CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

RETURNS TO SCHOOLING AND SKILLS IN MEXICO: A HETEROGENEOUS MARKET

TESINA

QUE PARA OBTENER EL TÍTULO DE

LICENCIADA EN ECONOMÍA

PRESENTA

ARELY ALEJANDRA GARCIA GARCIA

DIRECTOR DE LA TESINA: DR. JORGE OMAR MORENO TREVIÑO

CIUDAD DE MÉXICO 2021

Dedico esta tesis a mis padres:
no sólo por enseñarme a volar,
sino por jamás cortarme las alas
no importa cuán lejos vuele.

Acknowledgments

Quiero agradecer:

A mis padres, a quienes les debo gran parte de la persona que soy hoy; sin ellos nada de esto hubiera sido posible. A mis hermanos, que me esperaron con la misma anticipación y alegría cada semestre que me fui. Y a mi familia, que celebró a la distancia todos mis logros y me prestó su apoyo a cada paso.

Al Dr. Jorge O. Moreno, por su invaluable guía como asesor de este proyecto. Sin lugar a dudas, fue fundamental en mi decisión de entrar a esta disciplina que me permite comprender el mundo. Gracias por acompañarme hasta el final de esta etapa, que esperemos que esté lejos de ser el final del camino. Es admirable ver su vocación como profesor, pero todavía más su calidad humana.

Al Dr. Alejandro López Feldman, por sus comentarios en este proceso de investigación. Ha sido un profesor extraordinario, a pesar de los retos impuestos por la virtualidad. Al Dr. Gustavo A. Del Angel, por sus sugerencias a mi tesina. Gracias por darme la oportunidad de ser su asistente de investigación y aprender de sus proyectos.

A la Dra. Luciana C. Moscoso, por su dedicación, atención y apoyo para cada uno de mis compañeros a lo largo de nuestra licenciatura. A la Dra. Fernanda Márquez Padilla, que estuvo pendiente de mi progreso como profesora y tutora. Ambas han sido fuente de inspiración en una carrera que en ocasiones se siente faltante de mujeres. A la Mtra. Melissa Labougle, alentadora y crítica a partes iguales, quien me abrió las puertas de un espacio en el que el que aprendí tanto como enseñé y que fue mi refugio tantas tardes.

Al CIDE, que ha hecho más que formarme como economista. En esta institución he aprendido sobre esfuerzo, empatía y perseverancia. Me ha dado herramientas necesarias para ser mejor economista y persona. Esta etapa me ha permitido entender que el cambio es inevitable y que debo enfrentarlo con templanza y resiliencia. Cada uno de los profesores de la planta docente tiene mi admiración y respeto. Agradezco especialmente a los que me dieron la oportunidad de asistirlos como docente.

A los compañeros increíbles que esta etapa cruzó en mi camino; especialmente a dos. Gabriel, el primer amigo que conocí en esta aventura y que se quedó para dar conmigo hasta el último paso. Fernanda, quien ha sido amiga, confidente y hermana en los cuatro años pasados. Ustedes han sido mis cómplices, así como parte fundamental de mis éxitos. Es mi privilegio llamarlos amigos. Lograron algo que nadie hubiera imaginado: hacer de la CdMx un hogar.

Abstract

The characteristics of the Mexican labor market are changing. On the one hand, the population is increasingly educated, which means that skilled workers are less scarce. On the other hand, occupations are shifting requirements, giving greater weight to skills. The objective of this work is to estimate the returns to schooling and skills in different occupational markets. The present thesis estimates the returns to schooling and skills (cognitive and physical). For this purpose, the present work uses a model of endogenous factors with self-selection, inspired by Roy's model to compare returns across occupations and between genders. Both schooling and skills are found to have a significant effect on wages. In addition, there is evidence that people self-select in the market that presents higher returns to the skills they posses. Results indicate that wage premiums from schooling or skills are heterogeneous between genders and occupations, which should be accounted for in future estimations on returns.

Keywords: returns, schooling, skills, human capital, gender, factor
Clasificación JEL: J24

Contenido

List of figures

List of tables

Chapter 1

Introduction

According to the 2020 Population and Housing Census, the average schooling of the Mexican population increased from 8.6 to 9.7 years in the last decade. This result follows the trend, where there has been a constant increase in education. Not only the average schooling has increased, but also enrollment rates at all educational levels (INEGI, 2021). Since a positive relationship between individual schooling and wages has been proven, a better educated labor force could mean higher wages for the population. However, there is also an adverse effect from this circumstance: education is no longer a scarce good. The increase in the skilled workers supply has decreased the value or reward to schooling in wages. According to Caamal Olvera (2017), the increase in wage derived from an additional year of schooling has decreased from 10% to 8% in the same period.

Now, this is not the only changing aspect of the labor market. In recent years, several occupations have started to reinforce the importance of abilities. Several industries emphasize the relevance of creativity, willingness to work as a team, and critical thinking. Not only the hard and soft skills are relevant, but also health. The aforementioned characteristics, together with schooling, contribute to individual productivity and, in consequence, to wages. However, different occupations may value skills differently since it may depend on the job's specific requirements. In this sense, an understanding of wages should consider returns to schooling, but also returns to skills that potentially identify human capital of individuals (G. Becker, 1993).

The impact of education in wages has been extensively researched. Various studies have used diverse methodologies to estimate correctly the effect. It also has been done for different countries, and the comparison of results has provided some general tendencies. Nonetheless, the majority of these works consider the labor market as a homogeneous single entity. The objective of the present thesis is to determine if different occupations reward schooling and skills differently. The present work addresses the following questions: how do the returns to schooling and skills vary across occupations? What are the effects of schooling and skills in self-selection across occupations? The hypothesis proposed are: (1) Returns are heterogeneous across occupations, reflecting the specific market requirements, (2) People with greater skills self-select themselves in occupations that value their skills the most. To test this, I estimate an endogenous factor model to estimate the effects of cognitive and physical abilities on wage.

Human capital has been a topic of interest in the last decades since Schultz (1961) and G. Becker (1975) seminal works. According to microeconomic theory, wages are determined by the marginal productivity of individuals. First attempts to test this empirically focused solely on the relationship between schooling and earnings (Mincer, 1974). The marginal increase in earnings derived from an additional year of schooling was termed returns to schooling. Various authors have used different approaches to precisely estimate these returns (Griliches, 1977; Thomas & Strauss, 1997; Caamal Olvera, 2017). Recently, there has been a growing concern that education cannot capture an individual's productivity entirely. In this sense, several econometric approaches have been devised in order to capture latent skills, which are non-observable (Cunha & Heckman, 2007; Kejriwal, Li, & Totty, 2020; Ashworth, Hotz, Maurel, & Ransom, 2021).

My analysis differs from previous literature in two ways: First, most of the work on returns to schooling has focused on the labor market as a whole, neglecting the possible heterogeneity across occupations with specific characteristics. Even the approaches that have divided the workforce do not integrate occupations typically associated with informality, such as domestic workers and street sellers. It is important to include this sector, since it represents an important proportion of the market in Mexico. According to INEGI (2020) it represents around 28% of the labor force.

Second, the studies for Mexico have used single variables as a proxy for skills. The methodology proposed in this thesis attempts to construct the latent variables of cognitive and physical skills following a multidimensional perspective using a factor model approach as proposed by Cunha, Heckman, and Schennach (2010). This approach permits to include unobservable variables in the estimation that may be relevant determinants for wages. Additionally, the empirical strategy is relevant to help understand some aspects of labor market participation, as well as the distribution of wages in Mexico. Also, these results are useful to test empirically the impact of investments in human capital to economic outcomes.

Results agree with previous literature that identified schooling and skills as relevant determinants of wages. However, schooling and skills are priced differently in separate labor markets. While administrative workers market rewards more schooling and cognitive skills, the manual workers one recompenses physical skills. There is also relevant differences between genders, which may respond to the composition of occupations within each group. This can be taken as evidence of the heterogeneity of the labor market. Additionally, skills play an essential role in self-selecting in some occupations, specially for women.

The present thesis is organized as follows. Chapter 2 discusses the relevant literature on the topic. Chapter 3 details the data used to conduct the analysis and presents some descriptive statistics. Chapter 4 presents the endogenous factor framework and the estimated specifications. Chapter 5 displays the results obtained. Chapter 6 includes the final remarks.

Chapter 2

Literature Review

In the last four decades, the formalization of a theory of human capital has taken relevance in the economic literature. The concept of human capital refers to the increase in productivity due to personal investment. The central conception of human capital resides in individual abilities. In this sense, education and experience boost the skills of individuals, which in consequence make them more productive (G. S. Becker & Tomes, 1986). Since the beginning of the discussion on human capital, there has been particular attention to the relationship between education and productivity (Psacharopoulos, 1972). According to microeconomic theory, productivity is reflected by wages. Mincer (1974) was one of the first authors to establish the positive relationship between years of education and wages. The equation proposed by him takes the natural logarithm of wage as the dependent variable and schooling and work experience as independent ones. The estimated coefficient associated with education could be interpreted as the rate of return of an additional year of schooling.

However, early in the estimation several authors raised the concern that the returns to education are frequently over-estimated. They argued that a plausible explanation was the fact that ability was neglected, even when it may be independent from experience and schooling (Hause, 1972; Chamberlain, 1977). Then, some of the effect on wages due to skills would be attributed to schooling. Griliches (1977) explores some of the consequences of leaving out unobservable skills from the analysis. To address this issue, the author suggests using some observable variables as a proxy, such as a test score, to improve the estimations.

In Latin America, the problem of letting out ability has been evident. The returns to schooling first estimated were exceptionally high, which did not seem to match with the low intra-generational mobility observed. Therefore, several efforts have tried to extend the wage equations to reduce some of the upward bias. Because of the unobservable nature of skills, various approaches included some observable variables as substitutes. A recurrent method has focused on family background variables as a proxy for the omitted variable of skill. The inclusion of differences between siblings (Wolfe & Behrman, 1987) or parents' schooling (Heckman & Hotz, 1986) decreases the estimated returns by at least a third. Lam and Schoeni (1993) study the case of working males in Brazil. They propose a theoretical assortative mating model and intergenerational correlations. Building on this, they construct an econometric approach where the wage equation includes parental schooling, the parents-in-law and wife schooling. As a result, the coefficient associated with returns on education declined by one-fourth to one-third. In this regard, proxies seem to reduce upward bias, even though the cannot measure productivity accurately.

2.1 Skills

To include skills in the estimations, some authors have used achievement tests as an observable variable to measure cognitive skills. Nevertheless, Hansen and Mullen (2003) have argued that a common unobserved ability generates both schooling and test scores. In this sense, when test scores are added to a wage equation as a skill measure, they can be a poor proxy for two motives: First, the usual difficulty of causing endogeneity through its error term. Second, even if the error term is zero -meaning that test scores are a perfect proxy for cognitive skills- schooling determines test scores affects that the marginal effect of education will be biased. To be precise, if the ability has a positive repercussion on wages, the returns to schooling would be underestimated since their indirect effect on wages through increasing the measured ability would be ignored.

To solve this problem, there is an important distinction to make: general intelligence is composed of crystallized and fluid factors. Crystallized intelligence refers to the capacity to use acquired knowledge, and it is related to achievement test scores, such as vocabulary and general information. Fluid intelligence, in contrast, measures the ability to analyze and solve logical problems, like figural analysis and paired associates (Cattell, 1986). Carlsson, Dahl, Öckert, and Rooth (2015) consider this distinction and exploit conditional random variation in test dates to measure the effect of schooling in both kinds of skills. They find that formal education does increase the results in tests that measure crystallized intelligence. On the contrary, there is no statistically significant evidence to argue that schooling affects the fluid intelligence component. According to this result, the downward bias in returns to education may be avoided using fluid intelligence measures instead of crystallized ones.

Other authors have pointed out that human capital is multidimensional and that the link between productivity and health has been neglected, and several variables account for investments in that area (Maccini & Yang, 2009). For example, height is a variable that can account for early – i.e., long term-investments in human capital (Deaton, 2007). On the contrary, protein and caloric intake may reflect short-term investment (Martorell, 2017). A non-linear combination of both, the body mass index is a measure of the complexion of an individual and can be interpreted as a medium-term health component (Thomas & Strauss, 1997).

As another fundamental component of human capital, health indicators can be crucial in personal abilities. To address this, Thomas and Strauss (1997) use a cross-sectional household survey from urban Brazil. To make a comprehensive understanding of health, they use a wage equation that includes four health measures: height, body mass index, calorie intake, and protein intake. The results show that height has a significant effect on wages. However, body mass index (BMI) is more significant among the less educated, suggesting that strength is more relevant in low-skilled occupations. Even controlling for these measures, years of education are a significant determinant of wages, but the estimated returns also decrease by roughly one-third by using this approach.

The differentiated effect of skills across occupations has been suggested by Ingram and Neumann (2006). They use data from the dictionary of occupational titles to account for four broad skill categories: intelligence, fine motor skill, coordination, and strength. However, this means that they use a job description rather than an actual measurement recovered from the individual. Then, they use the Mincer equation augmented by these factors. They find that the returns to schooling decrease after including these factors and that skills can explain some of the wage dispersion.

2.2 Factors

More recently, instead of using indirect variables to decrease their bias, an extensive literature has been devoted to include abilities directly in the estimation of returns to schooling. As skills are self-productive and have dynamic complementarity, this has proven to be challenging. The first term refers to the fact that the skills produced at an early stage of life contribute to an easier acquisition of skills at a later stage: in other words, skills are self-reinforcing. The dynamic complementarity implies that skills raised at an early age make future investments in other skills more productive. Together, the two properties create a multiplier effect on abilities, which suggests that early investment in education would have a positive effect on the human capital of an individual (Cunha & Heckman, 2007).

In this context, factor analysis has been presented as a possible solution to determine the returns to skill apart from returns to schooling. Vivekananthan (2015), for example, employs job descriptions to determine the broad skills that determine assignment in the labor market. A more recent approach has suggested using an interactive fixed effects model with heterogeneous coefficients. This methodology permits to account for changes over time in the returns to skill and individual-level heterogeneity. However, although returns to skill components can be identified, it is not possible to separate skills from their prices (factor loadings). Using a panel dataset to address the problems of heterogeneity in returns to schooling and the endogeneity of schooling, Kejriwal et al. (2020) find greater returns to skills in recent years and considerable within-group heterogeneity.

The work of Heckman, Stixrud, and Urzua (2006) takes the task to use individual measurements for skill to account for differences in wages. They use the National Longitudinal Survey of Youth, a panel data, to include a cognitive and a non-cognitive (personal preferences and personality traits) variable to the analysis. To do so, they occupy an endogenous factor model. A relevant argument in this study is that skills are priced differently in different schooling markets. Given that, the authors cluster their regressions for every school level attained. They find that a model that accounts for at least two skill factors can explain diverse outcomes. Particularly they find that a low dimensional vector of cognitive and non-cognitive skills can explain schooling, wages, occupational choice, and social behaviors.

2.3 Mexico

In Mexico, the literature on returns to schooling has been somewhat scarce. Up until recently, many studies have used different datasets to address specific concerns. Mehta and Villarreal (2003) use the National Survey of Household Incomes and Expenditures (ENIGH) to address sheepskin effects in the distribution of earnings. However, they find no evidence for those effects, except on graduation from primary school. On the other hand, Kaufmann (2010) uses the survey "Jovenes con Oportunidades" to examine the effect of expected returns to schooling on college enrollment. However, these approaches have not focused on determining accurate returns to schooling but on identifying particular effects.

On the estimation of returns to schooling in Mexico, Cinthya Caamal has presented ample evidence using the National Survey of Employment in Mexico (ENOE), a quarterly repeated cross-sectional data about the labor market in Mexico. Specifically, the author has examined the significant difference in returns to schooling by gender (Caamal Olvera, 2013). Further, the author uses parametric, semi-parametric, and non-parametric methods to estimate the trends in returns to schooling for the general population (Caamal Olvera, 2017). This study controls selectivity bias and uses quantile regression to diminish the ability bias without precisely measuring it. The results provide evidence that there is a declining trend in the returns to schooling in Mexico.

To include specific measures of ability, there is a more comprehensive tool to study the characteristics of the Mexican labor market: the Mexican Family Life Survey (MxFLS). The inclusion of more individual-specific data permitted a more detailed analysis of the labor market and its determinants. Morales-Ramos (2011) uses a function control method to include variables such as natural ability index, mother's education, auto-perceived health, and home infrastructure. He finds that the returns to schooling in Mexico are between 8.2% and 8.4%. However, skills are not statistically significant in this specification. The survey has permitted, as well, to focus on more detailed aspects of health. Given that, Vogl (2014) concludes that height has a significant effect on wages in Mexico. A possible reason for this is the fact that growth and cognitive improvement share childhood inputs. The intuition is consistent with the evidence that taller workers have more education and are in occupations with higher intelligence requirements.

Following an aspect of Heckman et al. (2006), Valenzuela and Moreno (2018) make the first attempt to include both health and cognitive proxies in their specification. They estimate a self-selection model between two occupational groups and then construct a wage equation. They find that skills have a significant effect on wages, and their effect on the returns to schooling bias depends on the educational level attained. Although it seems that returns to schooling are always overestimated when skills are neglected, the magnitude of the bias increases at higher educational levels. This result can be pointed as evidence that skills are also determinant of the schooling level reached. Nevertheless, they only use one variable per ability (a test score for cognitive skills and height for physical skills) instead of a more comprehensive measure. Valenzuela and Moreno (2021) explore the role of skills in occupation allocation and then the effect of education mismatch in wages. They find that cognitive abilities have a crucial weight in self-selection and that white-collar occupations penalize and reward more educational mismatch.

The present work builds on the last and uses a low-dimensional vector of skills to explain heterogeneity in returns to schooling in Mexico. Instead of using cognitive and non-cognitive abilities, cognitive and physical skills are used to capture two of the most critical investments in human capital: education and health. The analysis is conducted for different occupational categories since their requirements for schooling and abilities differ. Also, the methodology is conducted separately for both genders.

Chapter 3

Data

3.1 Database

To estimate the relevant models and test the described hypothesis, I use data from the Mexican Family Life Survey (MxFLS). This dataset is a longitudinal survey that contains information for a 10-year period, collected in three rounds: 2002, 2005-2006, and 2009-2012. The first round reached 8,400 households and about 35,00 individuals; the following rounds aimed to relocate and interview the same people, even if they migrated. The last round successfully relocated and interviewed approximately 90% of the sample (MXFLS, 2013). The MxFLS is the most adequate source for the present analysis since it collects detailed information on schooling, labor force, adult labor income, and objective health measurements. Additionally, the survey contains a set of Raven's Progressive Matrices, which measures cognitive skills.

The present work will focus on the third (2009-2012) round, since it is the most recent available. The survey is divided into community, household, and individual level datasets; variables of interest for the present analysis are listed in the latter. The sample comprehends employed individuals between the ages of 15 and 65 and consists of 5,063 observations. The data analysis is carried out separately for females (34.97%) and males (65.03%) in order to contrast the relative impact of each factor in earnings between genders. Also, expansion factors will be considered (171,626 observations) since they provide a better fit for the population distribution. Table 3.1 presents descriptive statistics of the included variables. The data shows that mean age is similar between genders, and there are about three times more observations in urban than rural areas. Respect to marital status, the proportion of working men that are married is similar to the proportion that is not, while working woman have greater not-married proportion. However, wages are not homogeneous between genders. Although the mean hourly wage is slightly higher for women, it has significantly greater variance with respect to men. In the sample, employed women have, on average, almost one year of schooling more than men.

Table 3.1: Descriptive Statistics for demographic and employment variables with expansion factors

	Female				Male			
	Frequency		Central Tendency		Frequency		Central Tendency	
	Absolute	*Relative*	*Mean*	*Std Dev*	*Absolute*	*Relative*	*Mean*	*Std Dev*
Age								
Years			35.73	11.87			35.09	12.77
Marital Status								
Married	22,558	37.57			56,307	50.46		
Other	37,486	62.43			55,275	49.54		
Number of children								
Total number			1.33	1.48			1.40	1.51
Wage								
Hourly Wage			32.41	50.17			30.37	38.05
$log(Hourly\ Wage)$			3.00	0.94			3.02	0.88
Schooling								
Years			9.69	4.43			8.83	4.04
n	60,044				111,582			

Note: The original sample consists of 5,063 observations. After considering frequency weights the expanded sample comprises 171,626 observations, which correspond to a better fit for the population distribution. Source: Own elaboration with MXFLS (2013) data .

3.2 Skills

For cognitive ability, the MxFLS provides a set of 12 progressive Raven matrices. This test is an evaluation developed to measure the two components of general intelligence (eductive and reproductive ability) identified by Spearman (1927). Apart from its acceptance as a cognitive skill measure, this test is preferred since it does not require respondent literacy. Every item consists of an incomplete sequence that must be filled with one of eight options provided. There is only one correct answer per item, and usual scores are drawn on a scale from 0 to 100, increasing with correct items. Because of its characteristics, a Raven test can be interpreted as a measure of fluid intelligence, which is desirable since this intelligence component is not significantly influenced by schooling.

To measure physical ability, the survey presents a series of anthropometric measures. The available variables correspond to height, weight, waist and hip circumference, blood pressure, hemoglobin, blood sugar, cholesterol, and dry blood samples to analyze C-reactive protein. Mexico's National Institute of Public Health (INSP) representatives collected all of the latter measures. This is relevant since self-perception of health can be subjective and self-reported variables can be biased. Furthermore, the Body Mass Index (BMI) was calculated. Due to missing values on the sample and considering long, medium, and short-term effects in health, height, BMI, waist and hip circumference, pulse, and blood pressure are used (Thomas & Strauss, 1997). Table 3.2 presents descriptive statistics of skill-related variables. The cognitive test is not greatly different between genders; though, physical measures are slightly differentiated.

Now, a crucial assumption for the analysis is the independence of the two factors constructed. In order to test the relationship between cognitive and health measurements, Table 3.3 contains the correlation coefficients from the Raven score and anthropometric variables. Correlations among the health measures are high. However, the coefficient between these scales and the raven score is close to zero (except for height), which provides evidence that cognitive and health factors constructed from these measures will be independent.

Since variables related to the same factor are correlated, a Principal Component Analysis

Table 3.2: Descriptive Statistics for skill-related variables with expansion factors

	Female		Male	
	Central Tendency		Central Tendency	
	$Mean$	$StdDev$	$Mean$	$StdDev$
Cognitive skill				
$RavenScore$	55.77	24.77	54.62	24.48
Physical skill				
$Height(cm)$	153.86	6.95	166.35	7.05
BMI	27.38	5.33	27.15	4.90
$Waist(cm)$	88.39	12.41	92.34	13.12
$Hip(cm)$	102.15	12.80	99.52	11.08
$Pulse(bpm)$	76.79	44.47	72.74	26.47
$Bloodpressure\ \ (diastole)$	76.93	11.70	79.92	14.58
n	60,044		111,582	

Source: Own elaboration with MXFLS (2013) data.

Table 3.3: Correlation matrix of skill related variables

	Correlation Coefficients						
	Raven score	Height	BMI	Waist	Hip	Pulse	Pressure
Raven score	1.00						
Height	0.15*	1.00					
BMI	-0.02*	-0.05	1.00				
Waist	-0.03*	0.21*	0.89*	1.00			
Hip	0.06*	0.02*	0.66*	0.67*	1.00		
Pulse	0.01*	-0.04*	0.10*	0.11*	0.8*	1.00	
Blood pressure	0.05*	0.06*	0.22*	0.24*	0.15*	0.04*	1.00

Note: Significance at the 1% level is represented by $*$. Source: Own elaboration with MXFLS (2013) data.

(PCA) for each variable group and gender is performed. The PCA of the 12 cognitive items reveals that a single measure does not completely comprehend the intelligence factor. According to the proportion of variance captured by each component, at least four factors can be required to capture enough data dispersion. Similarly, when the analysis is performed for the physical measures, the first four components have an eigenvalue greater than one and explain more than 80% of the variance. Table 3.4 displays these results.

Table 3.4: Factor Analysis of Raven Test items and anthropometric measurements.

	Female		Male	
	Principal Components		Principal Components	
	Eigenvalue	*Proportion*	*Eigenvalue*	*Proportion*
Cognitive skill				
1st Component	3.36	0.28	3.41	0.28
2nd Component	1.32	0.11	1.30	0.10
3rd Component	0.95	0.08	0.94	0.08
4th Component	0.91	0.08	0.88	0.07
5th Component	0.83	0.07	0.81	0.07
6th Component	0.76	0.06	0.80	0.07
7th Component	0.65	0.06	0.74	0.06
8th Component	0.66	0.05	0.71	0.06
Total		0.79		0.80
Physical skill				
1st Component	4.63	0.46	4.35	0.43
2nd Component	1.58	0.15	1.41	0.14
3rd Component	1.11	0.11	1.09	0.11
4th Component	1.01	0.10	1.03	0.10
5th Component	0.98	0.09	0.98	0.09
6th Component	0.23	0.02	0.44	0.04
7th Component	0.20	0.02	0.39	0.03
8th Component	0.06	0.01	0.17	0.01
Total		0.99		0.99
n	60,044		111,582	

Note: Only the first eight components of each Principal Component Analysis are reported. Source: Own elaboration with MXFLS (2013) data.

3.3 Occupational categories

The present work proposes that heterogeneity in returns to schooling across occupations may be driven by differences in the return to cognitive and physical skills. Occupations in the used dataset are classified according to the Mexican Classification of Occupations (CMO) which recognizes 17 categories (INEGI, 2019). Following Heckman et al. (2006), the present analysis considers the employment classifications white-collar, blue-collar, and, additionally, elemental workers. The occupations classified as white-collar comprehend administrative tasks such as technicians, professionals, heads of department, arts, entertainment, and directives. The blue-collar classification comprises manual workers as salespersons, primary sector workers, personal services, fixed machinery operators, education workers, machinery and transport drivers, administrative workers, protection and armed forces, and artisans. Finally, elemental workers include occupations usually associated with the informal sector, such as domestic services, helpers and grooms, and street sellers.

Table 3.5 shows the distribution by gender of occupations in the sample. The aggregated proportion of occupational categories is somewhat similar. However, women report almost two times more frequently to work in elemental occupations; this is mainly explained by the number of women working in domestic services. The distribution of individual occupations is also different between genders. For women, the most frequent occupations in the sample are professionals and sales, while the least reported are machinery and transport drivers and arts; for men, the most common are technicians and primary activities, and the least frequent are directives and arts.

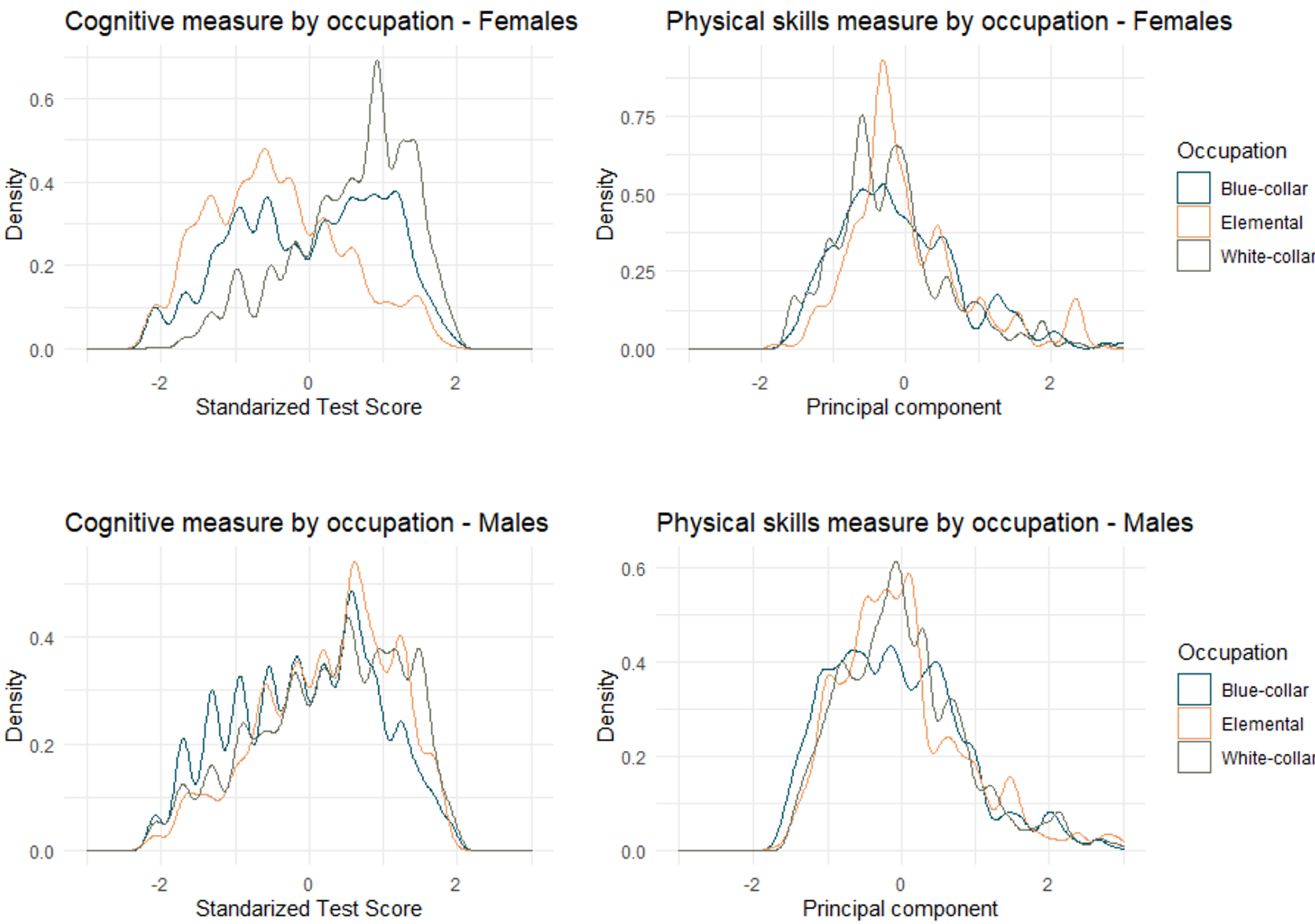
Table 3.5: Descriptive Statistics for occupations with expansion factors

	Female		Male	
	Frequency		Frequency	
	Absolute	*Relative*	*Absolute*	*Relative*
White-collar	**22,481**	**37.44**	**41,228**	**36.95**
Art, entertainment, and sports	190	0.32	439	0.39
Directives	955	1.59	100	0.09
Head of department and coordinators	1,128	1.88	1,313	1.18
Professionals	14,124	23.52	10,619	9.52
Technicians	535	0.89	24,390	21.88
Blue-collar	**27,348**	**45.55**	**60,836**	**54.52**
Administrative workers	2,980	4.96	2,764	2.48
Agriculture, cattle raising, hunting, and fishing	1,954	3.25	18,846	16.89
Artisans	392	0.65	1,069	0.96
Education workers	5,549	9.24	4,367	3.91
Fixed machinery operators	4,230	7.04	9,817	8.80
Machinery and transport drivers	133	0.22	7,326	6.57
Salespersons, commerce employees	10,015	16.68	10,051	9.01
Personal services	6,992	11.64	6,765	6.06
Protection, monitoring, and armed forces	652	1.09	4,198	3.76
Elemental	**10,215**	**17.01**	**9,518**	**8.53**
Domestic services	9,259	15.42	1,783	1.60
Helpers and grooms	617	1.03	6,160	5.52
Street sellers	339	0.56	1,575	1.41

Source: Own elaboration with MXFLS (2013) data.

Figure 3.1 illustrates the distribution of cognitive and physical measurements for each occupational category. There is evident heterogeneity in the case of cognitive skills, especially for women. For both genders, white-collar occupations present a distribution skewed to the right, while elemental occupations are skewed to the left just for women. Physical skills are less heterogeneous, but they do not follow an identical distribution as proven by pairwise t-statistics.

Figure 3.1: Distribution of skills measurements by gender and occupation.



Note: The cognitive measure corresponds to the standarized Raven test score. The physical measure presented is the first principal component of anthropometric measurements. Source: Own elaboration with MXFLS (2013) data.

Chapter 4

Empirical Strategy

The hypothesis of the present analysis states that cognitive and physical skills can affect wages due to unobserved heterogeneity. Following the model proposed by Carneiro, Hansen, and Heckman (2003), the empirical strategy proposed in this work is a factor model. To construct the specification, part from the linear system to estimate wages:

$$\boldsymbol{W} = \boldsymbol{X_W}\beta^W + \boldsymbol{U^W} \tag{1}$$

Where $\boldsymbol{W}$ is a $S \times 1$ vector of wages, $\boldsymbol{X_W}$ is a matrix of observable variables, and $\boldsymbol{U^W}$ is a vector of unobservables with the structure $\boldsymbol{U^W} = \boldsymbol{\Lambda^W}\boldsymbol{\theta} + \boldsymbol{e^W}$, so Equation 1 can be rewritten as:

$$\boldsymbol{W} = \boldsymbol{X_W}\beta^W + \boldsymbol{\Lambda^W}\boldsymbol{\theta} + \boldsymbol{e^W} \tag{2}$$

where $\boldsymbol{\Lambda^W}$ is a $S \times q$ matrix of factor loadings for each type of unobserved heterogeinity (q latent factors), $\boldsymbol{\theta}$ is a vector of unobserved factors (cognitive and physical skills in this specification), and $\boldsymbol{e^W}$ is a vector of error terms with distributions $f_{e^y s}(\cdot)$ for every $s = 1, ..., S$. Additionally, $\boldsymbol{\theta}$ has a distribution $f_\theta(\cdot)$.

For the present specification, the explanatory variables in Equation 2 correspond to those proposed by Mincer: schooling, experience, and squared experience. Besides that, some relevant controls are included, such as marital status and number of children. In the context of my two-factor model, Equation 2 becomes:

$$\boldsymbol{W} = \beta_0^W + \beta_1^W \boldsymbol{S} + \beta_2^W \boldsymbol{Exp} + \beta_3^W \boldsymbol{Exp^2} + \alpha^{W,C}\boldsymbol{\theta^C} + \alpha^{W,P}\boldsymbol{\theta^P} + \boldsymbol{e^W} \tag{3}$$

Nevertheless, the actual values of skills are not observed. In order to identify these factors, it is needed a measurement system:

$$\boldsymbol{M} = \boldsymbol{X_M}\beta^M + \boldsymbol{\Lambda^M}\boldsymbol{\theta} + \boldsymbol{e^M} \tag{5}$$

where $\boldsymbol{M}$ is a $L \times 1$ vector of measurements, $\boldsymbol{X_M}$ is a matrix of observable controls for each measurement, and $\boldsymbol{\Lambda^M}$ is a $L \times q$ matrix of factor loadings. $\boldsymbol{e^W}$ is a vector of error terms that have associated distributions $f_{e^l}(\cdot)$ for every $l = 1, ..., L$. For both equations, it is required the assumption that all of the elements in the vector e are independent. In the context of my two-factor model, Equation 4 becomes:

$$\boldsymbol{M} = \boldsymbol{X_M}\beta^M + \alpha^{M,C}\boldsymbol{\theta^C} + \alpha^{M,P}\boldsymbol{\theta^P} + \boldsymbol{e^M} \tag{6}$$

where θ^C is the factor associated to cognitive skills, θ^P is the factor associated to physical skills, and α is the loading associated to each factor.

Initially, the model is underidentified. For this model to be correctly specified, there are two conditions: cognitive skills must be independent from physical skills ($\theta^C \perp \theta^P$), and there is a minimum number of measurements for each factor. Carneiro et al. (2003) consider the elements of the diagonal on the matrix $COV(\boldsymbol{M}|\boldsymbol{X_M})$. For two factors, the rule is:

$$\frac{L(L-1)}{2(L+1} \geq 2$$

so, at least six measurements ($L \geq 6$) are needed to identify a model with two factors. The measurements chosen to identify the cognitive factor correspond to the first three principal components obtained from Raven's progressive matrix test. The measurements used to build the physical factor include the first three principal components created from anthropometric measures. Because the Body Mass Index is a non-monotonic measure, it has been redefined as the squared distance with respect to the adequate weight measure. All these measures were standardized. The explanatory variables that can affect these measures independently from the factors are age and squared age. Thus, the resulting equation is:

$$\boldsymbol{M} = \beta_0^M + \beta_1^M \boldsymbol{Age} + \beta_2^M \boldsymbol{Age^2} + \alpha^{M,C}\theta^C + \alpha^{M,P}\theta^P + \boldsymbol{e^M} \tag{6}$$

Using this measurement system, loadings, factor variances, and measurement residuals can be identified. Then, factors distributions can be estimated non-parametrically using the Kotlarski Theorem. The estimation can be carried out using maximum likelihood that will estimate $\hat{\beta^M}$, $\alpha^{M,C}$, $\alpha^{M,P}$, $\hat{F_{\theta^C}}(\cdot)$ and $\hat{F_{\theta^P}}(\cdot)$.

Once the distribution from which the skills are drawn from has been estimated, the factors are included in the wage equation. This can be computed using another Maximum Likelihood Estimation, identifying $\hat{\beta^W}$, $\alpha^{W,C}$, $\alpha^{W,P}$. Then, the effect of observable variables and unobservable factors in wages can be computed.

The previous model was later developed by Hansen and Mullen (2003). The extension of the model permits evaluating the effect of some treatment in the outcome of interest. On this specification, the treatment can also depend on observable factors and the estimated unobservable ones. The authors highlight that the advantage of using a factor structure is that potential outcomes are separable since selection in latent unobservable variables is already accounted for. In this sense, counterfactuals can be simulated.

In order to do so, the extension of the model considers a decision model inspired by the Roy model (Roy, 1951).

$$D \begin{cases} 1 & if \;\; \boldsymbol{X_D}\beta^{W_D} + \alpha^{W_D,C}\theta^C + \alpha^{W_D,P}\theta^P + e^D > 0 \\ 0 & otherwise \end{cases} \tag{7}$$

This is the most adequate specification for the present work since the occupation is a choice variable ($D = 1$ or $D = 0$), and an individual set of skills can influence it. The relevant regressors for this equation are schooling, experience, and a dummy indicating if the person resides in a rural municipality. Equation 7 can be represented as:

$$D \begin{cases} 1 \quad if \;\; \beta_0^D + \beta_1^D \boldsymbol{S} + \beta_2^D \boldsymbol{Exp} + \beta_3^D \boldsymbol{Rural} + \alpha^{W_D,C}\theta^C + \alpha^{W_D,P}\theta^P + e^D > 0 \\ 0 \quad otherwise \end{cases} \tag{8}$$

In this sense, the empirical strategy pursued in this analysis is a factor model with endogenous factor loadings estimated in two steps. First, cognitive and physical factors are non-parametrically estimated from the measurement system provided. In a second step, a wage equation corrected by occupational choice is computed, taking into account the effect of skills on both specifications.

Chapter 5

Results

The present section portrays the results obtained from the empirical strategy proposed earlier. First, a classical Mincerian equation is performed. Second, factors are constructed according to the measurement system proposed earlier. Third, factors are included in the wage equation for the general population. Finally, the extension of the factor model is implemented, where the specification includes a self-selection model, comprehending three occupational categories.

5.1 Mincerian estimation

Table 5.1 presents the results obtained for both females and males using a standard Mincerian equation (columns 1 and 3). For both genders, coefficients behave as expected. Schooling and experience have a positive effect on wages; meanwhile, squared experience does the opposite. Returns to schooling are higher than the ones found by previous literature (Caamal Olvera, 2017; Vogl, 2014; Morales-Ramos, 2011), which suggests that there may be an upward bias from letting out skills. The returns to schooling are greater for women than men, which is consistent with the trend for most countries (Montenegro & Patrinos, 2014) as well as Mexico (Caamal Olvera, 2013).

In columns 2 and 4, dummies for white-collar and blue-collar occupations are added. Occupational categories are statistically significant. In this specification, returns to schooling de-

Table 5.1: Wage equation by gender using a Mincerian Equation.

	Female		Male	
	(1)	(2)	(3)	(4)
$Schooling$	0.1248***	0.0861***	0.1012***	0.0913***
	(0.0021)	(0.0024)	(.0014)	(0.0015)
$Experience$	0.0191***	0.0185***	0.0225***	0.0197***
	(0.0193)	(0.0018)	(0.0014)	(0.0014)
$Experience^2$	-0.00007*	-0.00013***	-0.00024***	-0.00020***
	(0.00003)	(.00003)	(.00002)	(0.00002)
Occupation				
$White-collar$		0.3855***		0.2048***
		(0.0252)		(0.0199)
$Blue-collar$		-0.1790***		-0.0952***
		(.0211)		(0.0191)

Note: Columns (1) and (3) use a classic Mincerian equation. Columns (2) and (4) include occupational category dummies, elemental occupations are the baseline. Both specifications include individual controls. Standard errors are reported in parentheses. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Source: Own elaboration with MXFLS (2013) data.

crease, while the other coefficients remain stable. This can be considered evidence that the marginal effect of one year more of schooling is not correctly isolated with a classical Mincerian specification. For both genders, being employed in a white-collar occupation increases wages substantially in comparison to elemental workers. For men, the marginal effect of working as a white-collar in wages is 20%, while for women, the effect almost doubles. On the other hand, the estimated effect of being employed in a blue-collar occupation is negative. Again, the magnitude of the effect is almost double for women. This is consistent with the fact that variance in wages for women is greater than for men. These results present evidence that occupation has a significant effect on wages and that wages associated with the informal labor market may be greater than the non-qualified formal labor market.

Now, abilities must be added in order to estimate the returns to skill. However, figure 3.1 provided evidence that simply including occupational category in the factor model may cause endogeneity problems since abilities are differently distributed. Instead of including occupation as a regressor for the wage equation, the following model will be performed for each occupation and gender, which will permit identifying returns to schooling and skills in each occupation as a separate market. Nevertheless, since the occupation is a choice variable, the empirical strategy needs to consider that occupation is not exogenous but that individuals may select themselves into an occupation according to their skills.

5.2 Factor estimation

The first stage of the model corresponds to the construction of cognitive and physical latent variables. As introduced earlier, the non-parametrical estimation of factors uses a measurement system. The measurements included in this specification involve the first three principal components from Raven test score items for the cognitive factor and form the anthropometric measures for the physical factor. Tables 5.2 and 5.3 present the results from estimating equation 6 for women and men, respectively. As latent factors do not have a specific metric, one loading per factor is normalized. The rest of the loadings may be interpreted relative to the one used as numeraire.

Cognitive and physical factors are calculated separately for each gender since factor estimation relies on the position of a measure relative to the distribution. In this regard, anthropometric measures distribute differently between genders, and a single estimation with all population will not achieve convergence. This is consistent with the results presented. The factor loadings associated with the cognitive factor are more homogeneous between genders than the physical factor loadings. However, in order to have consistency in the direction of the effects, the same principal component is used as a numeraire. Age and squared age are appropriate controls for the cognitive measure, since it is a fluid intelligence measure. This is consistent with the findings of

Table 5.2: Non-parametric factor estimation for women

	Cognitive			Physical		
	$1stC$	$2ndC$	$3rdC$	$1stC$	$2ndC$	$3rdC$
Age	0.0050	-0.0279***	-0.0129***	0.0711***	-0.0240***	-0.0525***
	(0.0042)	(0.0023)	(0.0000)	(0.0029)	(0.0038)	(0.0036)
Age^2	-0.0003***	-0.0279***	0.0002***	-0.0006***	0.0006***	0.0007***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Loading factors						
θ^C	0.1376	1	0.0534***			
	(0.0099)		(0.0106)			
θ^P				1	-0.3334***	-0.1786***
					(0.0191)	(0.0183)

Note: Standard errors are reported in parentheses. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Source: Own elaboration with MXFLS (2013) data.

Table 5.3: Non-parametric factor estimation for men

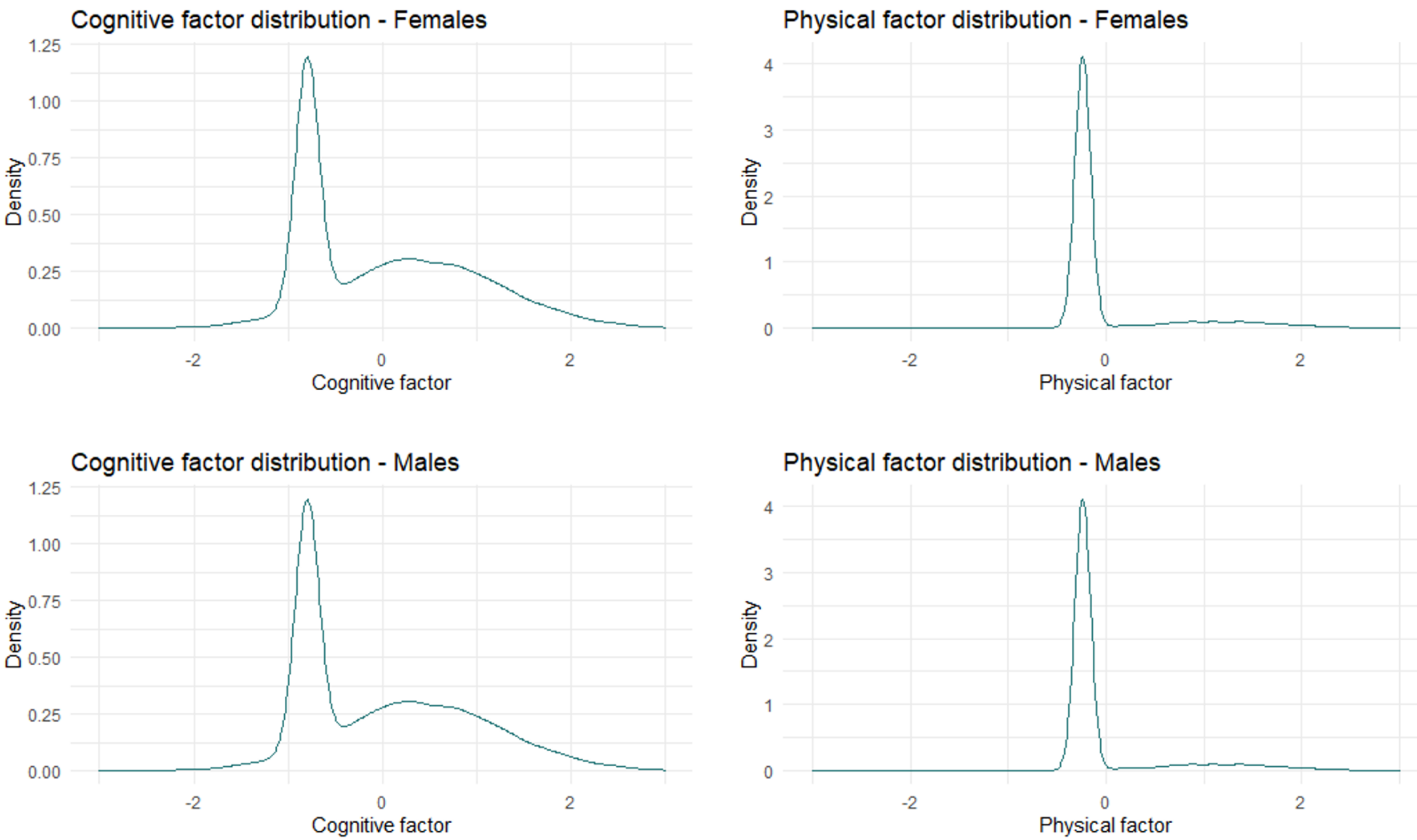
	Cognitive			Physical		
	$1stC$	$2ndC$	$3rdC$	$1stC$	$2ndC$	$3rdC$
Age	0.0173***	-0.0064***	-0.0002	0.1070***	-0.0118***	-0.0342***
	(0.0029)	(0.0025)	(0.0000)	(0.0027)	(0.0027)	(0.0020)
Age^2	-0.0004***	0.0001***	0.0000	-0.0011***	0.00007**	0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Loading factors						
θ^C	0.5806***	1	0.0251***			
	(0.0079)		(0.0083)			
θ^P				1	-1.3388***	2.5878***
					(0.0332)	(0.0538)

Note: Standard errors are reported in parentheses. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Source: Own elaboration with MXFLS (2013) data.

Carlsson et al. (2015). These controls are also statistically significant for health measures.

Finally, the measurement system is statistically relevant for the construction of factors. It is relevant to point out that the methodology does not uses the measurement system directly to construct the factor, as an index would do. In an opposite way, the strategy evaluates the impact of the estimated factors in the results presented in the observable measures. The fact that the coefficients are statistically significant suggests that a relevant set of measures has been chosen. Figure 5.1 illustrates the distribution of cognitive and physical skills for each gender, using the measurement system mentioned above.

Figure 5.1: Distribution of cognitive and physical factors by gender.



Source: Own elaboration with MXFLS (2013) data.

5.3 Factor model

The second stage of the factor model consists of the wage equation estimation. The specification on this factor model corresponds to the first one described in the empirical strategy section, without a decision model. Table 5.4 presents the results obtained for both females and males using the proposed factor model (columns 2 and 4), the standard Mincerian equation from Table 5.1 is provided as reference (columns 1 and 3).

The returns to schooling using the proposed model decrease compared to the standard model, as was expected, even though the magnitude of the change is small. The returns to schooling for women diminish from 12.48% to 12.16% and for men from 10.12% to 9.95%. This means that even after accounting for skills, returns to schooling are greater for women than men. Several studies have consistently found that the impact of schooling on wages is greater for females than males, even when females tend to earn less. Dougherty (2005) argues that education has a double effect on female wages. On the one hand, it increases productivity in the same fashion than males. On the other hand, additional years of schooling appears to reduce the gender gap in earnings attributable to factors such as discrimination, tastes, and circumstances.

Regarding the ability variables, their inclusion in the model results in statistically significant positive estimates. This can be interpreted as evidence that skills are, in fact, valued and rewarded by labor markets. As it was stated before, factors do not have a scale. As a consequence, the magnitude of the effect has no cardinal interpretation. However, factors can be compared between themselves. This analysis highlights the fact that cognitive skills are more rewarded than physical ones for women. The opposite is true for men; health-related variables have a greater effect on wages than intelligence-related ones.

5.4 Occupational categories

Finally, the extension of the factor model is estimated to retrieve the returns to schooling and skills for different occupational categories and the role of such variables in self-selection. Tables

Table 5.4: Wage equation by gender using a Mincerian Equation and factor model.

	Female		Male	
	(1)	(2)	(3)	(4)
Schooling	0.1248***	0.1216***	0.1012***	0.0995***
	(0.0021)	(0.0021)	(.0014)	(0.0015)
Experience	0.0191***	0.0194***	0.0225***	0.0225***
	(0.0193)	(0.0018)	(0.0014)	(0.0014)
$Experience^2$	-0.00007*	-0.00008**	-0.00024***	-0.00025***
	(0.00003)	(.00003)	(.00002)	(0.00002)
Factors				
Cognitive		0.0646***		0.0498***
		(0.0085)		(0.0063)
Physical		0.0416***		0.0723***
		(.0144)		(0.0178)

Note: Columns (1) and (3) use a classic Mincerian equation. Columns (2) and (4) include cognitive and physical factors. Both specifications include individual controls. Standard errors are reported in parentheses. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Source: Own elaboration with MXFLS (2013) data.

5.5 and 5.6 present the estimation for white-collar, blue-collar, and elemental workers. The estimations for the general population from Table 5.4 are kept as reference (column 1). Columns 2-4 presents the estimation for each category. The top panel documents the effect of the variables in the logarithm of wages. The second panel depicts the effect of the variables in an individual's choice to be part of a given group.

Table 5.5 presents the estimation for women. In the general model, the return to schooling for women was approximately 12%. The same is true for white-collar occupations. Nevertheless, returns decrease substantially for blue-collar and elemental occupations, where the coefficient is around 6%. This can be conflicting with Dougherty (2005) and the hypothesis that schooling has a double positive effect in female earnings. In this case, the analysis of Hannum, Zhang, and Wang (2013) may provide a better fitted explanation. They argue that in intra-household decisions women often trade their own income for their spouse income, specially less educated women. Thus, creating a composition effect, where the labor market has over-representation of educated women working in occupations where schooling is more valued. This explanation could be consistent with 3.1 where working women have approximately a year more of education and are mainly not married. Also agreeable, 3.5 presents that the most common occupation for women is professional, with almost a quarter of working women in this category.

Regarding returns to skill, there is a differentiated effect by occupation. For white-collar occupations, greater cognitive skills increase wages, while physical skills do not have a statistically significant effect. In blue-collar occupations, both skills have a significant positive effect, while neither of them seems to affect wages for elemental workers. In this sense, it seems that, for women, white-collar occupations reward credentials, while blue-collar ones rely more on ability. This provides evidence that, for women, returns to schooling and skills vary markedly across occupations.

In terms of self-selection, there are various effects to be noted. First, more years of schooling positively affect the probability of occupying an administrative type of work. By contrast, it has a negative effect on being employed in manual or informality-related occupations. So, individuals with more education self-select themselves in occupations where returns to schooling are greater. Second, being in a rural community also has a significant effect on occupational choice. Notably, living in a rural area increases the chance of laboring in a blue-collar occupation, while it has the inverse effect for the other two types.

Third, the significance of cognitive and physical skills seems to have an opposite pattern. On the one hand, cognitive skills appear not significantly to affect self-selection for white and blue-collar occupations. On the other hand, physical skills are more statistically significant for blue and white-collar occupations. In general, having greater abilities decrease the probability of being an elemental worker. This is also consistent with the view that people with specific abilities select themselves where returns to those skills are high.

Table 5.6 presents the estimation for men. Returns to schooling are positive and significant for all occupations. Similar to women, the returns decrease for blue-collar and elemental workers with respect to white-collar. Nevertheless, the magnitude of the change is smaller. As well, experience and squared experience are significant and have the expected direction. Returns to skills are also differentiated for each occupational category. For administrative workers, both skills are significant and positively rewarded in terms of wages. Also, cognitive skills have a slightly greater effect. In manual and elemental occupations, only cognitive skills are statistically important. These abilities have a negative effect on blue-collar occupations, but a positive one in elemental.

Concerning self-selection, years of schooling and rural background have the same effects in men as in women. Skills, in contrast, behave differently between genders. For men, having more skills, independently if they are cognitive or physical, results in a higher probability of being employed in a white-collar job. Both skills negatively impact the possibility of being a blue-collar worker. Finally, skills are statistically insignificant for self-selection in elemental occupations. As well as for women, the self-selection coefficients imply that workers self-select themselves in the occupational category that rewards their specific skills.

Finally, comparing genders, it is possible to appreciate several differences. The model es-

Table 5.5: Factor model wage equation by occupational classification for women

	General (1)	*White Collar* (2)	*Blue Collar* (3)	*Elemental* (4)
Returns				
Schooling	0.1216***	0.1209***	0.0625***	0.0564***
	(0.0021)	(0.0032)	(0.0042)	(0.0059)
Experience	0.0194***	0.0217***	0.0061**	0.0317***
	(0.0018)	(0.0027)	(0.0031)	(0.0042)
$Experience^2$	-0.00008**	-0.00004	-0.00005	-0.00037***
	(0.00003)	(0.00006)	(0.00006)	(0.00006)
Cognitive	0.0646***	0.0367***	0.1158***	-0.0108
	(0.0085)	(0.0107)	(0.0138)	(0.0207)
Physical	0.0416***	-0.0230	0.1339***	-0.0378
	(0.0144)	(0.0212)	(0.0206)	(0.03538)
Self-selection				
Schooling		0.2631***	-0.1205***	-0.1328***
		(0.0047)	(0.0035)	(0.0045)
Experience		0.0169***	-0.0198***	0.0034***
		(0.0012)	(0.0010)	(0.0012)
Rural		-0.11633**	0.2166***	-0.2319***
		(0.0411)	(0.0341)	(0.0421)
Cognitive		-0.0381	0.0207	-0.0960***
		(0.01605)	(0.0138)	(0.0179)
Physical		-0.1751***	0.1470***	-0.0491*
		(0.0280)	(0.0232)	(0.0288)

Note: Standard errors are reported in parentheses. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Source: Own elaboration with MXFLS (2013) data.

timated that returns to schooling were greater for women than for men for the general case. However, when the market is disaggregated, the relationship is not maintained. Women have greater returns to schooling for white-collar occupations, but for the other classifications, this is not the case. This is consistent with the hypothesis that women are over-represented in occupations were education is more valued. This is not only appreciated in greater returns to schooling, but also in the magnitude of the self-selection coefficients associated with schooling. These have greater magnitude for women, meaning that schooling has more impact in their labor market participation decisions.

In terms of returns to skills, since they were estimated separately, they cannot be compared by their absolute values, but relative to each other. For both genders cognitive abilities are more rewarded than physical skills in white collar occupations. In blue collar occupations, by contrast, physical skills have a greater coefficient than cognitive skills. In informality-related occupations, skills are less significant, with only the cognitive skills for men being significant. However, there are some relevant discrepancies. For example, for women, cognitive skills are rewarded the most in manual categories, while for men they seem to have a negative impact on wages for this category. This can also be explained by the composition of labor market. While most women in blue-collar occupations are salespersons and education workers, men are concentrated in primary activities and fixed machinery operators.

Comparing self-selection, skills have an opposite effect than schooling in terms of gender. While schooling is the driving force for market labor participation decisions for women, it seems that skills have a stronger effect on men. For them, an increase in abilities has a significant effect in diminishing the probability of working a blue-collar occupation and increases to be in a white-collar one.

Table 5.6: Factor model wage equation by occupational classification for men

	General	*White Collar*	*Blue Collar*	*Elemental*
	(1)	(2)	(3)	(4)
Returns				
Schooling	0.0995***	0.0982***	0.0868***	0.0774***
	(0.0015)	(0.0021)	(0.0021)	(0.0067)
Experience	0.0225***	0.0116***	0.0183***	0.0444***
	(0.0014)	(0.0026)	(0.0018)	(0.0044)
$Experience^2$	-0.00025***	0.00012**	-0.00019***	-0.00109***
	(0.00002)	(0.00005)	(0.00003)	(0.00008)
Cognitive	0.04989***	0.1285***	-0.0433***	0.0673***
		(0.0063)	(0.0083)	(0.0204)
Physical	0.0723***	0.0927***	0.0083	-0.0661
		(0.0178)	(0.0351)	(0.0501)
Self-selection				
Schooling		0.0784***	-0.0649***	-0.0315***
		(0.0025)	(0.0024)	(0.0036)
Experience		0.0064***	-0.0027***	-0.0092***
		(0.0007)	(0.0007)	(0.0009)
Rural		-0.1582***	0.2842***	-0.3777***
		(0.0214)	(0.0208)	(0.0321)
Cognitive		0.11472***	-0.1014***	-0.0226
		(0.1147)	(0.0103)	(0.0148)
Physical		0.1331***	-0.1639***	0.0459
		(0.0297)	(0.0321)	(0.0367)

Note: Standard errors are reported in parentheses. Significance at the 10% level is represented by $*$, at the 5% level by $**$, and at the 1% level by $***$. Source: Own elaboration with MXFLS (2013) data.

Chapter 6

Conclusion

Labor markets are evolving. Abilities are becoming more relevant in occupations as different as executives, physicians and drivers (Vivekananthan, 2015). The objective of the current work was to understand the effect of schooling and skills on wages for different labor markets in Mexico. To do so, I use an endogenous factor model with self-selection, using the Mexican Family Life Survey. The results present evidence that cognitive and physical abilities are determinants for economic outcomes, as suggested in early works by Schultz (1961) and G. Becker (1975). Nevertheless, these returns are differentiated across occupations and between genders. The skills possessed by individuals also play a significant role in their self-selection in an occupational category.

Returns to schooling and skills are significantly different in each labor market. For both genders, returns to schooling are greater in administrative occupations, while elemental ones have the smallest gain for an additional year of education. However, women in white-collar occupations have greater returns than men, and the opposite is true for blue-collar and elemental. My results coincide with previous literature, suggesting that there is a composition effect, where women are over-represented in occupations that value more education.

Cognitive and physical skills have mostly a positive effect on wages. Though, intelligence-related measures are more often significant than health-related. For women, cognitive skills are more rewarded in blue-collar occupations and for men in white-collar ones. In general, workers with greater abilities are less likely to be employed in occupations frequently associated with informality. While schooling has the greatest effect on labor market participation decisions for women, it seems that skills are determinant for occupation selection for men.

This thesis highlights the significant heterogeneity in labor markets. This discrepancies in returns to schooling and skills is evidence that a comprehensive analysis cannot condense the labor market in Mexico as a single homogeneous entity. Taking into account these differences can play a crucial role in understanding the wage distribution. A better comprehension of the labor market is vital for designing effective policies. For example, initiatives that permit less-educated women to access the labor market such as expand the offer of care services. Also, since skills are valued in different labor markets, study plans could incorporate a skill based approach in some areas.

Future research may not only replicate this analysis but also depict the evolution of returns to schooling and skill. In this sense, a more recent survey that measures abilities is needed to study the current state of the labor market.

References

Ashworth, J., Hotz, V. J., Maurel, A., & Ransom, T. (2021). Changes across cohorts in wage returns to schooling and early work experiences. *Journal of Labor Economics*, *39*(2). https://doi.org/10.1086/711851

Becker, G. (1975). *Human capital: A theoretical and empirical analysis, with special reference to education, second edition*. National Bureau of Economic Research, Inc.

Becker, G. (1993). *Human capital: A theoretical and empirical analysis, with special reference to education*. The University of Chicago Press.

Becker, G. S., & Tomes, N. (1986). Human capital and the rise and fall of families. *Journal of Labor Economics*, *4*(3), S1–S39. http://www.jstor.org/stable/2534952

Caamal Olvera, C. G. (2013, 07). Gender differences in the returns to education in mexico. *London Business Research Conference*. http://eprints.uanl.mx/7530/

Caamal Olvera, C. G. (2017, 06). Decreasing returns to schooling in mexico. *Estudios Económicos*, *32*, 27 - 63. https://doi.org/10.24201/ee.v32i1.2

Carlsson, M., Dahl, G., Öckert, B., & Rooth, D.-O. (2015). The effect of schooling on cognitive skills. *The Review of Economics and Statistics*, *97*(3), 533-547. http://www.nber.org/papers/w18484

Carneiro, P., Hansen, K. T., & Heckman, J. J. (2003, March). *Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college* (Working Paper No. 9546). National Bureau of Economic Research. http://www.nber.org/papers/w9546 doi: 10.3386/w9546

Cattell, R. B. (1986). *Intelligence: Its structure, growth and action*. Academic Press, Elsevier.

Chamberlain, G. (1977). Education, income, and ability revisited. *Journal of Econometrics*, *5*(2), 241-257. https://doi.org/10.1016/0304-4076(77)90027-6

Cunha, F., & Heckman, J. (2007). The technology of skill formation. *The American Economic Review*, *97*(2), 31–47. http://www.jstor.org/stable/30034418

Cunha, F., Heckman, J. J., & Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, *78*(3), 883–931. http://www.jstor.org/stable/40664515

Deaton, A. (2007). Height, health, and development. *Proceedings of the National Academy of Sciences*, *104*(33), 13232–13237. doi: 10.1073/pnas.0611500104

Dougherty, C. (2005). Why are the returns to schooling higher for women than for men? *The Journal of Human Resources*, *40*(4), 969–988. http://www.jstor.org/stable/4129547

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, *45*(1), 1–22. http://www.jstor.org/stable/1913285

Hannum, E., Zhang, Y., & Wang, M. (2013). Why are returns to education higher for women than for men in urban china? *The China Quarterly*(215), 616–640. http://www.jstor.org/stable/23510805

Hansen, K., & Mullen, K. (2003, 06). The effect of schooling and ability on achievement test scores. *Journal of Econometrics*, *121*, 39-98. doi: 10.1016/j.jeconom.2003.10.011

Hause, J. C. (1972). Earnings profile: Ability and schooling. *Journal of Political Economy*, *80*(3), S108–S138. http://www.jstor.org/stable/1831254

Heckman, J. J., & Hotz, V. J. (1986). An investigation of the labor market earnings of panamanian males evaluating the sources of inequality. *The Journal of Human Resources*, *21*(4), 507–542. http://www.jstor.org/stable/145765

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, *24*(3), 411-482. https://doi.org/10.1086/504455

INEGI. (2019). *Sistema nacional de clasificación de ocupaciones* (Tech. Rep.). Instituto Nacional de Estadística Geografía e Informática. www.snieg.mx/DocumentacionPortal/Normatividad/vigente/sinco_2019.pdf

INEGI. (2020). *Empleo y ocupación.* Instituto Nacional de Estadística y Geografía. INEGI. https://www.inegi.org.mx/temas/empleo/

INEGI. (2021, marzo). *Xii censo general de población y vivienda 2000.* Instituto Nacional de Estadística y Geografía (INEGI). https://www.inegi.org.mx/programas/ccpv/2020/

Ingram, B. F., & Neumann, G. R. (2006). The returns to skill. *Labour Economics*, *13*(1), 35-59. https://doi.org/10.1016/j.labeco.2004.04.005

Kaufmann, K. (2010, 01). *Understanding the income gradient in college attendance in mexico: The role of heterogeneity in expected returns* (Vol. 5; Working Paper). Innocenzo Gasparini Institute for Economic Research (IGIER) , Bocconi University. doi: 10.3982/QE259

Kejriwal, M., Li, X., & Totty, E. (2020). Multidimensional skills and the returns to schooling: Evidence from an interactive fixed-effects approach and a linked survey-administrative data set. *Journal of Applied Econometrics*, *35*(5), 548-566. https://doi.org/10.1002/jae.2759

Lam, D., & Schoeni, R. (1993, 08). Effects of family background on earnings and returns to schooling: Evidence from brazil. *Journal of Political Economy*, *101*, 710-40. doi: 10.1086/261894

Maccini, S., & Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *The American Economic Review*, *99*(3), 1006–1026. http://www.jstor.org/stable/25592491

Martorell, R. (2017, 01). Improved nutrition in the first 1000 days and adult human capital and health. *American Journal of Human Biology*, *29*, e22952. doi: 10.1002/ajhb.22952

Mehta, A., & Villarreal, H. J. (2003). *Returns to schooling, institutions and heterogeneous diploma effects: An expanded mincerian framework applied to mexico* (Tech. Rep. No. 1800-2016-142353). https://doi.org/10.22004/ag.econ.12608

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

Montenegro, C. E., & Patrinos, H. (2014, 9). *Comparable estimates of returns to schooling around the world* (Vol. 7020; Working Paper). World Bank Group.

Morales-Ramos, E. (2011). *Los rendimientos de la educación en méxico* (Working Paper No. 2011-07). México: Banco de México.

MXFLS. (2013). *Mexican family life survey.* Mexican Family Life Survey. http://www.ennvih-mxfls.org/english/introduccion.html

Psacharopoulos, G. (1972). Rates of return to investment in education around the world. *Comparative Education Review*, *16*(1), 54–67. http://www.jstor.org/stable/1186342

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, *3*(2), 135–146. http://www.jstor.org/stable/2662082

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, *51*(1), 1–17. http://www.jstor.org/stable/1818907

Spearman, C. (1927). *The abilities of man*. Macmillan.

Thomas, D., & Strauss, J. (1997). Health and wages: Evidence on men and women in urban brazil. *Journal of Econometrics*, *77*(1), 159-185. https://doi.org/10.1016/S0304-4076(96)01811-8

Valenzuela, N., & Moreno, J. O. (2018, 12). Asignación y retorno de habilidades en el mercado laboral en méxico. *Revista de Economía Laboral*, *15*, 1-33. doi: 10.21114/rel.2018.01.01

Valenzuela, N., & Moreno, J. O. (2021, 09). Asignación y desajuste en el mercado laboral de méxico: El rol de las habilidades bajo sesgos de selección múltiples en muestras complejas. *Revista de Economía Laboral*, *15*, 1-33. doi: 10.21114/rel.2018.01.01

Vivekananthan, S. (2015). *Returns to skill: Beyond returns to schooling* (Unpublished master's thesis). Queen's University, Kingston, Ontario, Canada.

Vogl, T. S. (2014). Height, skills, and labor market outcomes in mexico. *Journal of Development Economics*, *107*(C), 84-96. http://www.nber.org/papers/w18318

Wolfe, B. L., & Behrman, J. R. (1987). Women's schooling and children's health: Are the effects robust with adult sibling control for the women's childhood background? *Journal of Health Economics*, *6*(3), 239-254. https://doi.org/10.1016/0167-6296(87)90011-7