Stockholm School of Economics
Department of Economics
659 Degree Project in Economics
Spring 2021

An Academic Degree, What Is it Good For?

Using matching to investigate sheepskin effects in Chile

Hänninger, Sofia (24410)
Wägmark, Nils (24494)

Abstract:

The purpose of this study was to investigate signaling in higher education by looking at if Chilean students who receive their diploma earn more money than non-graduated students who otherwise share a common set of characteristics. We furthermore looked at how these signaling effects, also called sheepskin effects, develop over time by looking at years 5-11 after enrollment. The research design is based on a matching method where we use the diploma as treatment and include gender, program of enrollment, year of enrollment, cognitive ability and total time spent in higher education in the set of covariates. Our findings suggest that there are strong and statistically significant sheepskin effects on the Chilean formal labor market, however, it remains unclear whether these are in fact real effects or the result of composition effects. Finally, our research also suggests that there are higher levels of sheepskin effects for men than for women, and that the most prestigious universities in Chile (UCRUCH) show larger sheepskin effects than private universities and other higher institution types.

Keywords: sheepskin effects, signaling, educational economics, higher education, labor markets

JEL: E24, E29, I23, I26, J24

Supervisor: Alex Solis
Date submitted: May 17, 2021
Date examined: June 1, 2021
Discussants: Axel Hellbom Almström & Rebekka Andersson
Examiner: Johanna Wallenius

Acknowledgements

We would like to express our most sincere gratitude to our supervisor Alex Solis who has been a Nestor and loyal companion throughout this process. Thank you for your invaluable support.

Table of contents

1. Introduction

1.1 The Relevance of the Topic

The literature on the economics of education has long debated about the function of education. One predominant view considers education as the vehicle to obtain knowledge and skills that are key for productivity in the labor market. On the other hand, a different view argues that education serves as a mechanism to signal potential employers about the individuals' innate ability. Students with high ability engage in the costly process of education expecting to differentiate themselves from low ability individuals who are unable to finish a degree. Hence, the model suggests that education does not boost productivity but instead provides information that reduces employers' uncertainty about the ability of potential employees. Consequently, additional years of education translate into positive returns only when it conveys a certification or diploma, what is known as the "sheepskin effect."

1.2 What the Literature Is Missing

In today's literature many studies support that sheepskin effects are present in labor markets. Research has been done in several geographical areas including Colombia, the United States, Philippines, Sweden, Mexico and other. Most studies investigate sheepskin effects for high school degrees, but there are also some studies made on university degrees. Furthermore, something that many studies have in common is that the data used to conduct research is survey data. For example, studies made on sheepskin effects in the labor markets in the Philippines, Northern Ireland and the US uses survey data (Olfindo, 2018; McGuinness, 2003; Belman & Heywood, 1991). What is missing in the existing literature is research on data for full populations which would provide more precise results on sheepskin effects. There is no literature as of today on sheepskin effects for university graduates in Chile for the period 2007-2016, why our research provides this new geographical contribution to the literature. Furthermore, most research on the sheepskin effect today has only compared graduates to non-graduates, completely disregarding the subject's program of study as well as the overall academic ability. Another missing factor in much of the literature is controlling for the number of years the student has spent in higher education. Also, our research is implemented by running a matching methodology with graduation as the treatment variable which is something unique when comparing to previously made studies.

1.3 How This Is Investigated

In this paper, we estimate the effect of obtaining a higher education diploma using nearest-neighbor and propensity-score matching. We exploit a rich administrative dataset at the individual level that includes data on gender, program code and year of enrollment. The dataset also allows for extending the analysis to include number of years in university and academic ability.

1.4 Data

We combine three sources of publicly accessed data at the individual level. First, we use enrollment and graduation data from the Ministry of Education in Chile *(Ministerio de Educación)*. This data contains the annual enrollment registries from 2007 to 2016 and covers all higher education institutions in the country. We can follow the students though out their degrees until completion. Moreover, we use a PSU score data set that contains the admission test score (PSU scores) and a wide range of characteristics, such as gender, date of birth, parental education and parental work status. We base our outcome variable on a data sample on earnings collected by the Ministry of Labor in Chile showing observations at the individual level for all workers in the formal sector.

1.5 Key Results

We find large and significant results that there are sheepskin effects on the Chilean labor market when controlling for gender, program code and the year of enrollment. The results further hold up when adding past academic ability and total amount of years spent in higher education as controlling factors. In the heterogeneity analysis, we show that men consistently experience higher sheepskin effects than women, however, the effect seems to taper off with time. Furthermore, the level of signaling value you get from a diploma seems to be tightly connected to the type of institution it was provided from. Our analysis shows that students at the most prestigious universities in Chile (UCRUCH) as well as at private universities, experience significantly higher sheepskin effects than those who attend less prestigious ones.

2. Background

2.1 The Different Theories of Education

2.1.1 Human Capital Model

A common division of the income effects of higher education is on the one hand the human capital effect and on the other the signaling effect. Gary Becker’s research on human capital has its foundation in the income growth in the US, starting in the 1940s, where there was a larger increase in income than in physical capital and labor. Becker presents a theory that the residual in this case comes from improved labor force quality (Becker, 1993). Becker constructs a factor which he denominates human capital, where human capital is dynamic and is the collection of a person’s productive skills that can be used to generate earnings in the labor market, as stated by Yoram Weiss when describing Becker’s work (Weiss, 2015). The dynamic part of human capital is created by the possibility for people to invest in their own human capital by for instance choosing to pursue higher education. Since human capital is personal and cannot be transferred from one individual to another there is a decreasing rate of investment the older an individual gets.

Becker's explanation of how an individual's productivity increases from education and on-the-work training suggests that from each extra year of education, and the increase in productivity that comes with it, there will also be a proportional increase in salary. The human capital effect can in other words be described as the knowledge the student gains during his or her studies and that will then serve them useful at the workplace in which they end up. The higher amount of education resulting in stronger work-related skills would incentivize companies to hire and pay a higher salary to the well-educated workers.

2.1.2 The Signaling Model

The signaling effect on the other hand does not argue that education enhances the students' actual skills or competences, but rather states that it serves as a certification signaling productivity to the employer. This means that two people can have the same amount of education and be equally productive, but still be looked upon differently in hiring processes depending on whether they have a degree or not.

In 1973 Michael Spence presented research on Job Market Signaling (Spence, 1973), which became the foundation for signaling research in various fields such as financial markets and human resources management. Spence explains the concept of signaling with the example of a hiring process with asymmetric information for the two parties, the employer (the principal) and the employee (the agent). The asymmetry is that the employer does not have information about the employee's productivity. To simplify reality, Spence divides employees into two groups, good and bad, with either high or low productivity. The employer is willing to pay a higher salary to the good employees than to the bad ones. An assumption is that employers do not have information in advance regarding which employee is good and which one is bad. However, employees can invest in signaling to the employer to show that they belong to the group of good employees. This is made through investing in education. Spence assumes that one unit of education is cheaper for the good employee than for the bad one which causes positive correlation between the group that choses to invest in signaling and the group of good workers (Spence, 1973).

What can be concluded from Spence's research is that even in the case that education does not improve productivity or strengthens the human capital, it can still have a signaling value for both the employer and the employee. This effect, the increased wage due to education as a signal of higher productivity, not higher productivity in itself, is what can be called sheepskin effects.

2.2 The Education System in Chile

The education system in Chile is divided into preschool, primary school (Educación básica), secondary school (Enseñanza media) and technical or higher education (Educación superior). Where preschool is for children up to six years old, primary school which is compulsory for eight years, secondary school compulsory for four years and higher education that is not compulsory in Chile. The higher education is commonly divided into a bachelor's degree, master's degree, and doctorate. Higher education can be conducted at three main types of institutions, these are universities (Universidades), professional institutes (Institutos Profesionales, IP) and technical schooling centers (Centros de Formación Técnica, CFT). In the dataset used in this study universities are further divided into "Universidades CRUCH" and "Universidades Privadas". CRUCH stands for Consejo de Rectores de las Universidades Chilenas, and this category contains both state and non-state universities that are considered to have some of the highest quality of teaching in Chile (Ministerio de Educación, 2020). Univeridades Privadas are private universities that are not part of CRUCH. There are some differences between all four categories of higher education, mainly when it comes to entering the labor market after graduation (Solis, 2017). The CFTs are focused on delivering higher level technical training and to provide education that is specific for a certain area of work. Furthermore, the duration is lower than for university programs and the programs normally last for two to three years. IP offer technical careers that last between three and four years. Both types of universities offer education that deliver academic degrees, the different levels are undergraduate, graduate and postgraduate (Ministerio de Educación, Subsecretaría de Educación Superior, 2020).

The admission requirements for universities in Chile are a secondary school diploma as well as a PSU score (Prueba Seleccion Universitaria), which is a national exam that students take to apply for universities. However, most high school students take the PSU test before entering higher education regardless which type of higher education institute they are applying for. The PSU test is a way to measure students' ability, in this study we control for the language and mathematics scores on the PSU test since these are mandatory for all test takers. Note that there might be changes coming to the system with PSU tests as a necessary requirement for university enrollment in Chile (Ministerio de Educación, 2020). However, these changes are not relevant for the period 2007-2016 that our data is tracking.

This study will focus on higher education students, from all four above mentioned institutes for higher education in Chile, and a comparison between students obtaining a degree and students who initiate their studies but do not graduate will be conducted.

3. Literature Review

3.1 Previous Literature

In 1987 Hungerford and Solon were amongst the first to present evidence that in addition to the return for each year of education, there is also an additional significant return for the years of education where a degree or certificate is earned (Hungerford & Solon, 1987). These sheepskin effects can be presented with a discontinuous spline function in contrast to a linear or quadratic function for the natural logarithm of the wage which had previously been the standard when explaining returns to education. Because of this finding by Hungerford and Solon, the research on sheepskin effects experienced a boost.

In 2015 Rodríguez and Muro conducted a meta-analysis of 122 studies on the sheepskin effect for high school diplomas (Rodríguez & Muro, 2015). The meta-analysis looks at studies from 15 countries, including Colombia, the US, Philippines, Sweden, Mexico and other. What Rodríguez and Muro concludes is that both when using the fixed-effect method and the random-effect method there is a statistically significant effect of a high school diploma of 8% and 15% respectively. Rodríguez and Muro state that due to the heterogeneity from one study to another the random-effect method is the more suitable method for this meta-analysis (Rodríguez & Muro, 2015).

Another earlier study by Ferrer and Riddell examines the sheepskin effects in the Canadian labor market (Ferrer & Riddell, 2002). In this study both potential effects from high school degrees and university diplomas, after controlling for educational inputs, are investigated. Ferrer and Riddell test and compare the human capital earnings function with a "credentialist model", two models that show different views of the effect of education. Where the human capital earnings function expresses the logarithm of individual earnings as a linear function of years of completed schooling and a quadratic function of labor market experience. The "credentialist model" that they compare it with on the other hand measures the individual earnings as a step function of years of education. This model has discontinuities for the years where a degree or diploma is obtained. Ferrer and Riddell present several findings for different areas of the Canadian labor market. Amongst their findings are, that "across all fields, bachelor's degrees produce the highest returns, ranging from 14 to 47 per cent for men and 25 to 55 per cent for women" and that since the results indicate that there is an effect both for years of schooling and for obtaining a degree or diploma, neither the human capital model nor the "credentialist model" suffice for alone describing the reality of education effects. They conclude their research paper by raising the possibility that the underlying reason for the discrete increases in earning that comes when the students obtain their degrees could either be due to the students' skills then being more observable by employees or to the sheepskin effects/credentials effects (Ferrer & Riddell, 2002).

In 1997 Belman and Heywood did research on sheepskin effects, considering the time aspect. They present estimations of sheepskin effects across five age cohorts and compare the effects. The study uses data from the Current Population Survey in the US and limits the individuals to "non-agricultural non-black males between the ages 24 and 65" (Belman & Heywood, 1997). More specifically Belman and Heywood are looking into how returns to the educational signaling effects change over time. To do so they categorize workers into types with different maximum productivity levels and jobs into different types with different marginal product depending on the "worker-job match". Belman and Heywood include an element of matching in their research. Note that this matching does not refer to the matching method where a treatment group is matched with a control group, but instead a matching for employers and employees based on productivity. In their model workers can purchase productivity signals in the form of educational credentials which will help the workers to get a better job match, resulting in a higher wage. What is concluded in the study is that due to that workers being undermatched, i.e. the worker is more productive than thought, also have the lowest education signal. Therefore, this group of workers have the largest increase in productivity from one period to another. The overall result from the study implies that there are indeed sheepskin effects for different levels of education, with the strongest effects of 12% for the youngest age cohort and for college degrees. The results also show that the sheepskin effects are diminishing over time, meanwhile the returns to years of education remains almost constant (Belman & Heywood, 1997).

When it comes to sheepskin effects early research in the field that was conducted enabled researchers to confirm sheepskin effects in return to education for non-minority males. However, an early study that measured sheepskin effects for women and minorities could confirm sheepskin effects also for these groups, even if these effects were smaller when it came to lower-level education and greater for higher level education (Belman & Heywood, 1991).

3.2 Research Gap, Contribution and Research Questions

As per the literature review, sheepskin effects have been shown for many different labor markets before where the researchers often control for factors such as age and gender, however, most papers do not control for some type of IQ or ability variable. This is something we address by adding mathematics and language scores from the PSU exam to the set of covariates. Another contribution to the research field comes from that the data used in this study is administrative data for the entire population instead of aggregate data, which therefore gives more accurate results. Regarding the different markets that sheepskin effects have been shown for, the geographical distribution of these regions has been limited with a sparse exposure to South American labor markets. This is addressed by basing the analysis on Chilean data. Furthermore, previous research also shows that sheepskin effects change over time, but this has not been shown convincingly using newer data which we will try to demonstrate in our research.

Some previous research around sheepskin effects have been suffering from not knowing whether the individual actually graduated or not, since the studies are lacking a variable that controls this. The method has then instead been to look at potential non-linearities for certain years of schooling, the years for when the student is expected to graduate (Pons & Blanco, 2005; De Silva, 2009). However, when looking at years of schooling as a measurement for graduation there is no guarantee that the student is graduating that year or even at all. The dataset used in this study provides details about graduation, making it more suitable to investigating sheepskin effects.

Lastly, another contribution our research provides to the literature is investigating sheepskin effects through the method of matching, this is something that, to our knowledge, has not been done before. Also, most other studies investigate sheepskin effects on high school level data while our research investigates the phenomenon using higher education data with the ability to track the students' academic journey from high school all the way through higher education degree completion. We are also able to see how much time the students spent in their respective programs. The abovementioned gaps in the literature have all been considered when shaping the framework for this research on sheepskin effects. Leaving us with the questions:

1. Has there been sheepskin effects in the Chilean labor market the recent years?
2. How do the effects differ when controlling for gender, PSU-score, educational program, year of enrollment, years spent in higher education and type of institution?

4. Method

This section introduces a closer look at what the econometric method that was used as well as how the research was structured and executed.

4.1 The Method of Matching

The main concern when evaluating a treatment effect in a non-experimental study (where there is a lack of natural experiments) is the fact that the treatment variable is not randomly assigned. In this scenario there are several different methods that could be used. We chose to use a matching method.

The method of matching is about evaluating the effect of treatment by comparing observations that are similar on a set of variables and who differ only in treatment. The control group then serves as a counterfactual and an eventual discrepancy in the output variable would consequently be attributed to the treatment. The method rests upon two critical assumptions:

Assumption A.1: Unconfoundedness

"Selection into treatment is completely determined by variables that can be observed by the researcher, so conditioned on these observable variables the assignment to treatment is random."

It is therefore assumed that the covariates that are used will together constitute all explaining factors for the treatment enrollment. This means that when those variables are controlled for, the discrepancy in the output variable between the matched pairs will be solely attributed to the treatment effect.

Assumption A.2: Overlap

“There is a substantial overlap between the treatment and control on the propensity scores.”

This assumption ensures that there will be a corresponding control observation for each treatment observation.

For this research, the method of nearest neighbor matching was used in two separate ways where we:

1) Use the nearest-neighbor estimator with a Mahalanobis distance to calculate the average treatment effect.

2) Do a nearest-neighbor matching use propensity scores.

4.1.1 A Brief Mathematical Summary

In brief, the treatment effect ($\hat{\tau}_M$) is the average difference between individuals in the treatment group and their corresponding match(es)

The generic matching estimator for ATE:

$$\hat{\tau}_M = \frac{1}{N_T} \times \sum_{i \in \{D=1\}} [y_i - \sum_{j \in \{D=0\}} w_j(j) \times y_j]$$

Where y_i corresponds to the outcome variable, there are N_T treated units, and N_C control units. Then the counterfactual outcome (y_j) can be the outcome for the unit who is closer to *i*, or a linear combination of outcomes of observations that are similar to *i*, where $w_i(j)$ represents the weight in the linear combination.

$$w_i(j) : i : 1,2, \dots , N_T \; ; \; j : 1,2, \dots , N_C$$

And:

$$\sum_j w_j(j) = 1$$

4.1.2 Nearest-Neighbor Matching Estimator Using Mahalanobis Distance

Nearest neighbor matching is one of the most common matching methods used in the literature. Under the overlap assumption it provides a causal estimate under the selection on observables. The method pairs treated observations with untreated ones and bases the match on the shortest distance using some distance metric such as e.g. the Euclidian or the Mahalanobis distance.

Based on the brief notation from the mathematical summary outlined above the following apply for nearest-neighbor matching,

$w_i(j)$ equals one for the control unit with the closest X_j to X_j

This could be e.g. closeness in the Euclidean way however the Mahalanobis distance can also be used.$w_i(j)$ selects the “nearest (control) neighbor” j to the treated unit i and $\hat{\tau}_M$ computes the mean difference between each treated unit and its nearest control neighbor, producing a valid casual estimate under the selection on observables assumption, assuming that there is sufficient overlap between the treated and control groups.

The units of measurement for each element of X_i is arbitrary, so it may not make sense to weight each component equally when computing the distance between two points.

The Mahalanobis distance metric, that we used in our research is seen here below: $((X_i - X_j)' \sum_x^{-1}(X_i - X_j))$

Where $\sum_x^{-1}(X_i - X_j)$ is the covariance matrix of X (Solis, 2019).

4.1.3 Propensity-Score Matching Using Nearest-Neighbor Distance

For this alternative way of matching, instead of finding suitable matching candidates based on a set of characteristics by calculating either a Euclidian or Mahalanobis distance, treated and untreated observations are matched purely on the estimated probability of getting the treatment (propensity-score).

$$P(X) = \Pr\,(T|X)$$

Where T stands for treatment and X is the set of covariates that are controlled for. A key assumption for propensity score matchings is that participation in treatment is independent of outcome conditional on the covariates (see the assumption of unconfoundedness above).

To calculate the propensity scores the following process is followed. Firstly, a logit or probit model is used to estimate the program participation as a function of the covariates. A logistical regression is the most common way and further what we used in our analysis, and the regression model is displayed below:

$$ln\frac{e(x_i)}{1-e(x_i)} = \ln\frac{PR(z_i = 1|x_i)}{1-\Pr(z_i = 1|x_i)} = a + \beta^T x_i$$

Where:

$$e(x_i) = \Pr\,(z_i = 1|x_i)$$

$$e(X_i) = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_iX_i$$

And where:

b_0 is the intercept

b_i is the regression coefficient

X_i the treatment variables and covariates (random variables)

x_i observed value of variables

After this, the predicted values for each variable are used to calculate a propensity score for every observation in the treatment and control group.

The treatment variable is a binary in the logistic regression equaling 1 for treatment and 0 for non-treatment. The adjustment for the estimated propensity scores is then accomplished through matching. The coefficients are used together with the observation values for each covariate in the calculation of propensity scores in the following manner.

$$e(x_i) = \Pr\,(T_i = 1 | x_{1i}, x_{2i}, \dots, x_{ki})$$

$$\hat{e}(x_i) = \frac{1}{1 + e^{-(\hat{a} + s\hat{b}_k x_{ik})}}$$

Where $\hat{b}$ is the coefficient estimators for all covariates.

When propensity scores have been calculated for all observations, these are used to match all treated observations to a corresponding control observation by setting a tolerance level.

In the teffects psmatch STATA command that was used, the command implements a nearest-neighbor matching on the estimated propensity scores. The objective is to minimize the absolute difference between the propensity scores for the treatment and control groups. This could be done though a 1:1 matching or a 1:k matching where each treated observation is matched to either 1 or k (several) control observations. The accepted difference is set by declaring a tolerance level. Figure 1 below tries to capture the explained behavior.

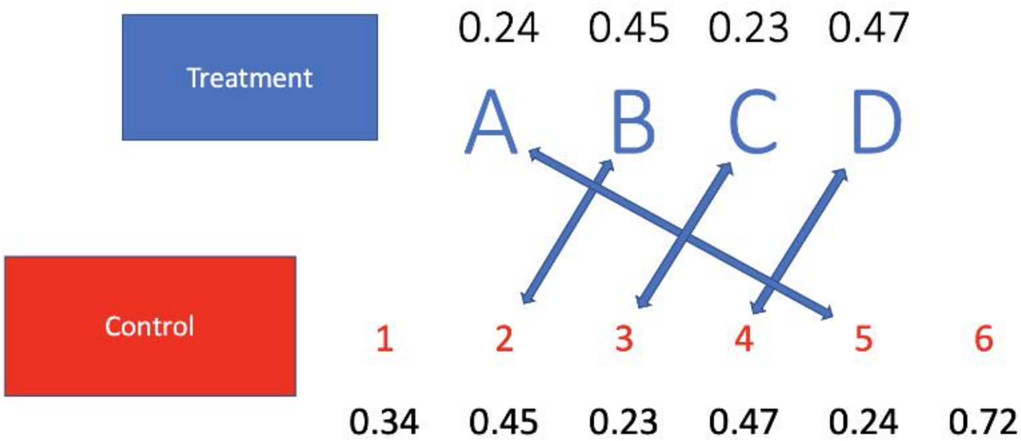

Figure 1: 1:1 nearest-neighbor matching based on propensity scores.

$$M(P_i) = \min\ |P_i - P_j|$$

M represents the set (one or several) of control observations *j* that are matched to the treated observation *i* based on the propensity scores. *P* refers to the propensity scores (Thavaneswaran, 2008).

Documentation about the STATA commands that we used can be found in the Appendix.

4.2 Research Design

On a fundamental level, we wanted to investigate the phenomenon of sheepskin effects by comparing graduated students to their non-graduated peers that otherwise shared the basic characteristics in the context of education and more broadly. After having conducted a thorough literature review, some common patterns were seen regarding what variables that were usually controlled for in similar studies. The typical variables were ethnicity and gender. The dataset used in this study, allows for controlling for the latter, but we also included additional variables based on our own assumptions of what variables would influence either the treatment or the outcome variable. In the initial research set up the covariates were gender, year of enrollment and program code with exact matches on gender and program code in the nearest-neighbor context. For the propensity-score matching, we introduce this method for control later in the analysis since the canned STATA command *teffects psmatch* did not allow for fixed matches on certain covariates (program code, gender etc).

When considering the Human Capital Model, described in the background, it is desirable to reduce the selection bias through finding a way of controlling for past academic success. Arguably, students who end up graduating will, on average, tend to be more motivated and intelligent than their peers, allowing them to complete their studies with more ease. To account for this, past scores on the university enrollment test in Chile (PSU) were used, which is mandatory for all applicants to universities. Particularly scores for the language and mathematics parts were used since these are done by all test takers and furthermore provides a well-rounded picture of the student's cognitive state prior to university enrollment. In addition to comparing students with a similar record of academic ability, another aspect that was considered to reduce the selection bias was making sure that the matched pairs had a similar history of limited labor market experience as well as no past exposure to higher education. This was carried out by excluding all observations in the PSU dataset that did not take the PSU test the same year as they graduated from high school and then merging the PSU data with the available enrollment data. This meant a sample that solely consisted of students who were newly graduated from high school when entering higher education.

To construct the treatment variable, two publicly available datasets were used where the first one included all enrolled students between the years 2007-2016 and the second one contained all graduated students over the same time. By merging these two datasets, a binary variable *graduated* could be created to serve as the treatment. With a treatment in place, as well as a set of suitable covariates to use for the matchings, the last piece was to find a suitable outcome variable. For the purposes of investigating sheepskin effects, this would have to be income or salary for one or several years after enrollment in higher education. Ideally, this income data would include years when the students who chose to obtain his or her degree had had time to accomplish that. When looking at the descriptive statistics in the results portion (Table 1 and Table 2) we see that the average number of years spent in higher education for graduates is 5.043 years why which we decided to only look at salary 5 to 11 years after higher education enrollment (11 being how far the data went). This average is biased by the existence of many programs not requiring that many years of schooling while more academically heavy programs such as Law, Medicine or Engineering often require 6 to 7 years of schooling. We discuss this further in relation to our results in the Discussion section.

Initially, a dataset describing formal sector earnings for Chilean individuals 2007-2016 was used. The data was assembled and put together by the Ministry of Labor in Chile. The raw data included number of active months in the formal labor sector for each year as well as the annual income registered in thousands of Chilean pesos (CLP). Using these two, the monthly income could be calculated and mapped to a corresponding year after enrollment for all observations.

Conducting our analysis using a couple of different matching methods (see former section for a more thorough review of the methods) the strategy was to first run the tests when controlling for a few fundamental covariates based on the literature background and then incrementally add more to see if the results would differ or remain the same. The first variables to control for were gender, year of enrollment, and the program of enrollment since they were all considered to be fundamental for distinct reasons in the research context. Gender typically has a large effect on income where several studies have shown that men consistently earn higher wages than their female counterparts thus creating a direct relationship with the outcome variable. Additionally, the student's year of enrollment was included as a covariate since labor market entry income levels can differ a lot depending on policies and the macroeconomic climate at the time the student leaves higher education and will thus influence the income levels. Lastly, matching individuals from the same programs of enrollment allows comparing individuals who would have the same education. This is necessary since income levels are dependent on the area of study, making it less interesting to compare e.g., a doctor to a middle school teacher. When running the nearest-neighbor matching there was an exact match requirement on gender and program code. For an exact program code match to be achieved for all treated observations, all graduated individuals who did not have any non-graduated program code counterpart were excluded from the analysis.

After having controlled for the variables mentioned above, the next step was to extend the analysis to include past academic performance and the total number of years spent in higher education. The former is interesting since those who follow through on their academic degree will also tend to be more gifted and motivated. By controlling for past PSU results in language and mathematics, the intention was to eliminate that variation and study how the effect might differ when doing so. This is an aspect that has often been overlooked or impossible to control for due to restricted data access for other researchers studying the signaling effect of education. The latter aspect is interesting when scoping in on the sheepskin effect versus the value of education more generally. If the time spent in the student's respective programs differ a lot this will inextricably imply variation attributed to e.g., human capital acquisition where students have more time to acquire competences for which they are later rewarded when entering the labor market. The variable of total amount of years spent in higher education was created by appending the enrollment info for all existing years and counting how many times each individual was enrolled to his or her program.

We also introduce the propensity-score matching with all the above-mentioned controls except for program code to see if the results of the two methods converge. After having faced indications of composition effects, we address it through running a regression on how the likelihood of entering the labor market is affected by graduation, defined as 1 if the individual had a positive salary for any of the 5-11 years after enrollment and 0 otherwise. We then ran a logistical regression using this variable as the dependent variable. Furthermore, we plot sheepskin effects in Figure 7 where the matches only include labor market participants for each year I.e., restricting the matching selection to observations with non-zero salaries in the outcome variable.

Finally, a heterogeneity analysis is performed where the robustness of the results is checked by stating conditions on gender, year of enrollment as well as the type of institution. By looking at how the effect differs for different sexes, a comparison with the existing literature can be made to see if past results are consistent with this research that is performed using newer data. Conditioning the year of enrollment will enable investigating how eventual sheepskin effects can differ over time. We will either see a temporal consistency or results showing differences for different cohorts. A comparison to the existing literature will be made. The thirdly added condition relates to the type of institution. This will enable investigating whether sheepskin effects differ for e.g., private, and public institutions as well as comparing the effect for professional/technical and more academically oriented institutions. The previous research on this topic is limited why the results will be an interesting contribution to the literature.

5. Data

This section provides a short description of the data used.

5.1 Overview of the Dataset

The data that we used in our analysis was put together using several different datasets, each contributing with its piece to the puzzle.

Variable name	Description
year_of_enrollment	The year the student enrolled in higher education.
language_score	PSU result on the verbal section.
math_score	PSU result on the quantitative section.
female	Gender of enrolled student. 1 means woman.
graduated	Treatment variable and equals to 1 if the student took his or her degree.
uni_years	Total number of years spent in higher education.
program_code	Code corresponding to the academic program the student entered.

type	The type of academic institution e.g. Private universities.
member_of_labor_market	Binary variable equaling 1 if the individual had a salary for one of the covered years after enrollment and 0 if not.
salary_year_X	Salary X year(s) after higher education enrollment.

Table 1: Variable definitions.

We combine three sources of publicly accessed data at the individual level. First, we use enrollment and graduation data from the Ministry of Education in Chile *(Ministerio de Educación)*. This data contains the annual enrollment registries from 2007 to 2016 and covers all higher education institutions in the country. We can follow the students though out their degrees until completion. Moreover, we use a PSU score data set that contains the admission test score (PSU scores) and a wide range of characteristics, such as gender, date of birth, parental education and parental work status. Finally, we applied for data on earnings to the Ministry of Labor that collects data at the individual level for all workers in the formal sector. Although our application was successful, we needed to travel to Chile to run our final regressions which was not possible due to health restrictions imposed because of the Covid19 pandemic. Instead, we used a sample data to estimate the effects in the labor market. We observe yearly earnings from 2006 to 2018 along with the number of months active in the labor force. This allows us to infer monthly earnings to base our regressions on.

6. Results

In this section we present the results of our matchings. Firstly, we present some descriptive statistics that summarize the treatment and control groups to then present some balance of covariates diagnostics showing that the matching estimators are working. The remaining parts of the section is dedicated to the actual results where we first control for some basic covariates to then expand the analysis to include more. Finally, we present the results from our heterogeneity analysis where we look at how the effects differ for gender, types of institution and year of enrollment.

6.1 Descriptive Statistics for the Control and Treatment Groups

Variable	Obs	Mean	Std. Dev.	Min	Max
female	352258	.475	.499	0	1

year of enrollment	352258	2011.106	2.109	2007	2014
uni years	352258	3.894	1.888	1	9
math score	352258	516.559	107.982	0	850
language score	352258	511.179	105.887	0	850
program code	352258	14278.677	5321.007	114	22086
type	352258	2.777	1.013	1	4
salary year 5	352258	172.254	268.803	0	4013.212
salary year 6	352258	183.037	318.678	0	8246.467
salary year 7	352258	175.107	349.444	0	7150
salary year 8	352258	144.12	351.468	0	8237.71
salary year 9	352258	107.669	332.251	0	10502.203
salary year 10	352258	68.944	286.246	0	6688.896
salary year 11	352258	31.893	209.375	0	7984.907

Table 2: Summary of control group

Variable	Obs	Mean	Std. Dev.	Min	Max
female	129821	.556	.497	0	1
year of enrollment	129821	2009.071	1.694	2007	2014
uni years	129821	5.043	1.486	1	9
math score	129821	522.578	104.231	0	850
language score	129821	517.994	101.617	0	850
program code	129821	14098.04	5362.1	114	22086
type	129821	2.762	1.014	1	4
salary year 5	129821	280.697	331.766	0	3446.437
salary year 6	129821	388.384	416.037	0	3151.93
salary year 7	129821	468.521	497.372	0	5461.6
salary year 8	129821	486.24	573.105	0	7983.5
salary year 9	129821	426.657	616.993	0	5555.75
salary year 10	129821	305.257	596.648	0	5610.356
salary year 11	129821	154.765	480.46	0	8053.908

Table 3: Summary of treatment group

The dataset consists of 482 079 observations in total where 129 821 observations are graduates and 352 258 observations are non-graduates. The enrollment span is 2007-2014.

6.2 Balance of Covariates

Tables 4, 5 and F*igures 1, 2* below show the similarity of the treatment and control group based on the set of characteristics we are controlling for. We compare the *Standardized mean differences* and the *Variance ratios* before and after the new control and treatment groups have been formed based on the different matching estimators. We used the canned *Stata* command *tebalance summarize* to produce the following two tables. When looking at the standardized mean differences after matching we see that SMD < 0.1 for all variables which indicates that the distributions of control and treatment are suitable for matching and with balanced covariates (Lakens, 2013).

$$SMD = \frac{Difference\ in\ mean\ outcome\ between\ groups}{Standard\ deviation\ of\ outcome\ among\ participants}$$

	Raw		Matched	
	SMD	Variance ratio	SMD	Variance ratio
female	.162	.989	0	1
program_code	-.033	1.015	0	1
year_of_enrollment	-1.063	.645	.989	.621
language_score	.065	.920	.070	.896
math_score	.056	.931	.080	.904
uni_years	.676	.619	.614	.614

Table 4: Balance of covariates for nearest-neighbor matching

	Raw		Matched	
	SMD	Variance ratio	SMD	Variance ratio
female	.162	.989	-.051	.996
year_of _enrollment	-1.063	.645	-.004	.994
language_score	.065	.920	-.044	.908
math_score	.056	.931	-.074	.927
uni_years	.676	.619	.003	.721
type	-.014	1.002	-.045	1.025

Table 5: Balance of covariates for propensity-score matching

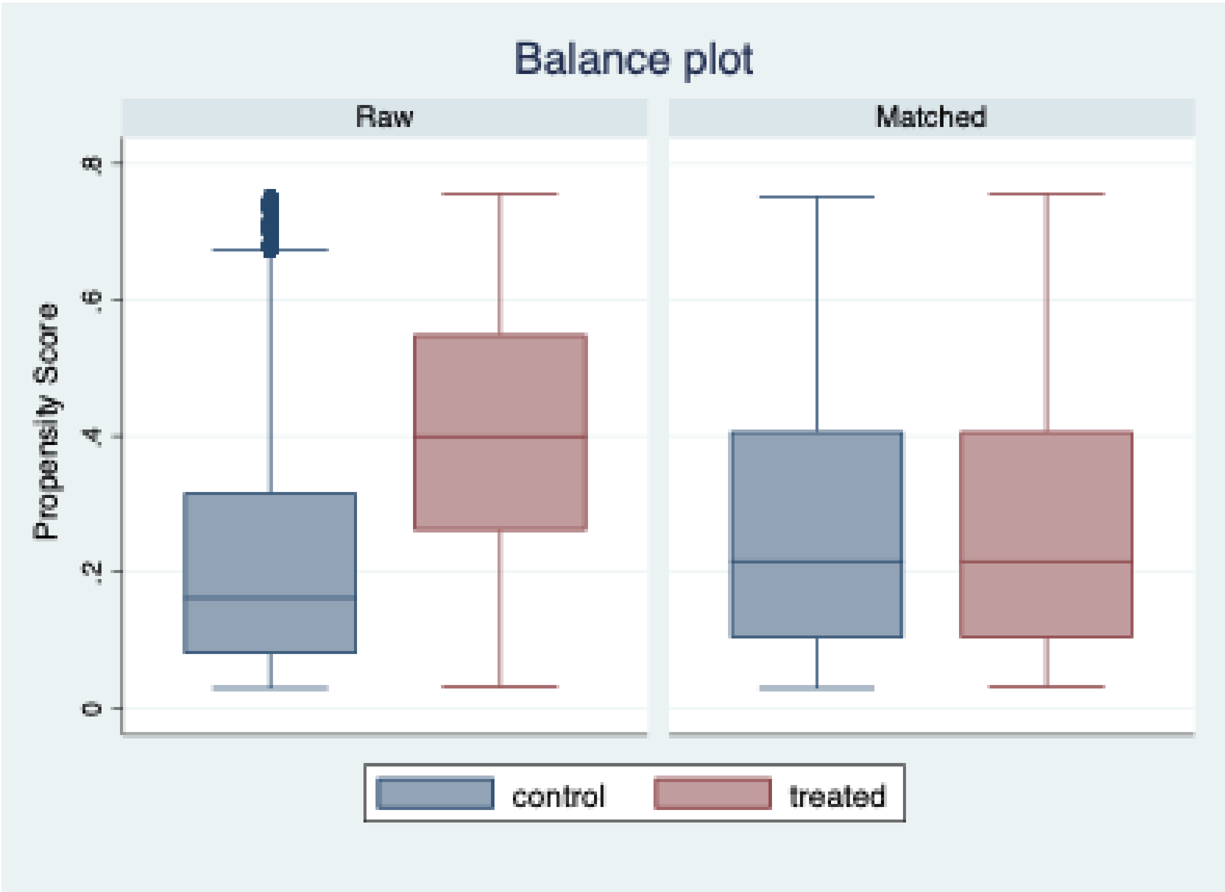

Figure 2: Box diagram for balance of covariates (propensity-score matching)

Figure 2, the balance plot, illustrates the distribution of the observations for the treatment group and the control group. It is evident that for the raw data the treatment group has higher propensity scores, meaning that they share a set of characteristics making them more likely to graduate. This could be e.g. more years of schooling, higher PSU scores etc. The matching estimators adjust for this as can be seen in the right half of the figure.

6.3 Main Results

This portion of the results introduce the basic set up where we run a nearest-neighbor matching method when controlling for gender, program code and the year of enrollment.

We observe a significant effect for all years considered. For earnings after 5 years since enrollment, we find that those who graduated enjoy higher earnings of 114 thousand CLP on average. The estimates are statistically significant with a z-score of 68. The same occur for years 6 to 11, where the effects on earnings increase up to the 7th year, to then decline to be close to zero. Although the effect is still significant at the usual significant levels.

Number of observations	482,079		
Distance metric	Mahalanobis		
Estimator	Nearest-neighbor matching		
Covariates	female, program_code, year_of_enrollment		
Exact matches	female, program_code		
Outcome variable	salary_year_5	salary_year_6	salary_year_7
Coefficient	114.486	182.689	231.457
Std. Err.	1.679	2.052	2.358
z	68.18	89.02	98.14
P>|z|	0.000	0.000	0.000
[95% Conf. Interval]	111.195 - 117.777	178.667 - 186.712	226.835 - 236.079

Table 6: Nearest-neighbor matching 5-7 years after enrollment.

Number of observations	482,079
Distance metric	Mahalanobis
Estimator	Nearest-neighbor matching
Covariates	female, program_code, year_of_enrollment
Exact matches	program_code, female

Outcome variable	salary_year_8	salary_year_9	salary_year_10	salary_year_11
Coefficient	213.449	136.104	63.917	27.475
Std. Err.	2.491	2.288	1.836	1.393
z	85.69	59.48	34.80	19.71
P>|z|	0.000	0.000	0.000	0.000
[95% Conf. Interval]	208.567 - 218.331	131.619 - 140.589	60.318 - 67.517	24.743 - 30.207

Table 7: Nearest-neighbor matching 8-11 years after enrollment.

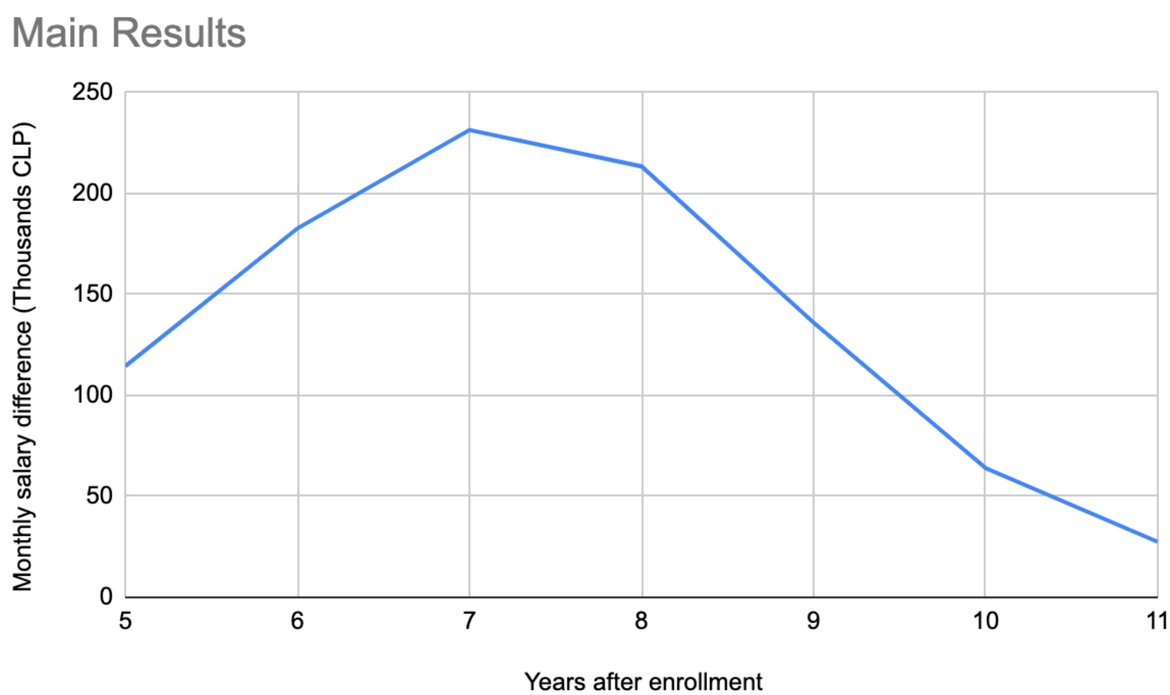

Figure 3: Sheepskin effects when controlling for gender, program code and year of enrollment. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

We summarize our results from the nearest-neighbor matching. Figure 3 and tables 6-7 suggest that there are, in fact, highly statistically significant sheepskin effects when controlling for gender, program code and the year of enrollment. The nearest-neighbor matching allowed for fixed matches, which was a necessary condition given our research design. Especially to be able to fix the program of enrollment. The effect initially increases for the first couple of years to then start converging towards zero after the point of 7 years after enrollment.

6.4 Sensitivity Analysis

This section investigates whether the effects shown in the original covariate set up remain when adding additional variables to the covariates mix. We start by adding mathematics and languages scores from the student's PSU test to then add total number of years spent in higher education. We plot the results over time when adding the two new covariates separately as well as collectively.

Number of observations	482,079		
Distance metric	Mahalanobis		
Estimator	Nearest-neighbor matching		
Covariates	female, program_code, year_of_enrollment, math_score, language_score		
Exact matches	female, program_code		
Outcome variable	salary_year_5	salary_year_6	salary_year_7
Coefficient	115.684	192.742	246.115
Std. Err.	1.534	1.886	2.176
z	75.40	102.15	113.10
P>|z|	0.000	0.000	0.000
[95% Conf. Interval]	112.677 - 118.691	189.044 - 196.440	241.850 - 250.380

Table 8: Nearest-neighbor matching 5-7 years after enrollment.

Number of observations	482,079
Distance metric	Mahalanobis
Estimator	Nearest-neighbor matching
Covariates	female, program_code, year_of_enrollment, math_score, language_score
Exact matches	program_code, female

Outcome variable	salary_year_8	salary_year_9	salary_year_10	salary_year_11
Coefficient	235.342	158.290	80.226	34.754
Std. Err.	2.321	2.177	1.802	1.391
z	101.35	72.68	44.52	24.98
P>|z|	0.000	0.000	0.000	0.000
[95% Conf. Interval]	230.791 - 239.893	154.021 - 162.558	76.694 - 83.758	32.028 - 37.481

Table 9: Nearest-neighbor matching 8-11 years after enrollment.

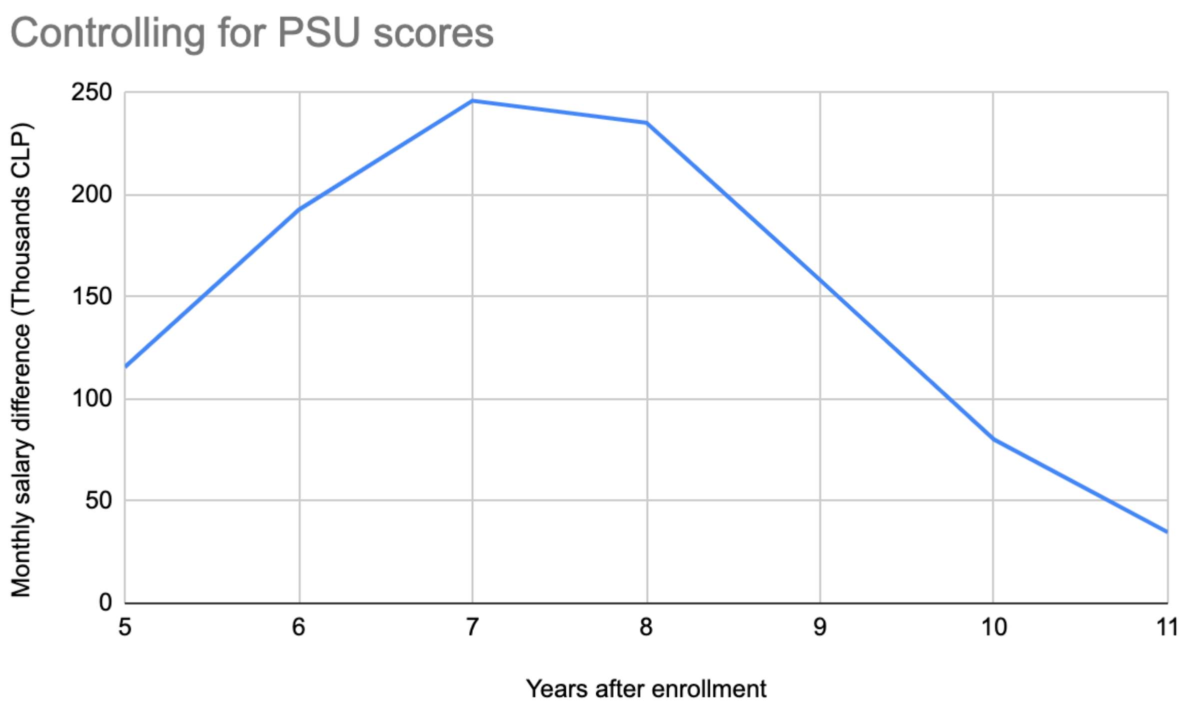

Figure 4: Sheepskin effects when controlling for gender, program code, year of enrollment and PSU scores. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

Figure 4 suggests that the effects remain similar after controlling for language and math scores on the PSU exam. The effect is even higher on average.

Number of observations	482,079		
Distance metric	Mahalanobis		
Estimator	Nearest-neighbor matching		
Covariates	female, program_code, year_of_enrollment, uni_years		
Exact matches	female, program_code		
Outcome variable	salary_year_5	salary_year_6	salary_year_7
Coefficient	105.244	172.664	222.816
Std. Err.	1.660	2.041	2.364
z	63.38	84.57	94.23
P>|z|	0.000	0.000	0.000
[95% Conf. Interval]	101.989 - 108.499	168.663 - 176.666	218.182 - 227.451

Table 10: Nearest-neighbor matching 5-7 years after enrollment.

Number of observations	482,079			
Distance metric	Mahalanobis			
Estimator	Nearest-neighbor matching			
Covariates	female, program_code, year_of_enrollment, uni_years			
Exact matches	program_code, female			
Outcome variable	salary_year_8	salary_year_9	salary_year_10	salary_year_11
Coefficient	214.716	144.528	74.124	33.915
Std. Err.	2.515	2.350	1.963	1.536
z	85.35	61.49	37.75	22.07

P>|z|	0.000	0.000	0.000	0.000
[95% Conf. Interval]	209.785 - 219.647	139.921 - 149.135	70.276 - 77.973	30.902 - 36.927

Table 11: Nearest-neighbor matching 8-11 years after enrollment.

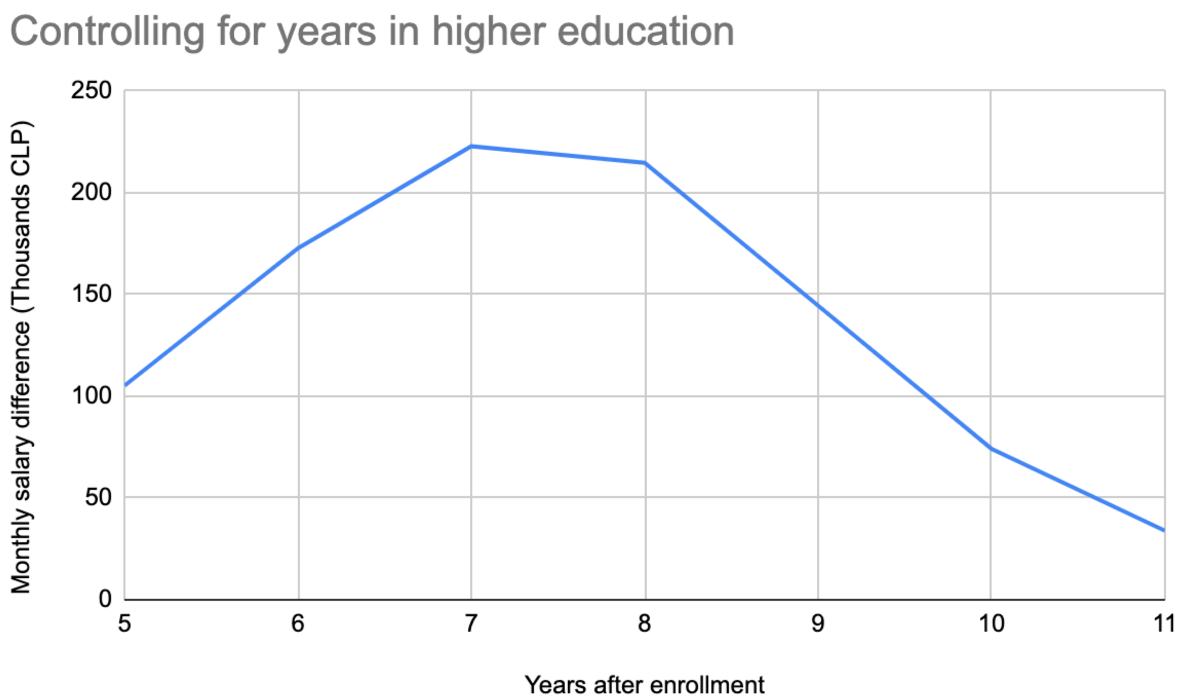

Figure 5: Sheepskin effects when controlling for gender, program code, year of enrollment and number of years in higher education. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

The effect is also unchanged when adding total number of years in higher education to the set of covariates, however, showing a slight decrease for all years.

Number of observations	482,079

Distance metric	Mahalanobis		
Estimator	Nearest-neighbor matching		
Covariates	female, program_code, year_of_enrollment, math_score, language_score, uni_years		
Exact matches	female, program_code		
Outcome variable	salary_year_5	salary_year_6	salary_year_7
Coefficient	110.517	185.712	240.826
Std. Err.	1.509	1.861	2.161
z	73.23	99.76	111.39
P>|z|	0.000	0.000	0.000
[95% Conf. Interval]	107.559 - 113.475	182.063 - 189.361	236.589 - 245.064

Table 12: Nearest-neighbor matching 5-7 years after enrollment.

Number of observations	482,079			
Distance metric	Mahalanobis			
Estimator	Nearest-neighbor matching			
Covariates	female, program_code, year_of_enrollment, math_score, language_score, uni_years			
Exact matches	program_code, female			
Outcome variable	salary_year_8	salary_year_9	salary_year_10	salary_year_11
Coefficient	238.098	166.227	89.450	39.988
Std. Err.	2.314	2.191	1.862	1.461
z	102.87	75.86	48.02	27.36
P>|z|	0.000	0.000	0.000	0.000

[95% Conf. Interval]	233.562 - 242.634	161.932 - 170.522	85.799 - 93.101	37.1239 - 42.852

Table 13: Nearest-neighbor matching 8-11 years after enrollment.

We also test the robustness of our results by introducing a propensity-score matching method that is summarized in tables 14 and 15.

Number of observations	482,079		
Estimator	Propensity-score matching		
Covariates	female, year_of_enrollment, math_score, language_score, uni_years		
Treatment model	logit		
Outcome variable	salary_year_5	salary_year_6	salary_year_7
Coefficient	113.059	114.296	121.621
Robust Std. Err.	3.347	1.572	1.602
z	33.77	72.66	75.89
P>|z|	0.000	0.000	0.000
[95% Conf. Interval]	106.498 - 119.621	111.213 - 117.378	118.480 - 124.763

Table 14: Propensity-score matching 5-7 years after enrollment using a logit treatment model.

Number of observations	482,079			
Estimator	Propensity-score matching			
Covariates	female, year_of_enrollment, math_score, language_score, uni_years			
Treatment model	logit			
Outcome variable	salary_year_8	salary_year_9	salary_year_10	salary_year_11

Coefficient	107.901	78.198	45.844	21.129
Robust Std. Err.	1.408	1.276	1.100	.856
z	76.60	61.27	41.64	24.66
P>|z|	0.000	0.000	0.000	0.000
[95% Conf. Interval]	105.140 - 110.662	75.696 - 80.699	43.686 - 48.001	19.450 - 22.809

Table 15: Propensity-score matching 8-11 years after enrollment using a logit treatment model.

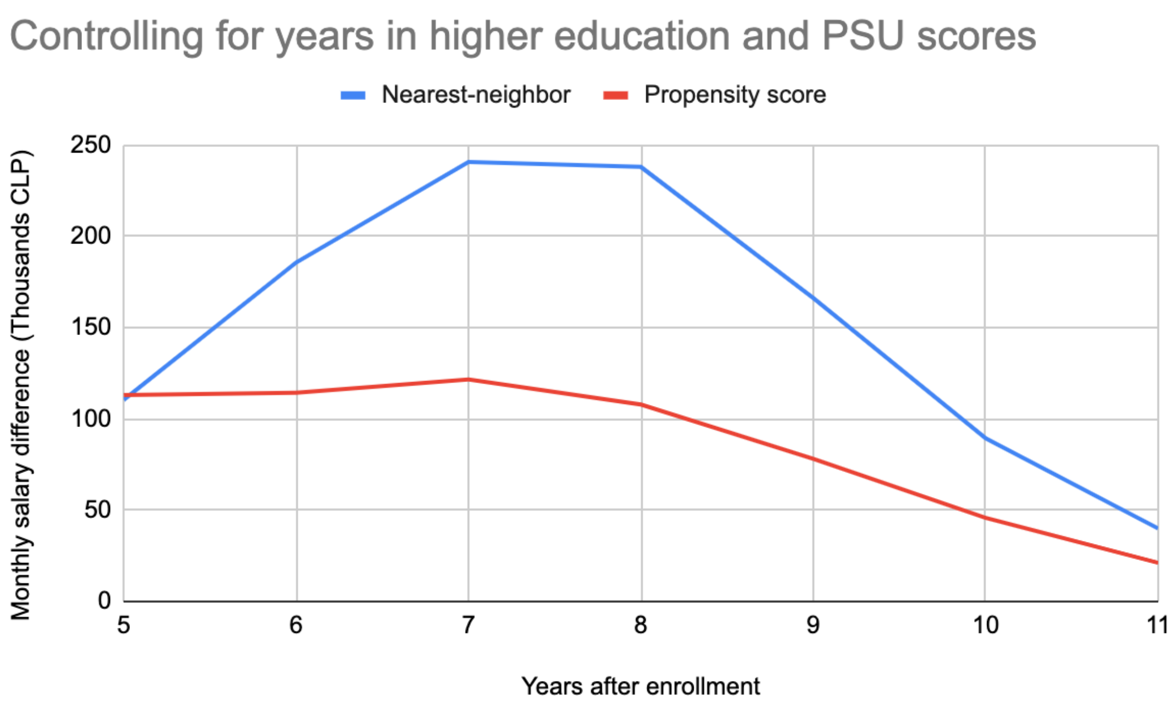

Figure 6: Sheepskin effects when controlling for gender, program code, year of enrollment, PSU scores and number of years in higher education. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

Figure 6 suggests that the inflating effect arising from the PSU scores slightly offsets the reduction in effect arising from the total number of years in higher education resulting in a slight net increase in effect compared to the original set up. In this graph, we further introduce a propensity score matching to see how those results hold up against the nearest-neighbor estimator. The propensity-score matching method did not allow for fixed matches which made us leave out the program code. This method shows a similar behavior with an effect that converges towards zero with time, however, without the initial overshoot.

Overall, the sheepskin effect originally demonstrated seems to be robust and insensitive to adding additional variables.

Below can be found a visual comparison of the main results and the sensitivity analysis.

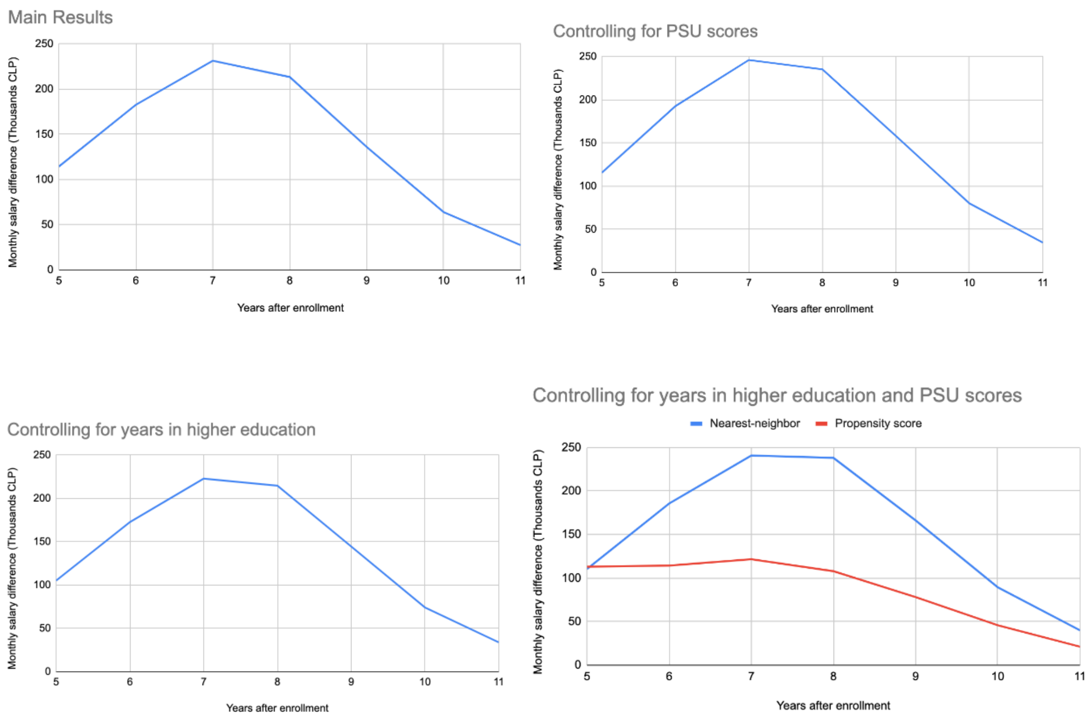

Figure 7: A comparison of the results from the main set up and the sensitivity analysis. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

<table>
<tr><td>Number of observations</td><td colspan="5">482,079</td></tr>
<tr><td>Prob > chi2</td><td colspan="5">0.0000</td></tr>
<tr><td>Entered labor market</td><td>Coefficient</td><td>Robust Std. Err.</td><td>Z</td><td>P>|z|</td><td>[95% Conf. Interval]</td></tr>
</table>

Graduated	.4132943	.0114912	35.97	0.000	.390772 - .4358166
Language score	-.0010847	.0000552	-19.64	0.000	.390772 - .4358166
Math score	-.0003295	.0000552	-5.97	0.000	.390772 - .4358166
Female	-.2504725	.007682	-32.61	0.000	.390772 - .4358166
Uni years	.0755775	.0041521	18.20	0.000	.390772 - .4358166
Type	-.1296266	.0042194	-30.72	0.000	.390772 - .4358166
Program code	-.0000204	7.82e-07	-26.04	0.000	.390772 - .4358166
Year of enrollment	-.6847212	.0036622	-186.97	0.000	.390772 - .4358166
Constant	1378.968	7.375297	186.97	0.000	.390772 - .4358166

Table 16: Logistical regression that looks at how the different variables explain the probability of entering the labor market any of the years 2007-2016.

Table 16 shows the output of a logistical regression that investigates how the different independent variables explain how likely an individual is to enter the labor market. When looking at the coefficient for *graduated* we see a statistically significant strong positive correlation.

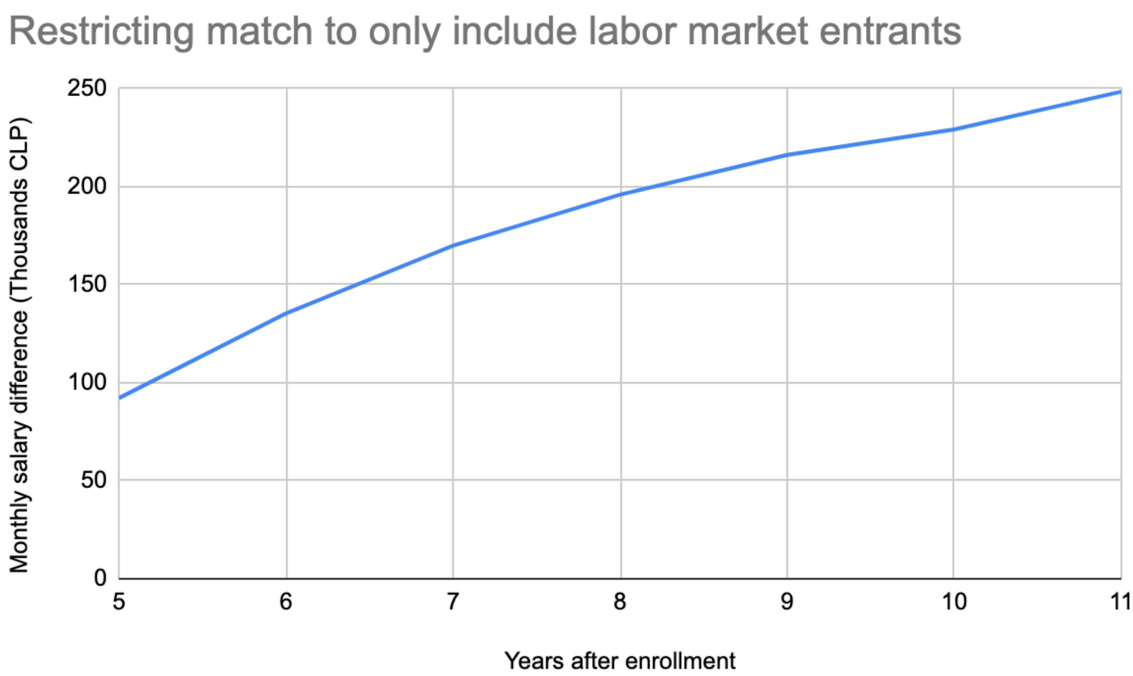

Figure 8: Sheepskin effects when controlling for all covariates and restricting the matching to only include labor market participants. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

6.5 Heterogeneity Analysis

This section introduces a couple of different conditions to investigate how the effects differ when comparing gender, different cohorts, and distinct types of institutions.

Estimator	Nearest-neighbor matching						
Distance metric	Mahalanobis						
Covariates	female, program_code, year_of_enrollment, math_score, language_score, uni_years						
Condition	female = 0	female = 1	enrollment year = 2007	enrollment year = 2008	enrollment year = 2009	enrollment year = 2010	enrollment year = 2011

Coefficient year 5	113.468	107.259	33.515	44.235	77.413	154.275	148.958
Coefficient year 6	199.664	171.249	84.812	125.938	190.088	210.790	136.530
Coefficient year 7	263.680	217.345	141.597	220.344	246.498	199.123	144.378
Coefficient year 8	267.450	208.002	201.728	238.271	224.133	202.543	-
Coefficient year 9	197.167	134.817	215.819	219.072	229.609	-	-
Coefficient year 10	104.650	73.869	201.882	223.197	-	-	-
Coefficient year 11	46.765	33.174	204.173	-	-	-	-

Table 17: Nearest-neighbor matching where we filter for gender and different years of enrollment.

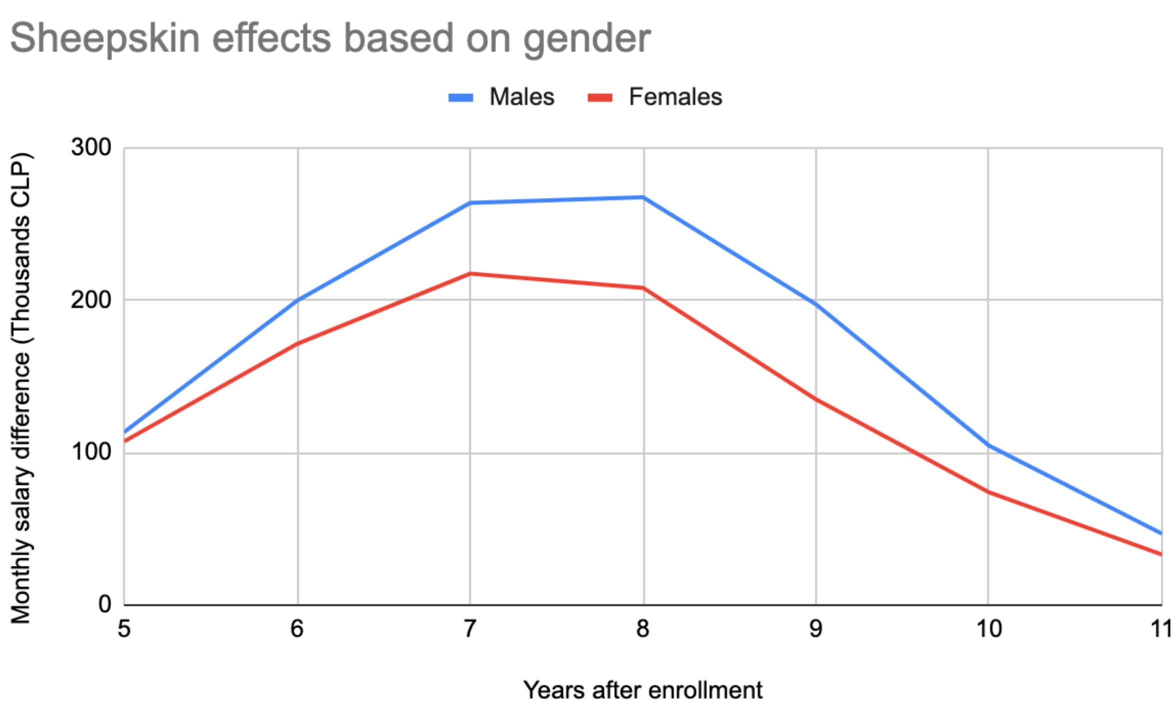

Figure 9: Sheepskin effects with all covariates added and looking at men and women separately. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

Figure 9 suggests that there is in fact a tangible gender difference in signaling value especially for the early years entering the labor market. The effect discrepancy then converges for each subsequent year after 8 years after enrollment.

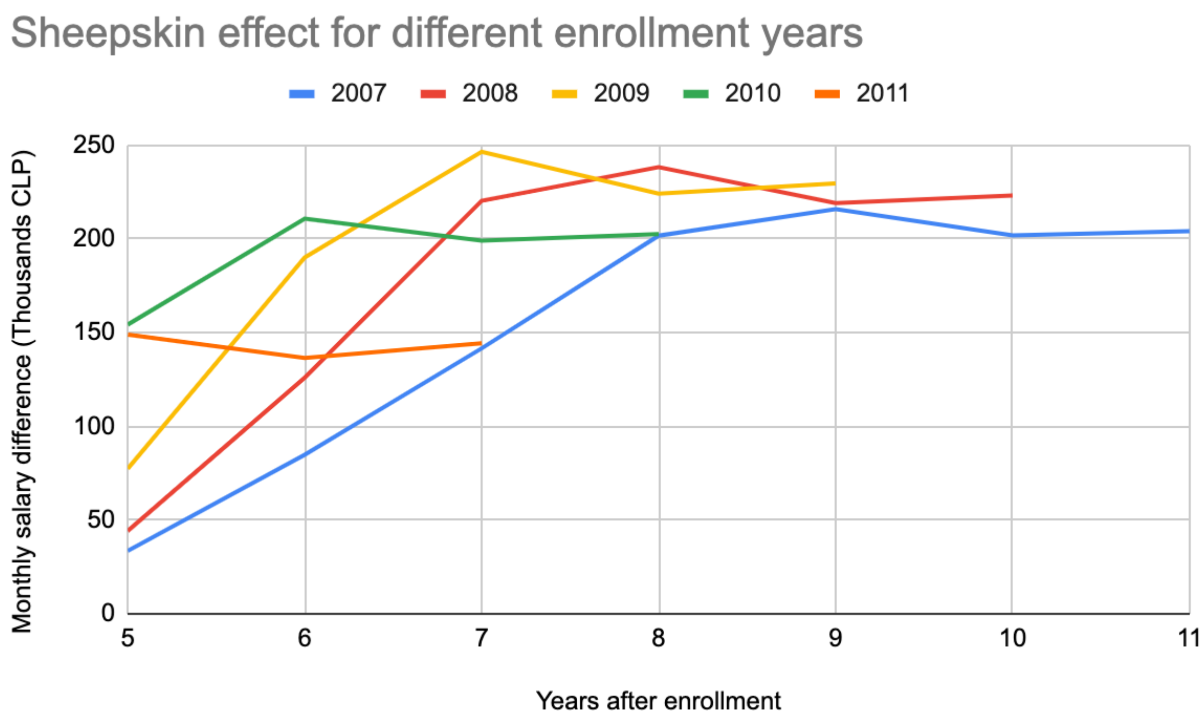

Figure 10: Sheepskin effects with all covariates added and looking separately at different years of enrollment. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

Figure 10 shows that sheepskin effects have increased in the initial stages for the past years to, in the later stage, stabilize at a level around 200K CLP.

Estimator	Nearest-neighbor matching			
Distance metric	Mahalanobis			
Covariates	female, program_code, year_of_enrollment, math_score, language_score, uni_years			
Year after enrollment	CFT	IP	UCRUCH	UPRIVADAS
5	134.040	127.478	96.452	93.220
6	133.049	147.640	237.147	188.012
7	109.550	149.256	356.886	267.883
8	90.099	113.962	405.338	257.225
9	62.686	77.455	303.411	164.779
10	42.380	43.856	162.737	85.521
11	19.984	18.464	74.467	36.892

Table 18: Nearest-neighbor matching where we filter for several types of institutions.

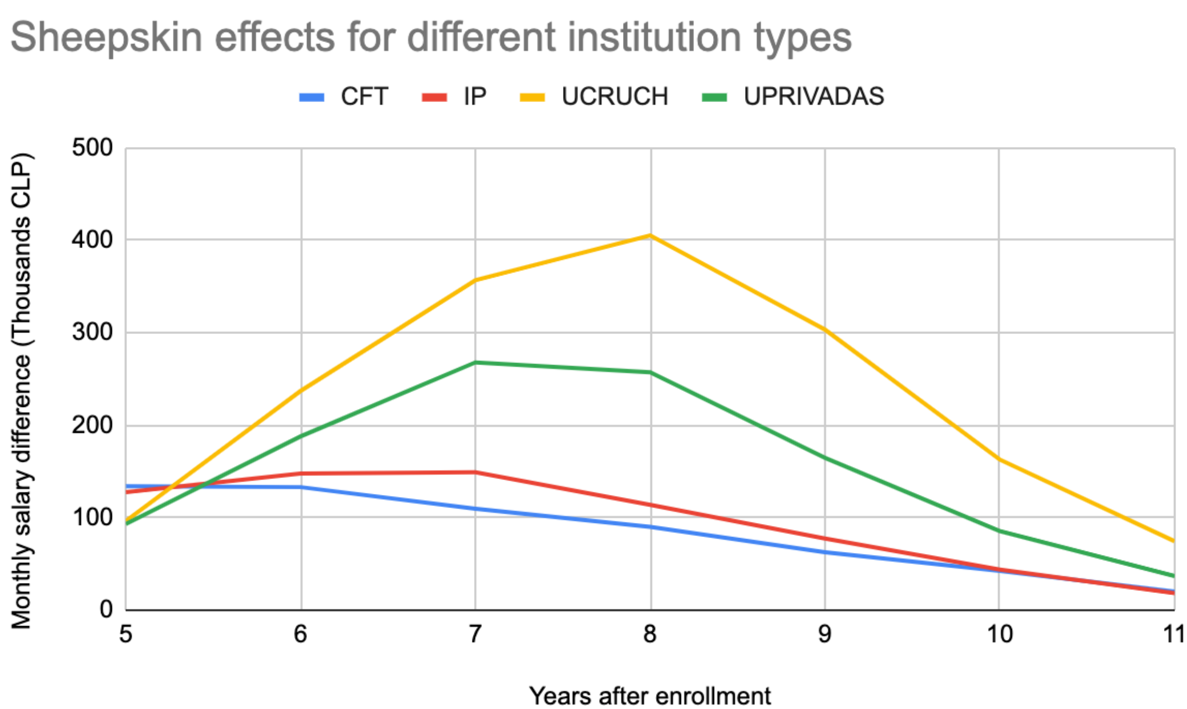

Figure 11: Sheepskin effects with all covariates added and looking separately at several types of higher education institutions. Based on enrollment, PSU and earnings data from the Ministries of Education and Labor in Chile (2007-2016).

Figure 11 suggests a substantial difference in signaling effect between the different types of institutions where UCRUCH (Consejo de Rectores) show the by far highest levels followed by private universities. CFT and IP show remarkably similar effects.

7. Discussion

The discussion is split into four parts. First, there is a results discussion that brings up the main points from the results of this paper, offering explanatory models as well as relating the findings to the existing literature. Secondly, there is a discussion about the main aspects about the research design and data used that speak in favor of the validity of the results when comparing to the modus operandi of other researchers investigating sheepskin effects in the past. Thirdly, there is a discussion about the main drawbacks with the research design and things that should potentially be added or improved for the findings to have more validity. Lastly, we also present a plan on what needs to be further done and explored on this topic to improve the research and enable a generalization of the results.

7.1 Result Discussion

7.1.1 Commenting on the Main Results

This study has significant results supporting that there are sheepskin effects from education in the Chilean labor market which potentially could be found when in the next step using real data instead of the limited sample that was used in this analysis. The results show an initial increase in sheepskin effects for the first couple of years after enrollment to later transition to a continuous decrease that slowly but steadily converges towards a zero-discrepancy state. The results are highly statistically significant, with a significance level lower than 0.001, and the sheepskin effects are furthermore large for all years in the dataset. The behavior for the propensity-score matching, as can be seen in figure 6, is different where there is no initial spike but a constant effect for the first two years that later tapers off in accordance with the nearest-neighbor behavior. We deemed the nearest-neighbor estimator to be the most suitable method for our purposes since it allows fixed matches for the program code. This is an important requirement since the variable is numerical and where the numerical closeness is unrelated to the similarity of the actual programs. We therefore put more emphasis on the results observed using the nearest-neighbor method.

The explanations for the initial increase in sheepskin effects could be manyfold, however, our main hypothesis is that many degrees that are more academically heavy such as Medicine, Engineering and Law require more than five years to completion which means it might be more accurate to measure the effects from e.g. year 7 or 8 after enrollment. In that case we would instead see diminishing sheepskin effects from the first year in the time series and a result that also better corresponds to the current literature. Belman and Heywood, for example, present results that support that sheepskin effects are diminishing over time (Belman & Heywood, 1997). Another potential explanation for the increase in sheepskin effects early on could be that some students with certain characteristics take more time to enter the labor market. For instance, students that do not graduate might face a lower probability for employment than their graduated peer. Furthermore, it could be due to that the research is made on different geographical areas, Chile and the US respectively, hence differences in the labor market such as different standard entry-level wages, or different wage trends in general, could affect the trends for sheepskin effects.

Disregarding the initial increase that is seen in the results, our graphs show sheepskin effects that are diminishing over time. This is consistent with the signaling model since the matching method controls for characteristics such as PSU scores and total amount of time spent in the subject’s program, making sure that the treatment and control observation have the same productivity. For the early years of employment there will be signaling effects of the academic degree. With time, however, employers will learn workers’ productivity levels leading to a decrease in salary differences when salaries are converging to a salary that corresponds to the accurate productivity level. This expected behavior is exactly what is reflected in our data.

7.1.2 Commenting on the Sensitivity Analysis

7.1.2.1 Controlling for PSU Scores

The results are remarkably similar after adding PSU scores to the set of covariates as can be seen in *Figure 4*. We still see the early increases in sheepskin effects to peak at around 8 years after enrollment after which the effect is reduced for each subsequent year. The coefficients are even a bit higher on average when comparing to the original covariate set up. This struck us as counterintuitive since the initial hypothesis was that controlling for cognitive ability would diminish the overall effects since graduates are typically more high achieving and motivated students, dedicated to follow through on their academic pursuits. The increase in effect is, however, marginal and we instead chose to focus on the larger picture which thus solidifies the findings of signaling value from academic degrees.

7.1.2.2 Controlling for Time Spent in Higher Education

When in the next step adding total number of years in higher education to the set of covariates, we did not, similarly to the case of PSU scores, observe much of a qualitative shift. A difference, however, is that the coefficients are now smaller compared to the original covariate set up which is more in line with what would be expected. This since many non-graduates will tend to exit higher education in a much earlier stage. By comparing graduates to non-graduates with a large discrepancy in time spent in higher education, there will logically be a tangible difference in competence attributed to the accumulated human capital. This variation is erased when adding the years spent as an additional covariate. The similar qualitative behavior to that of the initial set up further solidifies the existence of sheepskin effects.

7.1.3 Commenting on the Heterogeneity Analysis

7.1.3.1 The Gender Effect

The results suggest that men consistently, on average, experience stronger sheepskin effects than women. The discrepancy in effect initially increases to peak at 8 years after enrollment. After this point, the differences seem to converge to be almost negligible 11 years after enrollment. These results are interesting and could potentially be attributed to gender discrimination in the labor market. Despite equal competence, men are rewarded more for a degree completion in the initial stages of their career. Then, as time passes, women can demonstrate abilities by performing their job well and getting paid thereafter, consequently removing these initial differences in education signaling.

7.1.3.2 Comparing Cohorts and Commenting on Composition Effects

When running the analysis and looking at how sheepskin effects differ for different years of enrollment, the results surprisingly have a different quality than what was shown with the initial matching set up as well as when adding PSU scores and total time spent in higher education. Here, we see that sheepskin effects 5-7 years after enrollment have increased for every year to then converge to a stable level around the point of 8 years after enrollment. This complicates the conclusion since the earlier shown results of demonstrated sheepskin effects might instead be attributed to composition effects. An additional factor in favor of this potential argument is that graduates are significantly more likely to enter the labor market (see table 16 in *Results*), thus implying more zeros in the outcome variable for the control group.

Furthermore, it is important in all research on sheepskin effects to not automatically rule out other explanations to the wage increase for obtaining an academic degree. For instance, some educations could have a structure where the students initially have to learn mainly theoretical frameworks and gain background knowledge in the academic field and continue to do so for the most part of their education. For the last part they then learn how to apply this knowledge, giving them a concentrated human capital boost in the end of their education, resulting in an unproportioned productivity increase for this period.

Additionally, many professions such as M. Ds and lawyers, require some sort of certificate or diploma to be able to practice that profession. This effectively means that non-graduates will not be eligible to the jobs available to the graduates thus making the comparison less interesting from a signaling point of view.

7.1.3.2 Comparing Different Types of Institutions

What can be seen when testing for the type of institution is that the largest sheepskin effects occur for Universidades CRUCH, followed by Universidades Privadas. For CFTs and IPs the sheepskin effects are smaller, but still present. These differences in size of sheepskin effects could be explained by the fact that the universities are more prestigious than the IPs and the CFTs, leading to greater signaling effects to employers for obtaining a degree from the universities. These degrees could therefore be seen as more valuable on the labor market.

7.2 Key Quality Indicators of the Research Design

Regarding the data, the admin data (sample) used in this study, contrary to the commonly used aggregate data, provides several benefits when researching sheepskin effects. Firstly, looking at the full population is better than just looking at individuals participating in a survey since a sample like a survey might not be representative of the full population. Secondly, when people take a survey there is no guarantee for their honesty. For certain questions participants might have reason to not tell the truth. In this case it might be that people are not honest about salaries, for personal reasons, and hence the results might not show the accurate sheepskin effects.

Another big advantage with our research is the method that we are using. When investigating sheepskin effects, it is easy to misinterpret the results and attribute too much of the effects shown to signaling. Using the matching method is one way of reducing the risk of falsely confirming sheepskin effects due to late-stage human capital boosts, since this method matches the observations that are most similar in all aspects that are controlled for except for if they obtain the degree or not. Particularly important here is that the research design allows control for total number of years in higher education.

Lastly, a major advantage with our research design is that it allows for tracking the individuals throughout their academic career, controlling for PSU scores and program code allows for a richer analysis and adds a dimension that has often been disregarded in other research papers.

7.3 Improvement Areas and Future Action Plan

When redoing the analysis with full data usage authorization, the intention is to also further fine tune the research design as well as extending the sensitivity and heterogeneity analysis to dig deeper into the different underlying mechanisms for sheepskin effects. For the sensitivity analysis, it would be interesting to incorporate the student’s socio-economic background in the form of the parents’ education and profession as well as the students' GPAs from high school. For the heterogeneity analysis, the next natural step would be to see how the effects differ for different programs of enrollment. The main reason this was not done in this cycle was due to time constraints. The dataset contains over 20 000 program codes since these are unique for each institution and each program therein, furthermore, they have changed over time creating various program codes referencing the same program. An effective way forward would be to track these temporal changes and make sure that codes are consistent over time. The next step would be to cluster the program codes on category and then use these wider program categories as conditions. E.g., how does sheepskin effects differ between doctors and engineers?

An additional aspect we thought about was to instead find a way of displaying the coefficients in percentages to facilitate comparison to the existing literature. This is something that could be implemented when re-running the analysis with the full income dataset.

The richness of the dataset, as mentioned earlier, is one of the most important parts of our research design, however, there are some aspects that are not covered and that would be interesting to investigate when undertaking future work on this topic. One such thing would be looking more into the nature of the professions the subjects end up doing. It is likely that sheepskin effects will differ when comparing e.g., employees to entrepreneurs. The signaling of education would work differently where it, for employees, relates to conveying an image of themselves to the employer that shows their competence, for instance, through handing over a higher education diploma. For entrepreneurs, on the other hand, these are self-made to another extent and signaling would not affect salaries as much. However, the signaling value of education could be present for entrepreneurs in other settings, such as when trying to attract e.g., investors to an early-stage startup which would later have an impact on the net income of that founder/CEO.

Another improvement point for future research would be looking at data for hourly wages since this might give an even more precise estimate of the sheepskin effects. Since the salary data in this study is retrieved from a dataset, that provides information about yearly salary and number of months that the individual has was working per year, the final information that can be retrieved is the average monthly salary. Since individuals might not be working equally much, some might work part time and others may work full time and for certain periods of time some individuals might have parental leave. Their salaries would then differ more than what can be explained by sheepskin effects. The option to look at hourly wages instead of monthly salaries would therefore provide a better estimate for comparison between individuals.

In summary, there are some aspects that are left outside the analysis in the current research set up, even though we deem it to be near to complete. Furthermore, the limited sample size of the income data prevents us from generalizing our conclusions to apply to the overall population. However, since the data sample shares similar characteristics to the original, it is likely that many of the results would still apply. The next step would then naturally be to run all the tests again but this time using population wide data.

8. Conclusion

In conclusion, our research shows that there are significant sheepskins effects in the Chilean labor market when controlling for gender, program code and the year of enrollment. The effects also hold up when expanding the analysis to include cognitive ability and total time spent in higher education.

This is in line with the signaling model (Spence, 1973), since it could be explained by graduates initially getting an extra salary boost due to the signaling value of their diploma. This arises from the information asymmetry that exists between employers and new hires where an academic degree is regarded as a good way of measuring productivity. With time, however, light will be shed on the subject's real productivity level and the individual who did not graduate, however having a similar productivity level, will converge to the salary level of that of his or her graduated peer.

This also has support in the literature, where previous research has confirmed sheepskin effects in, for instance, Colombia, the US, Philippines, Sweden, Mexico and other (Rodríguez & Muro, 2015).

Furthermore, we make some interesting findings regarding how sheepskin effects differ between men and women as well as for different types of institutions. Men experience consistently higher sheepskin effect although the discrepancy decreases for each extra year spent in the labor market. Regarding types of institutions, students who attended the most prestigious universities in Chile (UCRUCH) experience superior signaling effects on their diplomas while as the signaling effect is not particularly strong for less known institutions who offer shorter programs more oriented towards technical and professional degrees.

Finally, we did discover that the found sheepskin effects could also be the result of composition effects where graduated students are significantly more prone to accessing the labor market thus implying a lot more zeros in the outcome variables for the control group. Furthermore, when separating the effects based on year of enrollment, we see another type of qualitative behavior over time where sheepskin effects 5 years after enrollment increase for every new cohort to later converge towards a more constant level.

References

Antelius, J., 2000. *Sheepskin Effects in the Returns to Education: Evidence on Swedish Data*.

Becker, G. S., 1993. *Human Capital : A Theoretical and Empirical Analysis with Special Reference to Education.* Third Edition ed. Chicago: The University of Chicago Press.

Belman, D. & Heywood, J. S., 1991. Sheepskin Effects in the Returns to Education: An Examination of Women and Minorities. *The Review of Economics and Statistics,* 73(4), pp. 720-724.

Belman, D. & Heywood, J. S., 1997. Sheepskin Effects by Cohort: Implications of Job Matching in a Signaling Model. *Oxford Economic Papers,* 49(4), pp. 623-637.

De Silva, I., 2009. Ethnicity and Sheepskin Effects in the Returns to Education in Sri Lanka: A Conditional Quantile Analysis. *International journal of development issues,* 8(1), pp. 61-79.

Ferrer, A. M. & Riddell, W. C., 2002. The Role of Credentials in the Canadian Labour Market. *Canadian Journal of Economics,* 35(4), pp. 879-905.

Hungerford, T. & Solon, G., 1987. Sheepskin Effects in the Returns to Education. *The Review of Economics and Statistics,* 69(1), pp. 175-177.

Lakens, D., 2013. Calculating and Reporting Effect Sizes to Facilitate Cumulative Science: A Practical Primer for t-tests and ANOVAs. *Frontiers in Psychology*.

McGuinness, S., 2003. Graduate Overeducation as a Sheepskin Effect: Evidence from Northern Ireland. *Applied Economics,* 35(5), pp. 597-608.

Ministerio de Educación, Subsecretaría de Educación Superior, 2020. *Ministerio de Educación (mineduc.cl).* [Online]
Available at: https://educacionsuperior.mineduc.cl/donde-estudiar/
[Accessed 12 May 2021].

Ministerio de Educación, 2020. *ayuda MINEDUC : Consejo de Rectores.* [Online]
Available at: https://www.ayudamineduc.cl/ficha/consejo-de-rectores-5
[Accessed 12 May 2021].

Ministerio de Educación, 2020. *Ministerio de Educación (mineduc.cl) Nuevo Sistema de Acceso Universitario reemplaza la PSU y aumenta reconocimiento de la trayectoria escolar de los jóvenes.* [Online]
Available at: https://www.mineduc.cl/principales-ejes-del-nuevo-sistema-de-acceso-universitario/
[Accessed 13 May 2021].

Olfindo, R., 2018. Diploma as Signal? Estimating Sheepskin Effects in the Philippines. *International Journal of Educational Development,* Volume 60, pp. 113-119.

Pons, E. & Blanco, J., 2005. Sheepskin Effects in the Spanish Labour Market: A Public–Private Sector Analysis. *Education Economics,* 13(3), pp. 331-347.

Preuniversitario Pedro de Valdivia, 2018. *Preuniversitario Pedro de Valdivia, ¿Cuáles Son Las Diferencias Entre Un CFT, IP y Una Universidad?.* [Online]
Available at: https://www.preuniversitariopedrodevaldivia.cl/blog/diferencias-entre-cft-ip-y-universidad/
[Accessed 10 May 2021].

Rodríguez, J. J. M. & Muro, J., 2015. On the Size of Sheepskin Effects: A Meta-Analysis. *Economics: The Open-Access, Open-Assessment E-Journal,* 9(1).

Solis, A., 2017. *Credit Access and College Enrollment.* Uppsala: s.n.

Solis, A., 2019. *Applied Econometrics, Lecture 4 Matching.* Uppsala: NEK - Uppsala University.

Spence, M., 1973. Job Market Signaling. *The Quarterly Journal of Economics*, 87(3), pp. 355-374.

Thavaneswaran, A., 2008. *Propensity Score Matching in Observational Studies,* Manitoba, Canada: Manitoba Centre for Health Policy.

Weiss, Y., 2015. Gary Becker on Human Capital. *Journal of demographic economics*, 81(1), pp. 27-31.

Appendix

Appendix A: Information About the STATA Commands

https://www.stata.com/manuals/teteffectspsmatch.pdf

https://www.stata.com/manuals/teteffectsnnmatch.pdf#teteffectsnnmatchMethodsandformulas