How substitutable are high-skilled workers? The case of expansion of tertiary education in Chile

Roberto Cases*

Comisión Nacional de Evaluación y Productividad

and

Adolfo Fuentes†

University of Edinburgh

September 15, 2024

Abstract

This paper explores the implications of tertiary education expansion in Chile from 2010 to 2019, mainly focusing on how large firms substitute workers with varying qualifications and experience. Despite a significant increase in the share of tertiary-educated workers, reaching 45 percent, there is no substantial decline in the wage premium associated with college-educated workers. Regarding occupations, we found a notable mismatch between educational attainment and job requirements, where most workers with higher vocational education find themselves overqualified, leading to a potential displacement of those workers by their college-educated counterparts. Then, we introduce a novel model estimated through administrative data, and we found close-to-perfect substitutability between workers with higher vocational and college education. Lastly, examining study plans at the technical, professional, and college levels reveals a relevant overlap, rationalizing the substitution between different educational levels. Finally, we emphasize the need to differentiate programs at each educational level to generate a more effective integration in the labor markets.

JEL Classifications: E24, J20, J31.

*Junior Researcher. Email: rcases@cnep.cl.

†PhD Student. Email: a.i.fuentes@sms.ed.ac.uk.

First version: October 2023. We thank the comments by Jesper Bagger, Diego Battistón, Axel Gottfries, Jan Grobovšek, Rafael Lopes de Melo, and David Mesa. We also appreciate comments from participants at the following events: Ph.D. Seminar at the University of Edinburgh, the Seminar at Centro de Estudios Públicos, and the conferences held by the Chilean Public Policy Society (SChPP), the Scottish Economic Society (SES), the Ruhr Graduate School of Economics (RGS), the Swiss Society of Economics and Statistics (SSES), and the Spanish Association of Labor Economics (AEET). Adolfo is thankful for the financial support from the National Agency of Research and Development (ANID) of the Chilean Government for his doctoral studies. All errors are our own.

1 Introduction

The expansion of tertiary education around the world is a well-documented fact, and the case of Latin America is particularly interesting given the speed at which the process has occurred, where the matriculation rate increase from 17 percent in 1991 to 40 percent in 2010 (Ferreyra et al., 2017). The significant growth in the labor market, coupled with rising real minimum wage rates, has consistently diminished the wage premium linked to education levels below tertiary education. Simultaneously, it has eroded the wage premium associated with work experience. These trends have collectively contributed to a widespread reduction in income inequality across the region over the past two decades (Fernández & Messina, 2018). The main mechanism used to explain this fact is that there is some degree of complementarity across workers with and without tertiary education, and therefore the increase of education levels is helpful to less skilled workers, since they are needed in the production process and they become more scarce, leading them to improve not only their wages, for also their formality status (Haanwinckel & Soares, 2021).

Regarding the Chilean case, the increase in tertiary education has been significant since 1990. In that year, the share of people with tertiary education in the CASEN[1] survey was 17 percent, while in 2022 it increased to 45 percent. This change could be classified in three stages. In the first stage (1990-2005) the support to students in tertiary education was relatively constant (with 30 percent of them having some kind of benefits) but the number of students more than doubled from 240.000 in 1990 to 637.000 in 2005, mainly lead from the expansion of private universities (Urzúa, 2012).

The second stage (2006-2015) was characterized by a surge in public expenditure on tertiary education, primarily channeled through scholarships targeted at university students. This funding increased substantially, rising from US$87 million in 2005 to US$701 million in 2015. Consequently, the student population also witnessed a significant upswing, reaching 1.15 million during this period (González & Ureta, 2015).

Finally, a third stage started in 2016 with the introduction of a law that guaranteed free education to students who came from the 90 percent of poorer households in the country. The law approved in 2016 only considered college students, but in 2017 it included vocational students also. At this point it is important to differentiate college and higher vocational programs. College programs, exclusively administered by universities, span a minimum of 5 years. In contrast, higher vocational degrees fall into two categories: technical degrees, typically lasting up to 3 years (though commonly 2), are offered by Technical Education Centers (CFT in Spanish), while professional

[1]The CASEN survey is the main household survey in Chile, conducting waves at regular intervals of 2 or 3 years since 1990.

degrees, with a 4-year duration, are conferred by Professional Institutes (IP in Spanish). The fiscal expenditure associated with the law passed from US$796 million in 2016 to an expected of US$1.4 billion in 2020 (Holz, 2019), however the student population did not increased as before, only achieving 1.2 million in 2021 (Mineduc, 2023).

As anticipated, the massification policy had a notable impact on the wage distribution within the country, consequently altering the observed wage premiums. In this context, a study by Beyer et al. (1999) examined the effects of the concurrent expansion of tertiary education and trade liberalization during the initial stage on wage premiums and income inequality. The authors found that while trade liberalization increased the wage premium for tertiary education, thanks to the reduced price of tradable goods produced by unskilled workers, the growth in tertiary education reduced the wage premium for this group. Their research uncovered an ambiguous effect, and the authors suggested that trade influences wage premiums not only by altering the relative prices of goods but also by prompting a restructuring of the production landscape within the economy.

Then, following the canonical model by Murphy & Welch (1990), Katz & Murphy (1992), and Katz & Author (1999), Gallego (2012) estimated the wage premiums in Chile for the period 1960-2000, showing an strong increase in it specially in the period between 1973 and 1990. Gallego (2012), in contrast with Beyer et al. (1999), proposes that the increase in the wage premium is due the imports of new technology and not because of the changes in relative prices between the tradable and non-tradable sectors. The author sustains this view showing that the wage premiums behave similarly in both sectors.

It is also interesting to note that the experimented expansion in tertiary education can lead to a premature de-industrialization process (Rodrik, 2016; McMillan et al., 2017) in which countries skip the industrialization phase and go directly to increase the production of services. Whether or not this process is detrimental to the economy will largely depend on whether the expansion of tertiary education induces firms to adapt their technology levels and become more productive. For example, Vu (2023) shows that this process in Vietnam was welfare improving because service firms increased their productivity.

While the empirical methodology employed has demonstrated its robustness and applicability across various contexts (Autor et al., 1998; Card & Lemieux, 2001; Gindling & Robbins, 2001; Battistón et al., 2014), certain initial practical challenges persist in many research papers. One such challenge pertains to the scarcity of extensive datasets, a particularly common issue in studies focused on developing economies. For instance, in Fernández & Messina (2018), the most recent paper estimating substitution elasticities for Latin American countries, the authors derived their results by amalgamating data from household surveys conducted in Argentina, Brazil, and Chile.

This data amalgamation became necessary due to the infrequent occurrence of household surveys, requiring the pooling of data from different countries to obtain a sufficient number of observations for the estimations. This approach introduces several complexities that can potentially distort the estimation effects, including addressing response bias in income surveys, standardizing education levels, and selecting an appropriate exchange rate to unify all income figures into a common currency.

A second challenge, prevalent in most studies, revolves around the reliance on survey data for income measures. Drawing income data from surveys gives rise to two primary concerns. Firstly, there is the potential issue of declaration bias, as highlighted by Moore et al. (2000), which can introduce a level of distortion into the resulting estimates. This concern has added significance, particularly in light of empirical evidence indicating that individuals with higher incomes tend to underreport their earnings (Hlasny, 2020). In this context, the issue of bias assumes heightened importance, directly impacting the precise estimation of wage premiums.

Finally, a second challenge inherent in studies reliant on survey data pertains to the nature of income measures. Even when these measures are adjusted to account for biases, they typically represent only workers' wages, excluding gross income. In practice, wages are subject to taxes, social security contributions, and other deductions, resulting in a discrepancy between a worker's income and the cost incurred by the employing firm. Consequently, when computing the elasticities of substitution among workers, it is more relevant to consider the firm's costs rather than the worker's wage alone. Concentrating solely on perceived wages can introduce biases into estimations, especially given that taxes and social security contributions tend to rise with higher wages, thus making it relatively more expensive for firms, for example, to hire skilled workers.[2]

Considering the above, this paper contributes to the literature studying the substitution across worker types in a developing country in three ways. First, we use the unemployment insurance (SC, in Spanish) administrative records in our calculations, and therefore, to the best of our knowledge, this is the first paper studying the elasticities of substitution in a developing country that uses administrative records. Using the SC database is useful for this case since we can follow individuals over time and, after some minor corrections that we will present later, have a good measure of the firm cost of having each worker. Second, we show that the abrupt change in the educational policy generated different growth rates in the tertiary educated population, leading to a period in which university education grows faster than technical vocational education and another period in which the opposite happens. Interestingly, in both periods, we see a decrease in the wage premium associated with technical vocational workers concerning workers with a college education. Also,

[2]The above is specially sensitive in the case of studies mixing data from different countries that will involve different tax and social security systems.

we observe a relevant overlap in the occupations in which college and technical vocational workers are employed. Finally, we develop a model to estimate the elasticities of substitution across worker types, especially differentiating technical-vocational workers from those with a college education. The results show that both types of tertiary educated workers are very close to being perfect substitutes, consistent with the evidence found in the wage premiums and the occupations taken by both types of workers.

The paper proceeds as follows: In the second section, we present descriptive statistics detailing the transformations in tertiary education within Chile. Moving on to the third section, we calculate wage premiums, utilizing data from the household survey (CASEN) spanning 1990-2022 and then examines the occupational overlap between college and technical vocational workers for the period 2010-2022. Subsequently, in the fifth section, we describe the administrative data of the unemployment insurance (SC). In the sixth section, we formulate a model and estimate the elasticity of substitution for various worker types during the period 2010-2019. The seventh section discusses the paper's primary findings, focusing on their implications for public policy. Finally, the eighth section presents the concluding remarks.

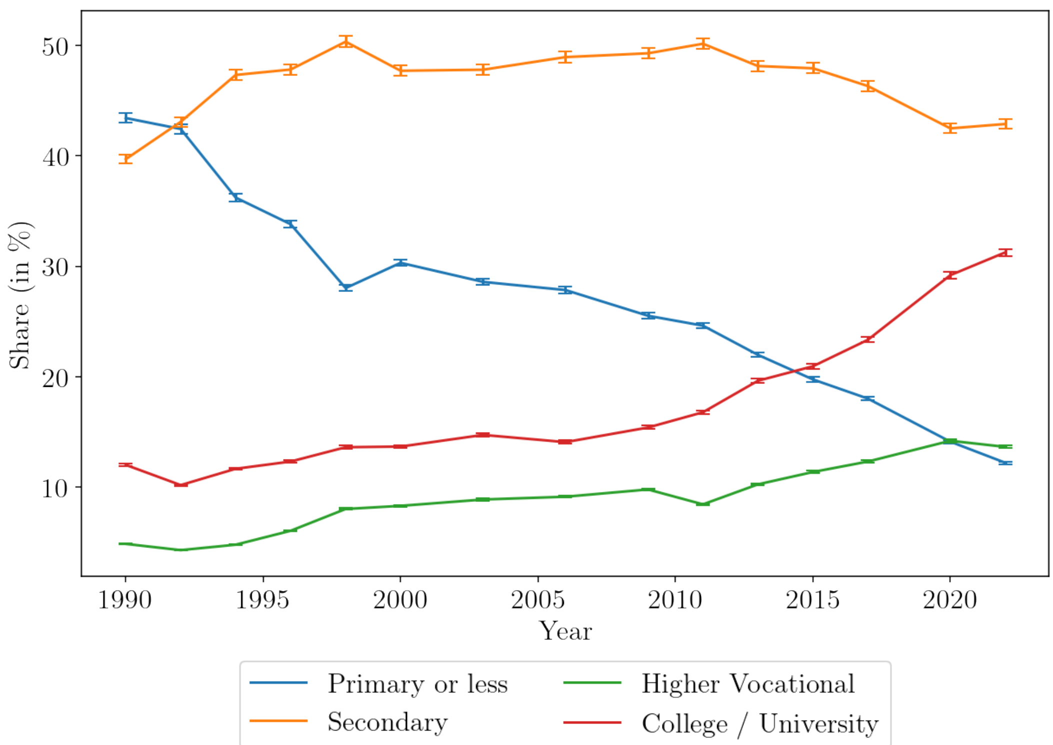
2 Expansion of tertiary education

To gain insight into the consequences of the policies mentioned earlier, Figure 1 illustrates the shifting skill composition among workers aged 25 to 59 during the period spanning from 1990 to 2022. The pronounced decline in individuals with primary education or lower stands out prominently. In 1990, this educational level was the most prevalent, encompassing 43 percent of the population. However, by 2022, it had become the least prevalent, accounting for just 12 percent of the population.

The data reveals a discernible upward trajectory in higher vocational and college/university education levels throughout the study period. In 1990, the proportion of workers with tertiary education stood at approximately 17 percent, soaring to 45 percent among the population aged 25-59 years old by 2022. Despite this overall increase, the dynamic interplay between these two educational tiers is far from constant. In 1990, the ratio between higher vocational and college/university education levels was 0.41, steadily climbing to 0.65 by 2006. However, a subsequent reversal ensued, leading to a regression to the 1990 levels by 2022, marked by a value of 0.43. This nuanced fluctuation underscores the evolving landscape of educational attainment trends over the specified timeframe.

Regarding wages, we computed the logarithm of real hourly wages for individuals belonging to four distinct educational groups. The series, that can be seen in Figure 2, reveals a noteworthy trend: the wage gap between individuals with higher vocational education and those with secondary or primary education has steadily narrowed. However, it is important to note that individuals with a college education continue to experience a significant wage disparity compared to their counterparts in the workforce with different educational backgrounds.

Figure 1: Skill composition for population between 25-59 years old



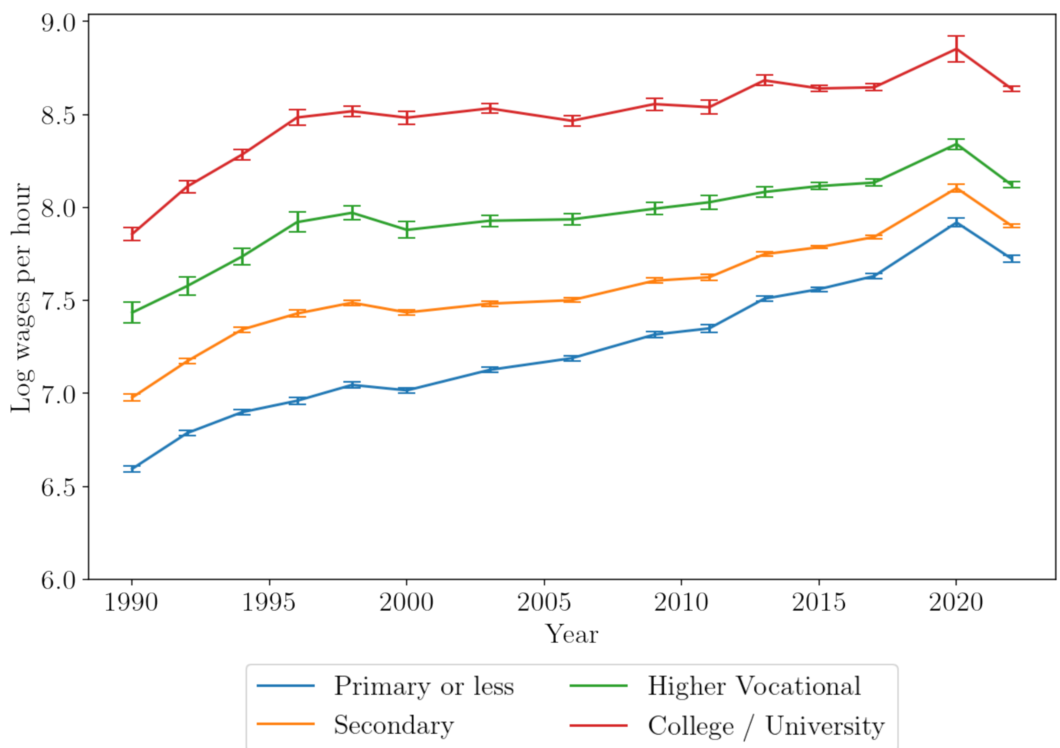
Source: Data from CASEN surveys correspond to waves 1990 to 2022. Each level considers people with complete and incomplete studies. The error bars represent the 95 per cent confidence intervals.

In relative terms, this signifies a consistent increase in hourly wages for individuals with a college education compared to those with higher vocational education. In 1990, individuals with a college degree earned an average of 42 percent more per hour than their counterparts with higher vocational education. This wage differential increased during the initial phase of tertiary education expansion, reaching 60 percent per hour in 2003. Subsequently, during the second phase of massification, this gap gradually narrowed to 52 percent, and in recent years, it has consistently hovered around this level.

The above fact is counter-intuitive since if something happens, the wage premium of college educated workers should decrease since there are substantially more of them. However, it is important to note that this shift cannot be attributed to working-hour alterations, as we control this variable. Instead, it is plausible that this phenomenon may be influenced by changes in job sectors, varying levels of work experience, or geographic regions where individuals are employed.

So, to understand that, we will use linear models to compute the wage premiums.

Figure 2: Logarithm of real wage per hour by education level, population between 25-59 years old



Source: Data from CASEN surveys correspond to waves 1990 to 2022. Each level considers people with complete and incomplete studies. The error bars represent the 95 per cent confidence intervals

3 Wage premiums and occupations

To compute the wage premiums, we will follow the canonical model of human capital accumulation by Mincer (1974), and estimate the following equation:

$$\ln wh = \beta_0 + \sum_{i=1}^{10} \text{educ}_i \beta_{ei} + \beta_2 \text{Female} + \beta_3 \text{exp} + \beta_4 \text{exp}^2 + \beta_5 \text{Region} + \beta_6 \text{Sector} \tag{1}$$

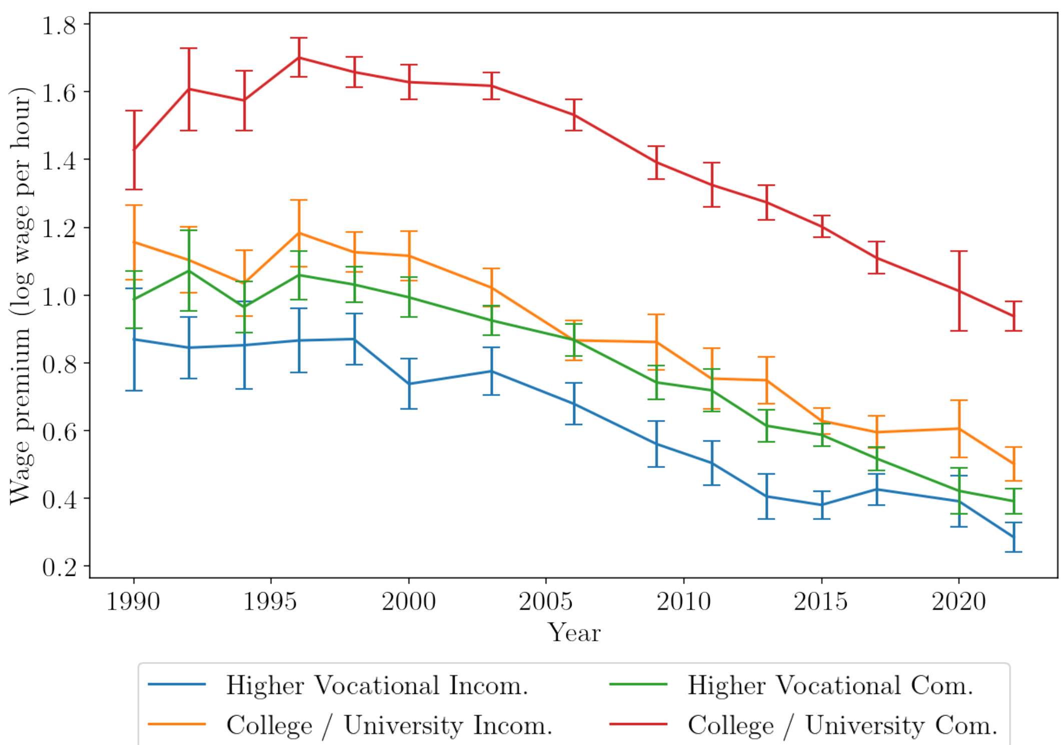
Where $\ln wh$ is the logarithm of the wage per hour, $\sum_{i=1}^{10} \text{educ}_i$ is a set of dummies, two for each one of the five educational levels,[3] "Female" is dummy that takes the value 1 if the worker is a woman, "exp" represents the potential labor experience and "exp^2" its square, "Region" represents the region in which the worker lives, and "Sector" the sector of the firm in which the individual

[3]In this case the baseline educational level is no education. There are two dummies for level because we are considering incomplete and complete levels.

works.

Figure 3 illustrates the coefficients corresponding to the higher vocational and college workers, differentiating between complete and incomplete levels. All educational wage premiums are declining over time, yet the premium associated with a complete college education remains notably distinct from the other premiums. Intriguingly, even after accounting for various controls, it is surprising that the wage premium for complete higher vocational education is not statistically distinguishable from the wage premium for incomplete college education.

Figure 3: Wage premium for educational level, population between 25-59 years old

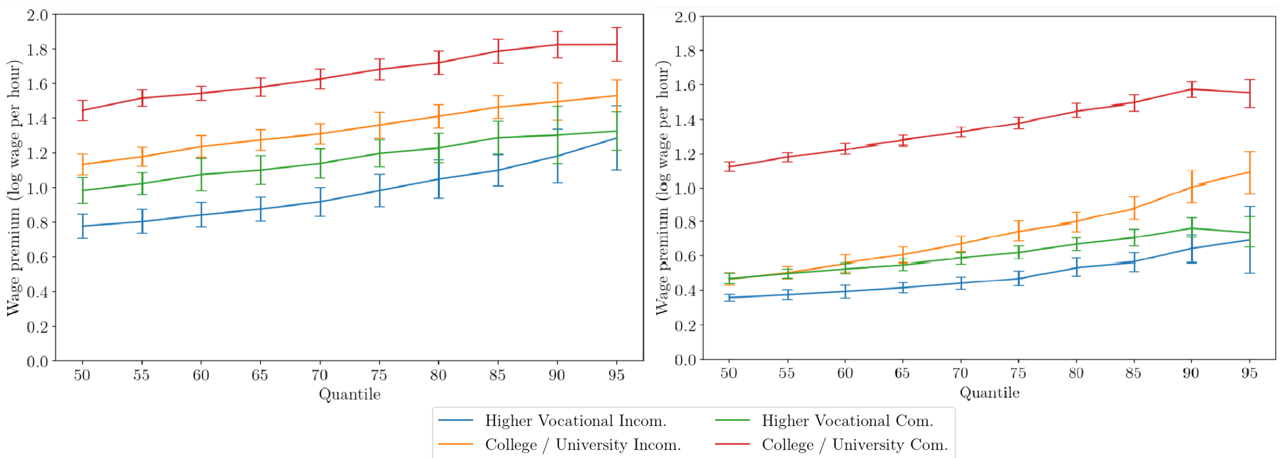


Source: Data from CASEN surveys correspond to waves 1990 to 2022. Each level considers people with complete and incomplete studies. The error bars represent the 95 per cent confidence intervals

Examining the variations in the wage premium across income quantiles and over time provides additional insights. To explore this, we conduct quantile regressions for the 50th to 95th quantiles, applying the same model described in Equation 1 to the 1990 and 2017 CASEN waves, respectively.[4] Figure 4 illustrates the coefficients of both estimations. Notably, a noteworthy shift emerged over the years; while in 1990, the wage premium to education displayed an increasing pattern across income quantiles, by 2017, this trend is evident primarily among workers with a complete college education. This observation hints at a potential divergence in the human capital accumulation process between these two worker categories.

[4] We opt for 2017 data instead of 2022 to mitigate the impact of the pandemic.

Figure 4: Wage premium for educational level and quantile 1990 (left) and 2017 (right), population between 25-59 years old



Source: Data from CASEN surveys correspond to waves 1990 to 2022.
The error bars represent the 95 per cent confidence intervals

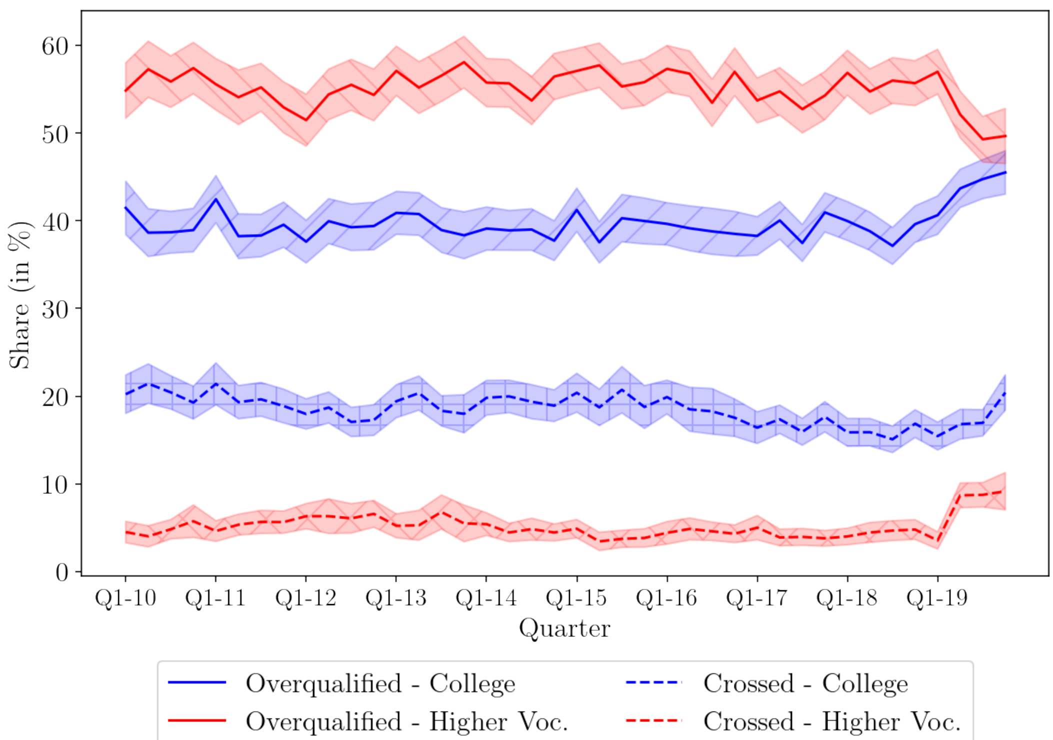
The results of the regressions imply that workers with higher vocational education are unlikely to attain comparable hourly wages to their counterparts with complete college education, even when employed within the same industry or possessing equivalent labor experience. Consequently, it is intriguing to examine the occupations held by these workers, as it is conceivable that having a high vocational education may not necessarily grant access to high-quality positions that allow to accumulate human capital over time. Then, this occupation divergence could account for the consistent disparities in wage premiums.

To conduct this analysis, we will employ the CIUO-88 classification available in National Employment Survey (ENE, in Spanish), which categorizes jobs on a scale ranging from (1) "Non-qualified workers" to (9) "Members of the government administration, legislative bodies, and senior personnel in the public administration and companies." This classification proves particularly valuable as its second tier corresponds to (2) "Professionals with a college education, scientists, and intellectuals", while the third tier is (3) "Professional technicians at the middle level", thus effectively distinguishing jobs suitable for individuals with a college education and those with higher vocational education, respectively.

Considering those above, we have established the following classifications: An individual with a college education is deemed overqualified for a position if the role falls within tier (3) "Technicians and middle level professionals" or below. Similarly, an individual with higher vocational education is considered overqualified if the job is categorized as tier (4) "Administrative support staff" or down. Furthermore, a worker possessing a college education is classified as being in a crossed occupation when the job is labeled as tier (3). In contrast, an individual with higher vocational education is classified as being in a crossed occupation when the job is designated as tier (2).

Figure 5 illustrates the distribution of overqualified and crossed occupations based on educational levels from 2010 to 2019 for the workers with higher vocational or college education, incomplete or complete.[5] Notably, a significant proportion of workers with higher vocational education find themselves in overqualified positions. Similarly, approximately 40 percent of college-educated workers are employed in roles for which they are overqualified.

Figure 5: Share of overqualified and crossed occupations
by educational level, population between 25-59 years old



Source: Data from ENE surveys correspond to the months of February, May, August and November 2010 to 2019. Each level only considers people with complete studies. The error bars represent the 95 per cent confidence intervals

Given this, it is interesting to see if there are crossed occupations between the two types of workers with tertiary education. The data reveals that nearly 20 percent of individuals with a college education, constituting about 50 percent of those overqualified for their roles, are engaged in occupations typically suited for workers with higher vocational education. Conversely, the occurrence is less frequent, with only 5 percent of workers with higher vocational education performing roles traditionally designated for college-educated workers.

[5] We computed the same shares only considering workers with the complete levels and the shares do not differ in a relevant manner

This finding is challenging to reconcile under the complementarity assumption between workers with college and higher vocational education. Instead, it suggests a contrary scenario where these two groups exhibit a certain degree of substitution, with workers possessing higher vocational education being displaced by those with college education. Consequently, the former are compelled to take on roles that necessitate being overqualified. With this hypothesis in mind, in the next section we build a model to estimate the elasticity of substitution across different worker types.

4 Data

We will employ two datasets to address our research question. The first dataset is compiled from three distinct Unemployment Insurance Affiliates database (SC) samples, representing 3, 5, and 12% of the total affiliates, respectively. Each sample is drawn from administrative records to ensure representation of the overall population while maintaining independence from the other two. This approach allows the combination of the datasets, resulting in a combined representation of 20% of the total affiliates.

The database comprises different components. One component pertains to affiliates, providing individual details such as an identification number, gender, birth date, educational level, years of schooling, marital status, county of residence, pension fund administrator (AFP in Spanish), and nationality. Another component pertains to wages, including an anonymised individual's ID number, the gross wage received in the month, payment date, contract type, economic sector of the employing firm, county of the firm, number of workers in the firm for the month, average and standard deviation of wages at the firm level for the month, as well as indicators identifying workers without wages, those receiving the minimum wage, and those whose gross wage exceeds the maximum wage considered for social security deductions.[6]

Having the gross wage is a relevant data improvement from previous papers that studied the elasticities of substitution in developing countries. This measure of labor income does not account for legal deductions such as taxes, social security contributions, and health contributions, among others, and therefore is a reasonable measure of the firm cost of hiring a worker. The above is a significant distinction from other databases that measure labor income, such as the Socioeconomic Characterization National Survey (CASEN) and the Supplementary Income Survey (ESI).[7].

Nevertheless, there are additional costs not accounted for in the SC database as there are not

[6]If a worker's wage is higher than the maximum wage considered for social security deductions, the gross wage presented in the SC data will be the former. The above implies a bias in our calculations, however the amount of workers that receive a wage higher than the maximum considered for payment of social security contributions is around 2-3%.

[7]For a more detailed exploration of the differences between these databases, refer to Abud et al. (2022)

integrated into the gross wage. This expense is an essential obligation for every employer hiring a worker—namely, workplace accident insurance, referred to as "*Seguro de accidentes laborales*" in Spanish. The responsibility for this insurance lies with the employer, and the payment amount is contingent on the firm's sector, with higher-risk sectors incurring elevated premiums. A fundamental premium of 0.93% of the gross wage applies universally across industries, complemented by a risk premium ranging from 0 to 3.4% depending on the specific sector. Given our access to information about the firm's sector, we can factor these insurance payments into calculating gross wages.

Furthermore, workers in the formal sector may be engaged through fixed-term or open-ended contracts. In each scenario, the firm must contribute a distinct percentage of the gross wage to the unemployment insurance[8]. The firm's contribution is 2.4% of the gross wage for an employee with an open-ended contract, whereas for a fixed-term contract, the premium is 3%. Since we observe the type of contracts workers hold, we can incorporate the employer's share of this insurance into calculating the gross wage.

The databases of sociodemographic characteristics and labor-related variables are merged to create a monthly panel of labor variables for each affiliate, identifying sociodemographic variables, of which we will use education and potential work experience to classify workers into different categories. Based on educational groups, workers are categorized into four groups depending on whether they have an educational level i) lower than or equal to complete basic education, ii) equal to complete secondary education, iii) equal to complete higher technical-professional education, or iv) equal to complete university education. On the other hand, four categories are established for experience levels: i) between 0 and 9 years of experience, ii) between 10 and 19 years, iii) between 20 and 29 years, and iv) 30 or more years of experience. Sixteen groups of workers are constructed by combining the categories of educational levels and experience.

It is important to note that, in a given month, some affiliates may appear in the database more than once. This could occur if an affiliate holds more than one job. In light of this, the decision is made to retain the gross wage from the main occupation for each affiliate, ensuring that all affiliates have the same relative weight in calculating aggregated variables for each education-experience group. In this context, the main occupation considered is the one that reports the highest gross wage or represents the longest employment relationship.

Although we have information since 2002 we only considered data from 2010 onwards. This is because the participation in unemployment insurance is mandatory for new hires after October 2002 and voluntary for workers with contracts before that date. This setup implies a progressive

[8]Note that this component differs from the portion charged to the worker, which is indeed factored into the gross wage

increase in insurance affiliates as new contracts emerge in the labor market, both from unemployed individuals who find a formal job as from individuals that moved from one formal job to another. In this context, Rojas & Sanchez (2014) highlight that by the end of 2008, approximately 80% of the target group (dependent workers in the private sector) were actively contributing to unemployment insurance. The authors attribute this trend to the notable workforce turnover in Chile, leading to a substantial influx of new contracts each month. Given the the above we opted to utilize data from 2010 onward, since we expect that by that year the share of formal workers covered by the unemployment insurance is close to the total of formal workers.

Regarding the type of workers covered by the unemployment insurance, the dataset encompasses dependent workers aged 18 and above whose jobs are ruled by the Labor Code. As such, the analyses conducted throughout this study exclude workers under apprenticeship contracts, those under 18 years old, domestic workers, retirees, independent or self-employed individuals, informal sector workers, and public sector employees.[9] Given the above, the sample allows to study the changes in the elasticities of substitution for an specific group of workers in the Chilean labor market: those formally employed.

Then, a natural question that arises is about the levels of informality in Chile. According to INE (2023) around 30% of workers are informally employed, and the rate is stable since 2017.[10]. Given the levels of informality, we will need to take the informal sector into consideration, particularly in the measurements of the labor supply that we will use in the model estimation. Since the SC database does not account for informal workers, we will rely on a second database which is the employment survey (ENE). This survey is a nation-wide representative survey that measures the labor force, unemployment, and occupied workers among other statistics at different levels. The survey works as a rotating panel built on six groups of households. In each quarter the group of households that has stayed in the panel for the longest time abandons it, and it is replaced by a new group of households.

5 Model

5.1 Model structure

Our empirical estimates of substitution elasticities rely on a labor market supply and demand model, assuming perfect competition. Specifically, we utilize an extended version of the model

[9] For public sector employees the access to the unemployment insurance will depend if they work for the central government administration or not. Those working in ministries and associated departments, and local government administrations are unlikely to have access to the UI. However those working on public firms, health and education sectors are likely to have access to UI.

[10] For years before 2017 we do not have data, since informality was not measured.

proposed by Fernández & Messina (2018).[11] This model posits that the economy's aggregate production function can be characterized by a nested Constant Elasticity of Substitution (CES) function at multiple levels. At the topmost level, the output is influenced by a variable input (labor), which is categorized into two types: high-skill workers (with complete tertiary education) and low-skill workers (with secondary education or less).

$$Y_t = \lambda_t \left(L_{B_t}^{\rho} + \alpha_t L_{A_t}^{\rho} \right)^{1/\rho}$$

Where Y_t corresponds to the total output at time t, L_{B_t} and L_{A_t} denote the total supply of low-skill and high-skill workers at time t, respectively. Additionally, λ_t represents a scale parameter, which is allowed to vary over time to capture general technological changes (independent of skill levels). The parameter α_t captures relative productivity differences between skill groups, and ρ is a function of the elasticity of substitution between high-skill and low-skill workers, denoted by the parameter σ_ρ, where $\sigma_\rho = \frac{1}{1-\rho}$.

At the next level, the total supply of low-skill and high-skill workers is subdivided into different subgroups based on their educational levels. Firstly, the supply of low-skill workers is divided into two subgroups. The first corresponds to workers with an educational level equal to or lower than complete basic education L_{P_t}. Simultaneously, the second group comprises workers who have attained complete secondary education L_{H_t}. As a result of these two subgroups, the total supply of low-skill workers is determined by the following function:

$$L_{B_t} = \left(L_{P_t}^{\delta} + \beta_t L_{H_t}^{\delta} \right)^{1/\delta}$$

The parameter β_t captures differences in relative productivity between workers with educational levels H and P. Analogously to the interpretation of ρ, the parameter δ is a function of the elasticity of substitution between the two subgroups of low-skill workers (σ_δ), where $\sigma_\delta = \frac{1}{1-\delta}$.

Secondly, the supply of high-skill workers is also divided into two subgroups. The first subgroup consists of workers who attained a higher vocational education L_{T_t}. Conversely, the second subgroup comprises individuals who have completed college education L_{U_t}. In aggregate terms, the total supply of high-skill workers is derived from the following expression:

$$L_{A_t} = \left(L_{T_t}^{\gamma} + \tau_t L_{U_t}^{\gamma} \right)^{1/\gamma}$$

Continuing in alignment with the approach taken for low-skill workers, the parameter τ_t is instrumental in capturing disparities in relative productivity between the two educational level subgroups

[11]In turn, the authors base their model on the theoretical frameworks of Katz & Murphy (1992), Murphy & Welch (1992) and Katz & Author (1999)).

(T and U). Simultaneously, the parameter γ captures the elasticity of substitution between the two subgroups of high-skill workers σ_γ, where $\sigma_\gamma = \frac{1}{1-\gamma}$.

At a final level, each labor supply within the educational group K (where $K = \{P,\ H,\ T,\ U\}$) is comprised of four experience groups i (where $i = \{1, 2, 3, 4\}$). The total supply of workers in the educational group K is determined by the following function:

$$L_{K_t} = \left(\sum_{i=1}^{4} \phi_{K_{it}} L_{K_{it}}^{\theta_K} \right)^{1/\theta_K}$$

Where L_{K_t} represents the labor supply of the educational group $K = \{P,\ H,\ T,\ U\}$ with the experience level $i = \{1, 2, 3, 4\}$ at time t. The parameter θ_K denotes a function of the elasticity of substitution between the experience subgroups $\sigma_{\theta_K} = \frac{1}{1-\theta_K}$ within the educational group K. Additionally, $\phi_{K_{it}}$ is a parameter associated with the relative productivity of each experience group within the educational group K. To reduce the number of parameters to estimate, it is assumed that for the low-skill worker group, $\phi_{P_{it}} = \phi_{H_{it}} = \phi_{B_{it}}$ and $\theta_P = \theta_H = \theta_B$. Similarly, for the high-skill worker group, it is assumed that $\phi_{T_{it}} = \phi_{U_{it}} = \phi_{A_{it}}$ and $\theta_T = \theta_U = \theta_A$.

The next step of the model consists of determining the equilibrium wages in this economy. Like Fernández & Messina (2018), the economy is assumed to operate along a competitive equilibrium demand curve. This assumption implies that the wages for each of the sixteen groups of workers are determined by their marginal labor productivity. Using the chain rule, the sixteen equilibrium wages are determined by the following expression:

$$W_{EK_{it}^E} = \frac{\partial Y_t}{\partial L_{E_t}} \frac{\partial L_{E_t}}{\partial L_{K_t^E}} \frac{\partial L_{K_t^E}}{\partial L_{K_{ti}^E}} \tag{2}$$

Where E denotes the skill groups of workers with $E = \{B, A\}$, K^E corresponds to the educational subgroup within skill group E. Note that $K^B = \{P, H\}$ and $K^A = \{T, U\}$. In logarithmic terms, equation 2 implies eight equilibrium wage conditions for each skill-level group of workers. The following functional form determines the eight conditions for the wages of low-skilled workers:

$$w_{BK_{it}^B} = \frac{1-\sigma_\rho}{\sigma_\rho} \log(\lambda_t) + \log(\psi_{K_{it}^B}) + \frac{1}{\sigma_\rho}(y_t - l_{B_t}) + \frac{1}{\sigma_\delta}(l_{B_t} - l_{K_t^B}) + \frac{1}{\sigma_{\theta_B}}(l_{K_t^B}) - l_{K_{it}^B}$$

On the other hand, the eight conditions for the wages of high-skilled workers are given by the following expression:

$$w_{AK_{it}^A} = \frac{1-\sigma_\rho}{\sigma_\rho} \log(\lambda_t) + \log(\psi_{K_{it}^A}) + \frac{1}{\sigma_\rho}(y_t - l_{A_t}) + \frac{1}{\sigma_\gamma}(l_{A_t} - l_{K_t^A}) + \frac{1}{\sigma_{\theta_A}}(l_{K_t^A}) - l_{K_{it}^B}$$

Where $w_{BK_{it}^{B}}$ is the logarithm of the wage for skill group B, with educational level K^B and experience level i at time t. For the high-skill group, $w_{AK_{it}^{A}}$ is defined similarly. Note that within the low-skill worker group, $\psi_{K_{it}^{B}}$ takes the following values: $\psi_{P_{1t}} = \psi_{B_{1t}}$, $\psi_{P_{2t}} = \psi_{B_{2t}}$, $\psi_{P_{3t}} = \psi_{B_{3t}}$, $\psi_{P_{4t}} = \psi_{B_{4t}}$, $\psi_{H_{1t}} = \beta_t$, $\psi_{H_{2t}} = \beta_t\phi_{B_{2t}}$, $\psi_{H_{3t}} = \beta_t\phi_{B_{3t}}$, and $\psi_{H_{4t}} = \beta_t\phi_{B_{4t}}$. In the same vein, for high-skill workers, $w_{AK_{it}^{A}}$ takes the following values: $\psi_{T_{1t}} = \alpha_t\phi_{A_{1t}}$, $\phi_{T_{2t}} = \alpha_t\phi_{A_{2t}}$, $\psi_{T_{3t}} = \alpha_t\phi_{A_{3t}}$, $\psi_{T_{4t}} = \alpha_t\phi_{A_{4t}}$, $\psi_{U_{1t}} = \alpha_t\tau_t\phi_{A_{1t}}$, $\psi_{U_{2t}} = \alpha_t\tau_t\phi_{A_{2t}}$, $\psi_{U_{3t}} = \alpha_t\tau_t\phi_{A_{3t}}$, $\psi_{U_{4t}} = \alpha_t\tau_t\phi_{A_{4t}}$. For the purpose of the estimations, the parameters for the lowest experience group are normalized, i.e., $\phi_{B_{1t}} = \phi_{A_{1t}} = 1$.

Our empirical strategy consists of estimating a three-stage model. The first stage is derived using the expressions for equilibrium wages for each education-experience group. In detail, for each educational group, we obtain the wage premium for belonging to a higher experience group. Within the low-skill worker group, $E = B$, twelve expressions relate the evolution of the wage premium per experience for educational groups $K = \{P, H\}$, with the relative supply among different experience groups within the same educational level. These expressions have the following general form:

$$w_{BK_{it}^{B}} - w_{BK_{jt}^{B}} = \phi_{B_{it}} - \phi_{B_{jt}} - \frac{1}{\sigma_{\theta_B}}\left(l_{K_{it}^{B}} - l_{K_{jt}^{B}}\right) \tag{3}$$

Where $K^B = \{P, H\}$; $i, j = \{1, 2, 3, 4\}$ denote the different levels of experience. Since we are interested in wage premiums by experience level $w_{BK_{it}^{B}} - w_{BK_{jt}^{B}}$, equation 3 considers that $i > j$. Similarly, for the high-skill worker group, $E = A$, twelve expressions are obtained, forming the first stage, which have the following general form:

$$w_{AK_{it}^{A}} - w_{AK_{jt}^{A}} = \phi_{A_{it}} - \phi_{A_{jt}} - \frac{1}{\sigma_{\theta_A}}\left(l_{K_{it}^{A}} - l_{K_{jt}^{A}}\right) \tag{4}$$

Where $K = \{T, U\}$; $i, j = \{1, 2, 3, 4\}$ with $i > j$.

Equations 3 and 4 relate the evolution of relative supplies by experience levels within the same educational group $\left(l_{K_{it}^{E}} - l_{K_{jt}^{E}}\right)$ to changes in the wage premiums obtained by higher experience groups $w_{AK_{it}^{E}} - w_{AK_{jt}^{E}}$. This relationship is given by the parameter $1/\sigma_{\theta_E}$, a function of the elasticity of substitution between workers of different experience levels within the same skill group.

In the second stage of our analysis, we explore the relationship between the changing wage premiums based on education levels within each skill group and the dynamics of relative supplies among various educational and experience groups. Specifically, within the low-skill worker group, we examine the logarithmic wage differential between workers with secondary education and those with primary education. The evolution of this wage premium is intricately connected to the shifting relative supplies within the low-skill worker group.

The second stage for low-skilled workers is determined by four expressions of the following general form:

$$w_{BH_{it}} - w_{BP_{it}} = \log(\beta_t) - \frac{1}{\sigma_\delta}(l_{H_t} - l_{P_t}) - \frac{1}{\sigma_{\theta_B}}\left((l_{H_{it}} - l_{H_t}) - (l_{P_{it}} - l_{P_t})\right) \tag{5}$$

Where $i = \{1, 2, 3, 4\}$. Similarly, the second stage for high-skilled workers is determined by four expressions of the following general form:

$$w_{AH_{it}} - w_{AP_{it}} = \log(\tau_t) - \frac{1}{\sigma_\gamma}(l_{U_t} - l_{T_t}) - \frac{1}{\sigma_{\theta_A}}\left((l_{U_{it}} - l_{U_t}) - (l_{T_{it}} - l_{T_t})\right) \tag{6}$$

Equations 5 and 6 establish a connection between the changing relative supplies based on educational levels within the same skill category and the shifts in wage premiums for higher educational groups. This relationship is governed by the parameters $1/\sigma_\delta$ and $1/\sigma_\gamma$, representing the elasticity of substitution between workers of different educational levels within the low-skill and high-skill groups, respectively.

Moreover, the wage premiums based on educational levels in the second stage are intricately connected to the relative supplies by experience levels within each educational group. As predicted in the equations of the first stage, this connection is characterized by the parameters $1/\sigma_{\theta_B}$ and $1/\sigma_{\theta_C}$. Notably, the model affords us the opportunity to validate the robustness of estimations for these latter two parameters between the first and second stages.

Finally, the third stage involves establishing a relationship for the evolution of wage premiums among workers of different skill levels. Considering that K^A denotes the educational levels comprising the high-skill worker group ($K^A = \{T, U\}$) and K^B denotes the educational levels comprising the low-skill worker group ($K^B = \{P, H\}$). Thus, the third stage is determined by 16 expressions with the following general form:

$$\begin{aligned}\log\left(\widehat{\frac{W^A_{AK_{it}}}{W^B_{BK_{it}}}}\right) &= \log(\alpha_t) - \frac{1}{\sigma_\rho}\log\left(\frac{L_{A_t}}{L_{B_t}}\right) - \frac{1}{\sigma_\gamma}\log\left(\frac{L^A_{K_t}}{L_{A_t}}\right) - \frac{1}{\sigma_{\theta_A}}\log\left(\frac{L_{K^A_{it}}}{L^A_t}\right) \\ &\quad - \frac{1}{\sigma_\delta}\log\left(\frac{L_{B_t}}{L_{K^B_t}}\right) - \frac{1}{\sigma_{\theta_B}}\log\left(\frac{L_{K^B_t}}{L_{K^B_{it}}}\right)\end{aligned} \tag{7}$$

Where $\log\left(\widehat{\frac{W^A_{AK_{it}}}{W^B_{BK_{it}}}}\right)$ represents the wage premium for high-skilled workers belonging to the educational group K^A with experience level i, compared to low-skilled workers belonging to group K^B with the same experience level. Note that $\log\left(\widehat{\frac{W_{AU_{it}}}{W_{BP_{it}}}}\right) = \log\left(\frac{W_{AU_{it}}}{W_{BP_{it}}}\right) - \log\left(\frac{\tau_t\phi_{A_{it}}}{\phi_{B_{it}}}\right)$, $\log\left(\widehat{\frac{W_{AU_{it}}}{W_{BH_{it}}}}\right) =$

$\log\left(\frac{W_{AU_{it}}}{W_{BH_{it}}}\right) - \log\left(\frac{\tau_t \phi_{A_{it}}}{\beta_t \phi_{B_{it}}}\right)$, $\log\left(\widehat{\frac{W_{AT_{it}}}{W_{BP_{it}}}}\right) = \log\left(\frac{W_{AT_{it}}}{W_{BP_{it}}}\right) - \log\left(\frac{\phi_{A_{it}}}{\phi_{B_{it}}}\right)$ and $\log\left(\widehat{\frac{W_{AT_{it}}}{W_{BH_{it}}}}\right) = \log\left(\frac{W_{AT_{it}}}{W_{BH_{it}}}\right) - \log\left(\frac{\phi_{A_{it}}}{\beta_t \phi_{B_{it}}}\right)$. These can be interpreted as the wage premium by skill level for experience group i, after subtracting the trends of various productivity parameters.

The third stage establishes a link between the dynamic evolution of wage premiums based on skill levels and the fluctuations in relative supplies between high-skill and low-skill workers. This relationship is quantified by the parameter $1/\sigma_\rho$, representing the elasticity of substitution between workers of differing skill levels. Importantly, Equation 7 serves as a robustness check for the estimated parameters $1/\sigma_\gamma, 1/\sigma_{\theta_C}, 1/\sigma_\delta$ and $1/\sigma_{\theta_B}$ in both the first and second stages for each skill group.

5.2 Production function and technological change

At this point, it is relevant to discuss two main points in the literature about elasticities of substitution. The first topic is related to the returns to human capital and the role of the production function in it. Assuming a production function can be seen as a complex assumption, and other ways of estimating the returns to human capital could be more suitable to this, such as the macro-Mincer approach (Klenow & Rodriguez-Clare, 1997; Hall & Jones, 1999; Caselli, 2005). In this approach, the aggregate human capital in a zone is given by aggregating different human capital levels and the number (or proportion) of workers who have them. Once we have the aggregate level of human capital in each zone, we can run Mincer regressions to obtain the wages associated with them. However, as pointed out in Jones (2019), this approach assumes implicitly that the different human capital levels are perfect substitutes. Therefore, this approach is not suitable for estimating elasticities of substitution. In addition to this, Jones (2014) shows that assuming a general enough production function (like the nested CES function used here) will deliver a lower bound of the human capital differences between zones, and that the only components needed to estimate the elasticities are those associated with human capital, and not physical capital or technology.

Another relevant point is the one related to technological change. It could be the case that improving the technology levels would make highly skilled workers even more productive, and then the substitution across different worker types would be different. This critique is precisely made by Caselli & Ciccone (2019) to the argument presented by Jones (2014). Regarding this, we offer three arguments to sustain our model. One conceptual defense to this critique is made by Jones (2019), who mentions that technological change not only implies changes in the component associated with the TFP but also other elements like division of labor and that human capital by itself could be improving. Then, mainly in contexts where the productivity is understood as what we cannot explain, is important to keep in mind that a general enough way of measuring the impact of human capital in the production function is also accounting by the changes in technology, which is precisely the case of the used production function.

Second, it is also possible to argue that the structure of the model allows to capture technological change (understood as a general boost of the production) and improvements in human capital (understood as changes in productivity achieved by specific groups) by the fact that, in the model presented, we allow both "technology" λ_t and the quality of human capital (captured by α_t, β_t and τ_t) to change over time.

Finally, from a practical perspective, it is also relevant to understand if technological change occurred in the country between 2010 and 2019. When observing the series, available in Figure 6 in the Appendix, we notice that in the ten years, only three years present a positive change in the productivity levels. Also, the productivity growth rate is negative in the aggregate, accounting for -0.4 percent per year. Therefore, we can think that there was little or no technological change to account in Chile during this period.

5.3 Model estimation

To estimate the model, we will leverage information from the aforementioned datasets. Initially, we will partition the data into three zones based on the geographical positioning of regions within the country. The first zone encompasses regions in the north of the country, namely Arica y Parinacota, Tarapacá, Antofagasta, and Atacama. These regions are distinguished by a pronounced presence of the primary sector, particularly the mining industry. Moving to the second zone, defined by regions in the center of the country—Valparaíso, Metropolitana, and O'Higgins—we observe a notable concentration of the national population and a prevalence of service-intensive activities. Lastly, the third zone comprises regions in the south of the country: Maule, Ñuble, Biobío, Araucania, Los Rios, Los Lagos, Aysen and Magallanes. These regions are characterized by sparse populations and a reliance on the primary sector, particularly in fishing and the forestry industry.

For wage computations, within each zone, we utilized the SC dataset to determine the average wage for males across the sixteen experience-education groups. We will only consider those workers earning at least the minimum wage and also those working on firms with at least 200 workers. The above assumptions are taken because we want to consider those workers who are actually affected by the labor market conditions and can be effectively substituted by others worker types, which is an assumption commonly used in the literature (Katz & Murphy, 1992; Autor et al., 1998; Card & Lemieux, 2001). Initially, our calculations focus on men's wages, influenced by the enduring gender roles prevalent in Chile (Contreras et al., 2012). Nevertheless, as a robustness check, we intend to incorporate data from women into our analysis. Furthermore, we plan to calculate the average hourly wage within each cell defined by zone, experience, and education. However, as the required hours data are not present in the SC dataset, we will resort to utilizing hours reported by formal workers engaged in 40 to 50 hours of work, as documented in the ENE surveys.

Concerning labor supply, we utilized the ENE survey data from 2010 to 2019 to calculate the supply for each of the sixteen groups within the three zones every month. In the computation, we consider both formally and informally employed workers and those unemployed. This approach allows us to accommodate variations in wages arising from changes in informality and fluctuations in the unemployment rate due to business cycle dynamics.

Regarding education, within the SC dataset, we have access to the highest completed educational level for each individual. Conversely, in the ENE data, we possess information about educational levels, regardless of whether they were completed or not. To construct the relative supplies of each type of worker, we will adopt a methodology in line with previous literature (Autor et al., 1998; Card & Lemieux, 2001). Individuals with incomplete levels of education will be divided into two equal parts. One half will be assigned to the completed level, and the other half will be assigned to the completed level below their current incomplete level. For instance, half of those with incomplete secondary education will be assigned to the completed level, while the remaining half will be assigned to complete primary education. Workers with incomplete primary education will be assigned to the primary level. It is noteworthy that both college and higher vocational workers with incomplete levels will be assigned to the secondary level.[12]

Ultimately, our approach involves conducting regressions on equations 3 and 4 to derive parameters related to the elasticity of substitution across various experience groups. Subsequently, equations 5 and 6 will be employed to obtain the elasticity of substitution between the two categories of low-skilled workers and between the two categories of high-skilled workers. Concurrently, this process will yield new estimates for the elasticities of substitution across diverse experience groups. In the final stage, regression on equation 7 will be undertaken to derive estimates for the elasticity of substitution between unskilled and skilled workers while simultaneously obtaining estimates for the previously mentioned elasticities of substitution. It is important to notice that each regression will include equation fixed effects (this means, different experience-labor groups are allowed to have different fixed effects), and we will also include the triple interaction of a time polynomial of order 9, the region, and the experience-education groups to allow different trends in each region and group across time.[13]

[12]It might be plausible to assume that workers with incomplete college education could transition to the higher vocational group, especially considering the similarity in wage premiums for both types of workers. However, adopting such an assumption could complicate the interpretation of the coefficient γ—our primary parameter of interest—since it would involve including workers with college education on both sides of the analysis..

[13]This procedure is similar to the one in Fernández & Messina (2018), since they use data from 3 countries

5.4 Results

Table 1 presents the regression estimates for equations 3 to 7. Concerning the estimates of substitution across various experience groups for unskilled workers (θ_B), we note that the values fall within the range of 0.20-0.24. This range is stable and is smaller than ranges from previous literature, which as expected comes from the fact of using administrative data instead of surveys to compute the coefficients. Additionally, the estimates from equation 7 closely resemble those in equation 3, enhancing the reliability of the latter estimates given the changes in specification. It is worth mentioning that estimates in equation 5 might differ due to the lower number of observations used to compute the coefficients. A similar pattern emerges when examining the coefficients of the elasticity of substitution for different experience groups among skilled workers (θ_A). The range of estimates is 0.14-0.15, even narrower than before, with the outlier being the one from equation 6, likely due to the lower number of observations used in its computation.

Table 1: Model estimates by equation (2010-2019)

Relative Supply	Equation 3	Equation 4	Equation 5	Equation 6	Equation 7
Experience: Unskilled ($-1/\theta_B$)	-0.204***		-0.235***		-0.204***
	(0.010)		(0.017)		(0.007)
Experience: Skilled ($-1/\theta_A$)		-0.141***		-0.151***	-0.142***
		(0.005)		(0.008)	(0.005)
Secondary / Primary ($-1/\sigma_\delta$)			-0.004		0.000
			(0.048)		(0.005)
College / Higher Vocational ($-1/\sigma_\gamma$)				-0.093***	-0.032**
				(0.034)	(0.016)
Skilled / Unskilled ($-1/\sigma_\rho$)					-0.232***
					(0.073)
N	4,263	4,284	1,421	1,428	5,698
R^2	0.96	0.96	0.81	0.80	0.94

Source: Own elaboration using SC and ENE data for the period 2010-2019. Robust standard errors in parenthesis. The observations in similar equations are not balanced due the lack of proper information in some groups at the zone level. The model was estimated through OLS with robust standard errors and included equation fixed effects and trends considering time, region and group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Concerning the substitution across unskilled worker types, we find that the estimates show values not statistically different from zero, which means that we cannot reject the hypothesis that the firms consider those workers perfect substitutes. The above is due to the following reasoning: if $1/\sigma_\delta$ is close to zero, it means that σ_δ tends to infinity, which in turn means that δ tends to 1.

(Argentina, Brazil and Chile) and in their estimations they use the triple interaction of a time trend polynomial, the country and the group, to allow different trends in each combination. The polynomial used by the authors is of order 3, and they have on average 16 observation by country. In our case, we have a balanced panel of 111 observations by zone.

[14] Given the above, it is direct to see that the CES function of unskilled workers with a parameter $\delta = 1$ will degenerate into a linear function.

A parallel observation emerges as in the previous case concerning the substitution between the two types of high-skilled workers. The estimation of equation 6 yields a small value for the substitution across both worker types. In contrast, the estimate of equation 7 indicates that the elasticity of substitution between both types is relatively modest. Given that σ_γ is close to zero, it suggests that college and higher vocational education workers can be considered close to perfect substitutes. This surprising result aligns with the evidence presented in the previous sections. Almost 40 percent of workers with a college education tend to occupy positions that do not necessitate their qualifications, and half assume roles designed for workers with higher vocational education. The above compels workers with higher vocational education to also take positions for which they are overqualified, reducing their wage premium and contributing to the observed tendencies in the data.

Then, we estimate the elasticity of substitution between skilled and unskilled workers using equation 7. The coefficient associated with the relative supplies is -0.23, which implies a elasticity of substitution $\sigma_\rho = 4.35$, which in turn, implies a parameter $\rho = 0.77$. The coefficient σ_ρ is less than a third than the obtained in Fernández & Messina (2018) for the period 1990-2013, and therefore it is consistent with the slower movements in inequality observed in Chile in the last decade.[15] In the Appendix we include some robustness checks regarding considering women into the sample, using formal part time workers to impute the average hours, different specifications for the time trend polynomial, and using only occupied workers as a measure of labor supply.

Finally, in terms of interpretation, it is relevant to mention that the elasticities computed here are micro elasticities at the zone level, i.e., if there was only one firm in each zone and the elasticity of substitution between worker types is common between them, then the parameters found here would be link to those elasticities. In case that we would interested in obtaining the aggregate or macro elasticity we would need additional structure in the model, specifically about the elasticity of demand for the good produced as illustrated in Oberfield & Raval (2021).

6 Discussion

The outcomes of our estimations paint a pessimistic picture of the massification of tertiary education, particularly for those pursuing higher vocational education. The data and our estimates indicate a trend where college-educated individuals displace these workers. This displacement

[14] This is because $\sigma_\delta = \frac{1}{1-\delta}$ by definition.

[15] Chile's Gini Coefficient decreased from 57.2 in 1990 to 46.0 in 2013. However, since then, the coefficient has been around 44-45.

forces them into positions for which they are overqualified, failing to attain a significant wage premium despite their tertiary education. This prompts the question of why this phenomenon is occurring. One hypothesis we propose is that higher vocational programs are not developed independently of college programs. Instead of equipping students with distinct skills that differentiate them from those with a college education, these programs might provide their students with just a fraction of the courses given in the college version.

As previously stated, a technical degree typically requires 1 to 3 years of study (with an average duration of 2 years), a professional degree entails a 4-year study period, and a college degree necessitates a minimum of 5 years of study. Our observation suggests that a noteworthy portion of technical degrees may encompass the initial two years of a professional degree. Furthermore, there appears to be a substantial overlap between the content of a professional degree and the corresponding college degree. It is relevant to mention that using this structure in both technical and professional degrees is developed to allowed the transition between higher vocational to college studies and it is highly regulated by the Ministry of Education, and it is used both by public and private universities and professional institutes.

In our analysis of the first group (i.e. technical and professional programs), we conducted a comparison of the curriculum structures of the two most prestigious institutions providing technical and professional education in Chile. Both institutions have received the maximum accreditation duration of 6 years from the National Agency of Accreditation (CNA) for their technical education programs. One of them offers 34 technical programs, and remarkably, 27 of these (79%) share identical study plans with the initial half of the associated professional degree. Similarly, the another one provides 41 technical degrees, with 30 of them (73%) having study plans that mirror the first half of the associated professional degree.

In our examination of the transition from professional to college degrees, we checked the possibilities at two universities, one with seven years of accreditation and the another with six years. Both universities, renowned for their emphasis on engineering and technology, stand as the most prestigious institutions offering both professional and college degrees. Notably, both universities facilitate a process for students to continue their studies from a professional degree to a college degree. One of them provides this pathway through seven professional programs associated with a corresponding college degree, while the other offers nine such degrees. Although it is intriguing to explore the similarities in study plans between professional and college degrees, this nuanced analysis falls beyond the scope of the present paper.

In light of the aforementioned observations, it becomes plausible to consider that the modular structure of technical and professional programs is contributing to the perception of workers with college education as possessing comprehensive training. Consequently, these individuals may be more likely to secure positions that allow them to accumulate additional human capital, thus contributing to the discernible differences in wage premiums evident in our data.

Regarding public policy design, a strategic approach involves tailoring study programs to cater to the distinct needs of technical and professional graduates versus college graduates. Specifically, a targeted training regimen for technical and professional graduates could be instituted, emphasizing skills in high demand within the labor market. In contrast, college graduates could hone their focus on more comprehensive and theoretical frameworks. This differentiation is pivotal, particularly when considering the synergies between higher vocational and college-educated workers. Such synergy empowers the former group to elevate their wages and enhance their overall labor conditions. This empowerment extends beyond qualification for specific job roles, encompassing access to formal sector employment and accumulating valuable human capital.

7 Conclusion

This paper delves into the repercussions of the massification of tertiary education in Chile throughout the 2010-2019 decade, with a specific focus on understanding how firms are substituting workers with different qualifications during this period. First, we showed that the massification of tertiary education implied increasing the share of workers with tertiary education from 17 percent to 45 percent in three decades. However, despite this changes, we do not observe a relevant drop in the wage premium associated with college educated workers, both at the average level as in the quantiles.

We focused on the requisite qualifications for various occupations, discovering a significant discrepancy between educational attainment and job requirements. Notably, nearly 60 percent of individuals with higher vocational education find themselves overqualified for their current roles, paralleled by a similar situation for approximately 40 percent of those with a college education. This issue gains heightened relevance when considering that nearly half of the overqualified individuals with a college education occupy positions initially designed for higher vocational graduates. This observation suggests an interplay where college-educated workers may be displacing their higher vocational counterparts into alternative roles.

Subsequently, we constructed a novel model to estimate the elasticities of substitution within the workforce, distinguishing between four distinct categories of workers. Our model breaks new ground by incorporating higher vocational workers as a differentiated group. What sets this model apart is its utilization of administrative data—an unconventional approach in developing countries currently undergoing processes of massification. The adoption of administrative data not only enhances the precision of our models but also affords us a more realistic assessment of the actual cost incurred by firms when hiring workers—an aspect often overlooked in the prevailing literature of developing countries, which tends to rely heavily on survey data. The model's estimations reveal a noteworthy trend during this timeframe, suggesting close-to-perfect substitutability between higher vocational and college-educated workers. This reinforces the consistency with the findings presented in earlier sections.

Finally, our discussion delved into technical, professional, and college degree study plans. Broadly, we noted that technical degree curricula are encompassed within the initial years of professional degree programs, furnishing technical workers with an education wholly embedded in the training of those pursuing professional education. Similarly, the relationship between professional and college degrees exhibited a comparable pattern, with universities facilitating the progression from professional studies to college education through affiliated study programs. These observations imply that the observed substitution in the market is rational. For policymakers aiming to leverage the complementarity between these two forms of tertiary education, a strategic approach would involve differentiation at each educational level. This approach ensures that the unique strengths of technical, professional, and college education are honed and optimized, fostering a more robust and productive integration within the broader workforce.

References

Abud, M. J., Cases, R., Fuentes, A., & Vergara, R. (2022). Salario mínimo en Chile: Un aporte a la discusión. *Puntos de Referencia*, *610*, Centro de Estudios Públicos.

Autor, D., Katz, L., & Krueger, A. (1998). Computing inequality: Have computers changed the labor market? *The Quarterly Journal of Economics*, *113*(4), 1169-1213.

Battistón, D., García-Domench, C., & Gasparini, L. (2014). Could an increase in education raise income inequality? evidence for Latin America. *Latin American Journal of Economics*, *51*(1), 1–39.

Beyer, H., Rojas, P., & Vergara, R. (1999). Trade liberalization and wage inequality. *Journal of Development Economics*, *59*(1), 103-123.

Card, D., & Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, *116*(2), 705-746.

Caselli, F. (2005). Accounting for cross-country income di§erences. *The Handbook of Economic Growth*.

Caselli, F., & Ciccone, A. (2019). The human capital stock: A generalized approach: Comment. *American Economic Review*, *109*(3), 1155-1174.

Contreras, D., Hurtado, A., & Sara, F. (2012). La excepcion chilena y las percepciones de genero en la participacion laboral femenina. *Documentos de Trabajo*, *374*, Departamento de Economía, Universidad de Chile.

Fernández, M., & Messina, J. (2018). Skill premium, labor supply, and changes in the structure of wages in Latin America. *Journal of Development Economics*, *135*, 555-573.

Ferreyra, M., Avitabile, C., Botero, J., Haimovich, F., & Urzúa, S. (2017). *Momento decisivo: La educación superior en América Latina y el Caribe.* Washington, DC: Banco Mundial.

Fuentes, A., & Vergara, R. (2018). Brecha salarial de género: evolución en el período 1990-2017. *Puntos de Referencia N° 490*, Centro de Estudios Públicos.

Gallego, F. (2012). Skill premium in Chile: Studying skill upgrading in the South. *World Development*, *40*(3), 594-609.

Gindling, T., & Robbins, D. (2001). Patterns and sources of changing wage inequality in Chile and Costa Rica during structural adjustment. *World Development*, *29*(4), 725-745.

González, L., & Ureta, C. (2015). Proyección del gasto fiscal en educación superior. *Documentos de Trabajo*, Dirección de Presupuestos, Gobierno de Chile.

Haanwinckel, D., & Soares, R. (2021). Workforce composition, productivity, and labour regulations in a compensating differentials theory of informality. *The Review of Economic Studies*, *88*(6), 2970-3010.

Hall, R., & Jones, C. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, *114*(1), 83-116.

Hlasny, V. (2020). Nonresponse bias in inequality measurement: Cross-country analysis using Luxembourg income study surveys. *Social Science Quarterly*, *101*(2), 712-731.

Holz, M. (2019). Implementación presupuestaria de las leyes N° 21.091 sobre educación superior, y N° 21.094, sobre Universidades Estatales. *Biblioteca del Congreso Nacional de Chile*.

INE. (2023). Estadísticas del mercado laboral en Chile. *Boletín Laboral*.

Jones, B. (2014). The human capital stock: A generalized approach. *American Economic Review*, *104*(11), 3752-3777.

Jones, B. (2019). The human capital stock: A generalized approach: Reply. *American Economic Review*, *109*(3), 1175-1195.

Katz, L., & Author, D. (1999). Changes in the wage structure and earnings inequality. In *Handbook of Labor Economics* (Vol. 3, p. 1463-1555). Elsevier.

Katz, L., & Murphy, K. (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, *107*(1), 35-78.

Klenow, P., & Rodriguez-Clare, A. (1997). The neoclassical revival in growth economics: Has it gone too far? *NBER Macroeconomics Annual*, 73-103.

McMillan, M., Rodrik, D., & Sepulveda, C. (2017). Structural change, fundamentals and growth: A framework and case studies. , *Technical report, National Bureau of Economic Research*.

Mincer, J. (1974). *Schooling, experience, and earnings.* New York: National Bureau of Economic Research.

Mineduc. (2023). Evolución matrícula educación superior en Chile. *Biblioteca del Congreso Nacional de Chile*.

Moore, J., Stinson, L., & Welniak, E. (2000). Income measurement error in surveys: A review. *Journal of Official Statistics*, *16*(4), 331-362.

Murphy, K., & Welch, F. (1990). Empirical age-earnings profiles. *Journal of Labor Economics*, *8*(2), 202-229.

Murphy, K., & Welch, F. (1992). The structure of wages. *The Quarterly Journal of Economics*, *107*(1), 285-326.

Oberfield, E., & Raval, D. (2021). Micro data and macro technology. *Econometrica*, *89*(2), 703-732.

Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth*, *21*, 1-33.

Rojas, E., & Sanchez, R. (2014). Do unemployment insurance saving accounts affect the quality of the job match?. a regression discontinuity aproach. *Working paper*.

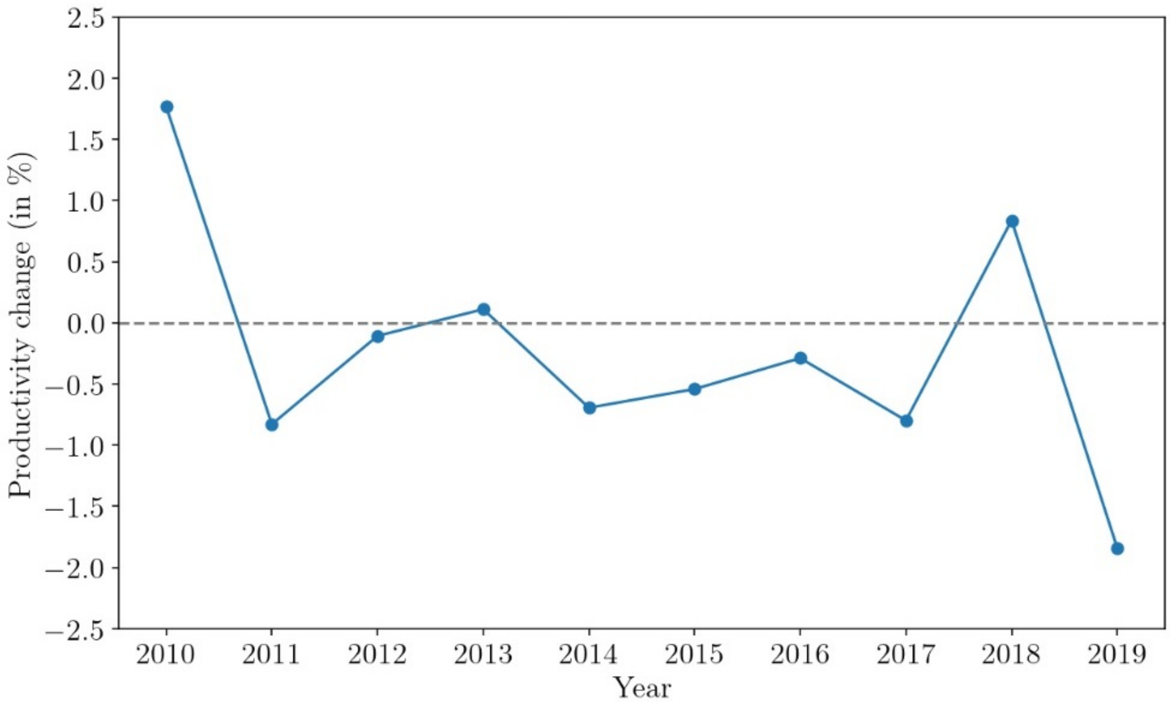
Urzúa, S. (2012). La rentabilidad de la educación superior en Chile: Revisión de las bases de 30 años de políticas públicas. *Estudios Públicos*, *125*, 1-52.

Vu, K. (2023). Higher education expansion and the rise of the service economy in vietnam. , *Job market paper*.

8 Appendix

8.1 Productivity change in Chile

Figure 6: Evolution of productivity in Chile (2010-2019)



Source: Data from Comision Nacional de Evaluación y Productividad (CNEP). The measure presented is computed at the aggregate level of the economy and it considers the adjustment in the capital levels developed by CNEP.

8.2 Robustness checks

important to assess the robustness of our conclusions under varying assumptions. To explore this, we begin by examining the inclusion of data from women in our analysis. In the main text, we highlighted the substantial influence of gender roles in Chile, indicating that female labor supply and wages might not solely respond to market conditions. Then, a first robustness check implies computing the elasticities of substitution, including women's data, and re-estimate all equations 3-7 by incorporating male and female observations to accomplish this.

Table 2 shows the results. As can be seen, we still can see that the parameter $1/\sigma_\gamma$ is close to zero, although in this case is relatively far from zero than in the male sample. Also the estimations for the coefficients $1/\theta_B$ and $1/\theta_A$ are similar to the ones using the male sample. However, a substantial difference is found when computing the coefficient for $1/\sigma_\rho$, where we found that in this sample there is a perfect substitution across workers with tertiary and secondary education. This could be driven because woman, even those with tertiary education, have a higher proportion of part time jobs and tend to work in low pay sectors as education or commerce (Fuentes & Vergara, 2018), and therefore there are missing an important component to accumulate human capital.

Regarding the imputation of working hours in each region-education-experience cell, we also estimate the model considering all individuals who work between 30 and 60 hours. the results can be seen in Table 3. Although the coefficient associated with $1/\sigma_\gamma$ increases to -0.05, is still rather small suggesting a relevant degree of substitution between the two types of tertiary educated workers.

Another relevant question is whether the decision of the polynomial order affects the coefficients or not. To assess this, we estimate the model considering a polynomial of one order less and more than the one we chose in the main specification. Tables 4 and 5 show the re-estimated coefficients of the model using the different polynomial orders. As can be seen, there are no significant differences between the models associated to the interest parameter $1/\sigma_\gamma$.

Lastly, we explore a model variant in which the relative labor supplies are calculated exclusively based on employed individuals, rather than considering the entire workforce. This represents a more stringent version of the model, presupposing that wage pressure originates solely from those who are currently employed. The outcomes are detailed in Table 6. Remarkably, all parameters maintain consistency with the previous estimations, underscoring the robustness of our estimation method to variations in defining the relevant population as a measure of labor supply.

Table 2: Model estimates by equation including women data (2010-2019)

Relative Supply	Equation 3	Equation 4	Equation 5	Equation 6	Equation 7
Experience: Unskilled ($1/\theta_B$)	-0.160***		-0.186***		-0.160***
	(0.009)		(0.015)		(0.006)
Experience: Skilled ($1/\theta_A$)		-0.157***		-0.168***	-0.158***
		(0.004)		(0.007)	(0.005)
Secondary / Primary ($1/\sigma_\delta$)			-0.007		0.003
			(0.042)		(0.004)
College / Higher Vocational ($1/\sigma_\gamma$)				-0.081***	-0.037***
				(0.031)	(0.014)
Skilled / Unskilled ($1/\sigma_\rho$)					-0.095
					(0.059)
N	4,263	4,284	1,421	1,428	5,698
R^2	0.96	0.96	0.81	0.80	0.94

Source: Own elaboration using SC and ENE data for the period 2010-2019. Robust standard errors in parenthesis. The observations in similar equations are not balanced due the lack of proper information in some groups at the zone level. The model was estimated through OLS with robust standard errors and included equation fixed effects and trends considering time, region and group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Model estimates by equation using workers working 30-60 hours (2010-2019)

Relative Supply	Equation 3	Equation 4	Equation 5	Equation 6	Equation 7
Experience: Unskilled ($1/\theta_B$)	-0.191***		-0.228***		-0.190***
	(0.010)		(0.017)		(0.007)
Experience: Skilled ($1/\theta_A$)		-0.104***		-0.116***	-0.101***
		(0.004)		(0.008)	(0.005)
Secondary / Primary ($1/\sigma_\delta$)			0.003		0.006
			(0.049)		(0.004)
College / Higher Vocational ($1/\sigma_\gamma$)				-0.081***	-0.051***
				(0.031)	(0.013)
Skilled / Unskilled ($1/\sigma_\rho$)					-0.142**
					(0.071)
N	4,263	4,284	1,421	1,428	5,698
R^2	0.95	0.94	0.79	0.75	0.91

Source: Own elaboration using SC and ENE data for the period 2010-2019. Robust standard errors in parenthesis. The observations in similar equations are not balanced due the lack of proper information in some groups at the zone level. The model was estimated through OLS with robust standard errors and included equation fixed effects and trends considering time, region and group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Model estimates by equation using a time trend polynomial of order 8 (2010-2019)

Relative Supply	Equation 3	Equation 4	Equation 5	Equation 6	Equation 7
Experience: Unskilled ($1/\theta_B$)	-0.199***		-0.232***		-0.197***
	(0.010)		(0.017)		(0.006)
Experience: Skilled ($1/\theta_A$)		-0.140***		-0.151***	-0.141***
		(0.004)		(0.008)	(0.005)
Secondary / Primary ($1/\sigma_\delta$)			-0.005		0.009**
			(0.041)		(0.004)
College / Higher Vocational ($1/\sigma_\gamma$)				-0.116***	-0.035***
				(0.033)	(0.015)
Skilled / Unskilled ($1/\sigma_\rho$)					-0.282***
					(0.066)
N	4,263	4,284	1,421	1,428	5,698
R^2	0.96	0.96	0.81	0.80	0.94

Source: Own elaboration using SC and ENE data for the period 2010-2019. Robust standard errors in parenthesis. The observations in similar equations are not balanced due the lack of proper information in some groups at the zone level. The model was estimated through OLS with robust standard errors and included equation fixed effects and trends considering time, region and group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 5: Model estimates by equation using a time trend polynomial of order 10 (2010-2019)

Relative Supply	Equation 3	Equation 4	Equation 5	Equation 6	Equation 7
Experience: Unskilled ($1/\theta_B$)	-0.208***		-0.236***		-0.207***
	(0.010)		(0.017)		(0.006)
Experience: Skilled ($1/\theta_A$)		-0.143***		-0.151***	-0.144***
		(0.004)		(0.008)	(0.005)
Secondary / Primary ($1/\sigma_\delta$)			-0.045		-0.041***
			(0.054)		(0.004)
College / Higher Vocational ($1/\sigma_\gamma$)				-0.088***	-0.032**
				(0.040)	(0.015)
Skilled / Unskilled ($1/\sigma_\rho$)					-0.493***
					(0.084)
N	4,263	4,284	1,421	1,428	5,698
R^2	0.96	0.96	0.81	0.80	0.94

Source: Own elaboration using SC and ENE data for the period 2010-2019. Robust standard errors in parenthesis. The observations in similar equations are not balanced due the lack of proper information in some groups at the zone level. The model was estimated through OLS with robust standard errors and included equation fixed effects and trends considering time, region and group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 6: Model estimates by equation using occupied workers as supply measure (2010-2019)

Relative Supply	Equation 3	Equation 4	Equation 5	Equation 6	Equation 7
Experience: Unskilled ($1/\theta_B$)	-0.208***		-0.248***		-0.207***
	(0.008)		(0.015)		(0.006)
Experience: Skilled ($1/\theta_A$)		-0.108***		-0.121***	-0.107***
		(0.004)		(0.007)	(0.005)
Secondary / Primary ($1/\sigma_\delta$)			-0.008		-0.005
			(0.043)		(0.004)
College / Higher Vocational ($1/\sigma_\gamma$)				-0.063**	-0.033***
				(0.026)	(0.011)
Skilled / Unskilled ($1/\sigma_\rho$)					-0.202***
					(0.064)
N	4,263	4,284	1,421	1,428	5,698
R^2	0.97	0.96	0.83	0.79	0.94

Source: Own elaboration using SC and ENE data for the period 2010-2019. Robust standard errors in parenthesis. The observations in similar equations are not balanced due the lack of proper information in some groups at the zone level. The model was estimated through OLS with robust standard errors and included equation fixed effects and trends considering time, region and group. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.