Universidad de Chile

Facultad de Economía y Negocios

Escuela de Economía y Administración

"LITERACY AND EDUCATION IN CHILE: 1840-1970 COHORTS"

Seminario para optar al título de Ingeniero Comercial, Mención Economía

Alumno: Mendel Steinsapir Carr-Rollit

Profesor guía: Javier Núñez Errázuriz

Santiago de Chile, Enero 2017

Agradecimientos

Me gustaría agradecer a mi familia por su apoyo, cariño y preocupación. A mis amigos, que siempre han estado conmigo, sobre todo cuando las cosas se han puesto cuesta arriba. También me gustaría agradecer a mi profesor guía, Javier Núñez, por su paciencia, comprensión y buenos consejos.

Abstract

La evolución del capital humano en la historia ha sido una materia compleja de estudio ya que carece de una definición absoluta y por la limitada información disponible al respecto. Sin embargo, el alfabetismo es uno de sus determinantes que tiene una relativa mejor medición histórica gracias a los censos de población. Esta tesis estudia la evolución del alfabetismo en Chile a través del análisis *age-period-cohort*, usando datos de censos nacionales desde 1907 hasta 2002, capturando cohortes nacidos desde la década de 1840 hasta el nacido en la década de 1970. La metodología utilizada se sobrepone al hecho de que históricamente el criterio para categorizar a un individuo como analfabeto ha cambiado, ya que se construye una corrección en los datos para estandarizarlos y así hacerlos comparables. La evolución del alfabetismo en Chile refleja el desarrollo del sistema educacional chileno en cuanto a cobertura y calidad, dado los esfuerzos realizados por el Estado durante este largo periodo. Partiendo de aproximadamente un 40 por ciento de gente alfabetizada en 1907, hasta los números vistos hoy en día, donde más del 97 por ciento de los últimos cohortes es capaz de leer y escribir, Chile ha convergido con países desarrollados.

The evolution of human capital has always been a complex matter of study since it lacks of an absolute definition and also because of the limited data available, more so in developing countries. Nevertheless, literacy is one of its determinants that has a relatively better historic measurement due to population censuses. This thesis studies the evolution of literacy in Chile by an age-period-cohort analysis approach, using data of censuses from 1907 to 2002, capturing cohorts born as early as the 1840 decade to the cohort born in the decade of 1970. The methodology used overcomes the fact that historically the criteria to categorize some individual as illiterate has changed, constructing a correction in order to standardize the results and making them comparable. The evolution of literacy reflects the development of the Chilean educational system, its coverage and quality due to the efforts made mostly by the State throughout this long period. Starting from roughly a 40 percent of literates in the 1907 census, to the numbers seen nowadays, where more than 97 percent of

the population of the latest cohorts is able to read and write, Chile has converged with developed countries.

Index

I.   Introduction

The development and accumulation of human capital has been a prior subject of study in a wide range of social sciences, especially in Economics and Economic History. Literacy is known to be a crucial element of human capital formation[1] and one of the main objectives of any educational system. This thesis tries to emphasize and study its impact.

The nature of human capital makes it a difficult subject matter to study, since it often represents unobservable characteristics, and therefore, research is usually relying on limited data or in assumptions far from reality. The literacy rate of a certain population, however, is often found in data series from censuses and population surveys. For the Chilean case, these are available since the beginning of the second half of the nineteenth century, specifically since the "III National Census of Chilean Population" of 1854, where the question of whether a person could read (and write) was introduced.

This thesis seeks to provide a historic consistent and comparable measurement of literacy rates in Chile since the cohorts born in the second half of the nineteenth century until the 1970s, using data available in the national censuses. Chilean censuses constitute a particular case in terms of robustness for the Latin American region because of their regular periodicity, wide coverage and official character. Also because of the political stability that Chile experienced,[2] in comparison with other countries, in periods where gathering nationwide information was a complicated aim to achieve due to the lack of resources and adequate technology.

Literacy census data are not directly comparable across censuses. This issue arises from the fact that between 1907 and 1940, the question about literacy was only about reading abilities, and since 1952 it included altogether a person's ability to read and write. In order to make both periods comparable, a correction for the first period had to be constructed using data of previous censuses that contained information for both reading and writing separately.[3]

---

[1] See Barro (1991).
[2] For a critical look on this subject see Valdivia (2009).
[3] This will be further addressed in the fourth section.

Data included in censuses since 1907, where the distinction of literacy rates by groups of age was introduced, until 2002, separated by ten year intervals approximately, takes account of cohorts born in the 1840 decade (the 60-69 years old group for the 1907 census), until cohorts born as recently as the 1970 decade (20-29 years old for the 2002 census), giving that way a broad perspective of the evolution of literacy in Chile, by birth cohorts.

The methodology applied takes into account three different effects on adult literacy: age, period and cohort. As for the first one, the results conclude that there is no statistical significance to include age in the final model, underlying that, since the study is made for people over 20 years old, yet a person can learn to read (and write) but is not likely to lose the ability throughout his or her life, not at least until the age this work follows up, 69 years old. For the other two effects, the statistical measures are conclusive of their importance in order to explain the rising literacy rates observed between the years of study. Although the period effect is significant, which captures effects for institutional changes and breaks, the predominant is the cohort effect. That means, the most important matter to determine if a person is likely to be literate is the moment this person was born, the socioeconomic environment in which he or she grew up and the education being imparted at the time.

Literacy is tightly related with schooling, being a necessary tool for the execution of elementary tasks with the subsequent repercussion on the labor market and the labor conditions to which most of the population is able to access.[4] However, until today there are no studies that thoroughly look up the Chilean case assessing the impact of schooling in the capacity to read and write. Therefore, the extension and evolution of the public educational system is important to understand the historic development of this matter. The Primary Instruction Law of 1860 (prior to this the state efforts were sporadic and of little significance), and later the Mandatory Primary Instruction Law of 1920 and the 1965 educational reform are the milestones of the public education system in the late nineteenth and twentieth centuries.

One of the main objectives is to relate the evolution of literacy rates with the public policy efforts made in educational matters during the period, as a portrait of the consolidation of the Chilean public educational system and its coverage. The central hypothesis of this thesis

---

[4] See Charette and Meng (1998).

is that the efforts made by the authorities through the educational system, in terms of coverage and quality of the education imparted, had a positive and significant impact on the literacy rates of Chilean population.

The thesis is structured as follows: The second section provides a snapshot of the Chilean educational system over time, particularly considering the landmarks that influenced the evolution of literacy. Section three is a literature review of previous studies about the relationship between education and literacy. Section four describes the data used and its treatment. The fifth section explains age-period-cohort analysis focusing on the context of this work, then shows the results of the study with the descriptive analysis and also with the econometric and statistical features with its interpretation, and a parallel between this thesis and a previous related study about numeracy is done. Finally, section six concludes.

II.   A snapshot of the Chilean educational system

Regarding literacy, one feasible assumption is that most people who have completed primary education should be able to read (and write). Although this is far to be an exact equivalence, some countries with poor data from censuses even use primary schooling attainment as a proxy for literacy.[5] Hence, this section takes into account principally the landmarks and evolution of Chilean primary education and especially the ones related with literacy.

Mass education is often seen as a crucial political issue, therefore, since the constitution of Chile as an independent country in the early nineteenth century, giving population the tool of education became a matter of public discussion. Before 1840 what was done in the field was very little because there was no educational system, there was little or none resources and little political experience. In sum, the country was giving its first steps with the challenges that this represents.

The development of education was, at first, a debate held by the political elites. They saw education mainly as a tool to give the population some basic abilities to maintain the relative order in which the country functioned since its independence. There was some mistrust about the role that increasing literacy rates on popular sectors may cause on their expectations.[6] But slowly the discussion moved into a more complex level, and the government, or at least some parts of the political sphere, became aware of the importance of education as a driving force of the country's development. The first evidences of these concerns were the creation of the Ministry of Justice, Cult and Public Instruction in 1837 and the Normal School of Preceptors in 1842. The state involvement and the formation of well-prepared teachers were two of the main subjects that arose from this initial diagnose.

The Primary Instruction Law of 1860 came to organize these efforts into a more orderly fashion. It detailed matters as administration of schools, depending on the municipalities under the state supervision. Also the financing, which mixed national treasury contribution and municipalities' budget. Duties of teachers and the schools visitors (who were government supervisors, figure existing since 1846 but by this law standardized) were also

---

[5] UNESCO (2004), pp. 127-128.
[6] For further explanation see Egaña (2000).

defined, and the number of schools that should exist given certain amount of children on an area, among other issues.

One key aspect of this law is that it stated: "The primary instruction will be given under the direction of the State"[7] and "Instruction given in accordance with this law will be free and will contemplate people of both sexes"[8]. It was, nonetheless, the culmination of an almost two decade period of discussion. This was mainly because the conservative party pressured against the liberal's financing proposal, which claimed that the resources should come from the state. Among the conservative parties there were two views. President Montt's supporters claimed that the resources should come from a direct taxation raised and managed by the municipalities. The ultra-conservatives, by the other hand, held that there was no sense on financing coming from the wealthiest because their children did not attend public schools, disguising their concerns as a disagreement about the proposed taxation. At the end, the law's financing was a mixture of the liberal's and Montt's proposals plus donations. The administration, financing, and teaching methods of private schools was left to their choice, under some state supervision relative to order and moral. Even though this law does not recognize education as a universal right, it shows an advance in Chilean history as it is the first time a discussion of such matter was held and a law acknowledging its importance was passed.

In the next decades some minor reforms were introduced in the primary educational system. One example of that is the change in 1881 of alternate schools (which taught boys in the morning and girls in the afternoon or in separate days) to mixed schools, giving especially to rural areas more students and therefore the possibility to keep their schools open and increase coverage.

As it can be seen in graph 1, the Primary Instruction Law of 1860 had a positive impact on enrollment rates. Though they do not show real attendance, it is a proper indicator when attendance rates are not available. The decade of 1870 had a high growth rate, as the law disposed the foundation of a school for every 2.000 people of each gender. In the following decade the rate decreased mostly due to the Pacific War and the subsequent male

---

[7] *Ley General de Instrucción Primaria* (1860), art. 1.
[8] *Ley General de Instrucción Primaria* (1860), art. 2.

enrollment in the army, with children either going to the war or dropping out of school to help in home labors or entering forcedly into the labor market.

Graph 1: Children between 7 and 15 years old enrolled in primary schools, 1855-1895
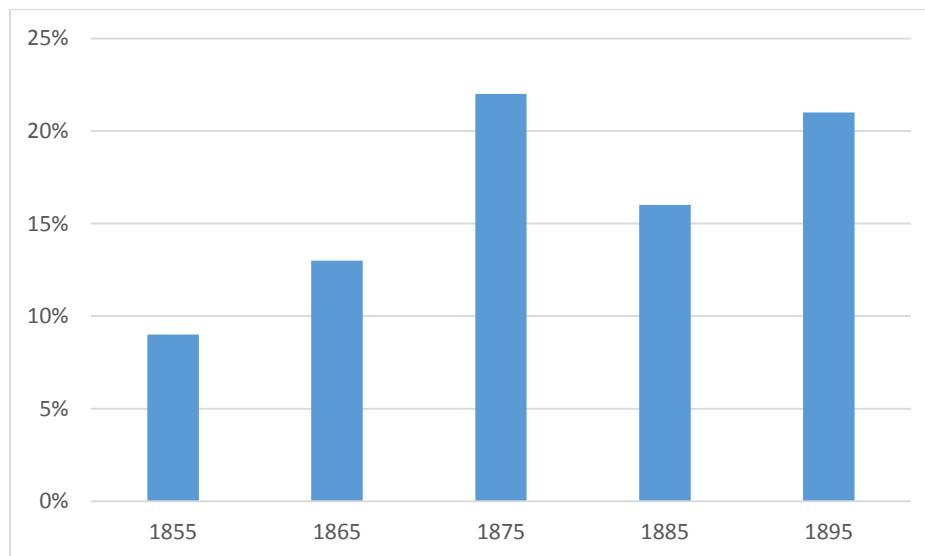
Source: Serrano, Ponce de León and Rengifo (2012), Vol. II., pp. 69.

Claudio Matte, educationist and rector of the Universidad de Chile in 1926-1927, published in 1884 in Leipzig the "*Nuevo método para la enseñanza simultánea de la lectura y escritura*" ("New method for the simultaneous teaching of reading and writing"). It was adopted in 1894 by the government as the official text for the learning of reading and writing, and was so for many decades. It summed up the experience Matte got from a long journey studying the primary educational systems of developed countries in Europe, modernizing the way literacy was introduced in schools all over the country.

His method, inherited from the German, had distinctions in aspects regarding the way reading and writing was taught. It introduced a way in which the student first had to be able speak properly, then had to learn the sounds (not the names) of the letters, and then the student was exposed to the words. Also, the writing was taught simultaneously with the reading on each letter, as a method to reinforce the acknowledgment of the letter and also not be left behind in the writing learning process. This was, by Matte's words, a faster and better method to teach reading and writing to Chilean children, especially to the lower

classes, since it relied less on memory and it made sure before teaching to read and write that the student was able to speak properly.[9]

Matte and other educationists from the period paid attention to a controversial issue. The fact that even when there were more schools did not mean that the attendance was higher. Some claimed that the mere fact of having more schools would make children to attend. Others argued that the existence of these schools, without the prior knowledge of the families of the importance of schooling and the guarantee of rights such as the regulation by law of children's labor, was irrelevant. Aspects of this debate existed since the discussion of the Primary Instruction Law, but only in the following decades it was taken into account seriously by the authorities.

These concerns were materialized into the creation of the Mandatory Primary Instruction Law in 1920. To stress how difficult was to pass this law in the parliament, it is enough to note that the first project was sent to congress in 1902, but as the Primary Instruction Law, almost two decades of discussion were needed. The conservatives, especially those tied with the Catholic Church and the Church itself, argued that the mandatory nature in hands of the State liberals asked for was against Freedom of Education, which was one of their most important premises for this period. In contrast, the liberals premise was the *"Estado Docente"* ("Teaching State") since the law of 1860, where the State was indeed in charge of public education, but their aims were to amplify its duties and to be able to educate a larger proportion of the population under the responsibility of the State. A main argument held in favor of Freedom of Education, besides the value of freedom itself, was that the progressive influence of the State in educational matters would lead to a secularization of the educational system and therefore of society, with the moral detriments that it would have for the population. This long debate finished only when the State and the Catholic Church were officially separated by the Constitution of 1925.

The Mandatory Primary Instruction Law of 1920, besides keeping the dispositions introduced by its predecessor, as gratuity and gender equality, had important new features. As the name says, primary schooling would be mandatory. It would be mandatory until the fourth grade of primary school, or until a child of thirteen years old or older was able to

---

[9] Matte (1884), pp. IV-VII.

pass the first two grades. Education was under the responsibility of parents or tutors, but this time regulated by the respective authorities, with monetary or even prison sanctions to those who did not obey the law. It forbid industrial child labor to children younger than sixteen years old if the primary schooling of the individual was not satisfactorily finished. Also, supplementary and complementary schools would be provided for adults who did not attend primary schools in early ages, essentially for the two first years of school (when basic skills of reading and writing were taught), apart from imparting in these schools lessons of a given profession *ad hoc* with the zone.

For the following decades, the educational system was ruled under the Mandatory Primary Instruction Law. Despite this, the new dispositions battled with the economic and political scenario of the nation. Chile was one of the most affected countries by the Great Depression, and the fall of the saltpeter industry in the 1920s deepened the crisis which triggered a period of political instability in the early years of the next decade.

To illustrate the profound impact of these issues on the educational system, graph 2 shows the enrollment and attendance of Chilean children to school in 1936. The persistence of children in schools was very poor, with only 9% attending sixth grade over the enrolled on first grade. The program of public schools consisted on teaching reading and writing basic skills in the first two years. There were 99.088 children attending second grade, against 241.779, this means only 40% of children attended second grade over the enrolled in first. One aspect that is important to point out is that the differences by gender are almost insignificant, with girls having slightly better rates than boys. This is consistent with the fact that the gender gap in literacy had narrowed down since the early twentieth century. Enrolling and keeping children in schools was a national issue, and it did not discriminate by gender.

The challenges that the educational system had at this point in order to solve the problem of illiteracy were two. It was imperative to do new efforts on primary schooling for children in terms of coverage and attendance, by a larger involvement of the State on this matter, with more resources allocated and political support behind. By the other hand, illiteracy became more and more a problem of adults, when primary schooling for children improved, but there was no specific policies about primary adult schooling.

Graph 2: Enrollment and attendance in primary schools, 1936
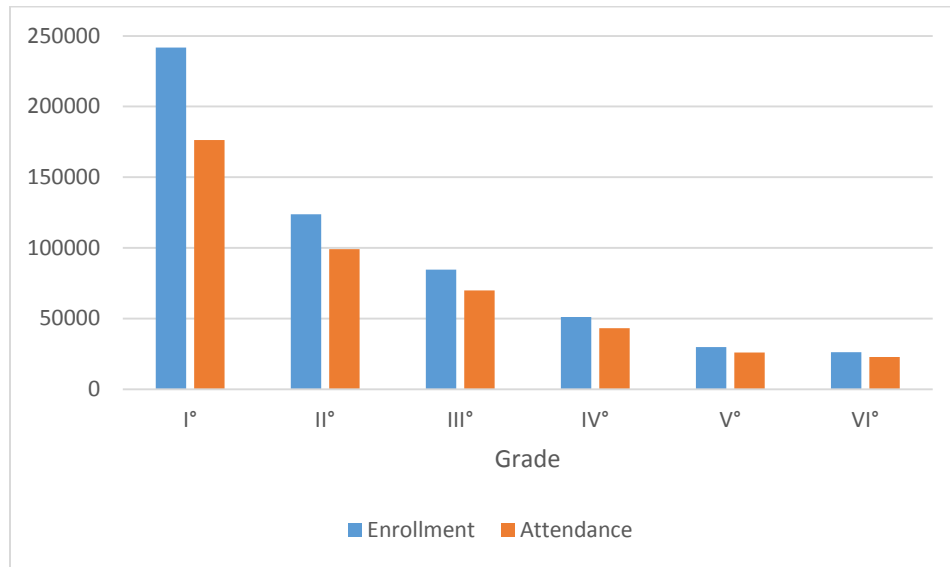
The efforts made by authorities in the decades after the Mandatory Instruction Law in order to overcome illiteracy were sporadic and decentralized. In 1942 a section for adult illiterates was created under the wing of the Ministry of Education, along with a decree in 1944 which created a Civic Group of Literacy, also aiming to teach adults. The decree was valid until 1968, when it was derogated. Sources say that this program taught around 170.000 people.[10]

The reference for this period still was the Mandatory Instruction Law of 1920. But in the 1960s, the State and private organizations like the Catholic Church got involved more directly. In 1965, under the administration of President Frei Montalva, the National Campaign for Literacy and Educational Recovery took place.

The reforms on education made by President Frei's government were not only regarding adult education. Even though the Law of 1920 was subject to some changes in the decade of 1940, the essentials remained untouched. Unfortunately, some dispositions of the Law, such as the compulsory character of primary schooling for children were not abided. Some children were not able to attend schools due to economic reasons, or because there were no schools nearby, especially in rural areas.

---

[10] Pérez and Silva (2013), pp. 90-91.

By the 1960s, there was consensus about the critical status of the educational system. The reforms included the expansion of primary schooling from six to eight years (reducing high school from six to four years). Over 2.500 schools were created, and enrollment rates also increased.

The last period that this study takes into consideration is 1970-2002. Literacy after the government of Frei Montalva became a less important issue because coverage of primary schooling for children improved greatly and the programs for adult illiterates also worked in this direction. Even so, it is important to state that the educational system had changes.

During the socialist government of President Salvador Allende, one of the iconic programs was the creation of the Unified National School. One of its objectives was to foster integral participation of social organizations in the educational system, aiming for more equity. This project never passed Congress as it arose suspicious of being a plan for implanting an ideological socialist type of education. With the dictatorship, a liberal influence came into the educational system. Therefore, the Ministry of Education delegated the administration of public schools to the municipalities. The return of democracy in the 1990s also motivated reforms in order to modernize the educational system and to align it with the global context. Proof of this is the creation of the Mandatory Secondary Instruction Law of 2003, which extended from eight to twelve the mandatory years of education.

III. Literature review

Literacy is a well-documented subject of research for it has relevance not only for economic purposes, but for other social sciences as well. The study of the evolution of literacy is related with important aspects such as human capital formation, human development and economic growth. Even though evidence in a matter like this is almost never conclusive nor categorical, there are many documents studying the relationship between literacy and progress, in a broad definition of the word.

In the 1960s, Anderson (1965) compared literacy rates and income per capita across and within countries finding a correlation between them. Although there was no causation, there was no evidence of steady economic growth with a literacy rate less than 40 percent, setting it as a threshold for the development of a country. Following his line of investigation, Blaug (1966) studied the economic benefits of adult literacy programs. He stated a trade-off between these programs and allocating resources to primary education, claiming that, though lacking reliable evidence, the economic benefits were greater for primary schooling.

Cipolla (1969) wrote about the relationship between literacy and development in western culture. Data and information was gathered from various sources[11] since the dawns of (and even before) the Industrial Revolution for a wide set of European and North American countries. He focused mostly on the nineteenth century, finding there was a strict relationship between literacy levels and the progress of education. This was motivated by the uprising economic sectors that came with the revolution, with the logical changes in the labor market due to the characteristics needed for the new jobs.

As for Latin-America, Thorp (1998) studied the economic growth and the evolution of living standards in the region in the twentieth century. Regarding literacy, he concludes that countries had the biggest increases in literacy levels on different times, with the ones of less indigenous population (and more European immigrants) having their leap in the early decades of the century. By the other hand, the ones with a larger percentage of indigenous population had it later. Another conclusion on this topic was that the gender breach on the 1990s was significantly larger in the second set of countries, which is consistent with the results seen in this thesis for Chile, qualifying in the first group. Also, for the region as a

---

[11] Such as national censuses, statistical yearbooks, recruitment papers, marriage registers, etc.

whole, the illiteracy levels on 1995 were about 11.5%, which was comparable to U.S.A's rates in 1900, constituting a picture of the relative underdevelopment of Latin-America. However, for the same decade, illiteracy levels in Chile were significantly lower (around five per cent), converging with evidence showing that nowadays, in terms of indexes regarding living standards, Chile is one of the countries with best performance in the region.[12] Núñez (2005) studied the evolution of literacy in twenty-two Latin American and Caribbean countries in the first half of the twentieth century. He gathered data from different sources[13] in order to make predictions of literacy rates that were not available in official censuses in some countries on the early decades of the century. Evidence shows that the path taken varied between countries as well as the starting points, also concluding that countries with larger proportions of Amerindian population had lower literacy rates.

Regarding gender differences, the breach has diminished in the twentieth century. This is not particular to Chile but to the whole Latin-American region, as Baten and Manzel (2009) state. Even though there still exists a gap between genders, it is low in comparison with other developing regions and it is actually near to the differences seen in Central and Eastern Europe. A possible explanation is that the region had a relatively lower historic subordination of women in the twentieth century comparing with African and some Asian countries, social constraints of the role of women that Stromquist (1992) argues as a fundamental reason for the lower female literacy rates. Also, Levy (2016) studied how the way reading and writing is taught in early childhood has a historic social construction associated with gender stereotypes and subsequent labor market discrimination. So even when literacy equity is a fit proxy of gender equality, it does not guarantee equality in the labor market in terms of wages and returns to education.

It is important to state that there has been some lines of research that try to refute the relationship between literacy and progress. Graff (1979), for instance, argues that the historic evolution of literacy in the nineteenth century was associated with spreading moral

---

[12] See appendix 1.

[13] Data was gathered from marriage registrations, crime statistics, etc. Also from urban (local) censuses.

restraints in favor of religious and political groups rather than a true interest in educating masses, putting a dubious note on the real value of literacy in that period.[14]

An important matter for this thesis is also to relate the evolution of literacy rates with numeracy, which are documented to have a close correlation, as with schooling. Manzel, Baten and Stolz (2012) studied numeracy by the age-heaping technique, which consists on measuring what percentage of a certain population rounds up its age[15] due to lack of knowledge about their exact year of birth or incapacity to calculate it. Then, this is indexed as a rate (like literacy) and used as a proxy for basic numeracy skills. Studying samples of Argentinian censuses on the second half of the nineteenth century they found out that the reports of exact-age amongst the literate were way higher than for the illiterates. With a correlation of the illiterate rates and the age-heaping index highly significant, they concluded that in fact age-heaping is an informative human capital indicator. For the Chilean case, Cardemil and Núñez (2015) studied age-heaping and education, relationship that will be further addressed and linked to this work in section five, sub-section four.

---

[14] Needless to say, these thoughts have been subject to major critique.
[15] Typically to numbers ending in 5 and 0.

IV.   Data and rates construction

The data used in this thesis belongs primarily to the ten nearly decennial official censuses conducted in Chile between 1907 and 2002.[16] The near decennial frequency of the censuses allows the construction of decennial birth cohorts that extend over several censuses, the first cohort being those born in the 1840s (who were about 60-69 years old in 1907) and the later cohort being born in the 1970s (about 20-29 years old in 2002). Although there is not always an exact ten-year interval between consecutive censuses, the differences are small. .

Chile has censuses since 1813, but only in 1854, on the third official one, the question of whether a person could read (and/or write) was introduced. Unfortunately, not until 1907 the division of literacy by groups of age was made, so it is impossible to incorporate previous censuses to the age-period-cohort analysis. However, the nineteenth century censuses had the characteristic of asking reading and writing abilities separately, being an important statistical difference between people who could only read and read and write.

Year	Reads	Writes	Reads +	Writes +	Total pop.	Total pop. +	Observations
1854 (+7)	13.47%	10.65%	16.94%	13.40%	1,439,120	1,144,393	No literacy by age. Reading and writing are separated.
1865 (+7)	17.00%	14.51%	21.14%	18.03%	1,819,223	1,463,334	Idem.
1875 (+7)	22.99%	20.29%	29.17%	25.73%	2,075,971	1,636,583	Idem.
1885 (+5)	28.93%	25.11%	35.69%	30.97%	2,527,320	2,048,987	Idem.
1895 (+5)	31.83%	28.83%	39.26%	35.56%	2,687,900	2,179,279	Idem.
1907 (+5)	39.95%	-	47.32%	-	3,249,279	2,743,747	Literacy by groups of age appear. Only reading.
1920 (+5)	50.40%	-	59.59%	-	3,753,619	3,174,685	Idem.
1930 (+5)	56.13%	-	69.64%		4,287,445	3,584,960	Only reading. Data available only for 60+ years old group (not for 60-69).
1940 (+5)	56.96%	-	66.89%	-	5,023,539	4,277,323	Idem.
1952 (+6)	-	62.92%	-	74.84%	5,932,995	4,987,964	Only reading and writing. Data available only for 60+ group.
1960 (+7)	-	65.40%	-	82.40%	7,374,115	5,852,768	Only reading and writing. Data available only for 65-84 years old group.
1970 (+10)	-	X	-	89.79%	8,884,768	6,518,004	Only reading and writing.
1982 (+10)	-	X	-	91.70%	11,329,736	8,916,323	Idem.
1992 (+10)	-	X	-	94.62%	13,348,401	10,648,647	Idem.
2002 (+10)	-	X	-	95.79%	15,116,435	12,648,761	Idem.

---

[16] The censuses were in 1907, 1920, 1930, 1940, 1952, 1960, 1970, 1982, 1992 and 2002.

Table 1: Reading and writing rates and population of Chile, 1854-2002 censuses

Source: Own elaboration from 1854-2002 Chilean national censuses.

Table 1 shows the evolution of reading and writing literacy rates for the total population since 1854 until 2002, and also for population excluding infants of certain age and below depending on how the census data was presented each time. There is a bias because there is the assumption that children under the age of the threshold are illiterates. That is why for 1970-2002 the rates for the total population are not present. Assuming that all children under the age of ten (which was the data available in those censuses) could not write (nor read) was too far from reality.

As literacy by age groups are introduced only in 1907, also the question regarding writing ability disappears until 1952. There, the criterion to consider a person literate changed to knowing both reading and writing, converging with the criterion that developed countries had been using by that time. With this methodological change arises an issue of particular significance. Comparing literacy rates since 1952 with the preceding would lack accuracy, due to the different measures. Relying on the data available from the nineteenth century, we can say these measures were statistically different.

Therefore, in order to make groups of age literacy rates comparable between these two set of censuses a correction had to take place for the first group. As mentioned before, the nineteenth century censuses had reading and writing data separately. So, using the 1895 census literacy rates separated by department, the correction consisted on regressing the writing rate against the reading rate and a quadratic term. This option was taken because it had a long range of rates, since in the large cities literacy would round up to 60 per cent, and in rural areas would be as low as 10 percent. That gave a wide perspective of the differences between these departments, being a reasonable forecaster of what should be the case for later periods. That is, it is likely that the different reading and writing rates in 1895 in Santiago, for example, which are available, are similar for the country as a whole in the period 1920-1930, since the rates that are comparable are indeed quite similar.[17]

Unfortunately, some of the censuses gather the 60 years old and up group as a whole without distinction for decades or lustrums. Therefore, there is a bias in the data impossible to disentangle, although the effect of it is uncertain. The population over 70 years old may be more likely to have health problems that make them unable to read and write. But, by

---

[17] For further explanation see appendix 2.

the other hand, the wealthier live longer so it is more possible that the fraction of age escaping the range is able to read than the others, at least in the periods in which the issue takes place.[18]

---

[18] See observations of table 1 for details.

V.   Results

V. I.   Age-period-cohort analysis

As the name indicates, there are three potential effects to take into account in analyzing long-term changes in some aspect of society: age, period and cohort. The Age (A) effect responds to any life-cycle effects associated with an individual's age and the moment of observation. The Period (P) effect is any effect associated with the moment in which the observation is executed. The Cohort (C) effect is any effect that on a group of people that share particular demographic and socioeconomic characteristics due to being born in a similar period.

Studying these effects has a well-known major problem, which is the linear dependence between them.[19] This is a matter of large research in order to have reliable results for each effect and altogether, but to this date there is no one-way-formula that provides a unique solution, even though from the literature available arises different methods to proceed.

Mason et al. (1973) studied the nature of the age-period-cohort problem of linear dependence. They argued that in cases where all three effects have distinct causal interpretations to some independent variable, an analysis omitting one of these variables would be futile and condemned to spurious results. The solution proposed was to constraint two adjacent values for each effect, relying on strong theoretical preconceptions or data-based impressions, in order to feasibly estimate differences between coefficients disentangling the linear dependence. The method is known as the Constrained Generalized Linear Model (CGLIM). This kind of model is subject to critiques due to the subjectivity of how the constraints are chosen. When they are chosen by intuitive reasons, or when they are chosen because of statistical consistency with the other models (A, P, AC, etc.), also because the variations introduced in the results by imposing constraints most of the times are not depreciable. Glenn (1976) argued that the *additivity* of these effects would lead to specification problems, stating against the assumed rigidity of each effect that underlies the CGLIM method for being rarely realistic.

---

[19] Cohort = Period + Age

Kupper et al. (1985) criticized both two-factor and APC models, claiming that the first ones are potentially misleading if the argument of "goodness of fit" was to dismiss one of the effects when there is an essentially linear pattern. By the other hand, it was argued APC models, with the various possible specifications, apart from the critiques previously pointed out, had not been subject to the proper statistical scrutiny required to assure their effectiveness. The conclusion was that any age-period-cohort analysis would be subject to objections. Even more, considering the fact that in the nature of cohort analysis there would always be missing observations that possibly hold relevant information. This is because the cohorts are in the diagonals of the data matrix and therefore the cohort in the extremes have only one observation linked to only one age group and period.

Relating age-period-cohort methodology with literacy, de Gabriel (1998) analyzed data for 1900-1950 in Spain for population from 11 to 100, capturing cohorts since the beginning of the nineteenth century until the 1930s. The author exposed that age was a relevant effect due to the biological needed maturation process for learning reading and writing, the social dimension which constitutes basically the stage in which a person goes to a primary school, and the psychological factors that produce children to have to be induced into learning literacy skills. As for period, the historical position of a country is argued to be important since the dynamics of literacy are mostly present in schools, therefore the evolution of society affects the possibilities of a certain individual to acquire literacy are also related with the period. For cohort effect, it is claimed that the characteristics of each cohort (such as size or education level) and the experienced events are also relevant to explain literacy rates. The analysis of the results reflected different patterns for cohorts and for genders as well, with a declining of literacy rates from the 20-29 years old group to the oldest, and with a rising of the rates for all periods. The findings detected influence of each effect on acquiring literacy skills, although only descriptive statistics were employed.

The way in which this problem is dealt is by comparing the results of different estimations, modeling by only one of the effects, two of them and then the three, leaving the others uncontrolled depending on each case. The APC model (the one with the three effects) was not satisfying for our purposes, because the inclusion of age effect in the estimations was counter intuitive and produced major statistic distortions. Relating this claim with the

exposed by de Gabriel, it is clear that age effect is important in early ages for all the arguments given. But, taking into consideration that this study works with data only for 20-69 years old individuals, the maturation process with the linked attendance to primary school needed to acquire literacy is already done (or not).

Also, it can be seen in the following sub-sections that the model only including age does not hold statistical significance. When associated with cohort, it changes its sign to positive, suggesting that the negative effect in the other models is due to the capture of cohort effect. Even more, taking into account theoretical reasons besides the ones regarding early ages, the effect of age on literacy should be positive. Wealthier people live longer and are more likely to be literate than the rest of the population. Adult education may be significant for the 20-29 or 30-39 age groups, then again increasing literacy rates with age. There was not found by the authors a solid reason to consider a model in which age effect is included with a negative sign, especially since the previous statistical reasons.

Therefore, in this thesis the model used is the PC, so it is not necessary to impose constraints on variables, although literature is reluctant about the utilization of models omitting one effect, as stated. Nonetheless, as it is shown afterwards, the PC model holds the best properties, follows intuition and it becomes quite clear that the inclusion of age effects deteriorate the results.

The period effect then reflects methodological variations and (or) institutional development. The methodological differences between censuses, which is the instrument used, must be treated in order to make periods comparable, so in this effect we just consider the institutional changes that may have had importance in the literacy levels of Chilean population.

The third effect is the cohort effect. It reflects the environment and socioeconomic context in which an individual was born, raised and educated, that for the case of literacy includes the public policies made in education throughout time and how population was affected by them. As it is shown in the next sub-sections, and following the previous intuition, this is the largest effect of them all.

V. II.   Descriptive analysis

	1907	1920	1930	1940	1952	1960	1970	1982	1992	2002
20-29	53,06%	65,25%	76,69%	75,56%	85,45%	88,29%	93,84%	96,51%	97,96%	98,83%
30-39	48,08%	60,89%	72,60%	71,45%	82,79%	85,35%	90,27%	93,41%	97,02%	97,65%
40-49	43,27%	56,01%	66,97%	66,17%	78,14%	82,25%	86,62%	88,65%	94,05%	96,67%
50-59	37,30%	51,39%	61,80%	60,39%	73,76%	77,29%	83,22%	84,62%	89,53%	93,77%
60-69	30,65%	43,64%	49,86%	49,28%	62,19%	68,49%	77,10%	79,69%	85,26%	88,72%

Table 2: Literacy rates by age group and year of census 1907-2002

Source: Own elaboration from 1907-2002 Chilean national censuses.

As it can be seen in table 2, literacy levels have evolved throughout time, cohorts corresponding to the diagonal elements, since censuses have roughly ten years intervals.[20] Each column shows that literacy rates are higher for younger individuals in any period given. Nonetheless, the effects that rule are the cohort and period effects rather than the age effect, meaning that older individuals have lower literacy rates because their cohort is associated with a range of literacy lower than the followings. Also, there is the distortion produced by mortality rates which is impossible to take into account.

The evolution within cohorts is not linear, although there is a trend of improvement in literacy rates from the 20-29 years old to the 30-39 years old in most of the cohorts, and there is no clear pattern for the next groups of age within cohorts. It comes to attention the diminishing of literacy levels from the 1930 to the 1940 censuses. This may correspond to period effect, such as methodological issues and measurement errors and (or) to the fact that Chilean economy was deeply affected with the Great Depression, plus a moment of political instability in the early years of the 1930s. These factors undermined the assistance and presence of middle class and poor children in schools with a subsequent impact over literacy levels.

---

[20] For example, the cohort born in the 1930 decade was 20-29 years old in the 1960 census, 30-39 years old in the 1970 census and so on.

It is relevant to point out that from the 1940 to the 1952 census there is the largest leap in literacy rates for almost every age group even though in 1952 the methodological change was introduced,[21] so this may shed light of bigger measurement problems in 1940.

Then again, visual examination does not consider the interactions between age, period and cohorts that are crucial in this kind of studies, so it gives a partial picture rather than a reliable result in terms of the true relevance of each effect.[22]

Graph 3: Literacy rates by age group and birth cohort in Chile 1840-1970
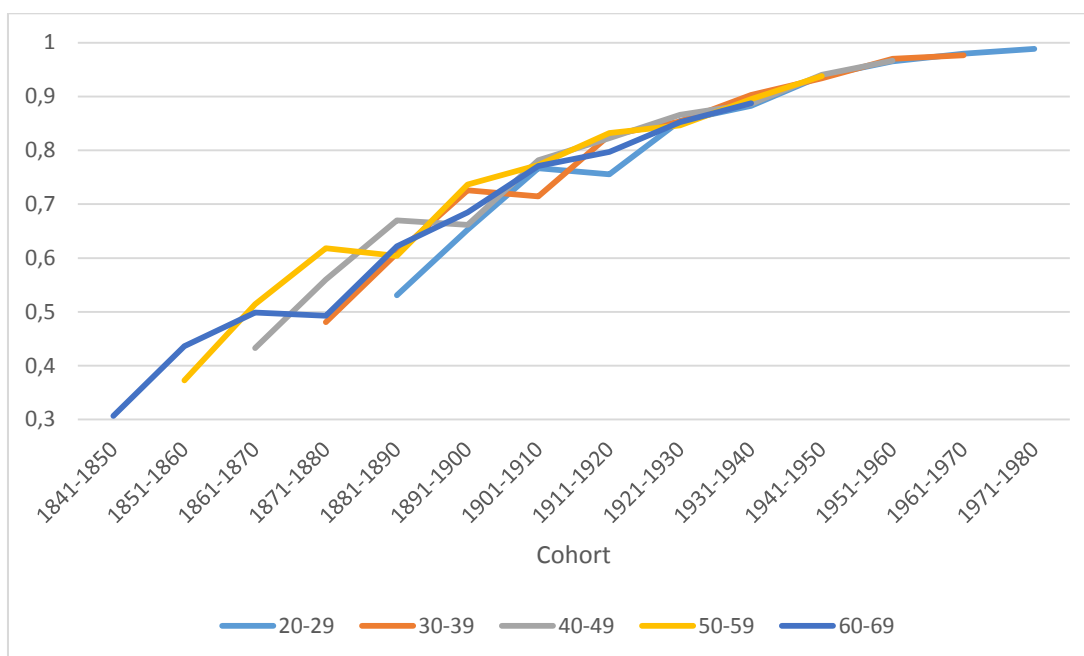

Graph 3 shows the evolution of literacy, and as pointed out before, we can see that the predominant effect is the cohort effect, presenting a concave curve. This means that the increase in literacy is faster for the older cohorts and then the trend gets smoother.

Graphs 4 and 5 show the evolution by gender, where it is clear that women start with much lower literacy rates but there is a catching up process which leads to a virtually insignificant difference in the cohort born in the 1930s. Evidence is consistent with the efforts made on female education since the last decades of the nineteenth century.

---

[21] And it is pretty far from reality to assume that the correction made in this study is exactly accurate.
[22] If taking only table 2 into account, age clearly has importance in literacy levels, which has been refuted.

Graph 4: Male literacy rates by age group and birth cohort in Chile 1840-1980
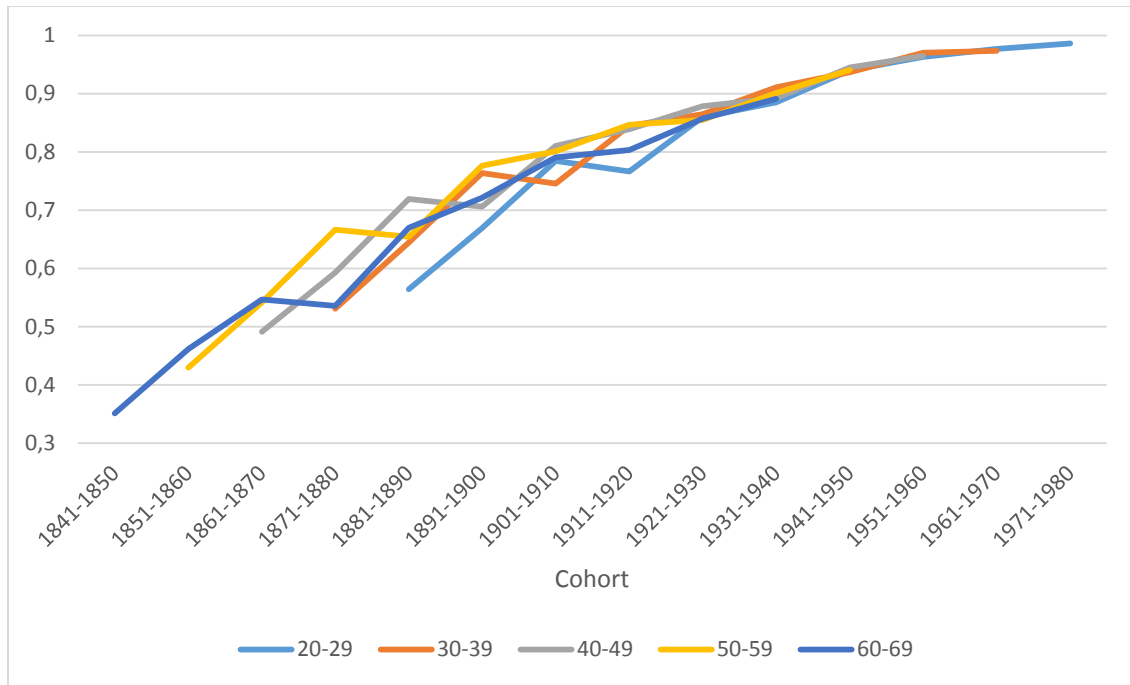
Graph 5: Female literacy rates by age group and birth cohort in Chile 1840-1980

X axis label: Cohort

Legend: 20-29 30-39 40-49 50-59 60-69
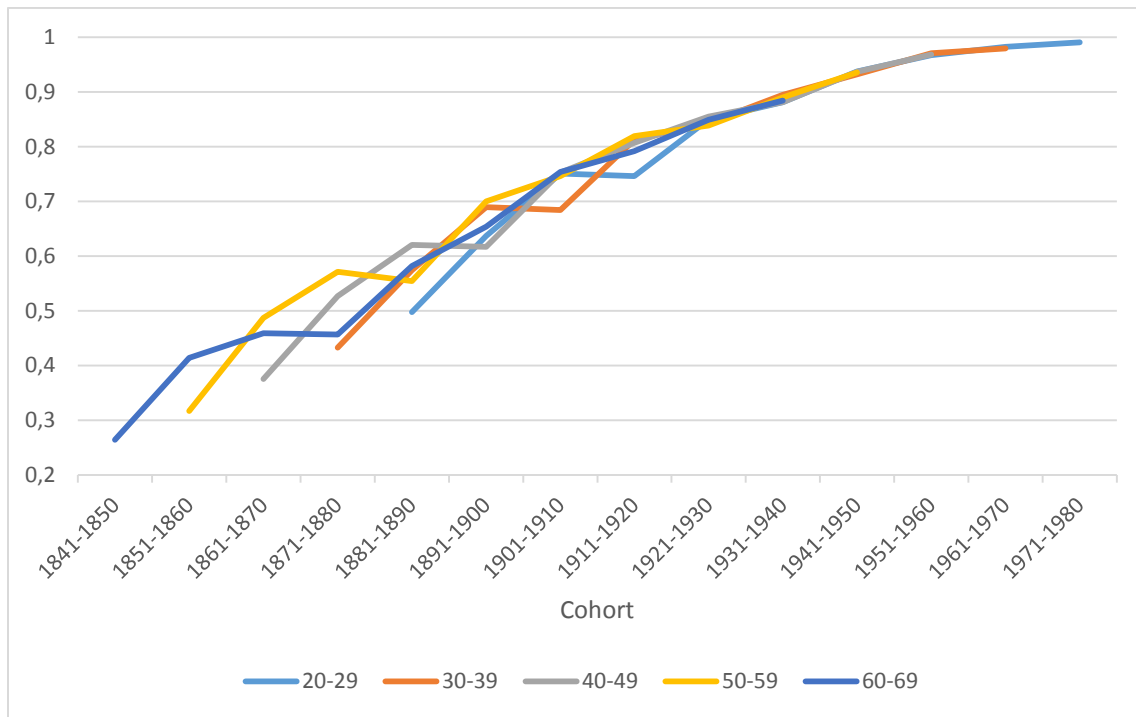
V. III.   Regressions and interpretation

Analyzing the evolution on literacy rates as it is done in the previous sub-section, though it is useful since it gives perspective of each effect, is incomplete. The interaction and correlation between the three effects has to be taken into account in order to understand the true relevance of each effect and the relation between them altogether, which is what is done in this sub-section.

As stated before, a major problem with APC methodology is to deal with the inherent linear dependence between the three effects. Nonetheless, the study done in this thesis concludes that the best model is the one including period effects and cohort effects, leaving apart the age effect. There are not few critiques on using a model with only two of the variables, but the related literature gives some concessions when the intuitive and statistical arguments go in favor of excluding one of the effects. Since the intuition is explained in the second to last sub-section, it is important to show now the results of the econometric models used and to explain further the results.

24

The models used correspond to OLS (Ordinary Least Squares) regressions, although the results were almost identical in every model for the equivalent GLM (Generalized Linear Models).[23]

Table 3 shows the results for each model, where the Models 1 to 3 (A, P and C) are reduced models where only one of the three effects is being controlled, giving the gross effect of each one. Models 4 to 6 (AP, AC, and PC) are controlling for two of the effects while leaving the other one uncontrolled. Finally, model 7 (APC) controls for the three effects altogether without constraints,[24] leaving the final cohort out due to collinearity.

---

[23] See appendix 3.

[24] Appendix 4 presents a table with several GLM models for APC with various constraints, showing that the results do not improve significantly although the most suitable empirical constraints are applied.

Table 3: Coefficients from OLS models of literacy rates for aggregated data

|	(1) A	(2) P	(3) C	(4) AP	(5) AC	(6) PC	(7) APC
Constant	0,828***	0,424***	0,307***	0,506***	0,277***	0,307***	0,489***
Age							
20-29	0			0	0		0
30-39	-0.0311			-0.0311**	0.0174		-0.0334***
40-49	-0.0708			-0.0708**	0.0299*		-0.0715***
50-59	-0.116			-0.116***	0.0432**		-0.111***
60-69	-0.193**			-0.193***	0.0300*		-0.182***
Period							
1907		0		0		0	0
1920		0.126**		0.126**		0.0672***	0.115***
1930		0.223***		0.223***		0.106***	0.198***
1940		0.213***		0.213***		0.0382**	0.175***
1952		0.340***		0.340***		0.112***	0.292***
1960		0.379***		0.379***		0.101***	0.326***
1970		0.438***		0.438***		0.113***	0.383***
1982		0.461***		0.461***		0.0960***	0.409***
1992		0.503***		0.503***		0.105***	0.462***
2002		0.527***		0.527***		0.101***	0.500***
Cohort							
1841-1850			0		0	0	0
1851-1860			0.0977*		0.0911**	0.0641**	0.00491
1861-1870			0.173***		0.169***	0.116***	0.00842
1871-1880			0.227***		0.227***	0.175***	0.0230**
1881-1890			0.295***		0.301***	0.230***	0.0365**
1891-1900			0.379***		0.385***	0.294***	0.0556***
1901-1910			0.0449***		0.454***	0.355***	0.0715***
1911-1920			0.497***		0.503***	0.405***	0.0774***
1921-1920			0.547***		0.553***	0.442***	0.0705***
1931-1940			0.584***		0.589***	0.480***	0.0653***
1941-1950			0.630***		0.638***	0.526***	0.0639***
1951-1960			0.660***		0.674***	0.559***	0.0558***
1961-1970			0.671***		0.692***	0.567***	0.0244*
1971-1980			0.682*		0.712***	0.580***	0
R-sq	0.141	0.849	0.974	0.990	0.979	0.995	0.998
Adj. R-sq	0.065	0.815	0.964	0.987	0.968	0.992	0.996
N	50	50	50	50	50	50	50

* Significant at the 10% level.   ** Significant at the 5% level.   *** Significant at the 1% level.

Graph 6: Age effects estimated from OLS models[2526]
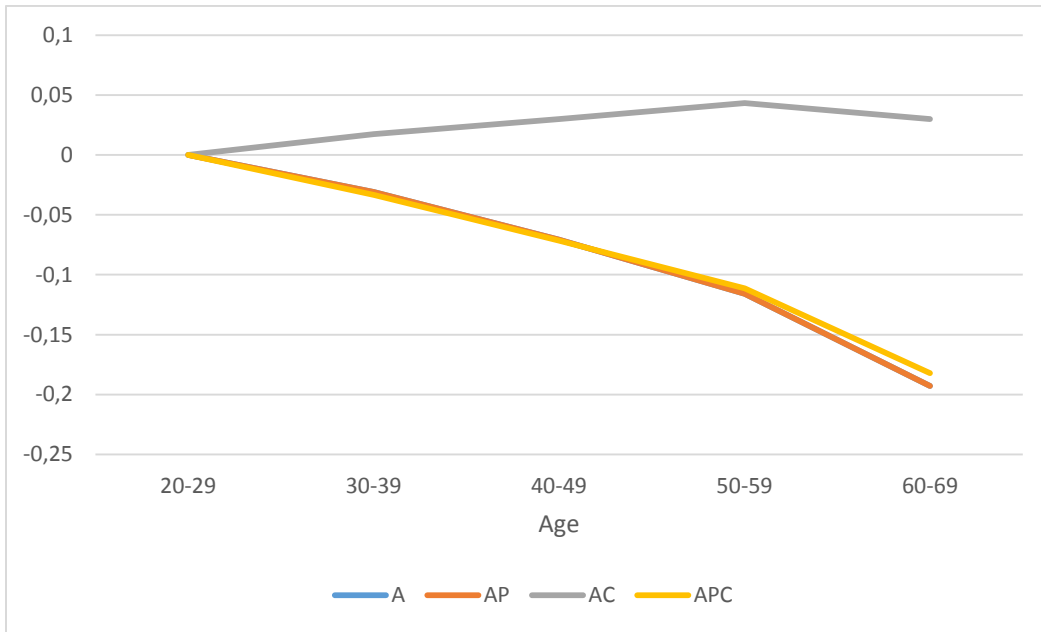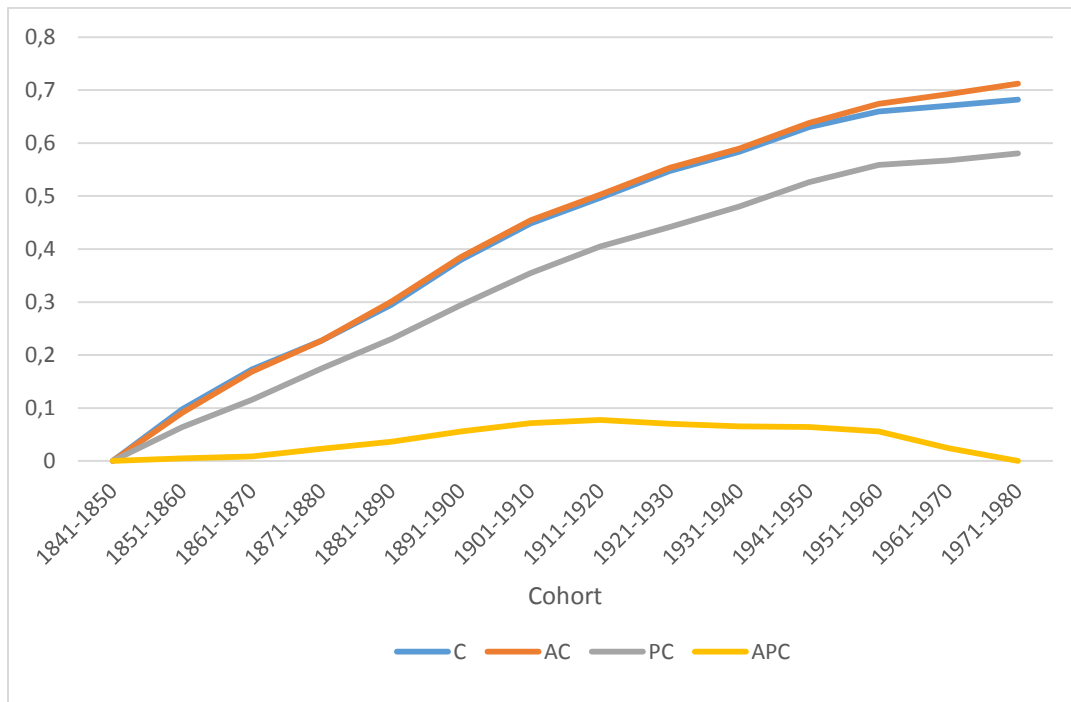
Graph 7: Period effects estimated from OLS models
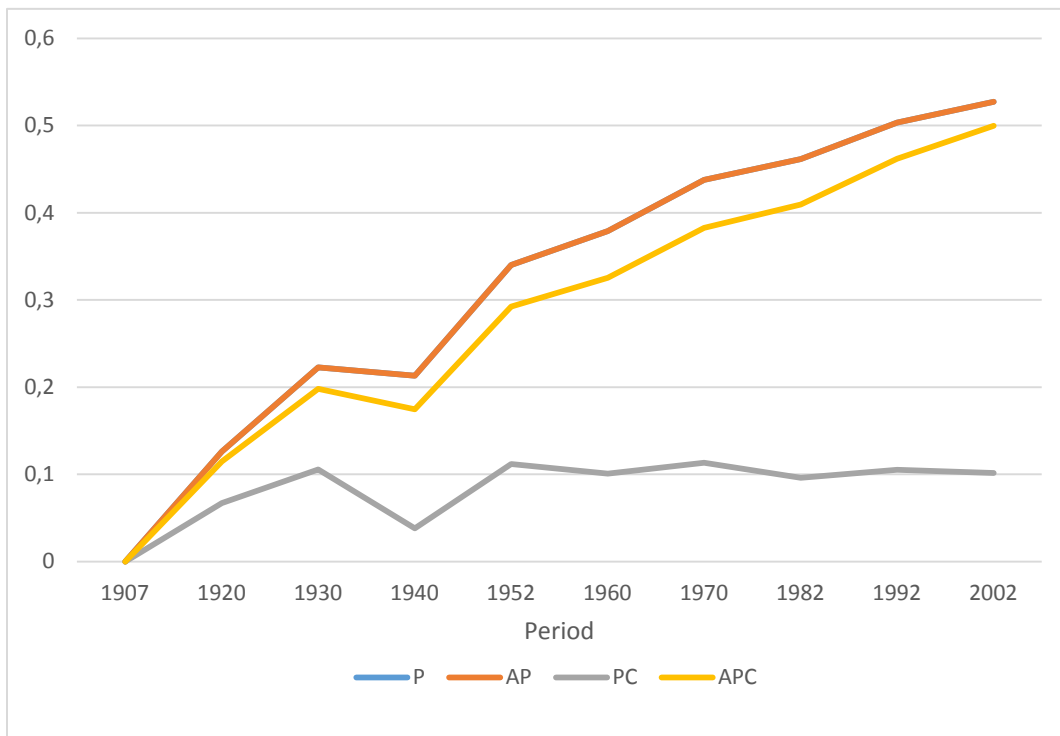
Graph 8: Cohort effects estimated from OLS models

---

[26] Trends for A and AP are almost identical, that is why the A line does not show. Same for P in Graph 5.

Graphs 6 to 8 show each effect separately by comparing the coefficient on the different models where each is present. Therefore it can be seen the difference between the gross effect of each dimension, presented in the individual models, in contrast to the net effect which comes from the interaction with the other effects in the two and three effects models.

Gross effects show that literacy levels diminish as age increases, this may be caused by physical or mental impediments that are more likely to occur in more advanced ages. Although, as it can be seen in table 3, none of the coefficients are significant at five per cent, and only the coefficient of the last age group is significant at ten per cent. Controlling by the other two effects, coefficients are quite the same for the AP model (but significant since it is capturing cohort effect by the linear dependence). Then, in the AC model coefficients turn positive but are not significant, while in the APC model they turn significant but distorting the values and significance of early cohorts. This inconsistency in the results is one of the major arguments to rule out the relevance of age, and therefore the utilization ultimately of the PC model for inferences.

The trend on period effect shows that the gross effect is significant and it holds the changes between censuses. After controlling for age the effect is exactly the same (another argument to consider age irrelevant). The PC model is the only one where the coefficients decrease,

but significance keeps intact. Then again, in the APC specification the numbers are closer to the previous models, which shed light on the distorting effect that has including age into the models, as it is also noticeable in the cohort effect.

The cohort effect by its own has a trend that is always increasing although it smoothens when reaching older cohorts, the same happening when controlling by age. When the PC model is introduced, the significance and trend of the effect remains the same but with lower coefficients. This may be due to the fact that some of the changes are attributable to period effects which in the other specifications are biasing the coefficients upwards. By including age effect, in the APC model, coefficients diminish considerably and significance of the earlier cohorts disappears. An overcompensation in the other two effects is possible, since the inconsistency in age effect patterns is already stated.

Graph 9 shows that, regarding gender differences, gross and net cohort effects are greater for women in the same specification. This installs larger evidence on the importance of female schooling that arose from the political discussion in the nineteenth and twentieth century. Though the difference is not clear for the earlier cohorts, for the ones since the beginning of the twentieth century it becomes noticeable. Results are consistent with the wider implementation of schooling policies focusing on girls, the rising of educational coverage for women, and also with the evidence that shows for the latest cohorts higher literacy rates for the female sub-group, having converged with men along the way.

Graph 9: Male and female gross and net cohort effects from OLS models[27]

[27] FC stands for Female Cohort model, FAC for Female Age-Cohort model, MC for Male Cohort model, etc.
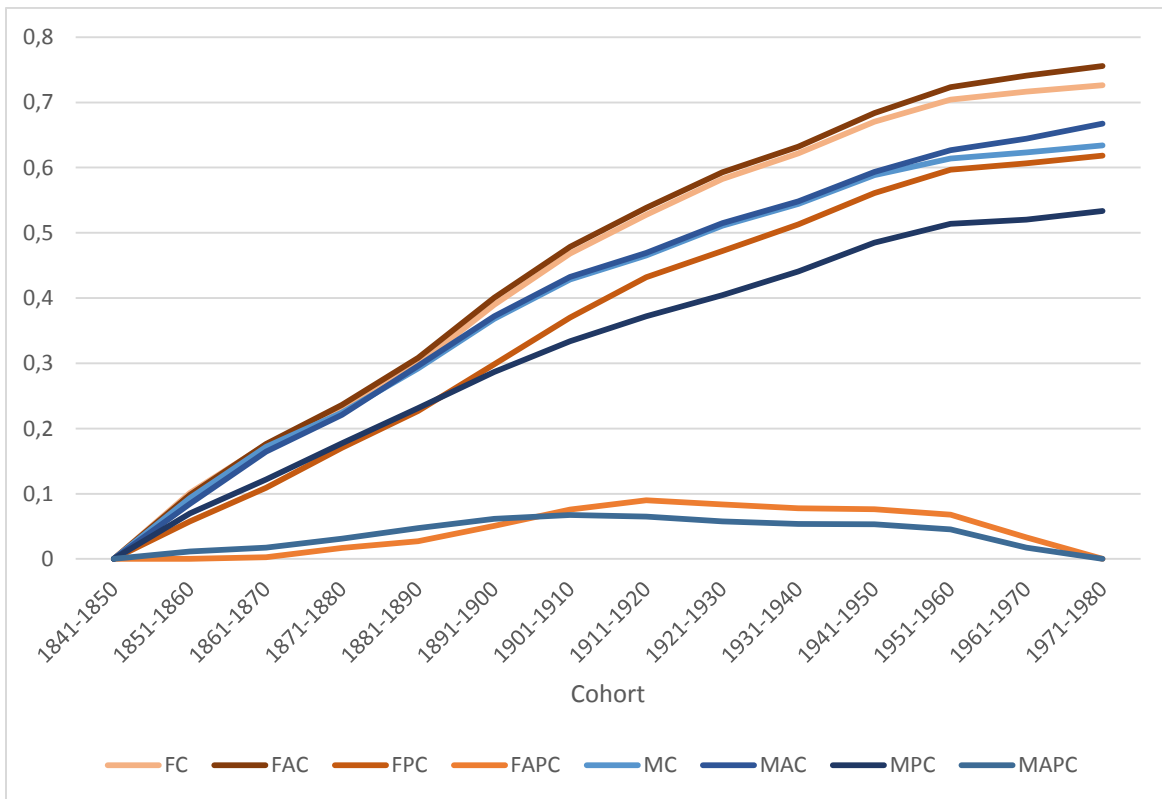
Graphs 10 and 11 show the first differences between cohorts. This way it is clearer the evolution of the cohort effect throughout time, since it is compared with the effect for the last cohort. Graph 10 shows the differences of coefficients for the aggregated data. By the other hand, graph 11 stresses the different patterns between genders, implying that the impact of cohort effect for the earlier ones is larger on the female sub-group, while for the latest cohorts the gap closes up.

Again, the period-cohort model is the most moderate one, as the differences in this model are more parsimonious. The fact that differences are smaller accounts for an overestimation of the cohort effect on the other models. The patterns for each specification do not differ when gender differences are taken into account.

Graph 10: Cohort effect first differences for C, AC and PC models
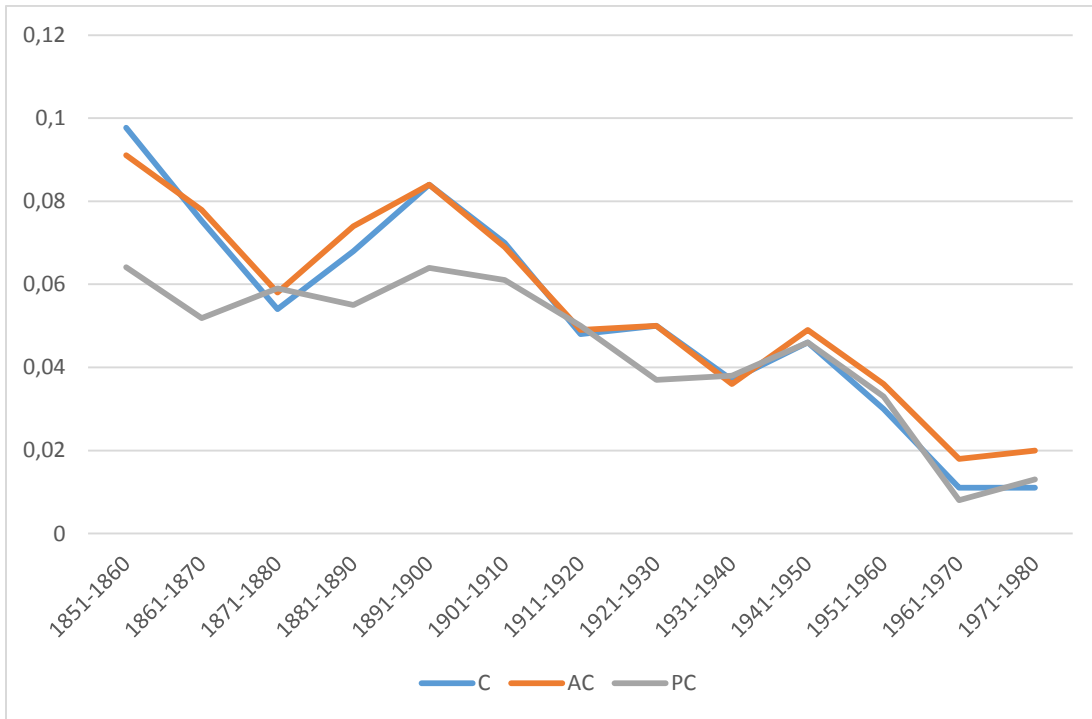
Graph 11: Cohort effect first differences for C, AC and PC models by gender[28]
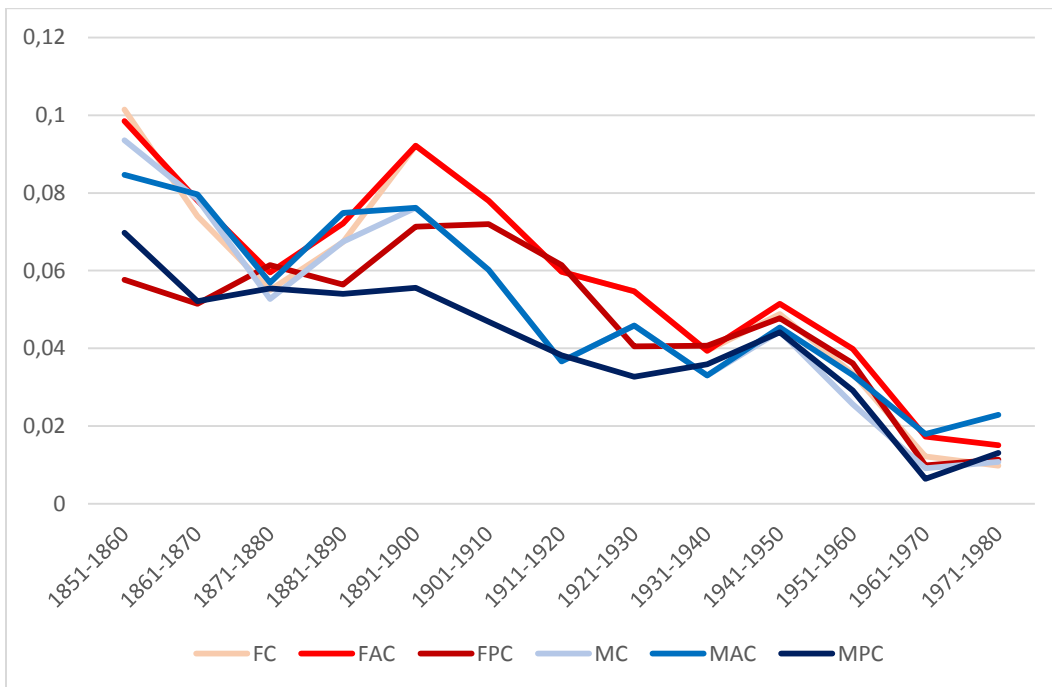
Studying these graphs, some inferences can be made. First of all, there is a positive difference for the 1851-1860 cohort, which can be explained by the fact that this cohort was

---

[28] F stands for female, M stands for male.

the affected by the Primary Instruction Law while the previous one born in 1841-1850 was not.

Then there is a decrease for the cohort effect, even though it is discrete for the PC model, it is significant. The negative effect of the Pacific War (1879-1883) may be the cause. The implementation of new educational methods and the real coming into force of the Primary Instruction law are the main reasons to explain the following positive effects seen in the graphs.

From the 1871-1880 cohort to the one born in the last decade of the nineteenth century the C and AC models show a sharp increase in the first differences for these cohorts, while the PC model only show slight variations. This may be due to period effects that took place on those years that the other models do not considerate, although even for the PC model there is a positive difference.

Then, for the cohorts comprehended between 1900 and 1940, the differences decline, meaning that cohorts born in the first decades of the twentieth century are less affected by their socioeconomic environment than the one preceding them, at least in terms of literacy rates. Though the Mandatory Primary Instruction Law was created in 1920, the Chilean educational system still had major issues concerning coverage and attendance. The fact that most children did not attend schools had its subsequent impact on literacy rates, as it is suggested by graph 10.

However, the PC model shows some impact for the cohort born in the 1930 decade, which is likely to be related with the reforms made in those decades and the implementation of adult schooling programs throughout the 1920-1960 period. The last break that can be seen is the one that involves the cohorts born in the 1940 and 1950 decades, as these were the most affected by the reform of 1965.

The 1965 school education reform is unlikely to have had a large impact on those individuals born in 1941-50, although presumably some of those born towards 1950 could have benefited from it, especially considering the expansion of compulsory education up to 8th grade as part of the 1965 reform. This suggests that the observed increase in literacy in the 1941-50 cohort is likely to be associated with the implementation of adult literacy

programs in the 1960 decade, many of which are part of the 1965 education reform. However, the 1951-60 cohort is expected to be affected by the 1965 reform, as also suggested by graph 10. As this reform allowed the country to reach near-universal coverage in primary education towards 1970, the room for further increases in literacy was accordingly reduced for cohorts born in the 1970 decade and after.

V. IV.   Literacy and numeracy

The relationship between literacy and numeracy has been a subject of study as it was said in the literature review section, having a strong correlation between each other and schooling. Cardemil and Núñez (2015) studied thoroughly numeracy also by an age-period-cohort analysis using the age-heaping technique, that is, the population that rounds up its age.

The age-heaping measure used was the ABCC index, which basically, as a literacy rate, shows from a value range from 0 to 100 the percentage of the population with the capacity to report exactly its own age. Therefore, with basic numeracy skills. The data used was from the same Chilean censuses from 1930 to 1970, using the annualized reported age information, capturing cohorts born in the 1860s until the ones born in the decade that started in 1940.

The results showed an improvement on the ABCC index coming from an 80.57% in the 1930 census to a 93.58% in the 1970 one, stating a decrease on the index while age rises (this may be because for numeracy and age-reporting memory is more important than for literacy).[29] Then an APC model with a GLIM specification was run in order to capture the interactions between the effects and get the net effects of age, period and cohorts.

Graphs 12 shows the patterns followed by the ABCC index in the commented study and literacy rates in this thesis. The graph shows that the starting point was higher for the ABCC index than for literacy rates. It also shows that they both followed trends upwards while the numeracy levels remained higher for any given age group and cohort until the last birth cohort studied (1941-1950) but with a catching up process which is likely to have finished, analyzing the patterns of both series. This sheds light on the fact that, even when for older cohorts numeracy skills were more usual than literacy skills, the efforts made on public education had a strong and significant impact on both numeracy and literacy rates. A subsequent catching up process due to the relative worse initial scenario regarding literacy took place, until nowadays, where both literacy and numeracy basic skills are almost totally acquired by the population.
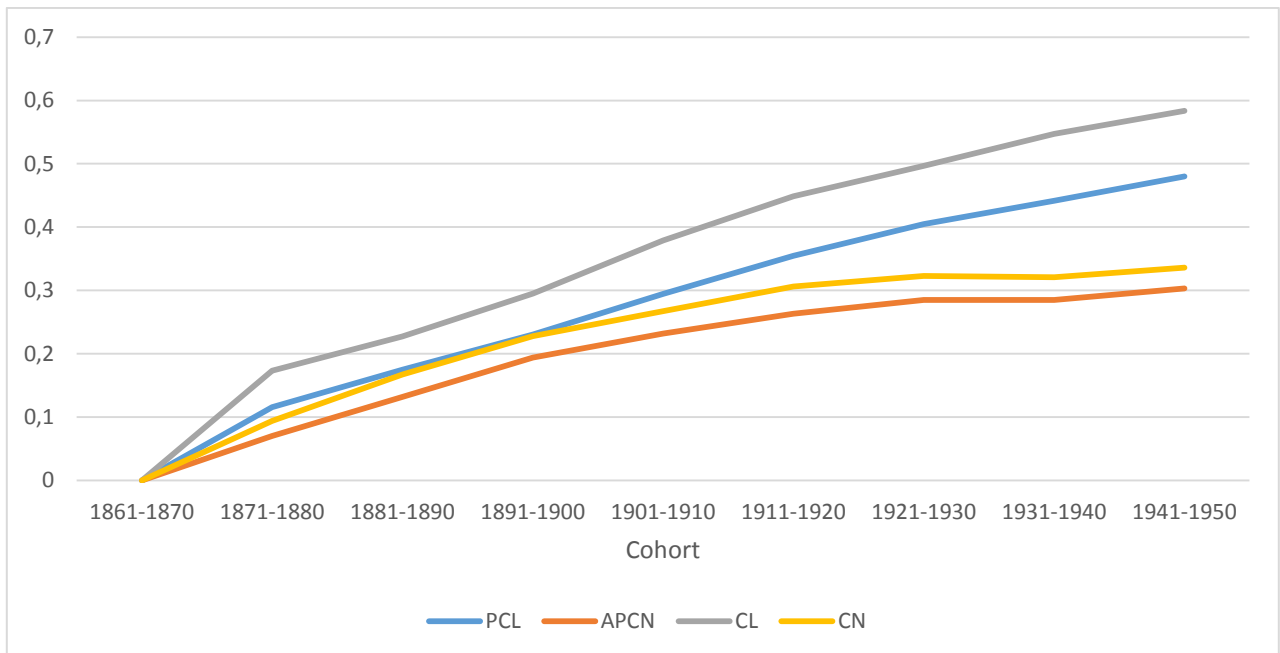
---

[29] Cardemil and Núñez (2015), pp. 29.

Graph 12: Evolution of Literacy rates and ABCC index by age group and cohort
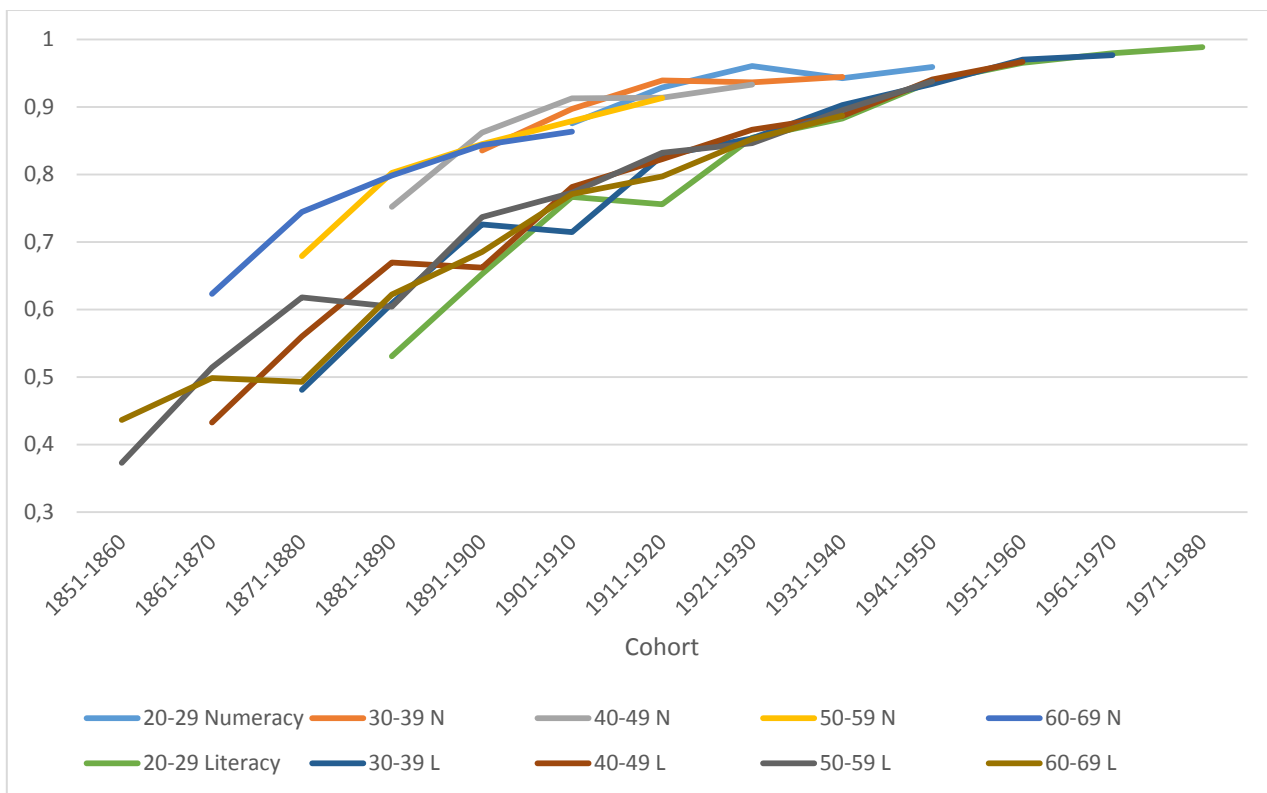
Graph 13: Gross and net cohort effects for literacy and numeracy for four specifications[30]

---

[30] PCL stands for Period-Cohort Literacy model, APCN for Age-Period-Cohort Numeracy model, CL for Cohort Literacy model and CN for Cohort Numeracy model.

Graph 13 shows the gross and net cohort effects for both studies by the reduced models (with cohort effect only) for the first case and by the best specification for each one (PC in this case, APC for numeracy) in the other. The graph exhibits that for both reduced models the cohort effects are overestimated comparing with the other specification, being the difference in literacy larger. It also shows that the cohort effect is ampler on literacy rates than on numeracy levels independently of the model used to compare.[31]

One reason that can be extracted from this analysis of the different initial paths is that for the earlier educational public policies the effect on basic numeracy skills were larger than in literacy. The fact that a substantial part of the children dropped out of school after the first or second year of primary school may be the explanation. Reading and writing, especially in adverse environments, were not learned by children in the few years of school they did attend. But, by the other hand, basic numeracy for age self-knowledge is more likely to be apprehended by the same children even though the attendance was only for one or two years. Considering that schools usually have registrations that include data of the children such as birth dates, it is plausible that almost null attendance had a larger impact on the age self-notion of children than in the capacity to read (and write), as children

---

[31] Also, a correlation test between the data of numeracy and the matching literacy rates obtained in this thesis was run, with a result of 0.8957, significant at the 1% level.

possibly heard in school their age and there may be a larger remnant for numeracy than for literacy.

VI. Conclusion

This thesis examines the evolution of literacy in Chile since the second half of the nineteenth century until the end of the twentieth. The national censuses give the necessary information to determine the relevance of the public policies regarding primary schooling. The impact of these policies is the most important aspect of this study, as they portrait the development of the Chilean educational system, which is highly related to the evolution of literacy.

Literacy rates in Chile show an increase from roughly 40 percent of literates in the beginning of the twentieth century to 97 percent for the latest cohorts. This evolution was mainly because of the public policy efforts made by the State. Although there were periods in which the evolution was discrete or even negative, in the broad picture, there is an improvement implying that the progress of the educational system affected the evolution of literacy in a positive and significant way throughout the whole 1860-2000 period.

The first major policy in educational matters was the Primary Instruction Law of 1860. Even though there were some efforts prior to this law, it came to organize the involvement of the State in education. It also recognized the gratuity of education, the nondiscrimination by gender or socioeconomic status, among others. The fact that the system imposed by this law has resemblance with the one seen nowadays is not irrelevant, since problems as

centralization and dubious quality are still present. The needs and expectations of the population have changed throughout time, and the educational system has not been able to solve these problems. To figure them out, looking back in time and see what was done wrong is almost always a suitable starting point.

The data is clear. During the first half of the twentieth century a significant leap on literacy rates takes place. This is due to a combination of the impact of the Primary Instruction Law and other minor policies of the period. The Mandatory Primary Instruction Law of 1920 organized the scenario of the first decades of the twentieth century. The coverage of the educational system was undermined in the 1930 decade due to the economic depression and political instability, with subsequent negative impacts on literacy rates seen in the census of 1940, the only period in which literacy rates significantly decreased.

Policies regarding literacy in the period 1920-1960 were sporadic, though it was the first time in which the State specifically took conscious of the problem of illiterates. The government of Frei Montalva in the mid 1960s focused these efforts to extend the coverage of primary schooling for children and also for adults, understanding that illiteracy in that period was mainly a problem dragged from previous cohorts (adult illiterates).

The convergence between genders on literacy rates that took place in the 1900-1950 period is also important to state. The difference in old cohorts for the 1907 census was around 10 percent, or even higher depending on the zone. Therefore, since the first educational policies, the instruction of girls and women was considered a relevant issue to improve. And it was. Literacy rates by gender show that for the cohorts born since the beginning of the twentieth century the difference is insignificant. For the latest cohorts, women have even higher literacy rates than men.

To our knowledge, this is the first attempt to link literacy and schooling using the age-period-cohort analysis with econometric techniques. It is significant because analyzing the evolution of literacy sheds light on the development of human capital in a country, even more considering the large period this study covers. We encourage further investigation on this subject.

The methodology used is innovative since it takes into account the different definitions of literacy prevailing throughout time. Prior to 1952, for Chilean censuses, a person was literate if he or she could read, and afterwards he or she was literate when being able to read and write altogether. This thesis corrects the first set of data predicting the writing rates for that period, therefore the information is standardized and comparable (being also a feature that we encourage to follow). The definition of literacy has been dynamic depending on time and place, so it is crucial to stress out the importance of having data with equivalent definitions, or when possible, standardized data.

It is suggested to extend this study for the urban-rural differences on literacy rates. In 1952 the literacy rates on rural zones was around 60 percent, while urban literacy rates were almost 90 percent. This line of investigation may delineate even further the true impact of public policies regarding education on the twentieth century, as the breach between literacy rates for urban and rural zones was present way after the one for gender.

VII.   References

Astorga, P., Berges, A., Fitzgerald, V. (2005): "The standard of living in Latin America during the Twentieth Century", The Economic History Review, Vol. 58, No. 4, pp. 765-796

Anderson, C. (1965): "Literacy and Schooling on the Development Threshold: Some Historical Cases," Education and Economic Development, eds. Anderson and Bowman (Chicago: Aldine Publishing Co.), pp. 347-362.

Barro, R. J. (1991) "Economic Growth in a Cross-Section of Countries", Quarterly Journal of Economics Vol. 106, No. 20, pp. 407–433.

Baten, J., Manzel, K. (2009) "Gender Equality and Inequality in Numeracy: The Case of Latin America and the Caribbean, 1880–1949", Journal of Iberian and Latin American Economic History, Vol. 27, pp. 37-73.

Baten, J., Manzel, K., Stolz, Y. (2012) "Convergence and divergence of numeracy: the development of age heaping in Latin America from the seventeenth to the twentieth century", The Economic History Review, Vol. 65, pp. 932–960.

Blaug, M (1966): "Literacy and Economic Development", The School Review, Vol. 74, No. 4, pp. 393-418.

Cardemil, M., Núñez, J. (2015) "Numeracy and education in Chile: 1860-1940 cohorts", Santiago, Chile: Universidad de Chile, Facultad de Economía y Negocios.

Charette, M., Meng, R. (1998) "The Determinants of Literacy and Numeracy, and the Effect of Literacy and Numeracy on Labour Market Outcomes", Canadian Journal of Economics, Vol. 31, No. 3, pp. 495-517.

Cipolla, C. (1969) "Literacy and Development in the West", London: Penguin Books.

de Gabriel, N. (1998): "Literacy, Age, Period and Cohort in Spain (1900-1950)", International Journal of the History of Education, Vol. 34, No. 1, pp. 29-62.

Egaña, Mª L., Núñez, I., Salinas, C (2003). "La educación primaria en Chile: 1860-1930. Una Aventura de niñas y maestras" LOM ediciones, Santiago de Chile.

Egaña, Mª L. (2000) "La educación primaria popular en el siglo XIX en Chile: Una práctica de política estatal", LOM ediciones, Santiago de Chile.

Fienberg, S., Mason, W., Eds. (1985). "Cohort Analysis in Social Research", New York: Springer-Verlag.

Glenn, N. D. (1976) "Cohort Analysts' Futile Quest: Statistical Attempts to Separate Age, Period, and Cohort Effects", American Sociological Review, Vol. 41, No. 5, pp. 900–904.

Graff, H. (1979) "The Literacy Myth: Cultural Integration and Social Structure in the Nineteenth Century", New York, Academic Press, Inc.

Ignacio, J., Larraín J. (1871) "Boletín de las principales leyes dictadas desde Enero 1° de 1860 hasta Enero 1° de 1871 y recopiladas por orden de materias", Imprenta del Mercurio, Valparaíso, Chile.

Kupper, L. L., Greenberg, B., Janis, J., Karmous, A. (1985) "Statistical age-period-cohort analysis: A review and critique", Journal of Chronic Diseases, Vol. 38, No. 10, pp. 811-830.

Labarca, A. (1939) "Historia de la Enseñanza en Chile", Santiago de Chile, Publicaciones de la Universidad de Chile.

Lei 3.654 sobre Educación Primaria Obligatoria. Diario oficial de la República de Chile Nª12.755, Santiago de Chile, 26 de agosto de 1920.

Levy, R. (2016) "A historical reflection on literacy, gender and opportunity: implications for the teaching of literacy in early childhood education", International Journal of Early Years Education, Vol. 24, No. 3, pp. 279-293.

Matte, C. (1884) "Nuevo método para la enseñanza simultánea de la lectura y escritura", Leipzig, Imprenta de F.A. Brockhaus.

Núñez, J. (2005) "Signed with an x: Methodology and Data Sources for Analyzing the Evolution of Literacy in Latin America and the Caribbean, 1900-1950", Latin American Research Review, Vol. 40, No. 2, pp. 117-135.

Oppenheim, K., Mason, W., Winsborough, H. and Poole, K. (1973) "Some Methodological Issues in Cohort Analysis of Archival Data", American Sociological Review, Vol. 38, pp. 242-258.

Pérez, C., Silva, C. (2013) "Educación y proyectos desarrollistas: discursos y prácticas de alfabetización popular en Chile, 1960-1970", Cuadernos Chilenos de la Historia de la Educación, Vol. 1, No. 1, pp. 87-111.

Serrano, S., Ponce de León, M. and Rengifo, F. (2012) "Historia de la Educación en Chile (1810-2010): Aprender a leer y escribir (1810-1880)", Editorial Taurus, Santiago de Chile, Vol. I.

Serrano, S., Ponce de León, M. and Rengifo, F. (2012) "Historia de la Educación en Chile (1810-2010): La educación nacional (1880-1930)", Editorial Taurus, Santiago de Chile, Vol. II.

Stromquist, N. P. (1992) "Women and Literacy: Promises and Constraints", Annals of the American Academy of Political and Social Science, Vol. 520, pp. 54-65.

Thorp, R. (1998): "Progreso, pobreza y exclusión: una historia económica de América Latina en el siglo XX", Washington, D.C., Banco Interamericano de Desarrollo (BID).

UNESCO (2004) "EFA Global Monitoring Report 2005: Education for All, the Quality Imperative", Paris: UNESCO.

Valdivia, V. (2009): "Estabilidad y Constitucionalismo: Las sombras de la excepcionalidad Chilena", Working papers ICSO-UDP, 1-2009.

VIII. Appendixes

Appendix 1: GDP per capita, Literacy rates and Life expectancy for LA countries (LA6)

Country		GDP per capita (1970 PPP dollars) Year			Literacy (percent) Year			Life expectancy (years) Year	
	1900	1950	2000	1900	1950	2000	1900	1950	2000
Argentina	497	811	1,459	51	88	97	39	61	73
Brazil	114	245	874	35	49	85	29	43	68
Chile	284	592	1,602	44	79	96	29	49	75
Colombia	290	389	921	34	62	92	29	49	71
Mexico	240	519	1,284	24	61	91	25	48	73
Venezuela	94	719	1,014	28	51	93	29	51	73

Source: Astorga et al. (2005), pp. 766.

Appendix 2: Methodology for 1907-1940 correction on literacy rates

In the census of 1952, the definition of whether a person was literate or not changed from being able to read, to being able to read and write. For the 1907, 1920, 1930 and 1940 census the official literacy rates are based on the self-reported capacity to read of the

surveyed population. Then, by the other hand, from 1952 and on, the literacy rates are a reflection of the people who stated that were able to read and write. This is important to explain because it is the period in which data was gathered. Therefore, the issue was to correct the data for 1907-1940 in order to make both sets of censuses properly comparable. In the 1854 census the question of literacy was introduced in the survey, and for the latter censuses of the nineteenth century, the question was divided in two: whether a person could read, and whether a person could write. There is no division by age groups in these censuses so it is impossible to incorporate them in the APC analysis. But, it became useful because, apart from having separated the information for reading and writing, there was also a gender differentiation and a department division (74 in total). So, for both reading and writing rates there was a not insignificant amount of data.

The population under five years old was subtracted from the total, under the assumption that the proportion of literates under five years old in 1895 was insignificant against their percentage in total population. After correcting for two department outliers most possibly due to major measurement errors in early age groups, there were 72 departments classified by gender. So, there were 144 observations with their related literacy numbers for reading and writing separately.

Literacy rates within departments differed greatly. For example, in Santiago 66.91% of male population over five years old was able to read, but in the rural province of Melipilla only 9.1% could. By gender, surprisingly, the difference in many of the less populated departments was rather insignificant. Although comparing them to the largest cities the overall rates were lower, and in these cities the gender difference was larger (over a four percent in Santiago and over a nine percent in Valparaiso). The difference between reading and writing literacy rates was between three and five percent in most departments. Taking into account that aggregated rates where about 30 percent, the relative difference between reading and writing was around a 10 percent, being this one of the major arguments and motivations to make the correction.

Having this data gave a broad perspective of the literacy situation in 1895. Forecasting the behavior of writing capacity rates became more suitable considering that, for example, the reading ability rate in Santiago in 1895 for population over five years old was only two

percent less than the total of the country in 1940 for the same group. Disaggregating by gender and department was sufficient data to regress the rates in an OLS model:

Where W corresponds to the writing literacy rates, R is the reading literacy rate, is a constant term, is the coefficient for reading rates, is the coefficient by controlling for the quadratic reading rate term to smooth the difference for latter censuses, and is the typical error.

The results were 0.00299 for , 0.8459 for and 0.1141 for , meaning that, for example, for a 50% literacy rate (only reading) in the 1920 census, the writing literacy rate according to this correction was a 45.4%. On these results the correction for the four following censuses was made and then compared with the set of 1952-2002.

Appendix 3: Coefficients from GLM models of literacy rates for aggregated data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A	P	C	AP	AC	PC	APC
Constant	0,828***	0,424***	0,307***	0,506***	0,277***	0,307***	0,489***

Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
20-29	0			0	0		0
30-39	-0.0311			-0.0311***	0.0174		-0.0334***
40-49	-0.0708			-0.0708***	0.0299*		-0.0715***
50-59	-0.116			-0.116***	0.0432***		-0.111***
60-69	-0.193**			-0.193***	0.0300*		-0.182***

Period

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1907		0		0		0	0
1920		0.126**		0.126***		0.0672***	0.115***
1930		0.223***		0.223***		0.106***	0.198***
1940		0.213***		0.213***		0.0382***	0.175***
1952		0.340***		0.340***		0.112***	0.292***
1960		0.379***		0.379***		0.101***	0.326***
1970		0.438***		0.438***		0.113***	0.383***
1982		0.461***		0.461***		0.0960***	0.409***
1992		0.503***		0.503***		0.105***	0.462***
2002		0.527***		0.527***		0.101***	0.500*

Cohort

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
1841-1850			0		0	0	0
1851-1860			0.0977**		0.0911**	0.0641***	0.00491
1861-1870			0.173***		0.169***	0.116***	0.00842
1871-1880			0.227***		0.227***	0.175***	0.0230**
1881-1890			0.295***		0.301***	0.230***	0.0365***
1891-1900			0.379***		0.385***	0.294***	0.0556***
1901-1910			0.0449***		0.454***	0.355***	0.0715**
1911-1920			0.497***		0.503***	0.405***	0.0774***
1921-1920			0.547***		0.553***	0.442***	0.0705***
1931-1940			0.584***		0.589***	0.480***	0.0653***
1941-1950			0.630***		0.638***	0.526***	0.0639***
1951-1960			0.660***		0.674***	0.559***	0.0558***
1961-1970			0.671***		0.692***	0.567***	0.0244*
1971-1980			0.682***		0.712***	0.580***	0

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
N	50	50	50	50	50	50	50
AIC	-27.18	-104.2	-183.7	-234.3	-186.9	-253.1	-294.3
BIC	-17.62	-85.09	-156.9	-207.5	-152.5	-209.1	-244.6

* Significant at the 10% level.   ** Significant at the 5% level.   *** Significant at the 1% level.

Appendix 4: CGLIM estimates for APC specifications (constraints)

	(1) A	(2) P	(3) C	(4) APC const.	(5) APC
Constant	0,495***	0,491***	0,489***	0,495***	0,489***
Age					
20-29	0	0	0	0	0
30-39	-0.0496a**	-0.0330***	-0.0329***	-0.0488a***	-0.0334***
40-49	-0.0496a***	-0.0703***	-0.0721***	-0.0488a***	-0.0715***
50-59	-0.104***	-0.109***	-0.112***	-0.103***	-0.111***
60-69	-0.174***	-0.179***	-0.183***	-0.173***	-0.182***
Period					
1907	0	0	0	0	0
1920	0.116***	0.116***	0.115***	0.117***	0.115***
1930	0.197***	0.199***	0.199***	0.198***	0.198***
1940	0.170***	0.175***	0.177***	0.174***	0.175***
1952	0.285***	0.291***	0.294***	0.287***	0.292***
1960	0.316***	0.324***	0.328***	0.317***	0.326***
1970	0.370***	0.392a***	0.385***	0.382a***	0.383***
1982	0.395***	0.392a***	0.412***	0.382a***	0.409***
1992	0.443***	0.455***	0.465***	0.441***	0.462***
2002	0.481***	0.493***	0.502***	0.479***	0.500***
Cohort					
1841-1850	0	0	0	0	0
1851-1860	-0.00837	0.0000720	0.00660	-0.0104	0.00491
1861-1870	-0.00909	0.00398	0.0176a	0.00122a	0.00842
1871-1880	0.0138	0.0191	0.0176a	0.00122a	0.0230**
1881-1890	0.0296	0.0333**	0.0366***	0.0266	0.0365**
1891-1900	0.0506***	0.0528***	0.0553***	0.0474**	0.0556***
1901-1910	0.0691***	0.0669***	0.0708***	0.0637***	0.0715***
1911-1920	0.0778***	0.0765***	0.0763***	0.0752***	0.0774***
1921-1920	0.0737***	0.0708***	0.0693***	0.0723***	0.0705***
1931-1940	0.0707***	0.0668***	0.0640***	0.0701***	0.0653***
1941-1950	0.0715***	0.0664***	0.0623***	0.0715***	0.0639***
1951-1960	0.0661***	0.0637***	0.0540***	0.0705***	0.0558***
1961-1970	0.0460**	0.0289*	0.0222	0.0472*	0.0244*
1971-1980	0	0	0	0	0
N	50	50	50	50	50
AIC	-237.1	-273.8	-290.7	-230.5	-294.3
BIC	-189.3	-226.0	-242.9	-186.5	-244.6

* Significant at the 10% level.   ** Significant at the 5% level.   *** Significant at the 1% level.

Note: a = adjacent coefficients constrained to be equal.