzada en Educación, Documentos de Trabajo, No. 20.

Anand, P., Mizala, A. and Repetto, A. (2009). 'Using school scholarships to estimate the effect of private education on the academic achievement of low-income students in Chile', *Economics of Education Review*, 28(3), pp. 370–381, [Online]. Available at: https://doi.org/10.1016/j.econedurev.2008.03.005 (Accessed: 11 June 2014).

Arum, R., Gamoran, A. and Shavit, Y. (2007). More inclusion than diversion: Expansion, differentiation, and market structure in higher education, in Shavit, Y., Arum, R. and Gamoran, A. (eds.) *Stratification in higher education: A comparative study*. Stanford, California: Standford University Press, pp. 1–35.

Asparouhov, T. and Muthén, B. (2010). 'Multiple imputation with Mplus', Technical note. Version 2. Unpublished paper. Available at: www.statmodel.com (Accessed: 03 February 2016).

Asparouhov, T. and Muthén, B. (2012). 'Bayesian structural equation modeling: A more flexible representation of substantive theory', *Psychological Methods*, 17(3), pp. 313–335, [Online]. Available at: http://dx.doi.org/10.1037/a0026802 (Accessed: 20 January 2018).

Asparouhov, T. and Muthén, B. (2017). 'Prior-posterior predictive p-values', Mplus Web Notes: No. 22. Version 2. Unpublished paper. Available at: www.statmodel.com (Accessed: 24 June 2017).

Bartlett, J. (2013) 'Lecture 1 - issues raised by missing data, a systematic approach, and missingness mechanisms', *Statistical Analysis with Missing Data using Multiple Imputation*. University of Manchester. Unpublished.

Bauer, D. J., Preacher, K. J. and Gil, K. M. (2006). 'Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations', *Psychological Methods*, 11(2), pp. 142–163, [Online]. Available at: http://dx.doi.org/10.1037/1082-989X.11.2.142 (Accessed: 9 July 2017).

Bell, D. (1973). *The coming of post-industrial society: A venture in social forecasting*. Reprint, New York: Basic Books, 1999.

Bellei, C. (2013). 'El estudio de la segregación socioeconómica y académica de la educación chilena', *Estudios pedagógicos (Valdivia)*, 39(1), pp. 325–345, [Online]. Available at: http://dx.doi.org/10.4067/s0718-07052013000100019 (Accessed: 13 June 2017).

Bellei, C. and Cabalin, C. (2013). 'Chilean student movements: Sustained struggle to transform a market-oriented educational system', *Issues in Comparative Education*, 15(2), pp. 108–123, [Online]. Available at: http://www.tc.columbia.edu/cice/pdf/28175_15_02_Bellei_Cabalin.pdf (Accessed: 19 March 2014).

Bellei, C., Cabalin, C. and Orellana, V. (2014). 'The 2011 Chilean student movement against neoliberal educational policies', *Studies in Higher Education*, 39(3), pp. 426–440, [Online]. Available at: http://dx.doi.org/10.1080/03075079.2014.896179 (Accessed: 21 November 2015).

Bernasconi, A. and Rojas, F. (2003).[IESALC, UNESCO Reports] Informe sobre la educación superior en Chile: 1980-2003 – August 2003. *DIGITAL OBSERVATORY FOR HIGHER EDUCATION IN LATIN AMERICA AND THE CARIBBEAN*

IES/2003/ED/PI/55 [Online]. Available at: http://www.iesalc.unesco.org.ve/ (Accessed: 08 February 2017).

Blau, P. M. and Duncan, O. D. (1967). *The American occupational structure*. New York: John Wiley & Sons, Inc.

Boudon, R. (1974). *Education, opportunity, and social inequality: Changing prospects in Western society*. New York: Wiley.

Boudon, R. (1976). 'Comment on Hauser's review of Education, Opportunity and Social Inequality', *American Journal of Sociology*, 81(5), pp. 1175–1187, [Online]. Available at: http://www.jstor.org/stable/2777566 (Accessed: 20 June 2015).

Bourdieu, P. and Passeron, J. C. (1977). *Reproduction in education, society and culture*. Reprint, Worcester: Sage Publications, 1990.

Breen, R. and Goldthorpe, J. H. (1997). 'Explaining educational differentials towards a formal rational action theory', *Rationality and Society*, 9(3), pp. 275–305, [Online]. Available at: http://dx.doi.org/10.1177/104346397009003002 (Accessed: 28 October 2013).

Breen, R. and Jonsson, J. O. (2005). 'Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility', *Annual Review of Sociology*, 31(1), pp. 223–243, [Online]. Available at: http://dx.doi.org/10.1146/annurev.soc.31.041304.122232 (Accessed: 13 April 2012).

Breen, R., Luijkx, R., Müller, W. and Pollak, R. (2009). 'Nonpersistent inequality in educational attainment: Evidence from eight European countries', *American Journal of Sociology*, 114(5), pp. 1475–1521, [Online]. Available at: http://www.jstor.org/stable/10.1086/595951 (Accessed: 24 March 2012).

Breen, R. (2010a). 'Social mobility and equality of opportunity', *The Economic and Social Review*, 41(4), pp. 413–428, [Online]. Available at: http://hdl.handle.net/2262/58542 (Accessed: 14 August 2017).

Breen, R. (2010b). 'Educational expansion and social mobility in the 20th Century', *Social Forces*, 89(2), pp. 365–388, [Online]. Available at: http://www.jstor.org/stable/40984537 . (Accessed: 05 February 2014).

Brint, S. and Karabel, J. (1989). *The diverted dream: Community colleges and the promise of educational opportunity in America, 1900-1985*. New York: Oxford University Press.

Buis, M. L. (2010). 'Direct and indirect effects in a logit model', *The Stata Journal*, 10(1), pp. 11–29.

Bulle, N. (2009). The actuality of education, opportunity and social inequality, in Cherkaoui, M. and Hamilton, P. (eds.) *Raymond Boudon: A life in sociology*. London: Bardwell Press, pp. 357–378.

Cameron, S. V. and Heckman, J. J. (1998). 'Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males', *Journal of Political Economy*, 106(2), pp. 262–333, [Online]. Available at: http://www.jstor.org/stable/10.1086/250010 (Accessed: 22 January 2014).

CASEN: Bases de datos Encuesta de Caracterización Socioeconómica Nacional (2003). Observatorio Social – Ministerio de Desarrollo Social. Santiago, Chile.

CASEN: Bases de datos Encuesta de Caracterización Socioeconómica Nacional (2011). Observatorio Social – Ministerio de Desarrollo Social. Santiago, Chile.

CASEN: Bases de datos Encuesta de Caracterización Socioeconómica Nacional (2013). Observatorio Social – Ministerio de Desarrollo Social. Santiago, Chile.

Contraloría General de la República (2012) *Financiamiento fiscal a la educación superior 2011*. Available at:

http://www.contraloria.cl/NewPortal2/portal2/ShowProperty/BEA%2BRepository/Merged/2013/Archivos/356_Financiamiento_2011 (Accessed: 31 January 2017). Santiago, Chile.

Contraloría General de la República (2013) *Financiamiento fiscal a la educación superior 2012*. Available at: http://www.contraloria.cl/NewPortal2/portal2/ShowProperty/BEA%20Repository/Merged/2013/Archivos/146_Financiamiento_Fiscal_2012 (Accessed: 31 January 2017). Santiago, Chile.

Contraloría General de la República (2014) *Financiamiento fiscal a la educación superior 2013*. Available at: http://www.contraloria.cl/NewPortal2/portal2/ShowProperty/BEA%20Repository/Merged/2014/Archivos/209 (Accessed: 31 January 2017). Santiago, Chile.

Contraloría General de la República (2015) *Financiamiento fiscal a la educación superior 2014*. Available at: https://www.contraloria.cl/portalweb/documents/451102/2367138/FFES_2014.pdf/0530a2b9-3444-7473-d499-271b96bd359a (Accessed: 31 January 2017). Santiago, Chile.

Contraloría General de la República (2016) *Financiamiento fiscal a la educación superior 2015*. Available at: http://www.contraloria.cl/NewPortal2/portal2/ShowProperty/BEA+Repository/Merged/2016/Archivos/0104 (Accessed: 31 January 2017). Santiago, Chile.

Correa, J. A., Parro, F. and Reyes, L. (2014). 'The effects of vouchers on school results: Evidence from Chile's targeted voucher program', *Journal of Human Capita*, 8(4), pp. 351–398, [Online]. Available at: http://www.jstor.org/stable/10.1086/679282 (Accessed: 11 June 2014).

Elacqua, G. (2012). 'The impact of school choice and public policy on segregation: Evidence from Chile', *International Journal of Educational Development*, 32(3), pp. 444–453, [Online]. Available at: http://www.sciencedirect.com/science/article/pii/S0738059311001283 (Accessed: 30 May 2014).

Elster, J. (1985). *Making sense of Marx*. Cambridge: Cambridge University Press.

Elster, J. (1986). *An introduction to Karl Marx*. Cambridge: Cambridge University Press.

Enders, C. (2015) 'Applied missing data analysis', *2015 Summer Quatitative Method Series*. Portland State University. Unpublished.

Epple, D. and Romano, R. (2008). 'Educational vouchers and cream skimming', *International Economic Review*, 49(4), pp. 1395–1435, [Online]. Available at: http://dx.doi.org/10.1111/j.1468-2354.2008.00516.x (Accessed: 11 June 2014).

Erikson, R. (1984). 'Social class of men, women and families', *Sociology*, 18(4), pp. 500–514, [Online]. Available at: https://doi.org/10.1177/0038038584018004003 (Accessed: 12 January 2018).

Erikson, R. and Goldthorpe, J. H. (1992). *The constant flux: A study of class mobility in industrial societies*. Oxford: Clarendon Press.

Erikson, R. and Jonsson, J. O. (1996). The Swedish context: Educational reform and long-term change in educational inequality, in Erikson, R. and Jonsson, J. O. (eds.) *Can Education Be Equalized? The Swedish Case in Comparative Perspective*. Colorado: Westview Press, pp. 65–93.

Erikson, R., Goldthorpe, J. H., Jackson, M., Yaish, M. and Cox, D. R. (2005). 'On class differentials in educational attainment', *Proceedings of the National Academy of Science*, 102(27), pp. 9730–9733, [Online]. Available at: www.pnas.org/cgi/doi/10.1073/pnas.0502433102 (Accessed: 4 November 2013).

Erikson, R. and Rudolphi, F. (2010). 'Change in social selection to upper secondary school--Primary and secondary effects in Sweden', *European Sociological Review,* 26(3), pp. 291–305, [Online]. Available at: http://dx.doi.org/10.1093/esr/jcp022 (Accessed: 27 September 2013).
European Social Survey Round 1 Data (2002). Data file edition 6.5. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 2 Data (2004). Data file edition 3.5. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 3 Data (2006). Data file edition 3.6. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 4 Data (2008). Data file edition 4.4. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 5 Data (2010). Data file edition 3.3. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 6 Data (2012). Data file edition 2.3. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
European Social Survey Round 7 Data (2014). Data file edition 2.1. NSD – Norwegian Centre for Research Data, Norway – Data Archive and distributor of ESS data for ESS ERIC.
Fried, B. and Abuhadba, M. (1991). 'Reforms in higher education: The case of Chile in the 1980s', *Higher Education,* 21(2), pp. 137–149, [Online]. Available at: http://www.jstor.org/stable/3447410 (Accessed: 22 March 2016).
Friedman, M. (1962). *Capitalism and Freedom*. Chicago: University of Chicago Press.
Gallego, F. (2006). 'Voucher-school competition, iniciatives, and outcomes: Evidence from Chile' Department of Economics, MIT. Unpublished paper. Available at: www.researchgate.net/profile/Francisco_Gallego/publication/228347124_Voucher-school_competition_incentives_and_outcomes_Evidence_from_Chile/links/02bfe50d5d220c262a000000.pdf (Accessed: 10 June 2014).
Gallego, F. and Sapelli, C. (2007). 'El financiamiento de la educación en Chile: Una evaluación', *Revista Pensamiento Educativo,* 40(1), pp. 263–284, [Online]. Available at: http://www.pensamientoeducativo.uc.cl/files/journals/2/articles/407/public/407-916-1-PB.pdf (Accessed: 12 June 2014).
Gallego, F. (2013). 'When does inter-school competition matter? Evidence from the Chilean "voucher" system', *The B.E. Journal of Economic Analysis & Policy,* 13(2), pp. 525–562, [Online]. Available at: http://dx.doi.org/10.1515/bejeap-2012-0032 (Accessed: 12 June 2014).
Ganninger, M. (2007).[European Social Survey] Weighting in the ESS cumulative data set – 23 January 2007. *HowTo, Final Version* [Online]. Available at: https://www.europeansocialsurvey.org/docs/methodology/ESS_weighting_cumulative_data.pdf (Accessed: 24 August 2016).
Ganzeboom, H. B. G. and Luijkx, R. (2004). 'More recent trends in intergenerational occupational class reproduction in the Netherlands 1970–2004: Evidence from an

expanded database', *The Netherlands' Journal of Social Sciences,* 40(2), pp. 114–142.
Gelman, A. (2007). 'Comment: Bayesian checking of the second levels of hierarchical models', *Statistical Science,* 22(3), pp. 349–352, [Online]. Available at: http://dx.doi.org/doi:10.1214/07-STS235A (Accessed: 20 January 2018).
Goldthorpe, J. H. and Marshall, G. (1992). 'The promising future of class analysis: A response to recent critiques', *Sociology,* 26(3), pp. 381–400, [Online]. Available at: https://doi.org/10.1177/0038038592026003002 (Accessed: 15 August 2017).
Goldthorpe, J. H. (2007a). *On sociology second edition volume two: Illustration and retrospect*. Stanford, CA: Stanford University Press.
Goldthorpe, J. H. (2007b). '"Cultural Capital": Some critical observations', *Sociologica,* (2), [Online]. Available at: http://www.sociologica.mulino.it/doi/10.2383/24755 (Accessed: 11 December 2013).
Goldthorpe, J. H. (2007c). '"Cultural Capital": A response to the comments', *Sociologica,* (2), [Online]. Available at: http://www.sociologica.mulino.it/doi/10.2383/24759 (Accessed: 11 December 2013).
González-Velosa, C., Rucci, G., Sarzosa, M. and Úrzua, S. (2015). 'Returns to higher education in Chile and Colombia', Inter-American Development Bank, Labor Markets and Social Security Unit, Working Paper No. IDB-WP-587.
Guidici, F. and Pallas, A. M. (2014). 'Social origins and post-high school institutional pathways: A cumulative dis/advantage approach', *Social Science Research,* 44, pp. 103–113, [Online]. Available at: http://dx.doi.org/10.1016/j.ssresearch.2013.11.004 (Accessed: 11 June 2014).
Haim, E. B. and Shavit, Y. (2013). 'Expansion and inequality of educational opportunity: A comparative study', *Research in Social Stratification and Mobility,* 31(March 2013), pp. 22–31, [Online]. Available at: http://dx.doi.org/10.1016/j.rssm.2012.10.001 (Accessed: 2 May 2016).
Halsey, A. H., Heath, A. F. and Ridge, J. M. (1980). *Origins and destinations: Family, class, and education in modern Britain*. Oxford: Oxford University Press (Clarendon).
Hastings, J. S., Neilson, C. A. and Zimmerman, S. D. (2013). 'Are some degrees worth more than others? Evidence from college admission cutoffs in Chile', NBER Working Paper, No. 19241.
Hauser, R. M. and Featherman, D. L. (1976). 'Equality of schooling: Trends and prospects', *Sociology of Education,* 49(2), pp. 99–120, [Online]. Available at: http://www.jstor.org/stable/2112516 (Accessed: 30 August 2017).
Hauser, R. M. (1976). 'On Boudon's model of social mobility', *American Journal of Sociology,* 81(4), pp. 911–928, [Online]. Available at: http://www.jstor.org/stable/2777610 (Accessed: 20 June 2015).
Hayes, A. F. and Preacher, K. J. (2014). 'Statistical mediation analysis with a multicategorical independent variable', *British Journal of Mathematical and Statistical Psychology,* 67(3), pp. 451–70, [Online]. Available at: http://dx.doi.org/10.1111/bmsp.12028 (Accessed: 19 August 2015).
Herrera, J. and Muñoz, D. (2013) 'Sólo un 17% de los ingresos de las universidades estatales provino de aportes fiscales', *La Tercera,* 18 December [Online]. Available at: http://www.latercera.com/noticia/solo-un-17-de-los-ingresos-de-las-universidades-estatales-provino-de-aportes-fiscales/ (Accessed: 30 January 2017).
Holm, A. and Jæger, M. M. (2011). 'Dealing with selection bias in educational transition models: The bivariate probit selection model', *Research in Social Stratification and*

Mobility, 29(3), pp. 311–322, [Online]. Available at: http://dx.doi.org/10.1016/j.rssm.2011.02.002 (Accessed: 10 August 2017).
Holm, A. and Jæger, M. M. (2013). Dentist, driver, or dropout? Family background and secondary education choices in Denmark, in Jackson, M. (ed.) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 228–252.
Hout, M., Raftery, A. E. and Bell, E. O. (1993). Making the grade: Educational stratification in the United States, 1925-1989, in Shavit, Y. and Blossfeld, H.-P. (eds.) *Persistent inequality: Changing educational attainment in thirteen countries*. Colorado, USA: Westview Press, pp. 25–49.
Hout, M. (2006). 'Maximally Maintained Inequality and Essentially Maintained Inequality cossnational comparisons', *Sociological Theory and Methods*, 21(2), pp. 237–252, [Online]. Available at: http://doi.org/10.11218/ojjams.21.237 (Accessed: 24 March 2012).
Hsieh, C. and Urquiola, M. (2003). 'When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program', NBER Working Paper, No. 10008.
Hsieh, C. and Urquiola, M. (2006). 'The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program', *Journal of Public Economics*, 90, pp. 1477–1503, [Online]. Available at: http://www.columbia.edu/~msu2101/Hsieh-Urquiola(2006).pdf (Accessed: 10 June 2014).
Hyman, H. H. (1953). The value systems of different classes: A social psychological contribution to the analysis of stratification, in Bendix, R. and Lipset, S. M. (eds.) *Class, Status and Power*. Reprint, New York: Irvington Publishers, pp. 426–442, 1993. Series: Reprint Series in Social Sciences.
Jackson, M., Erikson, R., Goldthorpe, J. H. and Yaish, M. (2007). 'Primary and secondary effects in class differentials in educational attainment: The transition to A-level courses in England and Wales', *Acta Sociologica*, 50(3), pp. 211–229, [Online]. Available at: http://dx.doi.org/10.1177/0001699307080926 (Accessed: 28 October 2013).
Jackson, M. and Jonsson, J. O. (2013). Why does inequality of educational opportunity vary across countries? Primary and secondary effects in comparative context, in Jackson, M. (ed.) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 306–333.
Jackson, M. (ed.) (2013a). *Determined to succeed? Performance versus choice in educational attainment.* Stanford, CA: Stanford University Press.
Jackson, M. (2013b). Social background and educational transitions in England, in Jackson, M. (ed.) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 253–278.
Karlson, K. B., Holm, A. and Breen, R. (2010). 'Comparing regression coefficients between models using logit and probit models: A new method', Aarhus University, CSER, Working Paper, No. 0003.
Karlson, K. B. and Holm, A. (2011). 'Decomposing primary and secondary effects: A new decomposition method', *Research in Social Stratification and Mobility*, 29(2), pp. 221–237, [Online]. Available at: http://dx.doi.org/10.1016/j.rssm.2010.12.005 (Accessed: 17 February 2014).
Karlson, K. B., Holm, A. and Breen, R. (2012). 'Comparing regression coefficients between same-sample nested models using logit and probit: A new method',

Sociological Methodology, 42(1), pp. 286–313, [Online]. Available at: http://dx.doi.org/10.1177/0081175012444861 (Accessed: 1 July 2015).

Kartsonaki, C., Jackson, M. and Cox, D. R. (2013). Primary and secondary effects: Some methodological issues, in Jackson, M. (ed.) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 34–55.

Kerckhoff, A. C. (1995). 'Institutional arrangements and stratification processes in industrial societies', *Annual Review of Sociology*, 21, pp. 323–347, [Online]. Available at: https://doi.org/10.1146/annurev.so.21.080195.001543 (Accessed: 6 June 2014).

Kerr, C., Harbison, F. H., Dunlop, J. T. and Myers, C. A. (1960). 'Industrialims and industrial man', *International Labour Review*, 82(3), pp. 1–15, [Online]. Available at: http://hdl.handle.net/10694/307 (Accessed: 16 August 2017).

Krull, J. L. and MacKinnon, D. P. (2001). 'Multilevel modeling of individual and group level mediated effects', *Multivariate Behavioral Research*, 36(2), pp. 249–277, [Online]. Available at: http://dx.doi.org/10.1207/S15327906MBR3602_06 (Accessed: 9 July 2017).

Kuha, J. and Goldthorpe, J. H. (2010). 'Path analysis for discrete variables: The role of education in social mobility', *Journal of the Royal Statistical Society*, 173(2), pp. 351–369, [Online]. Available at: http://dx.doi.org/10.1111/j.1467-985X.2009.00620.x (Accessed: 9 September 2015).

Lara, B., Mizala, A. and Repetto, A. (2010). 'The effectiveness of private voucher education: Evidence from structural school switches', Universidad de Chile, Centro de Economía Aplicada, Documentos de Trabajo Serie Economía, No. 263.

Larroucau, T., Ríos, I. and Mizala, A. (2013).[Departamento de Evaluación, Medición y Registro Educacional] Efecto de la incorporación del ranking de notas en la selección universitaria – August 2013. [Online]. Available at: http://www.psu.demre.cl/adjuntos/informe-efectos-ranking-seleccion-universitaria-2013.pdf (Accessed: 15 August 2014).

Larroucau, T. (2014).[Sistema Único de Admisión, CRUCH report on] Ranking de notas proceso de admisión 2014 – March 2014. [Online]. Available at: http://sistemadeadmision.consejoderectores.cl/public/pdf/publicaciones/Informe%20Ranking%2026-03-14.pdf (Accessed: 05 January 2017).

Leckie, G. (2013) 'Cross-Classified Multilevel Models - Concepts'. LEMMA VLE Module 12. Unpublished.

Letelier, M. (2014) 'Recursos públicos para educación superior: La histórica diferencia de las universidades estatales', *CIPER*, 07 November [Online]. Available at: http://ciperchile.cl/2014/11/07/recursos-publicos-para-educacion-superior-la-historica-diferencia-de-las-universidades-estatales/ (Accessed: 30 January 2017).

Lucas, S. R. (2001). 'Effectively maintained inequality: Education transitions, track mobility, and social background effects', *American Journal of Sociology*, 106(6), pp. 1642–1690, [Online]. Available at: http://www.jstor.org/stable/3081213 (Accessed: 20 March 2013).

Lucas, S. R., Fucella, P. N. and Berends, M. (2011). 'A neo-classical education transitions approach: A corrected tale for three cohorts', *Research in Social Stratification and Mobility*, 29(3), pp. 263–285, [Online]. Available at: https://doi.org/10.1016/j.rssm.2011.03.003 (Accessed: 29 August 2013).

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. New York: Lawrence Erlbaum Associates. Series: Multivariate Applications Series.

MacKinnon, D. P. and Valente, M. J. (2014). 'Mediation from multilevel to structural equation modeling', *Annals of Nutrition and Metabolism*, 65(2-3), pp. 198–204, [Online]. Available at: http://dx.doi.org/10.1159/000362505 (Accessed: 9 July 2017).

Mare, R. D. (1980). 'Social background and school continuation decisions', *Journal of the American Statistical Association,* 75(370), pp. 295–305, [Online]. Available at: http://www.jstor.org/stable/2287448 . (Accessed: 4 November 2013).

Mare, R. D. (1981). 'Change and stability in educational stratification', *American Sociological Review,* 46(1), pp. 72–87, [Online]. Available at: http://www.jstor.org/stable/2095027 . (Accessed: 4 November 2013).

Ministerio de Educación (2003) *Estadísticas de la educación 2002*. Available at: https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2017/07/EstadisticasdelaEducacion2002.pdf (Accessed: 25 August 2014). Santiago, Chile.

Ministerio de Educación (2010) *Estadísticas de la educación 2008*. Available at: https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2017/07/EstadisticasdelaEducacion2008.pdf (Accessed: 25 August 2014). Santiago, Chile.

Ministerio de Educación (2017) *Estadísticas de la educación 2016*. Available at: https://centroestudios.mineduc.cl/wp-content/uploads/sites/100/2017/07/EstadisticasdelaEducacion2016.pdf (Accessed: 25 August 2014). Santiago, Chile.

Ministerio de Desarrollo Social (2013) *Informe Política Social 2013*. Available at: http://www.ministeriodesarrollosocial.gob.cl/ipos-2013/media/IPOS_2013.pdf (Accessed: 14 September 2017). Santiago, Chile.

Mizala, A. and Romaguera, P. (2000). 'School performance and choice: The Chilean experience', *The Journal of Human Resources,* 35(2), pp. 392–417, [Online]. Available at: http://www.jstor.org/stable/pdf/146331.pdf (Accessed: 10 April 2014).

Mizala, A. and Torche, F. (2012). 'Bringing the schools back in: The stratification of educational achievement in the Chilean voucher system', *International Journal of Educational Development,* 32(2012), pp. 132–144, [Online]. Available at: https://doi.org/10.1016/j.ijedudev.2010.09.004 (Accessed: 5 December 2013).

Mizala, A. and Urquiola, M. (2013). 'School markets: The impact of information approximating schools' effectiveness', *Journal of Development Economics,* 103, pp. 313–335, [Online]. Available at: https://doi.org/10.1016/j.jdeveco.2013.03.003 (Accessed: 28 July 2014).

Muthén, L. K. and Muthén, B. (2011) 'Multilevel Modeling With Latent Variables Using Mplus: Cross-Sectional Analysis', *Mplus Short Courses Topic 7*. Muthén & Muthén. Unpublished.

Muthén, B. O., Muthén, L. K. and Asparouhov, T. (2016). *Regression and mediation analysis using Mplus*. Los Angeles, CA: Muthén & Muthén.

Muthén, L. K. and Muthén, B. O. (1998–2017). *Mplus User's Guide*. Eighth edn. Los Angeles, CA: Muthén & Muthén.

Neal, D. (2002). 'How vouchers could change the market for education', *The Journal of Economic Perspectives,* 16(4), pp. 25–44, [Online]. Available at: http://www.jstor.org/stable/3216913 (Accessed: 11 June 2014).

Neilsen, C. (2013). 'Targeted vouchers, competition among schools, and the academic achievement of poor students' Yale University. Unpublished paper. Available at:

http://faculty.chicagobooth.edu/workshops/micro/pdf/NeilsonJMP.pdf (Accessed: 11 June 2014).
Neugebauer, M., Reimer, D., Schindler, S. and Stocké, V. (2013). Inequality in transitions to secondary school and tertiary education in Germany, in Jackson, M. (ed.) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 56–88.
OECD (2003). *Education at a glance 2003: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2003-en.
OECD (2004). *Education at a glance 2004: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2004-en.
OECD (2005). *Education at a glance 2005: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2005-en.
OECD (2006). *Education at a glance 2006: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2006-en.
OECD (2007). *Education at a glance 2007: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2007-en.
OECD (2008). *Education at a glance 2008: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2008-en.
OECD (2009). *Education at a glance 2009: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2009-en.
OECD (2010). *Education at a glance 2010: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2010-en.
OECD (2011). *Education at a glance 2011: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2011-en.
OECD (2012). *Education at a glance 2012: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2012-en.
OECD (2013). *Education at a glance 2013: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2013-en.
OECD (2014). *Education at a glance 2014: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2014-en.
OECD (2015). *Education at a glance 2015: OECD indicators*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2015-en.
OECD (2016). *Education at a glance 2016: OECD indicators*. *Paris: OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/eag-2016-en.
OECD and World Bank (2009). *Tertiary education in Chile*. *OECD Publishing*.[Online]. Available at: http://dx.doi.org/10.1787/9789264051386-en.
Pallas, A. M. (2003). Educational transitions, trajectories, and pathways, in Mortimer, J. T. and Shanahan, M. J. (eds.) *Handbook of the life course*. New York: Kluwer Academic Publishers, pp. 165–184.
Parry, T. R. (1997). 'Decentralization and privatization: Education policy in Chile', *Journal of Public Policy*, 17(1), pp. 107–133, [Online]. Available at: http://www.jstor.org/stable/4007620 (Accessed: 27 July 2017).
Parsons, T. (1960). *Structure and process in modern societies*. New York: The Free Press.
Payne, G., Ford, G. and Robertson, C. (1977). 'A reappraisal of social mobility in Britain', *Sociology*, 11(2), pp. 289–310, [Online]. Available at: https://doi.org/10.1177/003803857701100203 (Accessed: 15 August 2017).
Pfeffer, F. T. (2008). 'Persistent inequality in educational attainment and its institutional context', *European Sociological Review*, 24(5), pp. 543–565, [Online]. Available at: https://doi.org/10.1093/esr/jcn026 (Accessed: 29 March 2012).

Preacher, K. J., Zyphur, M. J. and Zhang, Z. (2010). 'A general multilevel SEM framework for assessing multilevel mediation', *Psychological Methods*, 15(3), pp. 209–233, [Online]. Available at: http://dx.doi.org/10.1037/a0020141 (Accessed: 9 July 2017).

Preacher, K. J., Zhang, Z. and Zyphur, M. J. (2011). 'Alternative methods for assessing mediation in multilevel data: The advantages of multilevel SEM', *Structural Equation Modeling: A Multidisciplinary Journal*, 18(2), pp. 161–182, [Online]. Available at: http://dx.doi.org/10.1080/10705511.2011.557329 (Accessed: 9 July 2017).

Racynski, D. (2011).[Ministerio de Educación Fondo de Investigación y Desarrollo en Educación - FONIDE] El paso de la enseñanza básica a la media en estratos bajos: Un reto a la igualdad de oportunidades educativas – Diciembre 2011. *Proyecto FONIDE No. F511066*.

Raftery, A. E. and Hout, M. (1993). 'Maximally Maintained Inequality: Expansion, reform, and opportunity in Irish education, 1921-75', *Sociology of Education*, 66(1), pp. 41–62, [Online]. Available at: http://www.jstor.org/stable/2112784 (Accessed: 24 March 2012).

Ress, A. and Azzolini, D. (2014). 'Primary and secondary effects of social back ground on educational attainment in Italy. Evidence from an administrative dataset.', *Italian Journal of Sociology of Education*, 6(1), pp. 53–80, [Online]. Available at: http://ijse.padovauniversitypress.it/2014/1/3 (Accessed: 18 July 2017).

Riphahn, R. T. and Schieferdecker, F. (2012). 'The transition to tertiary education and parental background over time', *Journal of Population Economics*, 25(2), pp. 635–675, [Online]. Available at: http://dx.doi.org/10.1007/s00148-010-0347-3 (Accessed: 19 April 2013).

Rosas, R. and Santa Cruz, C. (2013). *Dime en qué colegio estudiaste y te diré qué CI tienes*. Santiago, Chile: Ediciones UC.

Rudolphi, F. (2013). Ever-Declining Inequalities? Transitions to upper secondary and tertiary education in Sweden, 1972–1990 birth cohorts, in Jackson, M. (ed.) *Determined to succeed? Performance versus choice in educational attainment*. Stanford, CA: Stanford University Press, pp. 185–227.

Santos, H. (2009). 'Dinámica de la deserción escolar en Chile', Centro de Políticas Comparadas de Educación Documento de Trabajo CPCE, No. N° 3.

Sapelli, C. and Vial, B. (2005). 'Private vs public voucher schools in Chile: New evidence on efficiency and peer effects', Pontificia Universidad Católica de Chile, Instituto de Economía, Documento de Trabajo, No. 289.

Schiefelbein, E. and Farrell, J. P. (1980). 'Women, schooling, and work in Chile: Evidence from a longitudinal study', *Comparative Education Review*, 24(2), pp. S160–S179.

Schiefelbein, E. and Farrell, J. P. (1982). *Eight years of their lives: Through schooling to the labour market in Chile*. Ottawa: IDRC.

Sewell, W. H., Haller, A. O. and Portes, A. (1969). 'The educational and early occupational attainment process', *American Sociological Review*, 34(1), pp. 82–92, [Online]. Available at: http://www.jstor.org/stable/2092789 (Accessed: 2 June 2014).

Sewell, W. H., Haller, A. O. and Ohlendorf, G. W. (1970). 'The educational and early occupational status attainment process: Replication and revision', *American Sociological Review*, 35(6), pp. 1014–1027, [Online]. Available at: http://www.jstor.org/stable/2093379 (Accessed: 2 June 2014).

Sewell, W. H. and Hauser, R. M. (1972). 'Causes and consequences of higher education: Models of the status attainment process', *American Journal of Agricultural*

Economics, 54(5), pp. 851–861, [Online]. Available at: http://www.jstor.org/stable/1239228 (Accessed: 2 June 2014).
Shavit, Y. and Blossfeld, H.-P. (1993). *Persistent inequality: Changing educational attainment in thirteen countries*. Colorado, USA: Westview Press.
SIES (2016). Compendio histórico de educación superior - Evolución de matrícula total de pregrado por tipo específico de institución 1983-2016, Servicio de Información de Educación Superior, Ministry of Education. Santiago, Chile.
Skrondal, A. and Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Boca Raton, FL: Chapman & Hall/CRC.
Snijders, T. A. B. and Bosker, R. J. (2012). *Multilvevel analysis: An introduction to basic and advanced multilevel modeling*. 2nd edn. London: SAGE Publications.
Sorokin, P. A. (1927). *Social mobility*. Reprint, NewYork: The Free Press, 1959.
Stromquist, N. P. and Sanyal, A. (2013). 'Student resistance to neoliberalism in Chile', *International Studies in Sociology of Education*, 23(2), pp. 152–178, [Online]. Available at: http://dx.doi.org/10.1080/09620214.2013.790662 (Accessed: 21 November 2015).
Sullivan, A., Parsons, S., Wiggins, R., Heath, A. and Green, F. (2014). 'Social origins, school type and higher education destinations', *Oxford Review of Education*, 40(6), pp. 739–763, [Online]. Available at: http://dx.doi.org/10.1080/03054985.2014.979015 (Accessed: 22 April 2015).
Swift, A. (2005). Justice, luck, and the family: The intergenerational transmission of economic advantage from a normative perspective, in Bowles, S., Gintis, H. and Osborne Groves, M. (eds.) *Unequal chances: Family background and economic success*. New Jersey: Princeton University Press, pp. 256–276.
Tang, W., He, H. and Tu, X. M. (2012). *Applied categorical and count data analysis*. Boca Raton, FL: CRC Press Taylor & Francis Group.
Tofighi, D., West, S. G. and MacKinnon, D. P. (2013). 'Multilevel mediation analysis: he effects of omitted variables in the 1–1–1 model', *British Journal of Mathematical and Statistical Psychology*, 66(2), pp. 290–307, [Online]. Available at: http://dx.doi.org/10.1111/j.2044-8317.2012.02051.x (Accessed: 9 July 2017).
Tofighi, D. and Thoemmes, F. (2014). 'Single-level and multilevel mediation analysis', *Journal of Early Adolescence*, 34(1), pp. 93–119, [Online]. Available at: http://dx.doi.org/10.1177/0272431613511331 (Accessed: 9 July 2017).
Tokman, A. (2002). 'Is private education better? Evidence from Chile', Central Bank of Chile Working Paper, No. 147.
Torche, F. (2005). 'Privatization reform and inequality of educational opportunity: The case of Chile', *Sociology of Education*, 78(4), pp. 316–343, [Online]. Available at: https://doi.org/10.1177/003804070507800403 (Accessed: 18 October 2013).
Treiman, D. J. (1970). 'Industrialization and social stratification', *Sociological Inquiry*, 40(2), pp. 207–234, [Online]. Available at: http://dx.doi.org/10.1111/j.1475-682X.1970.tb01009.x (Accessed: 15 August 2017).
Turner, R. H. (1960). 'Sponsored and contest mobility and the school system', *American Sociological Review*, 25(6), pp. 855–867, [Online]. Available at: http://www.jstor.org/stable/2089982 (Accessed: 6 September 2017).
Valenzuela B., J. P., Allende, C., Sevilla, A. and Egaña, P. (2013). 'La (ina)movilidad del desempeño educativo de los estudiantes chilenos: realidad, oportunidades y desafíos', Universidad de Chile, Centro de Investigación Avanzada en Educación, Documentos de Trabajo, No. 8.

Valenzuela, J. P., Bellei, C. and de los Ríos, D. (2014). 'Socioeconomic school segregation in a market-oriented educational system. The case of Chile', *Journal of Education Policy*, 29(2), pp. 217–241, [Online]. Available at: http://dx.doi.org/10.1080/02680939.2013.806995 (Accessed: 13 June 2017).
Villalobos, C. and Valenzuela, J. P. (2012). 'Polarización y cohesión social del sistema escolar chileno', *Revista de Análisis Económico,* 27(2), pp. 145–172, [Online]. Available at: http://www.rae-ear.org/index.php/rae/article/view/366/529 (Accessed: 13 June 2017).
Villarroel, G. (2014). *Mejoramiento en resultados académicos de la educación básica en Chile : ¿Primeros efectos de la ley de subvención escolar preferencial (SEP)?* Tesis para optar al grado de Magíster en Economía. Escuela de Postgrado, Economía y Negocios, Universidad de Chile.
Walters, P. B. (2000). The limits of growth, in Hallinan, M. T. (ed.) *Handbook of the sociology of education*. New York: Springer, pp. 241–261.
Waters, M. C., Heath, A., Tran, V. C. and Boliver, V. (2013). Second-generation attainment and inequality: Primary and secondary effects on educational outcomes in Britain and the United States, in Alba, R. and Holdaway, J. (eds.) *The Children of Immigrants at School: A Comparative Look at Integration in the United States and Western Europe*. New York: NYU Press, pp. 120–159.
Wormald, G., Flores, C., Sabatini, F., Trebilcock, M. P. and Rasse Figueroa, A. (2012). 'Cultura de cohesión e integración en las ciudades chilenas', *Revista INVI*, 76(27), pp. 117–145, [Online]. Available at: http://dx.doi.org/10.4067/S0718-83582012000300004 (Accessed: 13 June 2017).
Zibechi, R. (2012) 'A new Chile is possible', *Counterpunch*, 27 January [Online]. Available at: http://www.counterpunch.org/2012/01/27/a-new-chile-is-possible/ (Accessed: 21 November 2015).

Appendices

Chapter 5: Appendices

5A.1 Appendix

Table 5A.1 Odds ratios of the transition to higher education between parental education groups (GPA) (Buis, 2010)

Parental Education Groups Comparison	Total Effect	Different performance, similar transition probability	Similar performance, different transition probability
Degree/Low	4.58	1.70	2.69
	(.10)	(.01)	(.06)
Vocational/Low	2.69	1.25	2.15
	(.06)	(.01)	(.04)
Secondary/Low	1.75	1.12	1.56
	(.03)	(.00)	(.02)

Note: Bootstrapped standard errors in parentheses, from 200 replications.

Table 5A.2 Relative primary and secondary effects between parental education groups (GPA) (Buis, 2010)

Boudon's Model	Primary Effects	Secondary Effects
Degree/Low	34.8	65.2
Vocational/Low	22.6	77.4
Secondary/Low	20.3	79.7

Table 5A.3 Predicted and Counterfactual Transition Probabilities between parental education groups (SIMCE in grade 8)

If they had the log-odds of… / **If they had the SIMCE of…**	Degree	Advanced Vocational Q.	Secondary Education	Primary ed. or none
Degree	0.78	0.73	0.67	0.59
Advanced Vocational Q.	0.73	0.68	0.61	0.53
Secondary Education	0.70	0.64	0.58	0.49
Primary ed. or none	0.66	0.60	0.53	0.44

Table 5A.4 Predicted and Counterfactual Transition Odds Ratios between parental education groups (SIMCE in grade 8)

If they had the log-odds of… / **If they had the SIMCE of…**	Degree	Advanced Vocational Q.	Secondary education	Primary ed. or none
Degree	3.57	2.73	2.07	1.43
Advanced Vocational Q.	2.73	2.10	1.59	1.10
Secondary education	2.34	1.80	1.37	0.95
Primary ed. or none	1.93	1.48	1.12	0.78

Table 5A.5 Odds ratios of the transition to higher education between parental education groups (SIMCE in grade 8) (Buis, 2010)

Parental Education Groups Comparison	**Total Effect**	**Different performance, similar transition probability**	**Similar performance, different transition probability**
Degree/Low	4.60	1.85	2.48
	(.10)	(.02)	(.06)
Vocational/Low	2.69	1.42	1.90
	(.06)	(.01)	(.04)
Secondary/Low	1.75	1.22	1.44
	(.03)	(.00)	(.02)

Note: Bootstrapped standard errors in parentheses, from 200 replications.

Table 5A.6 Relative primary and secondary effects between parental education groups (SIMCE in grade 8) (Buis, 2010)

Boudon's Model	Primary Effects	Secondary Effects
Degree/Low	40.6	59.4
Vocational/Low	35.3	64.7
Secondary/Low	35.2	64.8

Table 5A.7 'Reduced' model (logit) of the transition to HE upon parental education and GPA residuals

Transition to Higher Education	Coefficients
Primary or none (reference)	-
Secondary Education	0.61 (.015)
Advanced Vocational Q.	1.07 (.021)
Degree	1.65 (.022)
Residuals (GPA)	0.65 (.007)
Intercept	-0.27 (.012)
McFadden's pseudo R-square	0.099
N	121,088

Note: Standard errors in parentheses.

All odds ratios are significant at p<.001 level.

Table 5A.8 Auxiliary (linear) model of GPA upon parental education

Std. GPA in Secondary	Coefficients
Primary or none (reference)	-
Secondary Education	0.19 (.007)
Advanced Vocational Q.	0.37 (.009)
Degree	0.88 (.009)
Intercept	-0.26 (.005)
R-square	0.084
N	121,088

Note: Standard errors in parentheses.
All coefficients are significant at p<.001 level.

Table 5A.9 'Reduced' model (logit) of the transition to HE upon parental education and SIMCE G8 residuals

Transition to Higher Education	Coefficients
Primary or none (reference)	-
Secondary Education	0.59 (.015)
Advanced Vocational Q.	1.05 (.021)
Degree	1.60 (.022)
Residuals (SIMCE G8 scores)	0.55 (.007)
Intercept	-0.26 (.012)
McFadden's pseudo R-square	0.083
N	114,585

Note: Standard errors in parentheses.
All odds ratios are significant at p<.001 level.

Table 5A.10 Auxiliary (linear) model of SIMCE G8 upon parental education

Std. SIMCE G8 scores	Coefficients
Primary or none (reference)	-
Secondary Education	0.38
	(.007)
Advanced Vocational Q.	0.67
	(.009)
Degree	1.17
	(.009)
Intercept	-0.43
	(.005)
R-square	0.149
N	114,585

Note: Standard errors in parentheses.

All coefficients are significant at p<.001 level.

5A.2 Methodological Appendix - Chapter 5

5A.2.1 Primary and secondary effects method - Erikson *et al.* (2005)

The method proposed by Erikson *et al.* (2005) extends Erikson and Jonsson (1996) method by finding the relative contribution of primary and secondary effects. Erikson and Jonsson (1996) method centres on deriving the transition probability curves and academic performance distribution by social class.

The first step, in both methods, is the estimation of separate logit models of the transition to higher education for each group of social background, as shown in equation (1). S refers to group j of social background, P to an standardised measure of academic performance, and T to a binary variable for the transition to higher education. The intercept of the model is $\propto$ and, the coefficient for performance is β.

$$logit(\Pr(T = 1|S = j)) = \propto_j + \beta_j P_j \tag{1}$$

In the second step a predicted probability of transition is estimated, conditional on academic performance by each group of social background. Then, a counterfactual analysis follows by combining the predicted probability with the normalised distribution of academic performance by a numerical integration[108], shown in equation (2).

$$\int_{-4}^{+4} \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-(P-\mu_P)^2/2\sigma_P^2}\right) \left(\frac{e^{(\propto_j+\beta_j P_j)}}{1+e^{(\propto_j+\beta_j P_j)}}\right) dP \tag{2}$$

In equation (2), μ_P and σ_P are the mean and standard deviation of the measure of academic performance, and the second term comes from the estimated coefficients in equation (1). The resulting counterfactual transition probability is a combination of different

[108] Just for illustrative purposes, the numerical integration is evaluated in the range of +4 to -4, as in Jackson et al. (2007).

social groups academic performance distribution and the coefficients estimated in equation (1).

Buis (2010) suggested a generalisation of the assumed normal distribution of academic performance. Consequently, the numerical integration allows any empirical distribution of the measure of academic performance.

The third step consist on deriving odds-ratios (Q) from the counterfactual transition probabilities, in equations (3) and (4). Where the first subscript of the transition rate P_{jj}, refers to the academic performance distribution of social group *j* and the second subscript to the transition probability of the same social group *j*. Likewise subscripts of social group *k* refer to academic performance and transition probability respectively.

$$Q_{jj.kj} = \frac{(P_{jj}/1-P_{jj})}{(P_{kj}/1-P_{kj})} \quad (3)$$

$$Q_{jj.jk} = \frac{(P_{jj}/1-P_{jj})}{(P_{jk}/1-P_{jk})} \quad (4)$$

Equation (3) reflects the counterfactual odds ratios of social group *j*, maintaining the transition probability of social group *j*, but the academic performance distribution of social group *k*. Similarly, equation (4) reflects the counterfactual odds ratios of social group *j*, maintaining the academic performance distribution of social group *j*, but the transition probability of social group *k*.

The total effect is the product of the counterfactual odds ratios in performance or alternately, in transition probability. These are shown in (5) and (6).

$$Q_{jj.kk} = Q_{jj.kj}Q_{kj.kk} \quad (5)$$

$$Q_{jj.kk} = Q_{jk.kk}Q_{jj.jk} \quad (6)$$

The first term in (5) and (6) isolate the effect of academic performance, that is to say the primary effects of the decomposition. The second term in the same equations isolate the effect of transition probability, the secondary effects.

Equations (5) and (6) can also be presented in their logarithmic form as

$$L_{jj.kk} = L_{jk.kk} + L_{jj.jk} \tag{7}$$

$$L_{jj.kk} = L_{jj.kj} + L_{kj.kk} \tag{8}$$

Therefore, the relative contribution of secondary effects is

$$\frac{L_{kj.kk}}{L_{jj.kk}} \tag{9}$$

or

$$\frac{L_{jj.jk}}{L_{jj.kk}} \tag{10}$$

The relative contribution of primary effects is

$$\frac{L_{jj.kj}}{L_{jj.kk}} \tag{11}$$

or

$$\frac{L_{jk.kk}}{L_{jj.kk}} \tag{12}$$

Primary and secondary effects in (9) to (12) respectively, apparently are equivalent. However, they are not the same. To overcome this discrepancy, Jackson *et al.* (2007) and Kartsonaki, Jackson and Cox (2013, p. 40) suggested using the average of (9) and (10) to derive secondary effects, expressed as the percentage of log-odds ratios. Likewise using the average of (11) and (12) to derive primary effects.

Buis (2010) generalisation method added bootstrapped standard errors to the odds-ratios and log-odds derived from (3) to (8).

Karlson and Holm (2011) method uses fewer steps to derive primary and secondary effects, and it is based on a regression modelling framework. The method is described in the next subsection.

5A.2.2 Primary and secondary effects method - Karlson, Holm and Breen (2010)

Karlson, Holm and Breen (2010) method compares estimated coefficients in two nested non-linear probability models. It explains to what degree academic performance mediates the association between social background and a particular education transition.

Let T^* be a continuous latent variable representing the propensity to make the transition to higher education. Let S be the social background variable of interest, in the selected student cohort it was parental education. The mediating variable P, is academic performance, which can be GPA or SIMCE scores.

The subscript F refers to the full model and R to the reduced model. The full model controls for academic performance, whereas the reduced model does not.

$$T^* = \alpha_F + \beta_F S + \gamma_F P + u, \qquad u \sim (0, \sigma_u) \qquad (1)$$

$$T^* = \alpha_R + \beta_R S + \varepsilon, \qquad \varepsilon \sim (0, \sigma_\varepsilon) \qquad (2)$$

The coefficients for the full model are α_F, β_F and γ_F. The coefficients for the reduced model are α_R and β_R. The residual terms for both specifications are u and ε.

The direct effect and the total effect of the social background upon the transition are β_F and β_R respectively. Following the decomposition in linear models, the indirect effect is the difference between the total effect and the direct effect. That is,

$$\beta_I = \beta_R - \beta_F \qquad (3)$$

Since, T^* is not directly observed but can be approximated by a binary variable T, such that

$$T = 1 \text{ if } T^* > \tau$$

$$T = 0 \text{ if } T^* \leq \tau$$

where τ is the threshold of making the educational transition.

To estimate the effect of the academic performance within the proposed linear model framework, equations (1) and (2) can be rewritten as logistic models to estimate the transition probabilities, by correcting the error terms $u = \sigma_F \cdot w$ and $\varepsilon = \sigma_R \cdot w$ (where $w \sim (0, \pi^2/3)$ follows a logistic distribution). The σ_F is a scale parameter that accounts for the effect of omitted variables independent of social background (S) and academic performance (P). In the case of σ_R, only omitted variables besides social background are accounted for.

The standard deviation of the error term in (1) is, $\sigma_u = \sigma_F \cdot (\pi/\sqrt{3})$, and in (2) is $\sigma_\varepsilon = \sigma_R \cdot (\pi/\sqrt{3})$.

Therefore, linear models (1) and (2) become logit models, corrected by scale parameters σ_F and σ_R.

$$logit(\Pr(T = 1)) = a_F + b_F S + c_F P = \frac{\alpha_F}{\sigma_F} + \frac{\beta_F}{\sigma_F} S + \frac{\gamma_F}{\sigma_F} P \qquad (4)$$

$$logit(\Pr(T = 1)) = a_R + b_R S = \frac{\alpha_R}{\sigma_R} + \frac{\beta_R}{\sigma_R} S \qquad (5)$$

From (4) and (5) only logit models coefficients can be estimated (i.e. a_F, b_F, c_F, a_R and b_R). Neither the scale parameters σ_F and σ_R, or the coefficients α_F, β_F, γ_F , α_R and β_R are directly estimated. Therefore, the estimated direct effect is b_F, total effect is b_R and indirect effect is $b_R - b_F$.

One solution, proposed by Karlson, Holm and Breen (2010)[109], is to estimate a model for academic performance upon social background, in order to extract from the academic performance the information not contained in social background. This is achieved by calculating the residuals from a linear regression of academic performance on social background. That is,

$$R = P - (\mu + \zeta S) \tag{6}$$

where μ and ζ are the coefficients from the linear regression model. R are the residuals, P is academic performance and S is social background.

Then the residuals (R) are included in (2) instead of academic performance in comparison to (1), such as

$$T^* = \hat{\alpha}_R + \hat{\beta}_R S + \ \hat{\delta}_R \ R + \varepsilon \tag{7}$$

Karlson, Holm and Breen (2010) assumed the equality of scale parameters from (7) and (2), $\hat{\sigma}_R = \sigma_R$. Their method also assumes that the parameters for the social background variable are equal in the aforementioned equations, $\hat{\beta}_R = \beta_R$. Moreover, they assume that the scale parameters in (7) and (1) are equal, $\hat{\sigma}_R = \sigma_F$, since they claim that the residuals in models (1) and (7) have the same standard deviation as a result of having similar predictability. (See Karlson, Holm and Breen (2010, pp. 10–11) for the mathematical proof of all these assumptions). By applying all the above assumptions to the indirect effect of models (4) and (5), it can be rewritten as

$$Indirect\ Effect = \hat{b}_R - \ b_F = \ \frac{\hat{\beta}_R}{\hat{\sigma}_R} - \ \frac{\beta_F}{\sigma_F} = \frac{\beta_R - \beta_F}{\sigma_F} \tag{8}$$

The direct affect comes from the estimated coefficient from the logit model in (4).

$$Direct\ Effect = b_F \tag{9}$$

The total effect is the sum of both direct and indirect effects.

[109] They call their approach the KHB method.

$$Total\ Effect = \hat{b}_R \qquad (10)$$

Primary and secondary effects can be rewritten as the indirect and direct effects relative contribution to the total effect, as percentages of the total effect

$$Primary\ Effects = \frac{\hat{b}_R - b_F}{\hat{b}_R} \qquad (11)$$

$$Secondary\ Effects = \frac{b_F}{\hat{b}_R} \qquad (12)$$

An effect that is over 50% is considered the dominant effect in the overall inequality of educational opportunities in the transition to higher education.

5A.2.3 Variance estimator for difference in proportions

This section describes the methodology used to calculate the estimates and the statistical significance of the difference between two groups.

Notation:

n_1 – sample size at group 1

n_2 – sample size at group 2

n_c – common sample across 2 groups

$\hat{y}_1 = \sum_{i=1}^{n_1} y_{i1}$ - the total number of individuals having the characteristic at group 1

$\hat{y}_2 = \sum_{i=1}^{n_1} y_{i2}$ - the total number of individuals having the characteristic at group 2

$\hat{p}_{y1} = \frac{\hat{y}_1}{n_1}$ - the proportion of individuals out of the total sample in group 1

$\hat{p}_{y2} = \frac{\hat{y}_2}{n_2}$ - the proportion of individuals out of the total sample in group 2

5A.2.4 Test of difference in proportions between groups

Wald test for the difference in proportions between groups

$$Z = \frac{\hat{p}_{y1} - \hat{p}_{y2}}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p} = \frac{\hat{y}_1 + \hat{y}_2}{n_1 + n_2}$.

Z is compared with 1.96 at 95% confidence level.

Chapter 6: Appendices

6A.1 Appendix

Table 6A.1 Linear SIMCE models from the mediation model accounting for sample-selection

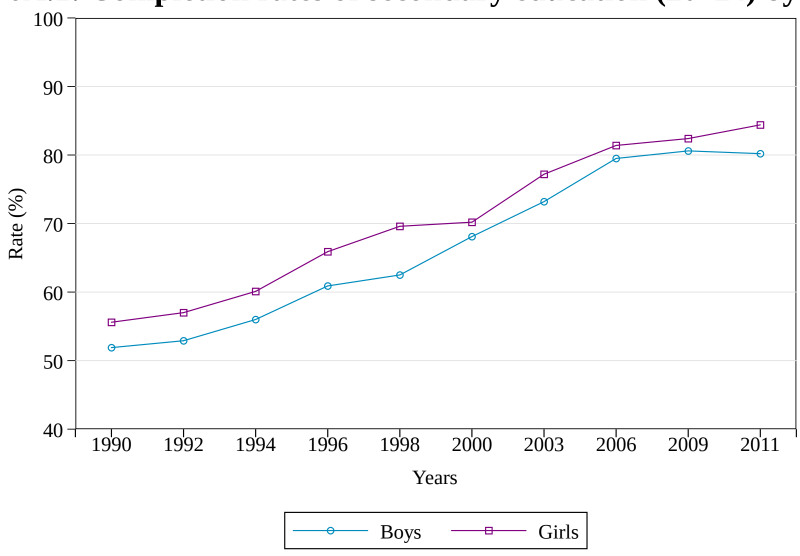
	Grade 4 SIMCE (Spanish+Maths)	Grade 8 SIMCE (Spanish+Maths)
	Coefficients	
Parental Education		
Primary ed. or none (reference)		
Degree	1.14	1.20
	(.01)	(.01)
Advanced Vocational Q.	0.74	0.69
	(.01)	(.01)
Secondary Education	0.46	0.39
	(.01)	(.01)
Girl(†)	-0.01	-0.01
	(.01)	(.01)
Intercept	-0.45	-0.43
	(.01)	(.01)
R-square	0.14	0.15
	(.00)	(.00)
N	170,856	

Note: Standard errors in parentheses.

All coefficients are significant (p<.001), except (†).

(†) is not significant at 95% confidence level in both models.

Figure 6A.1: Completion rates of secondary education (20–24) by gender



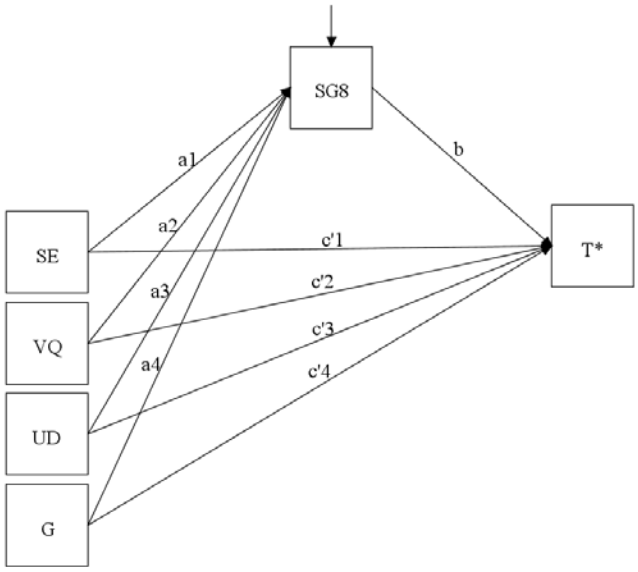
Note: The indicators in the figure were retrieved from ‘Informe Política Social 2013’ (Ministerio de Desarrollo Social, 2013, p. 140) using data from CASEN (1990–2011).

6A.2 Methodological Appendix – Chapter 6

6A.2.1 Non-Linear Mediation Model

In the proposed mediation model, the independent variables are the dichotomous transformation of the categorical variable reflecting the intermediate (secondary education and vocational qualifications) and the highest (university degree) parental education levels. The mediator is the academic performance, SIMCE or PSU average scores. The outcome of the model is the transition to higher education, whether or not the student is enrolled in a higher education institution after completing secondary education on 2013.

Figure 6A.2: Diagram of the non-linear mediation model



$$Probit(\mathrm{T}^*) = i_1 + c_1 \cdot SE + c_2 \cdot VQ + c_3 \cdot UD + c_4 \cdot G + \epsilon_1 \quad (1)$$

$$Probit(\mathrm{T}^*) = i_2 + c_1' \cdot SE + c_2' \cdot VQ + c_3' \cdot UD + c_4' \cdot G + b \cdot SG8 + \epsilon_2 \quad (2)$$

$$SG8 = i_3 + a_1 \cdot SE + a_2 \cdot VQ + a_3 \cdot UD + a_4 \cdot G + \epsilon_3 \quad (3)$$

Equations (1) to (3) describe the proposed mediation model in Figure 6A.2 for the analysis. T^* is an underlying continuous variable of the propensity of transition, approximated by a binary variable T. $SG8$ is the average SIMCE scores in grade 8 for Spanish and mathematics, and predictors are 3 dummy variables from the 4-categories of parental education for parents with secondary education (SE), advanced vocational qualifications (VQ) and university degree (UD). The reference category is parents with less than secondary education. G is a dummy variable that controls for student's gender, it takes the value of one for girls and zero for boys. Intercepts are i_1, i_2, i_3 and residual terms are ϵ_1,ϵ_2~ $N(0,1)$ and $\epsilon_3 \sim N(0, \sigma_{SG8})$.

It was not necessary to estimate equation (1) since direct and indirect effects can be derived from equations (2) and (3). Equation (1) is included for illustrative purposes only. Equation (2) is a probit model estimated by maximum likelihood, with robust standard errors regarding non-normality and non-independence of observations (Muthén and Muthén, 2015, p. 606). According to MacKinnon (2008, p.315), simulation studies suggested that probit models seem to be more accurate than logit models in a mediation-modelling framework. Therefore, a probit specification was preferred over a logit one.

Educational attainment treated as a categorical variable, rather than a continuous measure of years of education, seemed to be more appropriate, from a sociological perspective, to understand the effects of social stratification on performance and, subsequently, on the decision to continue into higher education. By the same token, it is of sociological interest to observe differences in levels of educational attainment, rather than differences in individual years of education. For an historical overview of the use of continuous and categorical measures of social status and education in sociological studies see Kuha and Goldthorpe (2010). However, a categorical independent variable, in the mediation-modelling framework, will translate into multiple indirect, direct and total effects. These multiple effects were analysed relative to the reference category of parental education (primary or none), following Hayes and Preacher (2014) study on categorical independent variables on a mediation model.

Moreover, in a non-linear mediation model, the two methods commonly used to derive the mediating effect in linear models are no longer equivalent (MacKinnon 2008, p. 305). Mackinnon (2008) demonstrates that the product of the slope in the mediation (a) with the slope of the mediator in the outcome (b) is the most accurate method in this scenario (i.e. $a_i \cdot b$ from the coefficients depicted in equations (3) and (2)). This is contrary to the alternative method, which uses the difference of the total effect (c_i) and the direct effect (c'_i) (i.e. $c_i - c'_i$) from equations (1) and (2).

The relative effects the reference group of parental education (primary or none) are

Relative Indirect Effects: $a_1 \cdot b, a_2 \cdot b, a_3 \cdot b$
Relative Direct Effects: c'_1, c'_2, c'_3
Relative Total Effects: $c_1, c_{2,} c_3$ or $a_1 \cdot b + c'_1, a_2 \cdot b + c'_2, a_3 \cdot b + c'_3$.

Consequently, primary and secondary effects relative to the reference group of parental education (primary or none) are

Primary Effects (secondary education): $\left(\frac{a_1 \cdot b}{a_1 \cdot b + c'_1}\right) \cdot 100$

Primary Effects (vocational qualifications): $\left(\frac{a_2 \cdot b}{a_2 \cdot b + c'_2}\right) \cdot 100$

Primary Effects (university degree): $\left(\frac{a_3 \cdot b}{a_3 \cdot b + c'_3}\right) \cdot 100$

Secondary Effects (secondary education): $\left(\frac{c'_1}{a_1 \cdot b + c'_1}\right) \cdot 100$

Secondary Effects (vocational qualifications): $\left(\frac{c'_2}{a_2 \cdot b + c'_2}\right) \cdot 100$

Secondary Effects (university degree): $\left(\frac{c'_3}{a_3 \cdot b + c'_3}\right) \cdot 100$.

The primary effects are the relative proportion mediated by academic performance in the association between social background and educational transition. Correspondingly, secondary effects are the relative proportion of direct effects (c'_i) relative to total effects. The total relative effects are the result of adding both direct and indirect effects.

The proportion mediated (PM) is subject to limitations in the data and the resulting estimates. The following scenarios should be considered: i) A large sample size is required (recommended N=1,000); ii) its values tend to be small; iii) when the mediator is distal (in time) to the outcome, the PM tends to be weaker, which might be indicating the effect of other mediating variables; iv) opposite signs of the direct and indirect effects might result in a PM over 100% or close to zero when both effects have similar size; and v) the measure used to compute the proportion mediated[110] might not yield similar results under non-linear models.

I argue that the results of the mediation model overcome all the aforementioned scenarios, which might impact the robustness of the proportion mediated as a valid estimate. To begin with, the sample size is significantly large (even when the effect of missingness on the sample is taken into account): over 100,000 students completing secondary education. The results of the model showed mediating effects of meaningful size, given that all estimated coefficients related to the proportion mediated are positive. It was part of the aim of the chapter to test two measures of academic performance. The first was PSU scores, which were close to the time the transition to higher education occurred, as the cohort transitioned a year after completing secondary education. The second was SIMCE scores, which were distant the time the transition to higher education occurred, as SIMCE tests were applied in grade 8, 4 years later students decided to go to higher education. The relative proportion mediated used the equation $\hat{a}_i\hat{b}/(\hat{c}'_i + \hat{a}_i\hat{b})$, and this measure was robust to the

[110] MacKinnon (2008, p. 307) specifies that there are three equivalent forms to compute the proportion mediated, such as $1 - (\hat{c}'/\hat{c}), \hat{a}\hat{b}/\hat{c}, \hat{a}\hat{b}/(\hat{c}' + \hat{a}\hat{b})$.

use of standardised[111] or unstandardized coefficients in its calculation (MacKinnon, 2008, p.315).

In the following section, we develop a methodology to correct for attrition in the students' cohort, as 64% of them graduated after 12-years of compulsory education.

6A.2.2 Sample selection on the non-linear mediation model

Sample selection was corrected within the mediation model by including a selection specification for the students graduating from secondary education in 2013, in equation (4).

$$Probit(S^*) = i_4 + d_1 \cdot SE + d_2 \cdot VQ + d_3 \cdot UD + d_4 \cdot G + d_5 \cdot LPB + d_6 \cdot LPG + d_7 \cdot TM + d_8 \cdot SG4 + \epsilon_4 \quad (4)$$

where S^* is the underlying continuous latent variable of the likelihood of completing school in 2013, approximated by a binary variable. A set of 3-dummy variables, as used in equations (1) to (3), accounted for parental education such as secondary education (SE), advanced vocational qualifications (VQ) and university degree (UD). The rest of the variables were introduced into the model in order to identify the system of regression models alongside to equations (2) and (3). G is a dummy variable for girls. The instruments are: LPB is the proportion of boys in the labour force, LPG is the proportion of girls in the labour force, and TM is the proportion of teen mothers. All of the instruments were measured in 2011 for teens aged 15–19, when the student cohort was aged 15–16. $SG4$ is the performance in SIMCE scores in grade 4, at age 9–10. The residual term is $\epsilon_4 \sim N(0,1)$ and i_4 is the intercept of the specification.

[111] Standardised coefficients are necessary in alternative measure to calculate the proportion mediated.

The linear model for SIMCE scores in grade 4 is:

$$SG4 = i_5 + o_1 \cdot SE + \ o_2 \cdot VQ + \ o_3 \cdot UD + \ o_4 \cdot G + \epsilon_5 \qquad (5)$$

which as in previous specifications is regressed upon parental education levels, controlling by student gender. The residual term is $\epsilon_5 \sim N(0, \sigma_{SG4})$ and i_5 the intercept.

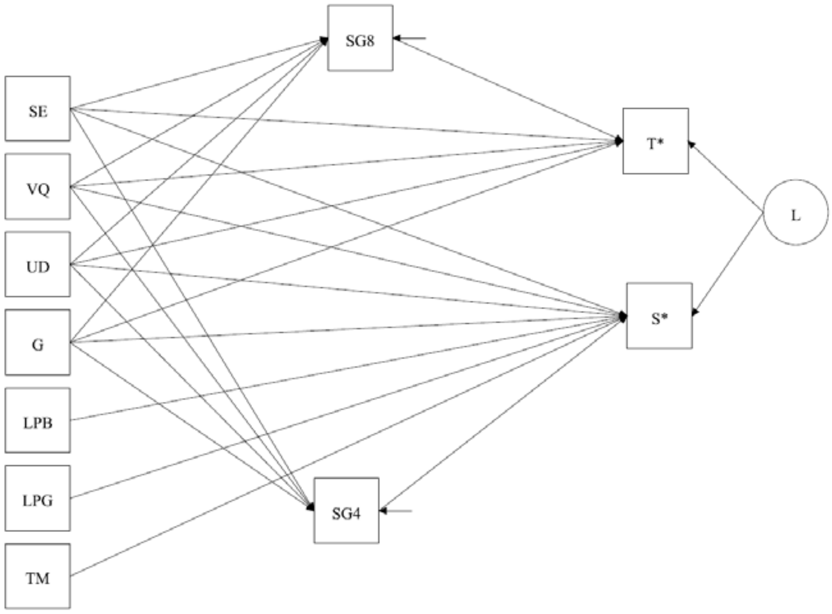
In order to be able to correct for sample selection, the residual terms in the transition to higher education (equation (2)) have to be correlated to the residual term in equation (4), which is secondary education completion (sample-selection). According to Skrondal and Rabe-Hesketh (2004, pp. 107-108); Muthén, Muthén and Asparouhov (2016), the correlation was induced by adding a common latent factor (L) among the residual terms in equations (4) and (2). In the selection equation (4), the factor loading (λ_{S^*}) was set to one to identify the specification.

$$\epsilon_2 = \lambda_{T^*} L + \varepsilon_2 \qquad (6)$$

$$\epsilon_4 = \lambda_{S^*} L + \varepsilon_4 \ , \ \ \lambda_{S^*} = 1 \qquad (7)$$

From equations (6) to (7), L, ε_2, and ε_4 follow a normal distribution of $N(0,1)$, because these terms came from probit models, in equations (2) and (4). Figure 6A.3 shows how sample selection, equation (4), was introduced into the mediation model along with equation (5), which mediates secondary education completion.

Figure 6A.3: Mediation model with sample-selection



Finally, a correlation among residuals, from equations (2) to (4), was calculated to evaluate whether the correlation was significantly different from zero, with the purpose to verify that the estimates found in the model were affected by sample selection.

$$\rho(\epsilon_2, \epsilon_4) = \frac{\lambda_{T^*}}{\sqrt{(\lambda_{T^*}^2+1)*(2)}}$$

The computation of primary and secondary follows the same coefficients used in the non-linear mediation model presented above.

6A.2.3 Two-level multiple imputation

The multiple imputation mechanism used an unrestricted two-level model for a continous variable, as it is SIMCE scores, and for a categorical variable in the case of parental education groups. The following imputation methods come from Asparouhov and Muthén (2010, pp. 10–11); it is a description of multilevel imputation implemented in Mplus.

The imputation of SIMCE scores as a continuous variable follows a two-level factor analysis model such as

$$SS_{ij} = \mu_j + \lambda_w \vartheta_{wij} + \lambda_b \vartheta_{bj} + \varepsilon_{bj} + \varepsilon_{wij} \quad (8)$$

where SS_{ij} is SIMCE scores for student i in school j. The latent factor ϑ_{wij} is the within level factor variable for student i in school j, and λ_w is the within level associated parameter. While the latent factor ϑ_{bj} is the between level factor variable for school j, and λ_b is the between level associated parameter. The residuals for the within and between levels are ε_{wij} and ε_{bj} respectively.

The imputation method relies on the unrestricted mean and variance-covariance two-level model.

The convergence of the imputation two-level model was corrected by looking at the potential scale reduction (PSR) for every 100 iterations for the first imputed dataset. When the PSR is between 1.0 and 1.10, the number of the iteration is selected to set the value of the thinning interval in the subsequent imputation, which will save a new data set at the selected n^{th} thinning cycle (Enders, 2015). Finally, after the adjustment of the thinning interval, the imputation creates the selected number of datasets for the imputation, i.e. 100 in this chapter.

The imputation of parental education groups as a categorical variable follows a two-level factor analysis model, described in Asparouhov and Muthén (2010, pp. 14–16), such as

$$PE_{ij}^{*} = \lambda_w \vartheta_{wij} + \lambda_b \vartheta_{bj} + \varepsilon_{bj} + \varepsilon_{wij} \qquad (9)$$

where PE_{ij}^{*} is the assumed unobserved continuous variable of ordinal groups of parental education, underlying a normal latent variable for student i in school j. Similar to equation (8), the latent factor ϑ_{wij} is the within level factor variable for student i in school j, and λ_w is the within level associated parameter. While the latent factor ϑ_{bj} is the between level factor variable for school j, and λ_b is the between level associated parameter. The residuals for the within and between levels are ε_{wij} and ε_{bj} respectively.

For the categorical variable PE_{ij} in the model, that takes the values $v = 1, \dots, 4$, for the each of the 4 parental education groups, under the assumption of the underlying continuous latent variable PE_{ij}^{*}, the threshold parameters ψ_{1v}, ψ_{2v} and ψ_{3v} are such that

$$PE_{ij} = v \iff \psi_{v-1} \leq PE_{ij}^{*} \leq \psi_v \qquad (10)$$

The imputation method is an unrestricted mean and covariance model for a categorical two-level variable. The thinning interval was adjusted for parental education groups for the computation of the imputed datasets in SIMCE scores, following Enders (2015).

Chapter 7: Appendices

7A.1 Appendix

Table 7A.1 Proportion of students that transitioned to HE by school characteristics in grade 8

School characteristics	Transition to HE (row %)	Total
School Funding		
Municipal	47.9	55,441
Voucher	61.6	66,361
Private	83.9	12,793
Location area		
Rural	43.3	12,618
Urban	59.6	121,977
Total students	**58.1**	**135,495**

Table 7A.2 Students' SIMCE scores by school characteristics in grade 8

School characteristics	N	Mean	Standard Deviation
School Funding			
Municipal	49,985	-0.34	0.95
Voucher	64,686	0.09	0.94
Private	12,163	0.92	0.81
Location area			
Rural	11,766	-0.32	0.91
Urban	115,068	0.03	1.00
Total students	**126,834**	**0.00**	**1.00**

Table 7A.3 Logistic mediation model

	Linear model (SIMCE scores)	Logit model (Transition)
	Coefficients	
Girl	-0.03	0.02
	(.01)	(.01)
SIMCE G8 scores (std.)	-	0.55
		(.01)
Parental Education		
Primary or none (reference)		
Degree	1.18	0.95
	(.01)	(.02)
Advanced Vocational Q.	0.67	0.68
	(.01)	(.02)
Secondary Education	0.38	0.39
	(.01)	(.02)
Intercept	-0.44	-0.04
	(.01)	(.01)
McKelvey and Zavoina's R-square		0.14
R-square	0.15	
N	121,088	

Note: Standard errors in parentheses. All coefficients are significant ($p<.001$), except 'girl' in the logit model ($p>0.05$).

Table 7A.4 Multinomial logistic mediation model by type of higher education institution

	Linear model (SIMCE scores)	Transition to Vocational college	Transition to Private university	Transition to Traditional university
	Coefficients			
Girl	-0.03	-0.06	0.29	-0.07
	(.01)	(.02)	(.02)	(.02)
SIMCE G8 scores (std.)	-	-0.05	0.68	1.37
		(.01)	(.01)	(.01)
Parental Education				
Primary or none (reference)				
Degree	1.17	-0.47	1.85	1.34
	(.01)	(.04)	(.03)	(.03)
Advanced Vocational Q.	0.67	0.07	1.44	1.05
	(.01)	(.03)	(.03)	(.03)
Secondary Education	0.38	0.15	0.82	0.66
	(.01)	(.02)	(.03)	(.02)
Intercept	-0.44	-0.75	-2.06	-1.55
	(.01)	(.02)	(.03)	(.02)
R-square	0.15	-		
N[112]	121,088	24,528	18,483	27,778

Note: Standard errors in parentheses. All coefficients are significant (p<.001).

[112] In the multinomial regression, institution-enrolled students add to the students that did not make the transition to higher education (50,299), yielding the total students (121,088) reported in the linear regression of SIMCE scores.

Table 7A.5 Multinomial logistic mediation model by type of higher education programme

	Linear model (SIMCE scores)	Transition to Vocational programme	Transition to Undergraduate programme
	Coefficients		
Girl	-0.03	-0.11	0.15
	(.01)	(.02)	(.02)
SIMCE G8 scores (std.)	-	-0.01	1.13
		(.01)	(.01)
Parental Education			
Primary or none (reference)			
Degree	1.18	-0.41	1.70
	(.01)	(.03)	(.03)
Advanced Vocational Q.	0.67	0.12	1.33
	(.01)	(.03)	(.03)
Secondary Education	0.38	0.16	0.80
	(.01)	(.02)	(.02)
Intercept	-0.44	-0.61	-1.28
	(.01)	(.02)	(.02)
R-square	0.15	-	
N[113]	121,088	27,619	43,170

Note: Standard errors in parentheses. All coefficients are significant (p<.001), except SIMCE scores in the transition to vocational programme (p>0.05).

Table 7A.6 Average partial effects of the transition to higher education institutions

	Average Partial Effects (APEs)		
	Vocational college	**Private university**	**Traditional university**
Parental Education			
Degree	-0.14	0.20	0.13
	[-.150, -.136]	[.189, .205]	[.119, .134]
Advanced Vocational Q.	-0.06	0.13	0.09
	[-.071, -.055]	[.121, .135]	[.085, .100]
Secondary Education	-0.02	0.06	0.06
	[-.023, -.011]	[.053, .062]	[.055, .066]

Note: Confidence intervals at 95% in square brackets. Relative to primary ed. or none.

[113] In the multinomial regression, programme-enrolled students add to the students that did not make the transition to higher education (50,299), yielding the total students (121,088) reported in the linear regression of SIMCE scores.

Table 7A.7 Average partial effects of the transition to higher education programmes

	Average Partial Effects (APEs)	
	Vocational programme	**Undergraduate programme**
Parental Education		
Degree	-0.15	0.33
	[-.160, -.144]	[.319, .337]
Advanced Vocational Q.	-0.06	0.22
	[-.071, -.054]	[.214, .231]
Secondary Education	-0.02	0.12
	[-.024, -.011]	[.115, .127]

Note: Confidence intervals at 95% in square brackets. Relative to primary ed. or none.

7A.2 Methodological Appendix - Chapter 7

7A.2.1 Cross-classified mediation model

The cross-classified mediation model with random intercepts is defined below. The description of the model considers model 1, which encompasses the variables of the mediation model seen in Chapter 6, without any school covariates.

$$Probit(\mathrm{T}^*) = i_1 + c_1' SE + c_2' VQ + c_3' UD + c_4' G + bSG8 + \upsilon_{grade8}^{(2a)} + \upsilon_{grade12}^{(2b)} + \epsilon_1 \quad (1)$$

$$\upsilon_{grade8}^{(2a)} \sim N(0, \sigma_{\upsilon(2a)}^2)$$

$$\upsilon_{grade8}^{(2b)} \sim N(0, \sigma_{\upsilon(2b)}^2)$$

$$\epsilon_1 \sim N(0, 1)$$

$$SG8 = i_2 + a_1 SE + \ a_2 VQ + \ a_3 UD + a_4 G + u_{grade8}^{(2a)} + u_{grade12}^{(2b)} + \epsilon_2 \quad (2)$$

$$u_{grade8}^{(2a)} \sim N(0, \sigma_{u(2a)}^2)$$

$$u_{grade8}^{(2b)} \sim N(0, \sigma_{u(2b)}^2)$$

$$\epsilon_2 \sim N(0, \sigma_{\epsilon_2}^2).$$

The notation for the cross-classified model follows Leckie (2013) citing Browne et al. (2001), who introduced subscripts to distinguish the cross-classified levels (primary and secondary schools) in the random effects. Here, *2a* denotes the random effect of primary schools in grade 8, and *2b* denotes the random effect of secondary schools in grade 12. In equation (1) the random effects follow a normal distribution with zero mean and an associated variance for primary schools and for secondary schools, $\sigma_{\upsilon(2a)}^2$ and $\sigma_{\upsilon(2b)}^2$ respectively. At the individual level, the residual term, ϵ_1, follows a $N(0, 1)$, as usual for probit specifications. In equation (2) random effects are defined akin to the random effects in equation (1). However, at the individual level, as $SG8$ is a continuous variable (SIMCE scores in grade 8), the residual term, ϵ_2, has an associated variance which follows a normal distribution $N(0, \sigma^2_{\epsilon_2})$.

The remaining variables in equations (1) and (2) are the same variables already defined in Chapter's 6 methodological appendix. T^* denotes the transition to higher education, from an underlying continuous variable, approximated by an observed binary variable T. $SG8$ is the academic performance in SIMCE tests in grade 8. The predictors are 3 dummy variables from the 4-categories of parental education in which low parental educational attainment (primary or none) is the reference group. The rest of the parental educational levels are secondary education (SE), advanced vocational qualifications (VQ) and university degree (UD). G is a binary variable for girls, which is added as a control at the individual level. Lastly, the intercepts are defined as i_1 and i_2.

Model 1 does not include school covariates as models 2 to 4b do. These covariates were not considered in the derivation of primary and secondary effects, neither were they considered in the derivation of direct and indirect effects of the mediation model[114]. Their role in the model is to operate exclusively as controls.

Therefore, the relative effects on the reference group of parental education (primary or none), from equations (1) and (2) above, are

Relative Indirect Effects: $a_1 \cdot b, a_2 \cdot b, a_3 \cdot b$

Relative Direct Effects: c_1', c_2', c_3'

Relative Total Effects: $c_1, c_{2,} c_3$ or $a_1 \cdot b + c_1', a_2 \cdot b + c_2', a_3 \cdot b + c_3'$

[114] In contrast, in counterfactual mediation the coefficients of the covariates are part of the derivation of direct and indirect effects in a mediation model.

From which primary and secondary effects, relative to the reference group of parental education (primary or none), are

Primary Effects (secondary education): $\left(\frac{a_1 \cdot b}{a_1 \cdot b + c_1'}\right) \cdot 100$

Primary Effects (vocational qualifications): $\left(\frac{a_2 \cdot b}{a_2 \cdot b + c_2'}\right) \cdot 100$

Primary Effects (university degree): $\left(\frac{a_3 \cdot b}{a_3 \cdot b + c_3'}\right) \cdot 100$

Secondary Effects (secondary education): $\left(\frac{c_1'}{a_1 \cdot b + c_1'}\right) \cdot 100$

Secondary Effects (vocational qualifications): $\left(\frac{c_2'}{a_2 \cdot b + c_2'}\right) \cdot 100$

Secondary Effects (university degree): $\left(\frac{c_3'}{a_3 \cdot b + c_3'}\right) \cdot 100.$

7A.2.3 Multinomial mediation model

The multinomial mediation for the three types of higher education institutions (i.e., vocational college, private university and traditional university) is

$$Multinomial\ logistic(\text{HEI}) = i_1 + c_1'SE + c_2'VQ + c_3'UD + c_4'G + bSG8 \quad (4)$$

$$SG8 = i_2 + a_1SE + \ a_2VQ + \ a_3UD + a_4G + \epsilon_2 \quad (5)$$

$$\epsilon_2 \sim N(0, \sigma_{\epsilon_2}^2)$$

where the reference category in (4) is the group of students that did not make the transition to higher education, after compulsory education. The independent variables are the exact same variables as in subsection 7A.2.1, which are parental educational levels, and gender (girls). As above $SG8$, is SIMCE scores in grade 8.

Primary and secondary effects were derived separately for each of the three types of higher education institutions. These effects, as in subsection 7A.2.1, are the percentage of the indirect and direct effects, in relation to the total effects for each of the higher education institutions.

Analogous, to the model of type of higher education institution, the mediation model for type of programme (vocational or undergraduate) is

$$Multinomial\ logistic(\text{P}) = i_1 + c_1'SE + c_2'VQ + c_3'UD + c_4'G + bSG8 \quad (6)$$

$$SG8 = i_2 + a_1SE + \ a_2VQ + \ a_3UD + a_4G + \epsilon_2 \quad (7)$$

$$\epsilon_2 \sim N(0, \sigma_{\epsilon_2}^2)$$

where the reference category in (6), is the group of students that did not enrol in a higher education programme.