RISAD
REVISTA INTERNACIONAL DE SALARIOS DIGNOS

The Relationship Between Education and Gender on Employability with Welfare Programs in Mexico: A Logit Analysis

David Cambron Jiménez[1]

Received: 01/09/2025
Accepted: 19/11/2025

Abstract

This paper analyzes the relationship that exists between level of education, gender, and employability within the framework of the social welfare program *Jóvenes Construyendo el Futuro (JCF)*, part of Mexico's *Becas del Bienestar*. A logit model was applied using date from the *Centro de Integración Juvenil (CIJ)* to estimate the probability of obtaining employment before the maximum twelve-month duration of the scholarship. The results highlight notable differences by sex and educational level: women, regardless of their education attainment, show negative slopes in their probability of employment; men with higher education also display a negative effect, though smaller in magnitude than the women, nevertheless, men with high school education obtain a positive effect. Overall, the study shows that gender and education are statistically significant variables, while the scholarship itself is marginally significant in the model. The global model is validated as significant against the null model, although the pseudo R^2shows low explanatory power in the model. The most relevant finding suggests that JCF scholarships fail to counteract structural inequalities in gender and education within labor market insertion, with women holding higher education degrees being the most disadvantaged by the scholarship, showing the lowest probability of employment in the model.

Key Words
Scholarships, Unemployement, Gender

Introduction

In Mexico, young people who are neither studying nor working pose a structural challenge to the country's economic and social development. The *Jóvenes Construyendo el Futuro* (JCF) scholarships are a Mexican public policy designed to mitigate this phenomenon, offering a monthly scholarship of 8,470.17 Mexican pesos in 2025 and training in workplaces to young people between the ages of 18 and 29 (Gobierno de México, 2025).

The program's overall objectives are: "to integrate young people into on-the-job training activities and thus equip them with the tools for a better life" and "to steer young people away from unemployment and antisocial behavior." (Secretaría del Trabajo y Previsión Social, 2023), which translates into a program that seeks to train young people for the labor market and thereby reduce unemployment.

Mexican government data shows that 6 out of 10 young people who participate in the program can join a productive activity or find employment (Gobierno de México, 2024).

[1] Graduate of the Bachelor's Degree in Economic and Financial Engineering, La Salle University, Mexico City

Five years after its initial operation in 2019, by the end of 2024, the investment had already reached 115 billion Mexican pesos (Gobierno de México, 2024); benefiting 3,311,387 people, including 1,382,284 men and 1,929,096 women (Gobierno de México, 2025). This scholarship also presents a social and ethical challenge, as there are several media outlets, such as Sol de Puebla, that show that these scholarships are used for purposes other than those intended for the scholarship recipients, such as in the case where several businessmen divert money in a simulated manner in exchange for remuneration of less than 10% (Juárez, 2021).

On the other hand, the gender pay gap in Mexico shows that only 46% of women participate in the paid economy, compared to 77% of men, indicating that women earn 35% less than men on average (IMCO, 2024), traduced to a program of this caliber aims to further close the wage gap between these populations.

Today, this social welfare program represents a new paradigm in Mexico because, although its objective is to reduce inequality, unemployment, and poverty, it has not been proven to guarantee entry into the labor market once the training process is complete (Miquel, 2022); despite the fact that several official Mexican government media outlets maintain this. These distortions of the program have raised a variety of questions regarding its functioning, one of which is the subject of this study: Can this scholarship really promote employability in Mexico among men and women in relation to their level of education? This article proposes a logit model, which highlights the correlations between this scholarship and both men and women finding employment in relation to their level of education before the 12-month duration of the scholarship, using a sample from the *Centro de Integración Juvenil* (CIJ) in Mexico until 2024, which is the only open data available from the Mexican government regarding the *Jóvenes Construyendo el Futuro* scholarships. The purpose of investigating this phenomenon is to lay the groundwork for the development of more robust public policies on social welfare in Mexico with a focus on gender and education, as the data and analysis on this subject should be studied in greater depth.

In addition to the introduction, the structure of this article is as follows: first, a section reviewing the existing literature; second, the data and methodology related to the logit model; third, the results and their discussion; and finally, the conclusions, explaining the main results and limitations of the analysis.

Literature review/Theoretical foundations

Becas Jóvenes Construyendo el Futuro and the youth population

The Welfare Scholarships entitled *Jóvenes Construyendo el Futuro* (JCF) are one of the Mexican government's main strategies to combat youth unemployment. Implemented since 2019, this scholarship aims to offer a monthly stipend and medical coverage to young people between the ages of 18 and 29 who are neither studying nor working, with the promise of training in workplaces for up to a maximum period of 12 months (Secretaría del Trabajo y Previsión Social, 2023). The program is based on the paradigm of learning by doing, with the aim of enabling participants to acquire skills and experience that will facilitate their subsequent entry into the formal labor market (CONEVAL, 2020).

Although the aim is for scholarship recipients to consolidate their position and remain in this first job, this is not always the case, as there are factors that prevent them from having an optimal working environment. In this regard, Zenteno-Hidalgo and Silva (2016), using a logistic

regression model of the Chilean labor market, showed that factors such as remuneration, teamwork, communication, and performance support, among others, prevent workers from building the confidence they need to perform their job duties optimally and thus remain employable.

In Mexico, the transition from the classroom (whether after high school, college, graduate school, or dropping out) represents a critical stage in the life plan of Mexicans. The program aims to fill the gap in formal work experience, providing a starting point for recent graduates or those who have dropped out of the education system. However, several academic studies have suggested that the type of training provided, as well as the profile of mentors and the lack of comprehensive program evaluation mechanisms, may actually skew the program and the employability of its beneficiaries (IMCO, 2022). Furthermore, in the tricolor nation, unemployment disproportionately affects the young population, as demonstrated by the Encuesta Nacional de Ocupación y Empleo (ENOE) (INEGI, 2023), which shows that the unemployment rate among young people aged 15 to 29 exceeds 7% and an even larger segment is underemployed or in informal employment, where they work without benefits or social security. Furthermore, Mexico has one of the highest rates of people who neither study nor work among OECD countries, which represents a structural challenge both for the country and for the development of its population (OCDE, 2022).

García, cited in the work of Muñoz Chávez, et. al in 2022, comments that most of these centers where scholarship recipients are trained are public, closing off the possibility of entering the business world through the private sector; thus, highlighting the difficulty for this group to find employment. This results in an additional difficulty for young people to enter the formal labor market.

Unemployment in Mexico

In April 2025, INEGI revealed that the unemployment rate in Mexico was 2.5%, while the underemployment rate reached 7.1%; meanwhile, informal employment stood at 54.7% (ENOE, 2025). These figures differ from those reported by the newspaper Milenio in May 2025, which stated that at least 23 million people in Mexico who are potentially productive are unemployed, according to information from *Acción Ciudadana Frente a la pobreza* (Citizen Action Against Poverty) (García, 2025). García states that 10.6% of the unemployment rate is made up of 14 million people who are unavailable for work due to household chores (an activity in which 95% are women). He also highlights that there are 19 million people out of work due to unavailability, such as the 8 million students between the ages of 15 and 24.

This phenomenon in Mexico has been evident for several decades, since in 1970 open unemployment represented only 3.8% of the workforce, but underemployment reached 44.8%, occurring more frequently in the agricultural sector and in urban services (Reyes, 1975). This is one of the main problems in Mexico, along with poor-quality employment, since the characteristics most associated with unemployment are gender and educational attainment, with incomplete and complete primary education (Hernández, 2020). According to Torres (1974), the causes of unemployment are mainly demographic and economic factors, resulting from the fact that Mexico's population growth exceeds the global rate and economic growth has not been sufficient to provide employment for the increase in the labor force. Mexico, a country with a Gini index of 43.5 in 2022 according to World Bank data, has asymmetries within its

economy that benefit only a small number of sectors and increase the social inequality gap. Yánez-Contreras and Cano-Hernández (2011) analyzed that age, socioeconomic status, educational level, marital status, rural or urban area, belonging to any minority group, crime, low-quality employment, innovation, technological development, government reforms, and labor regulations may be factors that influence unemployment in Latin America. Another important phenomenon, but one not addressed in this research, is that unemployment as a cause of migration and as a consequence of poverty causes a reduction in remittances, impacting Mexico due to the insufficient supply of jobs for immigrants in the United States (Figueroa et al., 2012).

Finally, García Viña in 2020 states: "There is clearly a clear connection between high youth unemployment rates, lack of education, and informal employment. If these people, who normally do not have completed compulsory education, manage to enter the labor market, the possibility that they will only find jobs in the informal economy and/or that the jobs they hold throughout their professional careers will be in this sector is extremely high." demonstrating that entering the labor market in Mexico is becoming increasingly complex.

Discrepancies in employability by gender

Pacheco and Parker (1996) mention that women continue to perform different jobs than men, as they continue to perform activities classified as “gender-specific.” Although this trend has been reversed in the last decade, with women joining jobs that were previously considered “masculinized,” such as entrepreneurship and employability, it is still necessary to recognize this group in order to achieve women's contribution to the productive process in Latin America (Mazuera Arias et al, 2017). This has led to the informal sector becoming an important destination for female employment (Loría et al., 2012, cited by Hernández, 2020) and has resulted in greater job insecurity for women, such as part-time jobs, excessive working hours, low wages, and a lack of social benefits (Brown and Domínguez cited by Murayama, 2011).

This phenomenon has resulted in 58% of the 10.8 million women in formal employment in Mexico not receiving a decent wage. Of this population, 37% earn poverty wages and 21% earn wages considered to be subsistence wages (Aquino, 2025). In addition, the report “*La justicia económica pendiente*” (Pending Economic Justice) in 2025, led by Paulina Gutiérrez, highlights that almost 2 out of 10 salaried women work excessive hours of more than 48 hours, and 47% of them do not have a stable contract that gives them job security (cited by Méndez, 2025). This document also argues that being indigenous increases the exclusion of women, since by 2025, more than 80% of indigenous women aged 15 and older had a secondary school education or less. The author shows that, added to the barriers of discrimination, this greatly affects their chances of finding decent work.

Data

Box 1. Data and Variables

Symbol	Variable	Source
Y_t	Dichotomous variable of staying in the scholarship or leaving	Author’s own elaboration based on data from *Centro de Integración*

		*Juvenil (*2024)
X_{1t}	Scholarship received monthly in the year it is awarded.	*Centro de Integración Juvenil* (2024)

Given the current context surrounding these *Jóvenes Construyendo el Futuro* welfare scholarships, this study aims to analyze whether these scholarships effectively promote integration into formal employment in the country or whether they contribute to unemployment, considering variables such as the amount received by the scholarship, educational level, and gender, adjusted and applied to the Mexican context in 2025. It should be noted that, although the model recognizes relevant variables, others such as region, level of informal employment, or social inequality index could not be included due to database limitations.

To this end, the following conducted:

1. Adjustment of scholarships based on the start and end year, knowing the years when they entered (Gobierno de México, 2018–2024). Considering that the years 2018 to 2021 are daily* and adjusted to 30 days, the others are monthly (see Table 1):

Table 1. Index of monthly scholarship amounts per year.

Year	Amount in Mexican pesos
2024	$8,480
2023	$7,572
2022	$6,310
2021	$172.87 * 30 días
2020	$141.70 * 30 días
2019	$123.22 * 30 días
2018	$88.36 * 30 días

Source: Prepared internally using data from the Mexican government.

2. The 2,534 available CIJ data were used to classify them as Female or Male,

3. The highest level of education was classified in the CIJ database (using only upper secondary and higher education),

4. Create a threshold (dichotomous variable) with the time they actually spent on that scholarship, creating the dichotomous variable and work assumption, where: based on whether or not they managed to complete the program before 12 months, inferring that if they manage to do so before 12 months, it means that they did enter the labor market (assigning 1 before 12 months, 0 in all other cases). It should be noted that the construction of the threshold does not consider other reasons for leaving, such as dropout, dissatisfaction, or change in personal circumstances, due to the limitations of the database,

5. Construct education and gender variables by multiplying them by the dichotomous variable (see Appendix 1 for the database) and finally,

6. The data were standardized uniformly by dividing all variables by 1,000 in order to capture marginal effects and analyze them under the logit model. (See Appendix 2 for the output in Stata).

In addition to the logit model specification, the possible presence of multicollinearity among the explanatory variables included was evaluated. To do this, a Pearson correlation matrix was constructed using the three variables included in the model: the standardized scholarship amount (becmil), as well as its interactions with education (becmiledu) and gender (becmilgen). Evans in 2025 suggests that correlations greater than 0.7 between regressors may indicate a serious problem of multicollinearity. In this paper, this threshold is considered to interpret the correlation matrix in Appendix 3 and 4.

Metohodology used

One of the main objectives of the study is to estimate the probability of finding employment while receiving welfare benefits, broken down by gender and level of education.

To begin with, the bases from linear regression are needed, with two conditions:

1) Understand the relationship between variables through linear regression and,

2) Statistical significance, which shows the minimum explanatory error made by the independent variable on the dependent variable. All this through the following linear expression:

$$Y_i = a_0 + a_1 X_{1i} + a_2 X_{1i} D_1 + a_3 X_{1i} D_2 + \cdots + a_k X_{ki} D_k + U_i \qquad (1)$$

where, Y_i represents the explained variable, $X_1, X_2, \dots, X_k$ y $D_1, D_2, \dots, D_k$ represent the k possible independent variables. On the other hand, the coefficients a_j represent the effects of changes in the independent variables on the slope of the theoretical model and finally, U_i represents uncontrollable and random variables, referred to as noise, where effects not correlated with the independent variables are attributed to these.

Seeking to see that the variables $X_1, X_2, \dots, X_k$ y $D_1, D_2, \dots, D_k$ are truly significant because of their individuality over Y_i, individual hypothesis tests are considered, $H_0: a_j = 0 \; vs \; H_a: aj \neq 0$, for j=1,2,...,k. To evaluate it, values are used p o p_{value} (probability of error). A $p_{value} < 0.05$ shows that variable X is statistically significant in explaining behavior of Y_i, at least at a 95% confidence level.

For the model, we can use the global likelihood ratio contrast using, $LR = -2[lnL_0 - lnL_1] \sim \chi_k^2$. As a way of assessing whether it is significant, if the probability of the model is $p < 0.05$ we will know that the complete model does improve on the null model.

Now, the explanatory variables in equation (1) are quantitative, but in some cases, it is of interest to include qualitative variables to consider other effects on the dependent variable. A particular case stands out in these variables known as dichotomous or binary. In this case, proposing the variable D_1:

$$D_1 = \begin{cases} 1 \; if \; Female \; (M) \\ 0 \; if \; \; Male \; (H) \end{cases}$$

As well as D_2:

$$D_2 = \begin{cases} 1\ if\ Higher\ Education \\ 0\ if\ Upper\ Secondary\ Education \end{cases}$$

This allows us to analyze the effect of X_{1i} over Y_i, depending on whether the case study is a woman or a man and also higher education or upper secondary education in Mexico. For the estimation of gender and education based on formula (1), the following expressions are proposed:

For women with higher education,

$\hat{Y}_i = E(Y_i|X_1, X_2, \dots\dots, X_k, D_1{=}1 \text{ y } D_2{=}1) = \widehat{a_0} + (\widehat{a_1} + \widehat{a_2} + \widehat{a_3})X_{1i} + \cdots + \widehat{a_k}X_{ki}D_k$

For men with higher education,

$\hat{Y}_i = E(Y_i|X_1, X_2, \dots\dots, X_k, D_1{=}0 \text{ y } D_2{=}1) = \widehat{a_0} + (\widehat{a_1} + \widehat{a_3})X_{1i} + \cdots + \widehat{a_k}X_{ki}D_k$

For women with upper secondary education,

$\hat{Y}_i = E(Y_i|X_1, X_2, \dots\dots, X_k, D_1{=}1 \text{ y } D_2{=}0) = \widehat{a_0} + (\widehat{a_1} + \widehat{a_2})X_{1i} + \cdots + \widehat{a_k}X_{ki}D_k$

For men with upper secondary education,

$\hat{Y}_i = E(Y_i|X_1, X_2, \dots\dots, X_k, D_1{=}0 \text{ y } D_2{=}0) = \widehat{a_0} + \widehat{a_1}X_{1i} + \cdots + \widehat{a_k}X_{ki}D_k$

Thus validating the difference between models, we propose the individual hypothesis test that inherently comes from the coefficient and the binary variable, $\widehat{a_{k+1}}$. That's, why Y_i represents a qualitative dependent variable, allowing a probability model to be applied to estimate the occurrence of the attribute in Y_i, through the explanatory variables of the model $X_1, X_2, \dots, X_k$. Giving form to Y_i, where it can take the value from 1 a 0; as for example,

$$Y_1 = \begin{cases} 1\ if\ Y_i\ meets\ the\ condition;\ prob = p \\ 0\ if\ Y_i\ don't\ meet\ the\ condition;\ prob = 1 - p \end{cases}$$

Testing in this way and estimating the average behavior of the binary Y constructed from the explanatory variable, which would be X_{1i},

$$Y_i = a_0 + a_1X_{1i}u_i \qquad (2)$$

The estimation would be:

$$\hat{Y} = E(Y|X_{1i}) \qquad (3)$$

Now, Y represents our qualitative variable of "success" and "failure," which translates into a Bernoulli variable, with values 0 and 1, where the expectation that it will happen is,

$$E(Y) = \sum_{i=0}^{1} iP(Y = i) = 0 * (1 - p) + 1 * p = p \qquad (4)$$

Therefore, according to (3) and (4), what is being estimated in the linear probability model (LPM) is the probability of possessing the attribute or not.

In this research, we decided to run a logit probability model, which solves the problems defined by Greene in 2002 as "logistic regression that models the probability that the binary dependent

variable Y take the value 1 as a logistic function of a linear combination of the explanatory variables." That model estimates the probability when Y=1, starting from,

$$Pi = E(Y = 1|X_1, X_2, \dots, X_k) = \frac{1}{1+e^{-z_i}} = \frac{e^{z_i}}{1+e^{z_i}} \qquad (5)$$

Where $z_i = a_0 + a_1X_{1i} + a_2X_{2i} + \cdots + a_kX_{ki}$. Causing the $\lim_{z_i \to \infty} Pi = 0$ y $\lim_{z_i \to \infty} Pi = 1$.

However, model (5), which is not linear in its explanatory variables $X_1, X_2, \dots, X_k$. To make it linear, applying algebra gives us the following expression: $\frac{P_t}{1-P_t} = e^{z_t}$ and by applying logarithms, we arrive at the logit model described by Greene:

$$L_t = \ln\left(\frac{P_t}{1-P_t}\right) = z_t = a_0 + a_1X_{1t} + a_2X_{2t} + \cdots + a_kX_{kt} \qquad (6)$$

Finally, in the results section, you can find the marginal effects through estimates on $X_1, \dots, X_k$ y $D_1, D_2, \dots, D_k$ and the probabilities of the four cases shown.

Statistical methods

In addition to the traditional logit model adjustment statistics, predictive capacity was evaluated using the ROC (Receiver Operating Characteristic) curve, using the model's predicted probabilities. The area under the curve (AUC) was also estimated, which summarizes the model's ability to discriminate between individuals who found employment (1) and those who did not (0). Values close to 0.50 indicate discrimination equivalent to chance, while higher values reflect greater predictive power.

Results

Initially, the aim was to find out whether there was any correlation between gender and level of education in terms of the relationship with the employability of welfare scholarships. To this end, the following dichotomous variables were defined:

$$D_1 = \begin{cases} 1 \; if \; Female \; (M) \\ 0 \; if \; Male \; (H) \end{cases}$$

$$D_2 = \begin{cases} 1 \; if \; Higher \; Education \\ 0 \; if \; Upper \; Secondary \; Education \end{cases}$$

Once these variables have been defined, the following adjusted models of formula (1) are proposed (see methodology used),

$$Y_t = a_0 + a_1X_{1t} + a_2X_{1t}D_1 + a_3X_{1t}D_2 + U_t \qquad (7)$$

$$Y_t = a_0 + a_1X_{1t} + a_3X_{1t}D_2 + U_t \qquad (8)$$

$$Y_t = a_0 + a_1X_{1t} + a_2X_{1t}D_1 + U_t \qquad (9)$$

$$Y_t = a_0 + a_1X_{1t} + U_t \qquad (10)$$

Model (7) explains the case of the sample being a woman with higher education; model (8) refers to a man with higher education; model (9) proposes a woman with upper secondary education;

finally, model (10) presents a man with upper secondary education in the sample of *Centro de Integración Juvenil*.

Running the logit model gives us the following estimate:

$$\widehat{Y_t} = 0.4466 + 0.0722X_{t1} - 0.0366X_{1t}D_1 - 0.0908X_{1t}D_2 + U_t \quad (11)$$

P_{value} (0.072) (0.024) (0.000)

The expression shows the gender variable (D_1) and the education variable (D_2) are significative (ver P_{value} 0.024 and 0.000), while (X_1) even that the P_{value} is greater than 0.05, is marginally explanatory, as will be discussed further below.

In order to avoid multicollinearity issues among the model regressors, Pearson's correlation matrix was estimated between the variables scholarship, education, and gender. The resulting matrix is presented below:

Table 2. Pearson's correlation matrix

	Scholarship	Education	Gender
Scholarship	1.000	0.258	0.281
Education	0.258	1.000	0.096
Gender	0.281	0.096	1.000

Source: Own compilation in R based on sample data.

As can be seen, all cross-correlations are below 0.30. The correlations between the scholarship and its interactions with education (0.258) and gender (0.281) are low and consistent with the way the variables were constructed, while the correlation between education and gender (0.096) is practically nil.

These results indicate that there is no strong linear relationship between the regressors in the model, so no multicollinearity problem is identified that could skew the coefficient estimates. Therefore, these variables do provide information to explain the probability of being employed.

Returning to equation (11) gives rise to the four cases described above:

Case 1: Woman with higher education

It is obtained by doing $D_1 = 1$ y $D_2 = 1$

$$\widehat{Y_t} = 0.4466 - 0.0552X_{1t} + U_t \quad (12)$$

Case 2: Man with higher education

It is obtained by doing $D_1 = 0$ y $D_2 = 1$

$$\widehat{Y_t} = 0.4466 - 0.0186X_{1t} + U_t \quad (13)$$

Case 3: Woman with upper secondary education

It is obtained by doing $D_1 = 1$ y $D_2 = 0$

$$\widehat{Y_t} = 0.4466 - 0.0356X_{1t} + U_t \quad (14)$$

Case 4: Man with upper secondary education

It is obtained by doing $D_1 = 0$ y $D_2 = 0$

$$\widehat{Y_t} = 0.4466 + 0.0722X_{t1} + U_t \quad (15)$$

Through these estimates, we can see that women tend to have a negative slope and that, although men with higher education also have a negative slope, it is less pronounced than that of women. On the other hand, when analyzing the case of men with upper secondary education, we find that their slope is positive. When analyzing these four cases marginally, we find the results shown in Table 2:

Box 2. Estimation of the logit model through its marginal effects based on scholarship, gender, and level of education (see Appendix 2).

Acceptance	Marginal Effects	P_{values}
Scholarship	0.0176	0.072
Gender	-0.0366	0.024
Education	-0.0222	0.000
	Pr(acceptance probability)	
Marginal effects after logit	(predict)=0.5774	
Tests	Pseudo R^2 = 0.0092, LR chi^2 (3) =31.66, $p > \chi^2$ = 0.000	

Source: Own elaboration based on the logit model.

Therefore, we can describe what we found in the four cases:

(1) Women with higher education (see equation 12): when income increases by 1,000 Mexican pesos (one monetary unit in the model), the probability of being employed (0.5774) decreases by -0.2412, so that their $\hat{p}$ of employment is: 0.5362.
(2) Men with higher education (see equation 13): when income increases by 1,000 Mexican pesos (one monetary unit in the model), the base probability of employment (0.5774) decreases by -0.0046, so that their $\hat{p}$ of employment is: 0.5728.
(3) Women with upper secondary education (see equation 14): when income increases by 1,000 Mexican pesos (one monetary unit in the model), the probability of being employed (0.5774) decreases by -0.019, so that their $\hat{p}$ of employment is: 0.5584.
(4) Men with upper secondary education (see equation 15): when income increases by 1,000 Mexican pesos or (one monetary unit in the model), the base probability of employment (0.5774) increases by +0.0176, so that their $\hat{p}$ of employment is: 0.5950.

The results show that the case that benefits the least is, first, women with higher education, who have the lowest probability (see probability matrix in Table 2), followed by cases three, two, and four in ascending order.

Table 3. Probability matrix based on the logit model ($\hat{p}$) in percentage

The columns of the matrix show gender. The rows of the matrix show level of education. The subscripts indicate which case they belong to.

($\hat{p}$)	Women	Men
Higher Education	$53.62\%_{1}$	$57.28\%_{2}$
Upper Secondary Education	$55.84\%_{3}$	$59.50\%_{4}$

Source: Own elaboration based on the logit model

Running significance tests, the overall fit of the model is indeed significant compared to the null model. As defined by UCLA and OARC Stats in 2006, the likelihood ratio test compares the "null model" (with a single intercept) with the "full model" (with all explanatory variables), evaluating whether the set improves the fit, since the statistic is based on the difference between the likelihood logarithms of both models and the chi-square distribution (χ^2). This means that at least one of the explanatory variables provides information for predicting the dichotomous variable.

The pseudo R^2 (0.0092) corresponds to the McFadden index, which should not be interpreted as a R^2 linear, but rather as a relative measure of improvement of the model with respect to the null model. Therefore, even though it is low (common due to the structural variables not considered and the social contexts with high unexplained variability), the global likelihood test validates that the model does have a predictive capacity statistically different from zero. This is because it showed a value of LR=31.66 with 3 degrees of freedom, which is statistically significant due to (p < 0.001). Based in the next hypotheses:

H_0: All coefficients of the explanatory variables are equal to zero, meaning that the complete model does not improve on the null model,

H_1: At least one of the coefficients is different from zero, meaning that the complete model does improve on the null model,

Thus, rejecting the null hypothesis that all explanatory variable coefficients are equal to zero, accepting the alternative that at least one of the variables, whether education, gender, or scholarship, provides relevant information to explain the probability of employment.

In addition, two predictors, both the education variable and the gender variable, are significant, being below $p < 0.05$ and, although the scholarship is not significant at 5%, it is marginally significant, implying that not all of the model is noise, since at least two predictors show consistent effects due to the Wald test, which has the following hypothesis:

H_0: The coefficients are equal to zero and the variables are not important for the model,

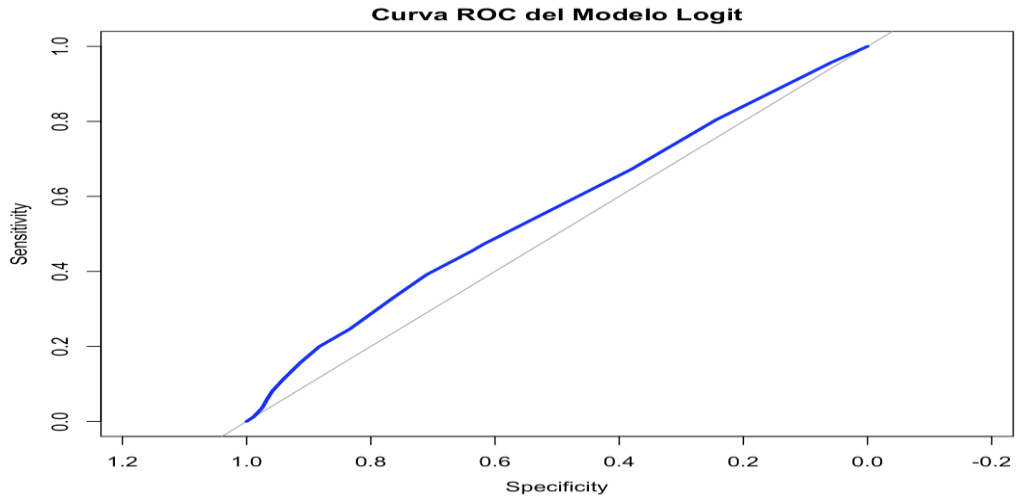
H_1: The coefficients are different from zero and the variables are important for the model.

Thus, giving results where gender and education reject H_0 as long as the scholarship does not reject H_0 (See Box 2). That is why this model, rather than providing perfect predictions, serves to detect close statistical relationships between education, gender, and scholarships. The low explanatory power is reflected in the fact that more relevant variables are missing in the effect

of being employed through a *Jóvenes Construyendo el Futuro* scholarship, as described in the literature, but this does not invalidate the findings in the model.

When evaluating the model's discriminatory capacity, the ROC curve was estimated using the predicted probabilities. The area under the curve was AUC: 0.5598, indicating low predictive ability and barely above chance (AUC=0.50). The ROC curve was very close to the diagonal (see image 1), showing that the model has limited ability to distinguish between young people who found employment and those who did not.

Image 1. ROC curve of the logit model and AUC (Title in Spanish)



Area Under the Curve (AUC)= 0.5598

Source: Prepared internally based on the logit model. Graph created in R.

Discussion

Two variables included in the analysis, education and gender, were found to be significantly accurate. Dichotomous variables were incorporated into this analysis in order to distinguish and analyze two factors related to how a specific welfare grant, namely the *Jóvenes Construyendo el Futuro* (Youth Building the Future) grant, can be used.

The results showed that, of the four cases, women with higher education benefited the least. This can be contrasted with Flores' 2024 research cited by Ramírez Carrera in the same year, which showed that there are precarious working conditions for women with high academic qualifications, many of them with postgraduate degrees, aged between 25 and 40. This shows that, although there has been an increase in women's access to higher education and the labor market, this does not guarantee decent conditions.

On the other hand, we see that the slope in both cases for women is negative, which shows that, in terms of employability in Mexico, being a woman can make it much more difficult than being a man, as evidenced by the model. According to INEGI, as of July 2020, the employment rate in Mexico was 34.7% for women, compared to 71.8% for men.

Returning to the approach of Levy and Székely (2016) and García Viña (2020) described in the literature review, there is a reality in Mexico regarding a variable that the model does not study

but which has an impact on all areas of employability: labor informality. This leads to a lack of correlation between educational attainment and employability, as demonstrated in the study by Yánez-Contreras and Cano-Hernández (2011), which found that in Latin America, improving schooling and maintaining or increasing the degree of formality leads to high employability among citizens qualified for employment.

Welfare scholarships as part of a social assistance program in Mexico promote Mexico's employability in the short term. However, as Muñoz Chávez argues in 2022, he concludes that in the long term, these scholarships are not sufficient to ensure successful employability in the labor market, as there is a lack of bridges with the private sector to enable integration into the corporate and business world.

On the other hand, the most benefited case is that of men with a high school education. This is due to the findings of Pacheco and Parker (1996), who almost 30 years ago found that socially and culturally in Mexico, women are assigned activities "belonging to their sex" and that at the time could be included in a world of formal employment. Today, although many barriers have disappeared, some still remain, as shown by the research of Mazuera Arias et al. in 2017, who recognize that despite efforts to break down the barrier of male-dominated jobs in Latin America, there is still a long way to go before women are recognized and included in formal employment.

In the probability matrix (Table 2), it can be seen that the probabilities of being male exceed 57% and reach almost 60%. Meanwhile, for females, they exceed 53% and reach almost 56%. Testing with the model, we see that even at their maximum probability (women with upper secondary education = 55.84%), women do not reach the minimum probability for men (men with higher education = 57.28%). On the other hand, if we analyze it from the perspective of education, the model shows that the variation between women and men in higher education is 366 basis points, the same as for upper secondary education. This shows that the intervals between both classes are equal; however, those in the upper secondary education class have higher probabilities, so we can determine that the *Jóvenes Construyendo el Futuro* scholarships for the *Centro de Integración Juvenil* sample help people who have an upper secondary education degree over those with a higher education degree.

This finding may contribute to the creation of more accurate, fair, and equitable public policies tailored to the Mexican population, since, as can be seen in the results regarding gender and education, there is a group in the sample that is less advantaged.

Finally, the multicollinearity assessment shows that the three explanatory variables included in the logit model have low correlations with each other, which supports the idea that each one captures different dimensions of the phenomenon analyzed. However, it only analyzes three key variables: education, gender, and social welfare scholarships. However, it is considered important to analyze other key variables such as the degree of informality, tax benefits, region, age, household characteristics, and measurements of inequality variables, among others, in order to build a more refined model with a high degree of prediction.

Regarding to the model tests, it was validated that the overall fit of the model is indeed significant compared to the null model, validated by the alternative hypothesis. Meanwhile, two of the three variables were significant according to the Wald test, which means that the variables are still statistically relevant. The absence of additional controls may reduce the explanatory

power of the model and contribute to the low pseudo value. In future extensions of the work, the incorporation of these socioeconomic controls would allow for the estimation of a more complete model with greater predictive power.

The low discriminatory capacity observed in the ROC curve (AUC=0.5598) is consistent with pseudo R^2 reduced logit model. Although the model is statistically significant and shows consistent relationships between the variables included (scholarship and mainly education and gender), its predictive power is low. This suggests that the phenomenon of employability depends on additional factors not found in the database, thus limiting the model's ability to adequately separate cases of job success and failure.

Conclusions

A logit analysis was applied with the main objective of determining the probability of being employed through a *Jóvenes Construyendo el Futuro* scholarship and leaving the program before 12 months, which is the maximum period for using this scholarship. These scholarships are characterized as a social welfare program in Mexico for young Mexican men and women who are neither studying nor working, in exchange for which they receive a scholarship and are trained for a job so that they can enter the formal labor market. The variables used for this model were the amount of the scholarship, their level of education (higher or upper secondary), and their gender.

What stood out about the model was that in all four cases among men and women, women with both higher and upper secondary education had a negative slope, while men with higher education also had a negative slope, but men with upper secondary education had a positive slope. This shows that men with this scholarship are more likely to find employment in Mexico than women.

The results of the analysis also show that it is not particularly beneficial for women with higher education to enter this scholarship program, as they are less likely to leave the program to take up employment. Conversely, men with upper secondary education are more likely to find employment and leave the program (below the threshold of the dichotomous variable defined in the article).

The factors found complement the literature (see literature review), which has identified not only gender and education as causes, but also the degree of labor informality, region, age, and many other variables in Mexico that are among the main causes of difficult labor market integration within that country.

On the other hand, the analysis has limitations, considering that the model presents a low pseudo R^2, The overall plausibility test shows that the model is better than the null model and, finally, that two of the three predictors are significant and consistent for the dichotomous probability through the Wald test (gender and education). The sample is limited, as only 2,534 data points available from the Mexican government's Youth Integration Center are used, and this model may not be representative of the entire population of the Youth Building the Future program. Furthermore, the model's low explanatory power, evidenced by the AUC of 0.5598, suggests that the program operates in a context where unaccounted structural factors limit the individual prediction of labor market insertion.

Finally, the analysis seeks to be an approximation that explains how narrow and explanatory these variables can be, but not sufficient to provide a conclusive conclusion about the reality of

the program at the national level. Therefore, the ultimate goal of this model is more focused on detecting significant relationships that could lead to further investigation of this phenomenon, such as the labor sector, informality, region, or age, which are and could be key variables for a more robust model.

Greetings

Special thanks to the entire community at Universidad La Salle México and Johannes Kepler Universität, Linz, Austria. For believing in me by opening the doors of your institutions, training me, and developing me as a great professional.

References

Agresti, A. (2021). *Categorical data análisis*. (4th ed.). Wiley.

Aquino, E. (2025, March 7). *Desigualdad económica de las mujeres: enfrentan desempleo por cuidar, informalidad laboral y empleos con salarios de pobreza.* Animal Politico. https://animalpolitico.com/sociedad/desigualdad-economica-mujeres-desempleo-informalidad-salarios

Bruin, J. (2006). *Newtest: command to compute new test.* UCLA: Statistical Consulting Group. https://stats.oarc.ucla.edu/stata/ado/analysis/

Consejo Nacional de Evaluación de la Política de Desarrollo Social. (2020). *Ficha de monitoreo de programas y acciones federales 2020: Jóvenes Construyendo el Futuro*. https://www.coneval.org.mx/Evaluacion/IEPSM/Documents/Fichas/2020/Ficha_Jovenes_Construyendo_el_Futuro_2020.pdf

Evans, R. (2025). *Detecting and Dealing with Multicollinearity: An Introduction to Econometrics*. econometricstutor.co.uk. https://www.econometricstutor.co.uk/multicollinearity-detecting-and-dealing-with-multicollinearity

Figueroa Hernández, E., Ramírez Abarca, O., González Elías, J. M., Pérez Soto, F., & Espinosa Torres, L. E. (2012). *Análisis del desempleo, la migración y la pobreza en México.* Revista Mexicana de Agronegocios, *30*(enero-junio), 835-847. Sociedad Mexicana de Administración Agropecuaria A.C. http://www.redalyc.org/articulo.oa?id=14123097006

García, Á. (2025, May 13). *Desempleo en México repunta con 23 millones de personas en exclusión laboral.* Grupo Milenio. https://www.milenio.com/negocios/desempleo-mexico-repunta-23-millones-personas

García, A. K. (2018, december 26). *Educación en México: insuficiente, desigual y la calidad es difícil de medir.* El Economista. https://www.eleconomista.com.mx/politica/Educacion-en-Mexico-insuficiente-desigual-y-la-calidad-es-dificil-de-medir-20181225-0028.html

García Viña, J. (2020). *La situación del empleo de los jóvenes y una propuesta de políticas públicas para afrontar el desempleo juvenil.* Revista Latinoamericana de Derecho Social, 30, 65–94. https://doi.org/10.22201/iij.24487899e.2020.30.14072

Greene, W.H. (2002) *Econometric Analysis*. 5th Edition, Prentice Hall, Upper Saddle River, 802.

Hernández Pérez, J. (2020). *Desempleo en México por características sociodemográficas*, 2005-2018. Economía UNAM, 17(50), 166-181.

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

INEGI. (2023). *Encuesta Nacional de Ocupación y Empleo (ENOE), cuarto trimestre 2023.* https://www.inegi.org.mx/programas/enoe/15ymas/

Instituto Mexicano para la Competitividad. (2022). *Jóvenes Construyendo el Futuro: ¿aprenden los jóvenes?* https://imco.org.mx/jovenes-construyendo-el-futuro-aprenden-los-jovenes/

Instituto Nacional de Estadística y Geografía (n.d.). *Características educativas de la población.* https://www.inegi.org.mx/temas/educacion/

Juárez, V. H. (2021, february). *Fraude y corrupción, desvían dinero de Jóvenes Construyendo el Futuro.* El Sol de Puebla. https://www.elsoldepuebla.com.mx/local/fraude-y-corrupcion-desvian-dinero-de-jovenes-construyendo-el-futuro-6299984.html

Levy, S., & Székely, M. (2016). *¿Más escolaridad, menos informalidad? Un análisis de cohortes para México y América Latina.* El Trimestre Económico, 83(332), 499–548. https://doi.org/10.20430/ete.v83i332.232

Loría, Eduardo; Libreros, Catalina; Salas, Emmanuel (2012). *La ley de Okun en México: una mirada de género, 2000.2-2011.1.* Investigación Económica, 71(280), 121-140.

McFadden, D. (1974). *Conditional logit analysis of qualitative choice behavior.* En P. Zarembka (Ed.), Frontiers in econometrics (pp. 105–142). Academic Press.

Martínez, F. (2002). *Nueva visita al país de la desigualdad.* La distribución de la escolaridad en México, 1970-2000. Revista mexicana de investigación educativa, 7(16).

Mazuera Arias, Rina & García, Marisela & Quintero, María & Santafe Rojas, Akever Karina & Carreño Paredes, Myriam Teresa & Peraza, Miguel & Martínez-Rozo, Camila-Andrea & Vera, Miguel & Tinoco-Guerra, Antonio-Enrique & Ruscheinsky, Aloisio & Danúbia, Cará-Treis & Benites-Lopes, Silva & Aldrete-Glez, Ana-Paola & Pocovi-Garzón, Patricia & Robles, Josefina & Morales-Carrasco, Lilian-Victoria & Lascano, Nelson & Roberto, Edison & Albornoz-Garzón, María-Augusta. (2018). *Mujer, emprendimiento y empleabilidad: una mirada interdisciplinaria.* Ediciones Universidad Simón Bolívar. https://doi.org/[si existe] ISBN: 978-958-5430-90-7

Méndez, R. (2025). *8M 2025: "Justicia económica para mujeres" es el gran pendiente de México, según organización.* MVS Noticias. https://mvsnoticias.com/economia/2025/3/5/8m-2025-justicia-economica-para-mujeres-es-el-gran-pendiente-de-mexico-segun-organizacion-681468.html

Miquel Hernández, M. (2022). *Una visión del funcionamiento del programa "Jóvenes Construyendo el Futuro".* Palimpsesto. Revista Científica de Estudios Sociales Iberoamericanos, 12(20), 111–125. https://doi.org/10.35588/pa.v12i20.5248

Morales-Ramos, E. (2011). *Los rendimientos de la educación en Méxic*o, Working Papers, No. 2011-07, Banco de México, Ciudad de México

Muñoz Chávez, R. L., Muñoz Chávez, J. P., Valle Cruz, D., & Barrios Quiroz, H. (2023*). Jóvenes Construyendo el Futuro: Análisis y recomendaciones sobre el diseño de la política pública.*

RECAI. Revista de Estudios en Contaduría, Administración e Informática, 12(33), 1–19. https://doi.org/10.36677/recai.v12i33.19330

Murayama, C. (2011). *Reseña de México: desigualdad económica y género, de Flor Brown y Lilia Domínguez (Coords.).* Revista de Economía Mundial, (29), 327–330. Sociedad de Economía Mundial. Disponible en: http://www.redalyc.org/articulo.oa?id=86622169013

Organisation for Economic Co-operation and Development. (2021). *Perspectivas del empleo 2021: Reforzando la resiliencia del mercado laboral tras la crisis del COVID-19*. https://www.oecd.org/employment-outlook/ Centro de Estudios Espinosa Yglesias. (2021). ¿Y después qué? Inserción laboral de jóvenes egresados del programa Jóvenes Construyendo el Futuro. https://ceey.org.mx/insercion-laboral-jovenes-jcf/

Organisation for Economic Co-operation and Development. (2022). *Panorama de la juventud en México 2022.* https://www.oecd.org/mexico/panorama-de-la-juventud-en-mexico-2022.htm

Pacheco, E., & Parker, S. (1996). *Participación económicamente activa femenina en el México urbano. Un breve recuento y algunos hallazgos recientes.* Problemas Del Desarrollo, 27(106), 21–33. http://www.jstor.org/stable/45480800

Ramírez Carrera, S. (2024, 17 de junio). *Mujeres profesionistas y precarización laboral: El CIEG continúa ciclo de conferencias en colaboración con la Universidad de la Coruña.* Gaceta Digital UNAM. Recuperado de https://www.gaceta.unam.mx/mujeres-profesionistas-y-precarizacion-laboral/

Reyes, J. A. (1975). *El desempleo en México.* Universidad Nacional Autónoma de México, Facultad de Economía.

Secretaría del Trabajo y Previsión Social. (2023). *Informe Anual del Programa Jóvenes Construyendo el Futuro 2023.* Gobierno de México. https://jovenesconstruyendoelfuturo.stps.gob.mx

Torres Gaitán, R. (1974). *Desempleo y subdesempleo en México.* Problemas del desarrollo, 5(18), 13-18.

World Bank Open Data. (n.d.). *Indicador GINI.* https://datos.bancomundial.org/indicador/SI.POV.GINI?locations=MX

Yánez-Contreras, M., y Cano-Hernández, K. del C. (2011). *Determinantes del Desempleo: Una revisión de la literatura.* Panorama Económico, 19(19), 135–148. https://doi.org/10.32997/2463-0470-vol.19-num.19-2011-352

Zenteno-Hidalgo, Á. C., y Silva, C. A. D. (2016). *Factores y prácticas de alto desempeño que influyen en el clima laboral: análisis de un caso.* Innovar: Revista de Ciencias Administrativas y Sociales, 26(59), 119–136.

Appendix 1. Data base.

To access the database, access to the following link:

https://docs.google.com/spreadsheets/d/1IjOowzwhy9FnJse2SPzOat6XsQpTVxGw/edit?usp=drive_link&ouid=113101130806851783168&rtpof=true&sd=true

Appendix 2: Results in Logit Model.

```
. logit dico becmil becmiledu becmilgen

Iteration 0:   log likelihood = -1726.9208
Iteration 1:   log likelihood = -1711.1179
Iteration 2:   log likelihood = -1711.0924
Iteration 3:   log likelihood = -1711.0924

Logistic regression                                     Number of obs     =      2,534
                                                        LR chi2(3)        =      31.66
                                                        Prob > chi2       =     0.0000
Log likelihood = -1711.0924                             Pseudo R2         =     0.0092

------------------------------------------------------------------------------
        dico |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     becmil |   .0721976   .0400767     1.80   0.072    -.0063513    .1507465
   becmiledu |  -.0908367   .0181023    -5.02   0.000    -.1263166   -.0553567
   becmilgen |  -.0366372    .016272    -2.25   0.024    -.0685298   -.0047447
       _cons |   .4467616   .2048432     2.18   0.029     .0452764    .8482469
------------------------------------------------------------------------------

. mfx

Marginal effects after logit
      y  = Pr(dico) (predict)
         =  .57739797
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.      z    P>|z|   [    95% C.I.   ]       X
---------+--------------------------------------------------------------------
  becmil |   .0176169      .00978     1.80   0.072  -.001546   .03678   5.42947
becmil~u |   -.022165      .00441    -5.02   0.000  -.030814 -.013516   4.23678
becmil~n |  -.0089398      .00397    -2.25   0.024  -.016721 -.001159    3.8704
------------------------------------------------------------------------------
```

Appendix 3: R code used for correlation matrix.

```
library(readxl)

datos <- read_excel('RUTADELOSDATOS.xlsx')

vars_logit <- datos[, c("becmil", "becmiledu", "becmilgen")]

vars_logit_num <- data.frame(lapply(vars_logit, as.numeric))

mat_cor <- cor(vars_logit_num, use = "pairwise.complete.obs")

round(mat_cor, 3)
```

Appendix 4: Correlation matrix result.

```
> round(mat_cor, 3)
          becmil becmiledu becmilgen
becmil     1.000     0.258     0.281
becmiledu  0.258     1.000     0.096
becmilgen  0.281     0.096     1.000
```