Cairo University

Faculty of Economics and Political Science

Statistics Department, English Section

Fourth Academic Year

Disparities in Egyptian Labor Force

Graduation Project 2024

Prepared by:

Nour Mohamed Eletreby 5200917	**Haneen Ahmed Elgabry** 5200199
Nada Hussein Mohamed 5200699	**Mareez Maher Mawaad** 5200524

Under the supervision of:

Dr Wafaa Ibrahim.

Acknowledgment

We would like to express our sincere gratitude to Professor Fatma Elzanaty for her insightful feedback, scholarly wisdom, and steadfast support. We are deeply appreciative of her mentorship and the invaluable contributions she has made to our academic development.

We extend our deepest appreciation to Dr. Wafaa Ibrahim for her invaluable guidance, unwavering support, and expert advice throughout the duration of our graduation project. Her commitment to our development as scholars has left an indelible mark on our academic journey. We are profoundly grateful for her dedication, encouragement, and willingness to invest her time and expertise in our project.

In addition to our academic mentors, we owe a debt of gratitude to our beloved family and friends. Their unwavering encouragement, unconditional love, and unwavering support have sustained us through the highs and lows of our academic journey. Their belief in our potential has been a constant source of inspiration, motivating us to persevere in the face of challenges and setbacks. We are profoundly grateful for their steadfast presence in our lives and for the countless sacrifices they have made to support our academic pursuits. Their encouragement has not only enriched our academic experience but has also shaped us into the individuals we are today.

Abstract

The Egyptian labor force serves as a vital component of economic productivity and development, yet disparities persist, affecting individuals based on gender, urban/rural residence, and disability status. This study explores these disparities, with a particular focus on wage gaps, aiming to provide a comprehensive understanding of their nature and extent. Recognizing the significance of these inequalities not only for social justice but also for fostering inclusive economic growth, this research aims to inform policymakers and stakeholders about the urgency of addressing labor market inequities. By uncovering the root causes and implications of these disparities, this study seeks to contribute to the formulation of evidence-based policies aimed at promoting a fairer and more prosperous labor force landscape in Egypt.

“Every journey begins with a step towards equality.”

Table of Contents

List of Tables

List of Figures

Chapter 1
Introduction

1.1 Introduction

In Egypt, as in many other countries, the labor force serves as a vital component of economic productivity and development as well. However, there exists various disparities that affect individuals based on factors such as gender, urban/rural residence, and disability status.

When thinking about the pattern of these disparities, naturally we thought of the disparities of wages and employment status. This is done by measuring the wage gap between gender, urban/rural residence, and disability status, and in another model measuring the disparities in employment status between gender, urban/rural residence, and disability status. Understanding labor force disparities is crucial not only from a point of social justice or equality, but it is also important for achieving economic development by allowing policymakers and researchers come up with suitable plans for a more inclusive labor force landscape in Egypt. In summary, measuring disparities in the labor market serves critical roles in uncovering inequality, improving economic productivity, and meeting legal and ethical responsibilities. This study aims to empirically examine disparities in labor force according to gender, urban/rural residence, and disabilities in labor force.

1.2 Literature Review

This section includes previous studies about labor force disparities in Egypt. Zeitoun et al., (2023) studied the impact of social norms on women's labor force participation in Egypt. It finds that women's employment is significantly impacted by strong gender role attitudes, which are particularly influenced by empirical expectations and norms surrounding men's position as providers. The study examined female labor for participation "FLFP" social norms in Egypt through both qualitative (focus groups) and quantitative (phone survey) methods. Qualitative findings informed the design of the survey, which collected data from over 6,600 respondents across Egypt's governorates. Statistical techniques such as descriptive analysis, inferential tests, regression, and factor analysis were used to analyze the data and understand FLFP norms and barriers.

Zaghlool (2022) used quantile regression on Egypt Labor Market Panel Survey "ELMPS 2018" data, the study reveals varying salary gaps across distributions, with equity sensitivity addressing the impact on contextual performance and attributing a large share to discrimination.

Al Shalkamy (2021) conducted a survey with Baseera Organization on a sample of 2016 Egyptians from all governorates of the Republic. Data was collected by phone during the period from March 22 to April 4, 2020. All estimates in this policy brief are subject to a less than 3% margin of error. The survey suggested that 5% of respondents have a family member with a disability, which is below the national average of 10%. This discrepancy may be due to respondents choosing not to disclose such information or being unaware of the broad definition of disability.

Krafft et al., (2019) proved that from 1988 till 2018: supply and demand-side factors continue to influence decisions, and despite educational improvements, female involvement in Egypt's labor market remains low due to post-revolutionary consequences. The paper used descriptive analysis and trend analysis using longitudinal data from the Egypt Labor Market Panel Survey (ELMPS) and the Labor Force Survey (LFS).

Said et al., (2018) proved that gender differences in Egypt's labor market and finds that, in knowledge-intensive service industries, gender diversity and productivity/wages are positively correlated, whereas in less knowledge-intensive sectors, the correlations are negative or neutral. The paper used linear econometric model with fixed effect, regression analysis and Herfindahl index calculation using Labor Force Survey (LFS).

Biltagy (2018) proved that despite advancements over time, research indicates persistent female salary discrepancies in Egypt, which are assigned to elements including education, experience, discrimination, and unobservable traits. The paper used methods that rely on estimating log earnings functions based on “Human Capital Theory” for each gender separately as well as the standard Oaxaca-Blinder procedure is used to estimate to what extent the overall wage gap between males and females exists using data from Egypt Labor Market Panel Survey “ELMPS"2006 and 2012.

Nazier and Ramadan (2018) proved that in Egypt, low female participation rates are a result of maternal employment status and education levels, even in the face of overall labor force expansion. Using the Egypt Labor Market Panel Survey ELMPS, 2012, the paper used probit regression modeling and instrumental variable method.

Hendy (2015) proved that from 1998 till 2012: despite educational advancements, female participation remains low in Egypt, exacerbated by post-revolution effects, with supply and demand-side factors influencing decisions. The data used in this paper comes from the three rounds (1998, 2006 and 2012) of the Egypt Labor Market Panel Survey “ELMPS”, as well as the 1988 round of the labor force sample survey (LFSS). The research employs descriptive statistics to investigate the complexities of labor force trends in Egypt.

Helmy (2011) proved that following 1994, there was a noticeable increase in the rural–urban income difference in Egypt, especially among middle-class individuals. Assaad (2000) proved that demographic trends have a big impact in rural and urban areas. In rural regions, there are higher birth rates and limited migration to cities as well; this has led to a fast-growing labor force and obviously higher youth unemployment rates. These demographic pressures are expected to last for about another ten years as a larger group of young people enter the workforce. The paper used data from Labor Force Sample Survey, carried in October 1988 (LFSS 1988) by the Central Agency for Public Mobilization and Statistics (CAPMAS), and the Egypt Labor Market Survey 1998 (ELMS 1998), carried by the Economic Research Forum for the Arab Countries, Iran, and Turkey (ERF).

1.3 Study Problem

Prior research has documented substantial differences in employment patterns between men and women in Egypt. Figure (1) shows that the female labor force participation rate is only around 14.9% compared to over 85.1% for males (The Central Agency for Public Mobilization and Statistics “CAPMAS”, 2022). also figure (2) shows that Unemployment rates are also higher for females at 18.4 million unemployed women versus only 5 million for males (CAPMAS, 2022).

Figure (3) shows that gaps also exist across geographic regions, with labor force participation rates lowest in urban with 42.3% compared to 57.7% for rural in Egypt, Figure (4) shows that unemployment rates are also higher for urban areas with 10.7 million unemployed in urban versus only 4.6 million in rural areas (CAPMAS, 2022).

Egyptians with disabilities shape about 10.7% of the total Egyptian population. In terms of gender, males shape about 53.9% of people with disabilities while 46.1% for females.

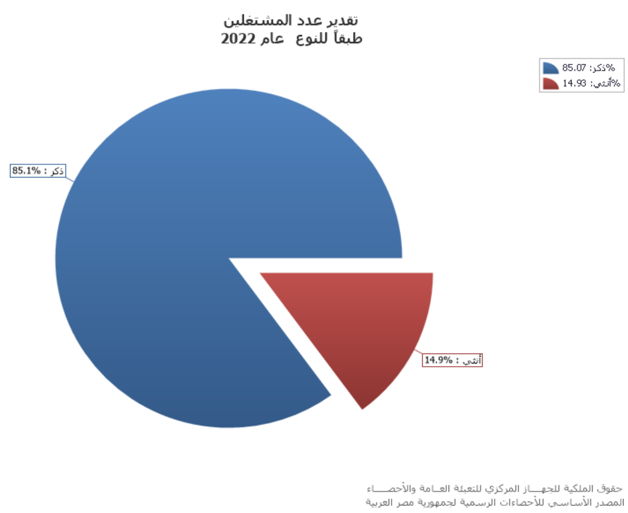

Figure 1: labor force participation rate in terms of gender.

Source: The Central Agency for Public Mobilization and Statistics “CAPMAS”, 2022

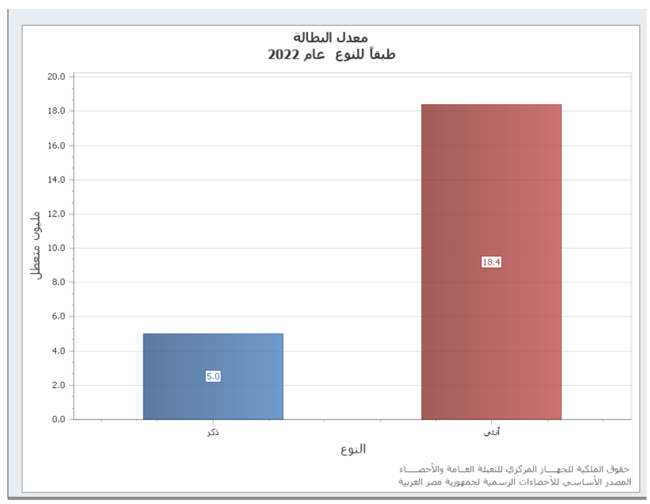

Figure 2: Unemployment rates in terms of gender.

Source: The Central Agency for Public Mobilization and Statistics “CAPMAS”, 2022

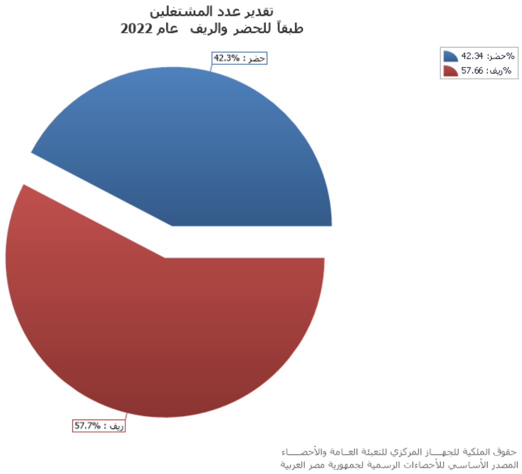

Figure 3: labor force participation rate in terms of urban/rural.

Source: The Central Agency for Public Mobilization and Statistics "CAPMAS", 2022.

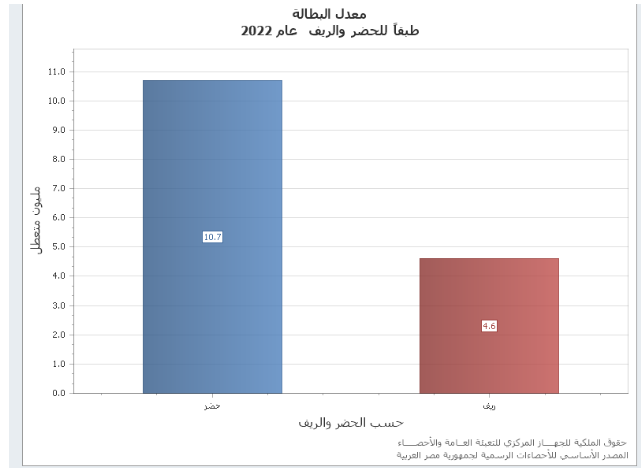

Figure 4: Unemployment rates in terms of urban/rural.

Source: The Central Agency for Public Mobilization and Statistics "CAPMAS", 2022

After reading previous studies, we found that most of the statistical studies related to wage gap and labor force participation between urban and rural areas are relatively old and the new studies are focusing more on the wage gap between males and females. Moreover, statistical studies related to people with disabilities concerning wages and participation are relatively few. Additionally, there are few papers concerned with labor force with the 3 fields intersected together (gender, urban/rural residence, and disability status).

1.4 Objectives

The main objective of this study is to fill the literature gaps and examine labor market disparities in Egypt according to gender, residence, and disabilities. The following are the sub-objectives and research questions of this study.

1- Determine disparities in Egyptian labor force according to gender.
2- Determine disparities in Egyptian labor force according to residence.
3- Investigate how disability can affect labor force outcomes.
4- Explore how gender, urban/rural residency and disability status intersect to create labor force disparities.
5- Identify subgroups (disabled women in rural areas for example) as they experience multiple challenges.

1.5 Research Questions

1- What is the gender-based disparities within the Egyptian labor force, considering wage gap and employment status?
2- How do labor force disparities vary based on residence in Egypt, examining differences in wages gaps, and employment status between urban and rural areas?
3- To what extent does disability impact labor force outcomes in Egypt, including wage gaps and employment status?
4- How do gender, urban/rural residency, and disability status intersect to contribute to labor force disparities in Egypt, and what are the specific patterns and dynamics associated with these intersections?
5- Which subgroups (e.g. disabled women in rural areas) experience the greatest labor market disparities when considering the intersection of gender, residence, and disability status?

1.6 Methodology

In this study we use The Labor Force Survey 2022 dataset. Our study will adopt a quantitative approach to investigate labor force disparities in Egypt, using data from the Labor Force Survey LFS 2022 dataset. Through quantitative analysis, we aim to statistically examine the factors influencing labor force participation and wage differentials.

We will use statistical software packages such as SPSS, R and STATA for data analysis. These software tools offer a wide range of statistical functions and data manipulation capabilities necessary for analyzing large datasets like the LFS 2022.

Descriptive analysis: In our report, descriptive analysis is used as a main method. By This way, we could break down and understand the main parts of our data, giving us a full view of our dataset. We did a detailed descriptive analysis for each variable in our study, this step was key in spotting any patterns or odd things that could affect our later work. Also, we looked at how our two response variables relate to their explanatory variables. This helped us see how strong and in what direction the relationships between our variables are. This is very important for making good guesses and drawing useful conclusions from our data. Descriptive analysis gives us a strong starting point for any statistical work, it lets us understand our data before we do more complicated analyses; by looking at the relationships between variables, we can spot important connections that might not be easy to see. This can guide us in choosing the right statistical models later in our research.

Measures of association: Measures of association help us understand whether and how strongly two variables are related. First, we will examine the relationship between our first response variable (employment status) and the explanatory variables. Second, we will examine the relationship between our second response variable (wage) and the explanatory variables keeping in mind with categorical explanatory variables we will get the average of wage in each category and t-test or analysis of variance "ANOVA" will be computed.

For **categorical variables**, we will be using the following measures of association based on the types of variables:

1. **Odds Ratio**: This measure is used to assess the strength of association between two categorical variables, particularly when they are binary (having two categories). The odds ratio quantifies the likelihood of an event occurring in one group compared to another group.
2. **Gamma**: Gamma is a measure of association used for ordinal categorical variables. It assesses the strength and direction of the relationship between two ordinal variables. The gamma coefficient ranges from -1 to 1, where values closer to -1 or 1 indicate a stronger association.
3. **Contingency Coefficient**: The contingency coefficient is a measure of association for nominal categorical variables. It indicates the strength of association between two nominal variables. The coefficient ranges from 0 to 1, where 0 represents no association and 1 represents a perfect association.

For **continuous variables**, we will be using the Pearson correlation coefficient. This measure assesses the strength and direction of the linear relationship between two continuous variables. The Pearson correlation coefficient ranges from -1 to 1, where values closer to -1 or 1 indicate a stronger linear association, and 0 represents no linear association.

It is important to note that these tests are conducted under a significance level of $\alpha = 5\%$, which means that the results will be considered statistically significant if the probability of obtaining the observed association by chance is less than 5%.

Binary logistic regression model: In our study, we utilized a binary regression model for our first model. This model was chosen because our response variable, employment status, is binary in nature - a person is either employed or not. The binary regression model, also known as logistic regression, is particularly suitable for this type of data. It allows us to predict a binary outcome based on a set of independent variables, providing us with the probability of a particular outcome. In our case, we used the binary regression model to understand how different factors influence a person's employment status. The model helped us identify which variables have a significant impact on employment and estimate the size of these effects. This approach is powerful and flexible, allowing us to handle both numerical and categorical variables, and to account for complex relationships between variables. By using a binary regression model, we were able to gain valuable insights into the factors that influence employment status, thereby enhancing the depth and validity of our analysis.

Equation of binary logistic regression: $\log\left(\frac{p(Y)}{1-p(Y)}\right) = \beta 0 + \beta 1(X1) + \cdots + \beta n(Xn)$

Definition of the symbols:

Y: The dependent variable, representing the binary outcome. It can take on two values, usually coded as 0 and 1 (e.g., employed/unemployed, yes/no).

p(Y): The probability of the outcome Y occurring. In logistic regression, this probability is estimated based on the values of the predictor variables.

X1, X2, ... Xn: The independent variables, or predictor variables. These are the factors that might influence the outcome Y. They can be continuous (e.g., age, income) or categorical (e.g., gender, education level).

β0: The intercept term, representing the log odds of the outcome Y when all predictor variables are equal to zero.

β1, β2, ... βn: The regression coefficients, representing the change in the log odds of the outcome Y for a one-unit change in the corresponding predictor variable, holding all other variables constant.

Blinder-Oaxaca decomposition: In this study, we utilized the Blinder-Oaxaca decomposition method to analyze wage gaps based on three different models: gender, urban/rural residence, and disability status. The Blinder-Oaxaca decomposition is a statistical technique that allows us to break down the wage gap into two components: the explained part and the unexplained part.

We examined the wage gap between gender, residence, and disability status. The decomposition allowed us to understand how much of the wage gap is due to differences in observable characteristics (explained part), and how much is due to discrimination and other factors no included in the explained part (unexplained part).

Given are two groups, A and B; an outcome variable, Y; and a set of predictors. For example, think of a group of males and a group of females, (ln) wages as the outcome variable, and human capital indicators such as education and total weekly working hours as predictors. The question now is how much of the mean outcome difference,

$$R = E(Y_A) - E(Y_B)$$

where E(Y) denotes the expected value of the outcome variable, is accounted for by group differences in the predictors.

Based on the linear model

$$Y_\iota = X_\iota{}' \beta_\iota + \varepsilon_\iota \, , \; E(\varepsilon_\iota)=0 \, , \; \iota \in (A, B)$$

where X is a vector containing the predictors and a constant, β contains the slope parameters and the intercept, and is the error, the mean outcome difference can be expressed as the difference in the linear prediction at the group-specific means of the regressors. That is,

$$R = E(Y_A) - E(Y_B) = E(X_A)' \beta A - E(X_B)' \beta_B \qquad (1)$$

because

$$E(Y_\iota) = E(X_\iota{}' \beta_\iota + \varepsilon_\iota) = E(X_\iota{}' \beta_\iota) + E(\varepsilon_\iota) = E(X_\iota)' \beta_\iota$$

where $E(\beta_\iota) = \beta_\iota$ and $E(\varepsilon_\iota) = 0$ by assumption.

To identify the contribution of group differences in predictors to the overall outcome difference, (1) can be rearranged, for example

$$R = \{E(X_A) - E(X_B)\}' \beta_B + E(X_B)' (\beta_A - \beta_B) + \{E(X_A) - E(X_B)\}' (\beta_A - \beta_B)$$

Then this method is decomposed using “twofold” decomposition as follows:

Let β^* be such a nondiscriminatory coefficient vector

The “twofold” decomposition,

$$R = Q + U$$

where the first component,

$$Q = \{E(X_A) - E(X_B)\}' \beta^*$$

is the part of the outcome differential that is explained by group differences in the predictors (the “quantity effect”), **and the second component,**

$$U = E(X_A)^{\prime} (\beta_A - \beta^{*}) + E(X_B)^{\prime} (\beta^{*} - \beta_B)$$

is the unexplained part. The latter is usually attributed to discrimination.

By using the Blinder-Oaxaca decomposition, we were able to quantify the extent to which these wage gaps are due to differences in characteristics (explained part) versus other factors (unexplained part). This methodology provided valuable insights into the nature of wage disparities in our sample, paving the way for targeted policy interventions. The **explained part** of the wage gap, also known as the "endowment effect", is the portion of the wage gap that can be attributed to differences in observable characteristics. In our models, these characteristics vary based on gender, urban/rural residence, and disability status.

The **unexplained part** of the wage gap, also known as the "coefficient effect", is the portion of the wage gap that cannot be explained by differences in observable characteristics. This component captures the effects of potential discrimination. Discrimination, in this context, refers to unequal treatment of individuals based on their gender, place of residence, or disability status, which is not justified by differences in their qualifications or job performance.

Chapter 2
Data Preparation

2.1 Data Source

As we said before, our primary data source for this study is the LFS 2022 dataset obtained from the Economic Research Forum (ERF) Data Portal. It covers a comprehensive array of structured survey questions, covering various demographic characteristics such as age, gender, and education. Additionally, it explores employment-related questions including employment type, occupation, and industry. Furthermore, the dataset contains valuable insights into income levels and other relevant variables. These variables will form the basis of our analysis of labor force outcomes, enabling us to identify and assess disparities within the labor market. The original sample size in the Labor Force Survey 2022 dataset designed for each quarter is 22500 households with a total of 90000 households per year (Designed Sample 89982, Implemented Sample 77002), allocated 45.4% for Urban and 54.6 % for rural over all governorates (urban / rural) in proportion to the size of each governorate. In our study we will focus on specific variables that will be stated later in the report.

2.2 Data Screening

For this study, the dataset used is sourced from the Labor Force Survey conducted in 2022; aiming to investigate the disparities in labor force across various factors such as gender, urban/rural residence, and disabled people. This dataset is divided into 2 parts: households and individuals. However, for this study, the analysis focuses on individual-level.

Our sample contains 45453 observations, and it is restricted to individuals aged 18 and above as we wanted our sample to start from youth individuals and chosen from round 3 only as round 3 uniquely includes questions about the disability status unlike the other rounds.

This data suffers from missing data and outliers. Missing data is handled by imputation, as for continuous variables, these missing values were treated using mean imputation. It is well mentioning that some of the categorical variables contain not stated category, so we treated the 'Not Stated' category as missing data and imputed by the mode of the variable.

Regarding the outliers, it is treated through "Winsorization" technique using 2nd and 98th percentiles; this technique works on limiting extreme values in the data to reduce the effect of outliers, here data below 2nd percentile is set to the 2nd percentile and data above 98th percentile is set to the 98th percentile.

2.3 Conceptual Design and Framework

2.3.1 For the first model

In the first model the response variable is the employment status and 6 explanatory variables which are gender, urban/rural residence, disability status, age, marital status and educational level. The sample size in this model is 45453.

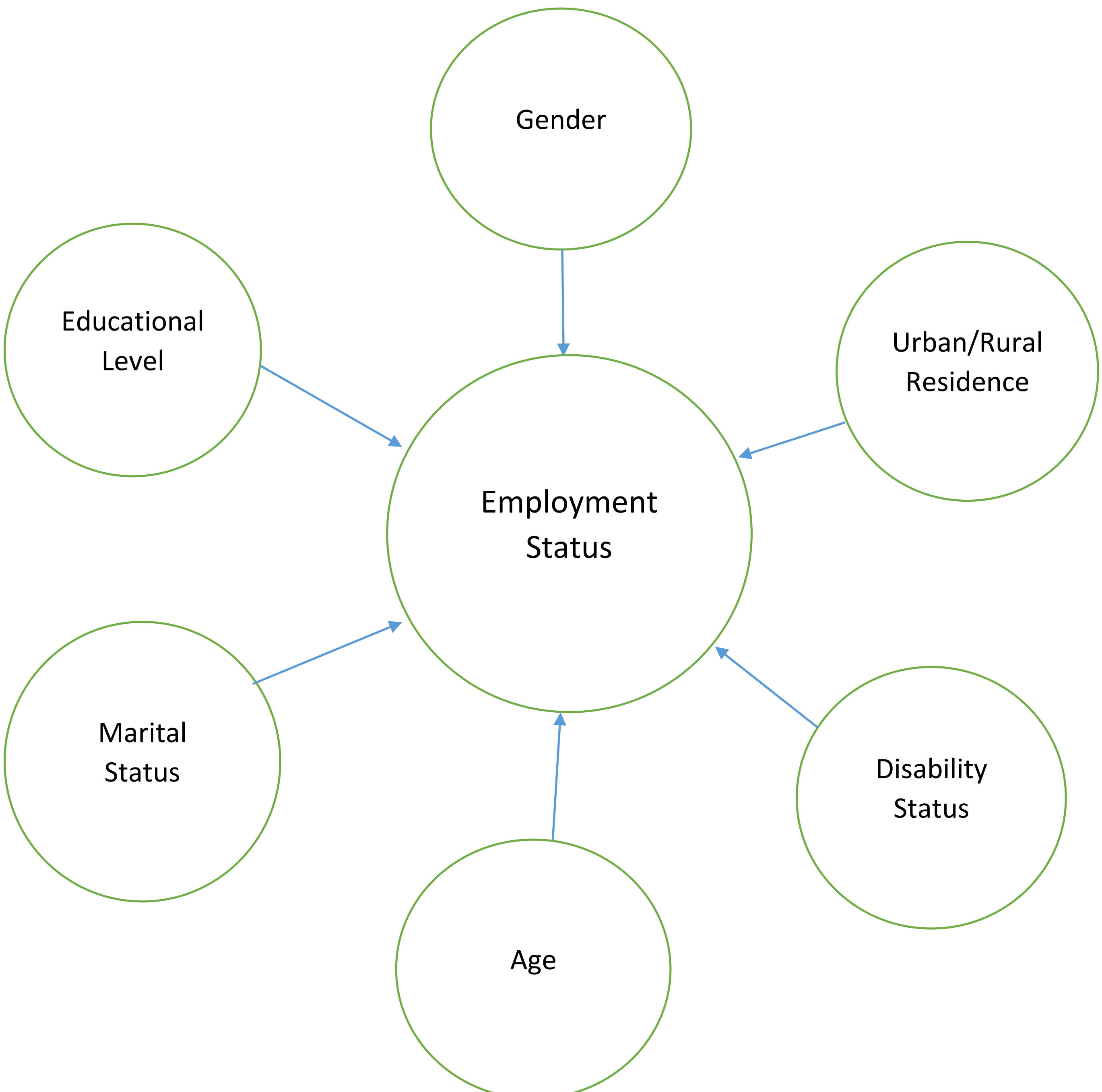

Figure (5): Conceptual Design for the First Model

2.3.2 For the second model

In the second model the response variable is the wage and 9 explanatory variables which are gender, urban/rural residence, disability status, age, marital status, educational level, employment stability, total weekly working hours in the main job and coronavirus caused a decrease in wage. The sample size in this model is 18735

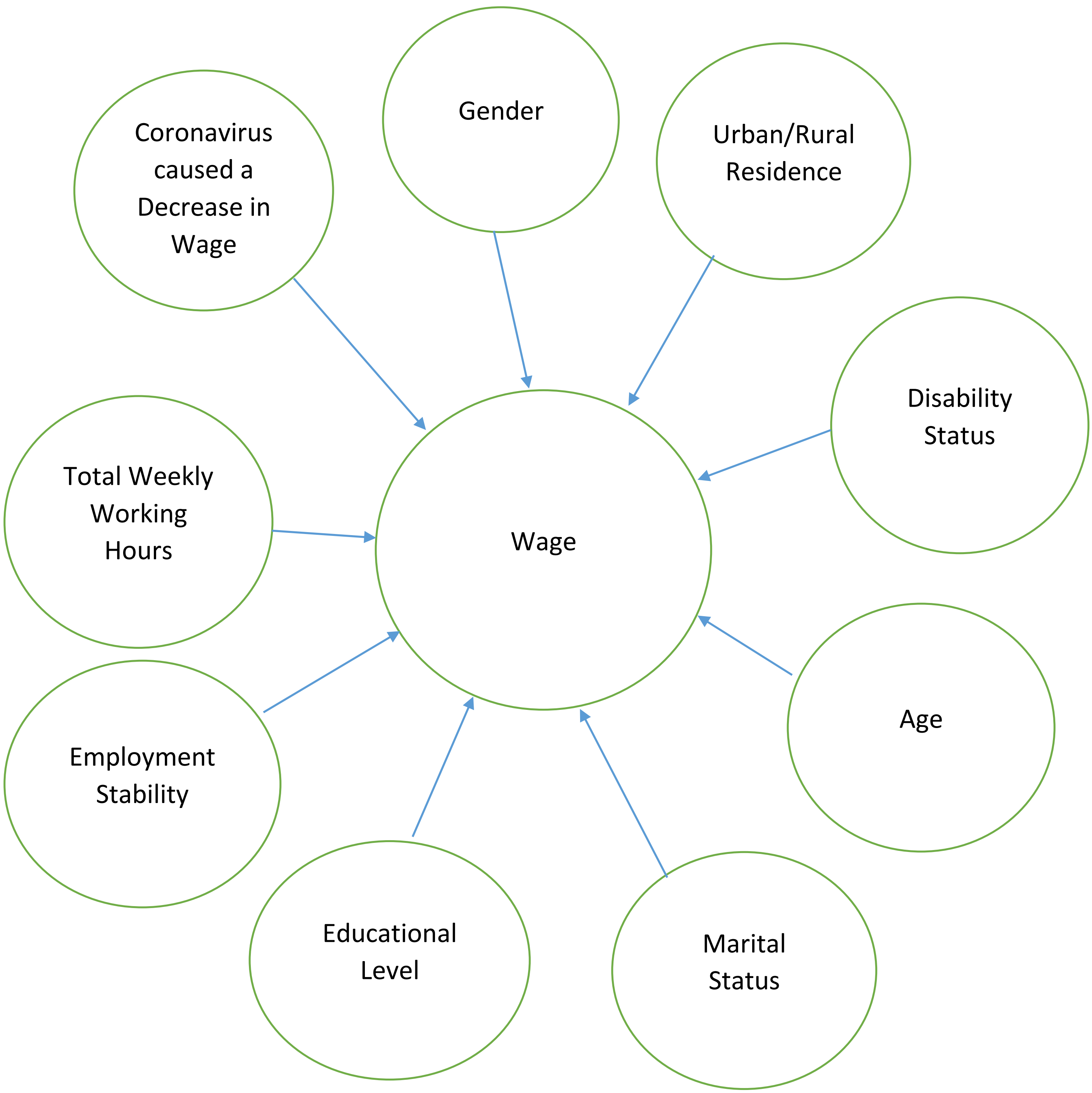

Figure (6): Conceptual Design for the Second Model

Chapter 3
Descriptive Analysis

The primary goal of this study is to investigate and comprehend the differences in the Egyptian labor force based on gender, urban/rural residency, and disability. The project's goal is to assess the scope of these gaps, investigate their causes and consequences.

We have 2 response variables which are wage and employment status, and 9 explanatory variables which are gender, urban/rural residence, disability status, age, marital status, educational level, employment stability, total weekly working hours in the main job, and coronavirus caused a decrease in wage/income.

3.1 Response Variables

- **Employment Status**

This variable includes 2 categories which are employed and unemployed; the unemployed category involves unemployed, homemaker, student, pensioners, retired, disabled and others. The following figure illustrates that the percentage of employed participants is 41.20% which is lower than their unemployed counterparts which is 58.80%.

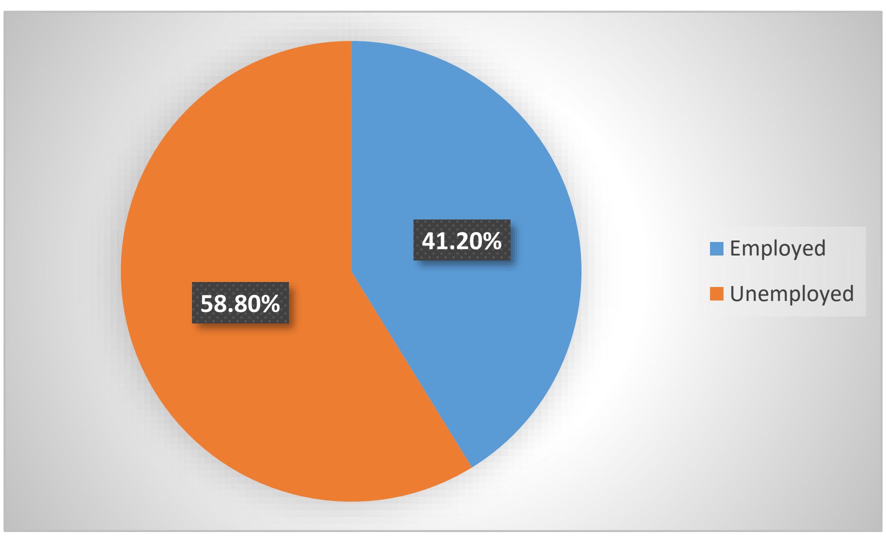

Figure (7): Distribution of participants according to their employment status

- **Wage**

Table (1): Descriptives of wage variable along with total monthly income/wage from various job types variables.

	N	Minimum	Maximum	Mean	Std. Deviation
Wage	18735	1000.00	6000.00	3158.2481	1081.75053
Total monthly wage from the regular main job	10373	1200.00	6500.00	3235.6154	1147.15701
Total monthly income(for employers) from the main job	565	1000.00	8000.00	3602.1894	1376.72757
Total monthly income (for self-employed) from the main job	3297	600.00	5304.00	2678.3497	1056.98684
Total monthly wage from seasonal/irregular jobs	3490	1500.00	6000.00	3292.4912	974.62485

Wage has 18,735 observations that corresponds to participants answered employed in the employment status variable, with a minimum wage of 1000.00 pounds and a maximum of 6000.00 pounds. The mean wage is approximately 3158.25 with a standard deviation of 1081.75. This variable was created using responses from related questions. Specifically, if a respondent answered one of the questions related to total monthly income/wage from various job types, their response was used to compute the variable.

The variables and there descriptives:

1. Total monthly wage from the regular main job: With 10,373 observations, this variable ranges from 1200.00 to 6500.00 pounds, with a mean of approximately 3235.62 and a standard deviation of 1147.16.
2. Total monthly income (for employers) from the main job: This variable has 565 observations, ranging from 1000.00 to 8000.00 pounds. The mean income is around 3602.19, with a standard deviation of 1376.73.
3. Total monthly income (for self-employed) from the main job: With 3,297 observations, this variable varies from 600.00 to 5304.00 pounds. The mean income is approximately 2678.35, with a standard deviation of 1056.99.

4. Total monthly wage from seasonal irregular job: This variable encompasses 3,490 observations, with values ranging from 1500.00 to 6000.00 pounds. The mean wage is approximately 3292.49, with a standard deviation of 974.62.

We then categorized wage variable as follows: 1000-2500, 2501-4500, 4501-6000. The following figure illustrates that participants that has wages from 2501-4500 represents the highest wage group (61.5%) followed by participants that has wages from 1000-2500(30.2%) and the lowest wage group is 4501-6000(8.3%).

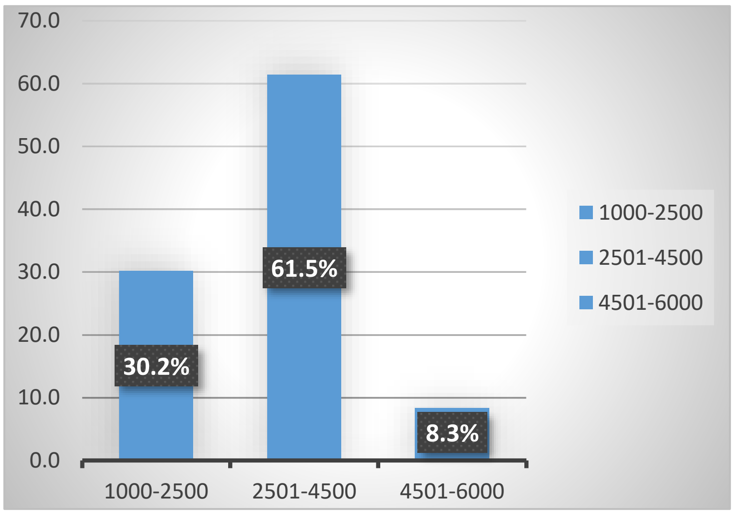

Figure (8): Distribution of participants according to wage categories

3.2 Explanatory Variables

This part will include the descriptives of the 9 explanatory variables which are gender, urban/ rural residence, disability status, age, marital status, educational level, employment stability, total weekly working hours in the main job and coronavirus caused a decrease in wage.

- **Gender**

The following figure illustrates that the percentage of female participants is 50.40% which is higher than their male counterparts which is 49.60%.

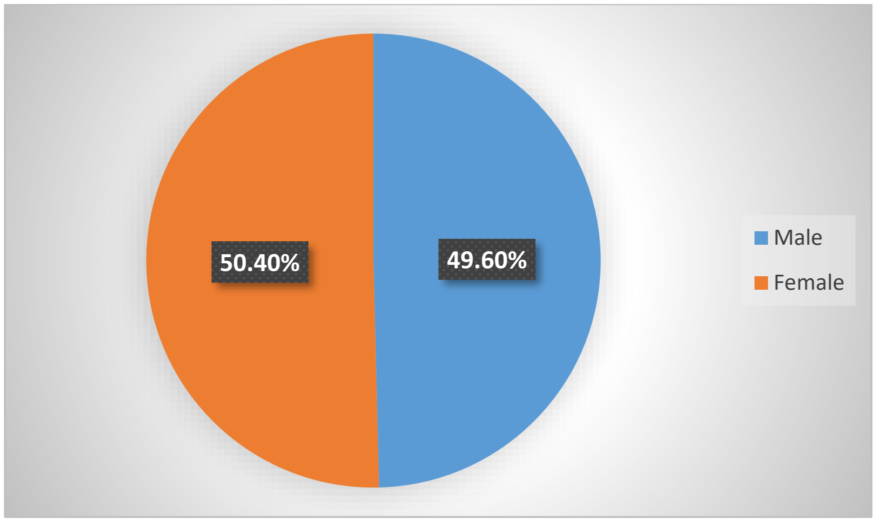

Figure (9): Distribution of participants according to their gender

- **Urban/Rural Residence**

The following figure illustrates that the percentage of urban participants is 58.40% which is higher than their rural counterparts which is 41.60%.

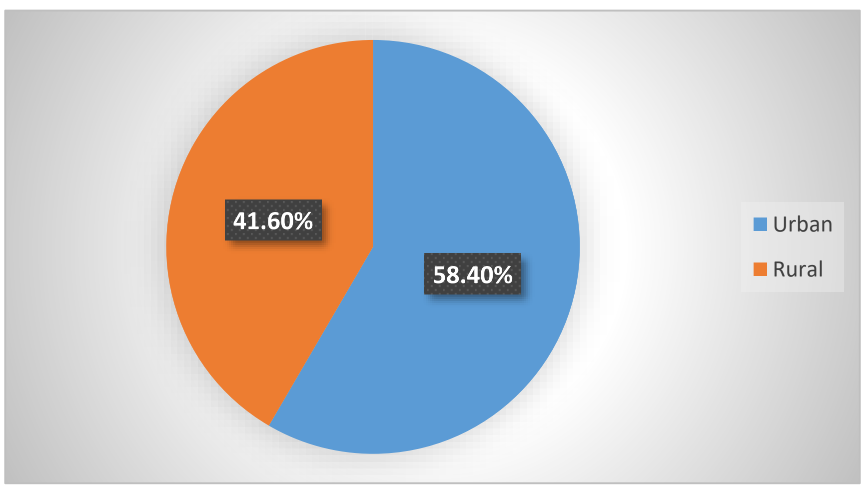

Figure (10): Distribution of participants according to their urban/rural residence.

- **Disability Status**

The following figure illustrates that the percentage of disabled participants is 8% which is lower than their non-disabled counterparts which is 92%.

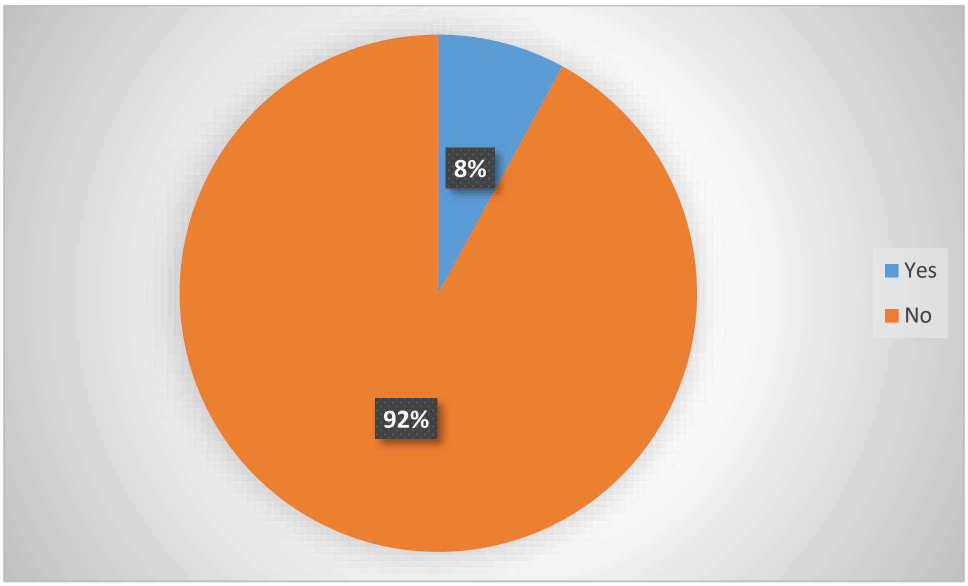

Figure (11): Distribution of participants according to their disability status

- **Age**

Table (2): Descriptives of age variable.

N	Minimum	Maximum	Mean	Std. Deviation
45453	18	101	40.61	16.036

There are 45,453 observations, age ranges from 18 to 101 years old, with a mean of 40.61 years and a standard deviation of 16.04.

We then categorized the age variable as follows: 18-30, 31-50, 51-70, 71-101.The following graph depicts that middle adults between the ages of 31-50 represent the highest age group (40.04%), followed by people between the ages of 18-30 (32.61%) while the elderly group represents the lowest age group (4.3%).

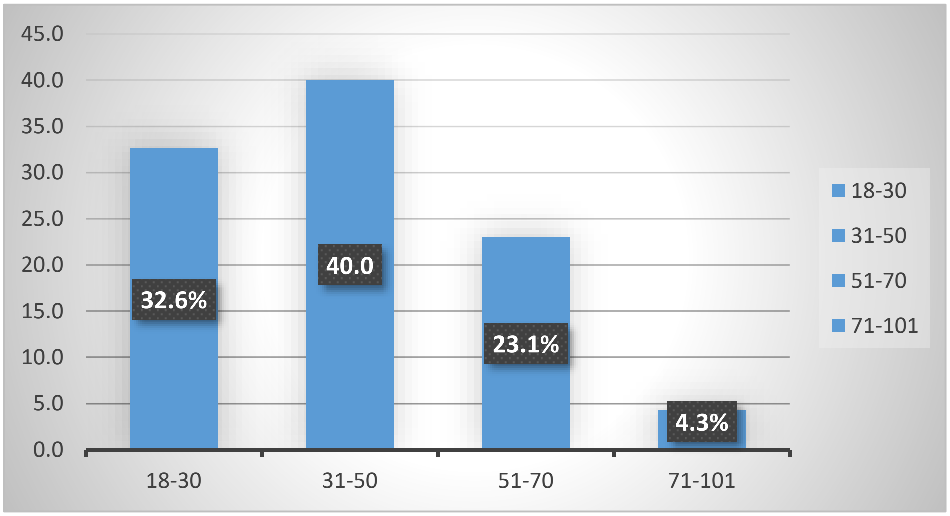

Figure (12): Distribution of participants according to their age groups.

- **Educational Level**

The following figure illustrates that 29.90% of the participants chose none, 15.50% among survey participants are from primary/lower secondary, 37.20% of the participants are from secondary, 3.20% of the participants are from post-secondary or equivalent, 14.60% of the participants are from university, while 0.30% of the participants are from postgraduate.

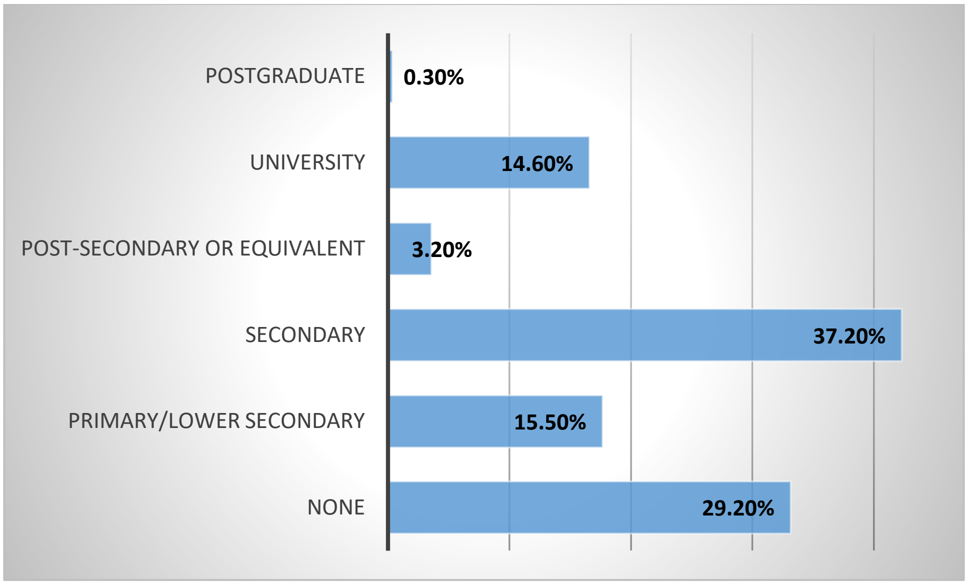

Figure (13): Distribution of participants according to their educational level.

- ## Marital Status

The following figure illustrates that 23.60% among survey participants are never married, 66.6% of the participants are married, while 9.80% of the participants are widowed or divorced.

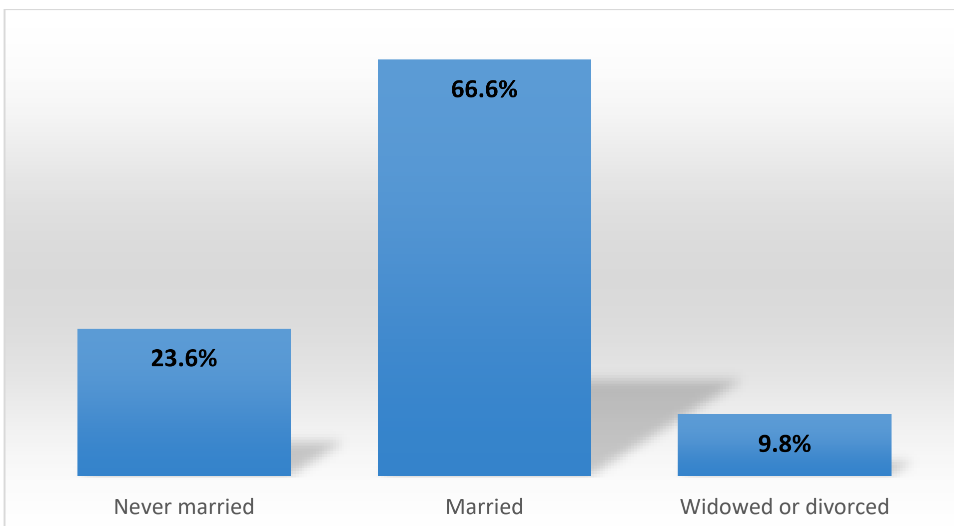

Figure (14): Distribution of participants according to their marital status

- ## Employment Stability in the Main Job

This variable is answered among only 18735 of the survey participants as these participants answered employed in the employment status so they are applicable for this question. The following figure illustrates that 72% among employed participants are involved in full time/regular job, 7.8% of the employed participants are involved in part time/temporary job, while 20.3% of the employed participants are involved in seasonal/irregular job.

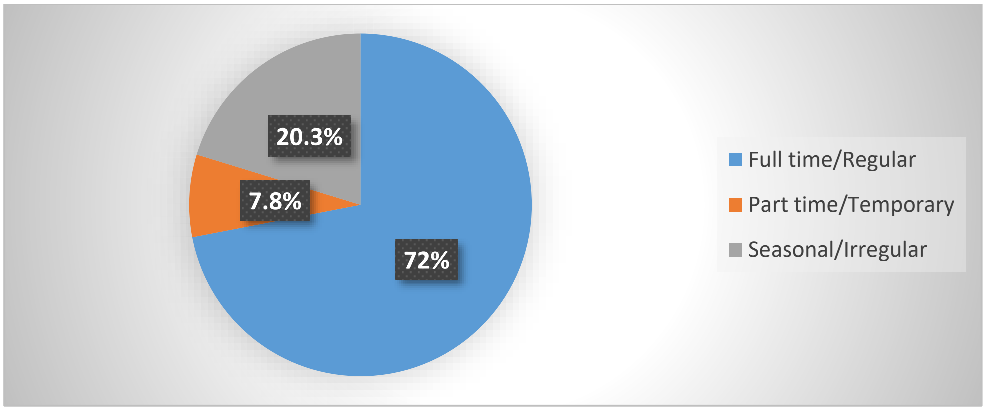

Figure (15): Distribution of participants according to their employment stability in the main job.

- **Total Weekly Working Hours in the Main Job**

Table (3): Descriptives of total weekly working hours variable.

N	Minimum	Maximum	Mean	Std. Deviation
18735	15.00	84.00	45.4777	12.64842

There are 18735 observations for this variable that corresponds to participants answered employed in the employment status variable, with working hours ranging from 15.00 to 84.00. The mean is approximately 45.47 hours per week, with a standard deviation of 12.64.

- **Coronavirus caused a decrease in wage/income.**

The following figure illustrates that the percentage of participants answered employed in the employment status variable participants who chose yes is 13.7% which is lower than their employed counterparts who chose no which is 86.3%.

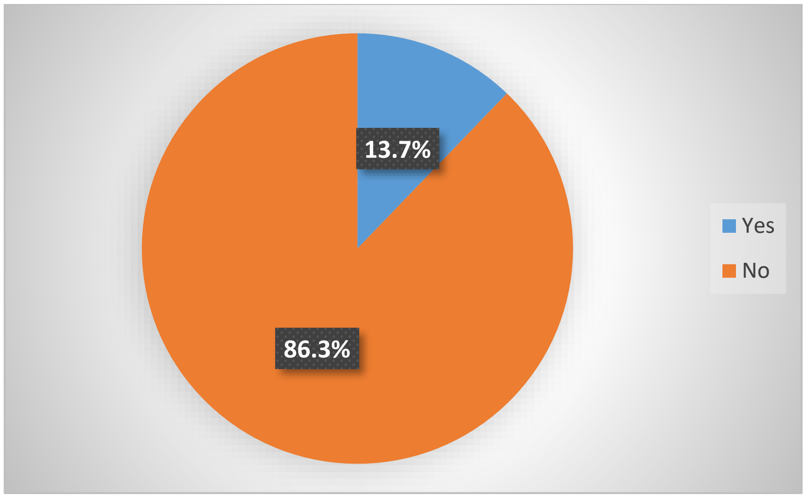

Figure (16): Distribution of participants according to coronavirus cased a decrease in their wage/income or not

Chapter 4
Bivariate Relationships

Measures of association are statistical tools used to examine the relationship between two variables, providing insights into the strength and degree of their association.

This section will be divided into 2 parts:

First, we will examine the relationship between our first response variable (employment status) and the explanatory variables.

Second, we will examine the relationship between our second response variable (wage) and the explanatory variables keeping in mind with categorical explanatory variables we will get the average of wage in each category and t-test or analysis of variance "ANOVA" will be computed.

Measures of association we will be using between two categorical variables in our report: odds ratio, gamma and contingency coefficient according to the types of variables (binary, nominal or ordinal)

(The tests are conducted under significance level of $\alpha = 5\%$)

Measure of association we will be using between two continuous variables in our report: Pearson correlation coefficient.

(This test is conducted under significance level of $\alpha =5\%$)

4.1 Measures of Association for Employment Status and its Determinants

This part will include the measures of association of the response variable employment status and its 6 explanatory variables which are gender, urban rural residence, disability status, marital status, educational level and age

- **Employment status and gender**

The figure and table show that, for males, 6,615 were unemployed with 24.8%, while 15,935 were employed with 85.1%. On the other hand, for females, 20,103 were unemployed with 75.2%, while only 2,800 were employed with 14.9%.

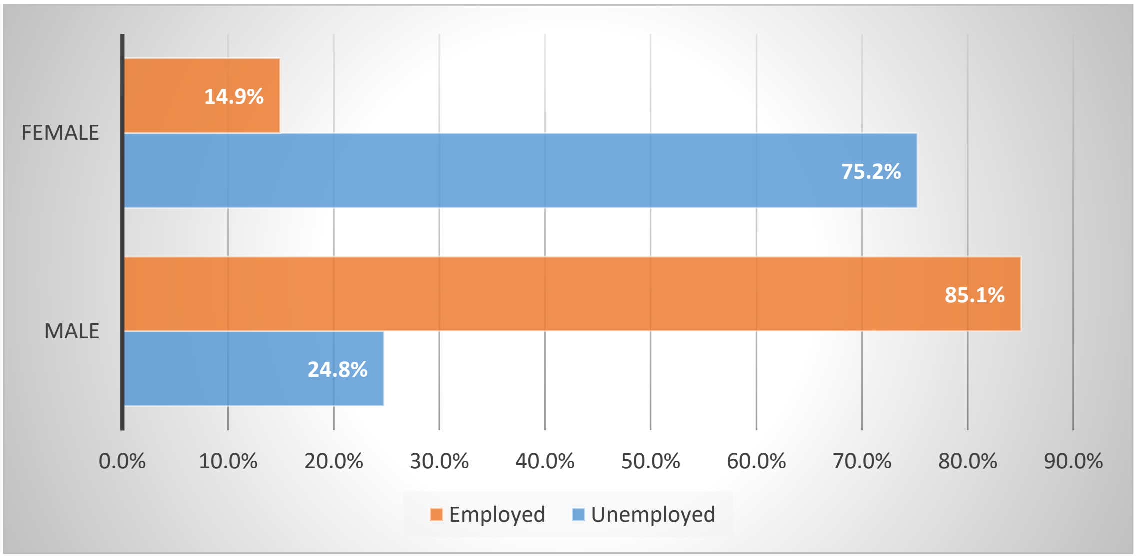

Figure (17): Bar chart for gender by employment status

Table (4): 2-way contingency table between employment status and gender

		Employment Status		
		Unemployed	Employed	Total
Gender	Male	**6615**	**15935**	**22550**
	Female	**20103**	**2800**	**22903**
Total		**26718**	**18735**	**45453**

- There is a significant relationship between employment status and gender from the confidence interval of odds ratio while it does not include 1 (0.055,0.061)

- **Odds ratio=0.058:** this means that the unemployment rate for females is higher than males.

- **Employment status and urban/rural residence**

From this figure and table, it can be observed that out of the rural residents 15,237 were unemployed with 57% and 11,309 were employed with 60.4%. For urban residents, 11,481 were unemployed with 43% and 7,426 were employed with 39.6%.

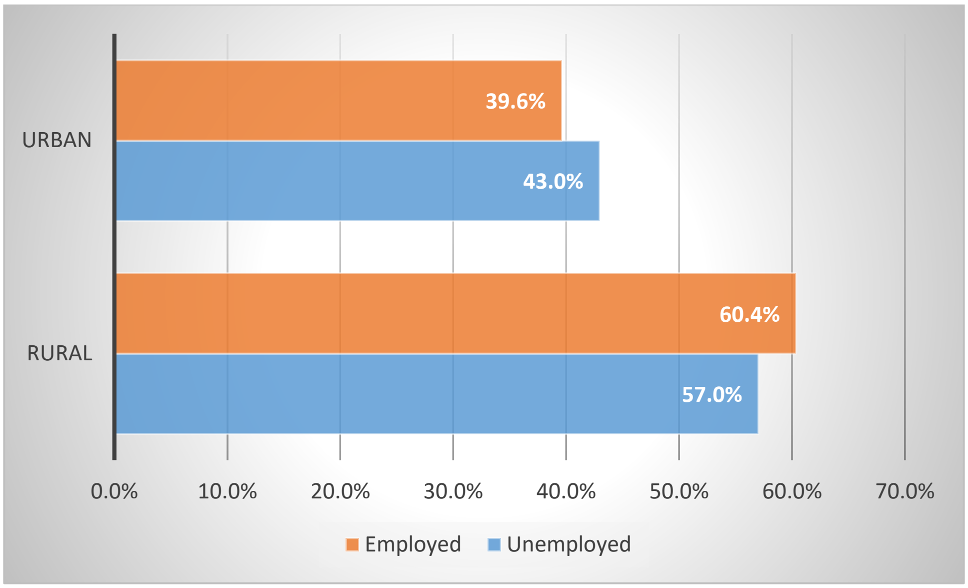

Figure (18): Bar chart for urban/rural residence by employment status.

Table (5): 2-way contingency table between employment status and urban/rural residence

		Employment status		
		Unemployed	Employed	Total
Urban/Rural residence	Rural	**15237**	**11309**	**26546**
	Urban	**11481**	**7426**	**18907**
Total		**26718**	**18735**	**45453**

- There is a significant relationship between employment status and urban/rural residence from confidence interval of odds ratio while it does not include 1 (0.839,0.905)
- **Odds ratio=0.871:** The unemployment rate for people living in urban areas is higher than those living in rural areas.

- **Employment status and disability status**

The figure and table shows that for individuals without disabilities, 23,884 were unemployed with 89.4% and 17,951 were employed with 95.8%. On the other hand, for individuals with disabilities, 2,834 were unemployed with 10.6% and only 784 were employed with 4.2%.

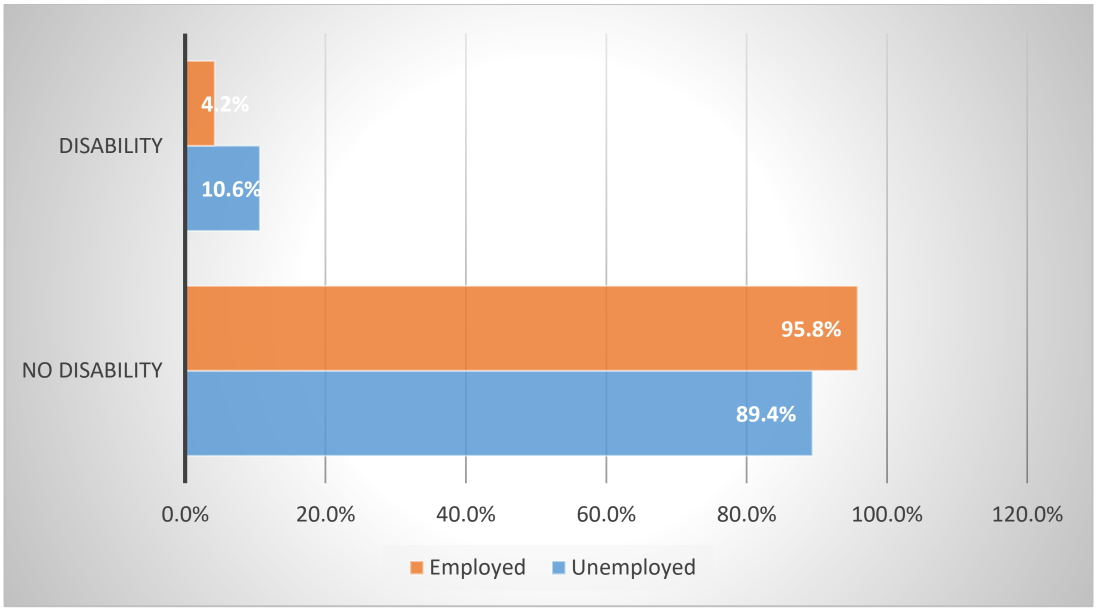

Figure (19): Bar chart for disability status by employment status

Table (6): 2-way contingency table between employment status and disability status

		Employment status		
		Unemployed	Employed	Total
Disability status	No	**23884**	**17951**	**41835**
	Yes	**2834**	**784**	**3618**
Total		**26718**	**18735**	**45453**

- There is a significant relationship between employment status and disability status from confidence interval of odds ratio while it does not include 1 (0.339,0.399)

- **Odds ratio=0.368:** The unemployment rate for people with disabilities is higher for those without disabilities.

- **Employment status and educational level**

Table (7): 2-way contingency table between employment status and educational level

		Employment status		Total
		Unemployed	Employed	
Educational level	None	9025	4233	13258
	Primary/Lower secondary	4342	2723	7065
	Secondary	9553	7362	16915
	Post secondary or equivalent	655	779	1434
	University	3094	3542	6636
	Postgraduate	49	96	145
Total		26718	18735	45453

- There is a significant relationship between employment status and educational level as p-value of gamma= 0.000 < 0.05.
- **Gamma=0.231:** this indicates a positive weak relationship; more educated individuals tend to be employed.

- **Employment status and marital status**

Table (8): 2-way contingency table between employment status and marital status

		Employment Status		Total
		Unemployed	Employed	
Marital Status	Never married	6535	4170	10705
	Married	16437	13834	30271
	Widowed or divorced	3746	731	4477
Total		26718	18735	45453

- There is a significant relationship between employment status and marital status as p-value of contingency coefficient= 0.000 < 0.05.
- **Contingency coefficient =0.174:** this indicates a weak association between employment status and marital status.

4.2 Measures of Association between Wage and its determinants

This part will include the measures of association of the response variable wage and its 9 explanatory variables which are gender, urban rural residence, disability status, age, marital status, educational level, employment stability, total weekly working hours in the main job and coronavirus caused a decrease in wage.

- **Wage and gender**

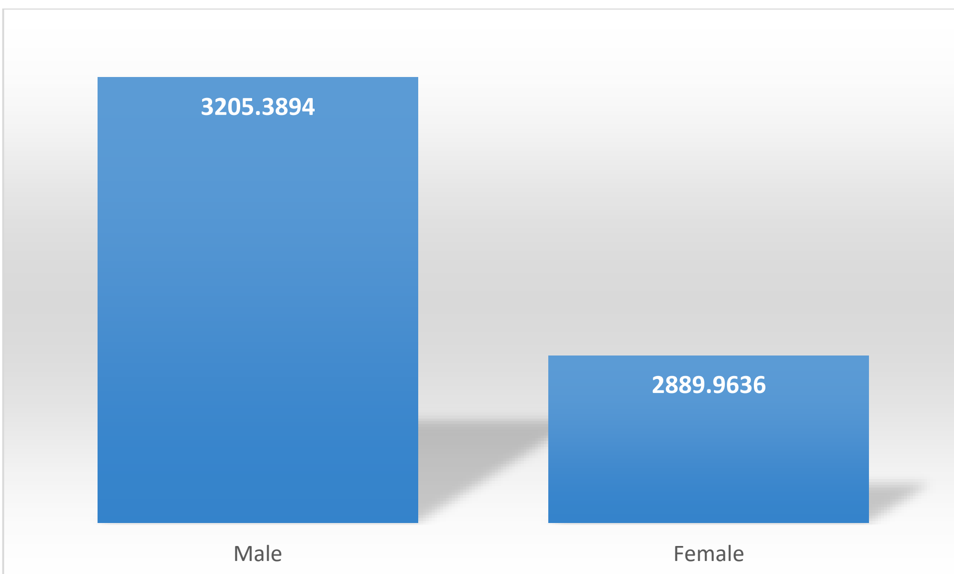

Figure (20): Bar chart for average wage by gender

- The average of wage for males is 3205.4 EGP, while for females is 2889.96 EGP, and this difference is significant such that t-test: 14.307, and p-value associated with the t-test = 0.000<0.05.

- **Wage and urban/rural residence**

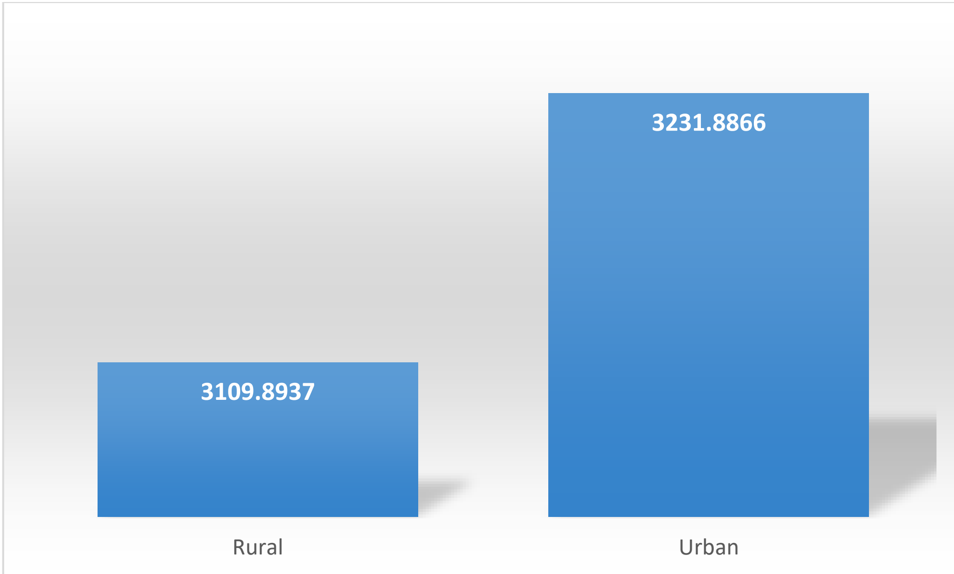

Figure (21): Bar chart for average wage by urban/rural residence

- The average wage for rural is 3109.8937 EGP, while for urban is 3231.8866 EGP, and this difference is significant such that t-test: -7.313, and p-value associated with the t-test=0.000<0.05

- **Wage and disability status**

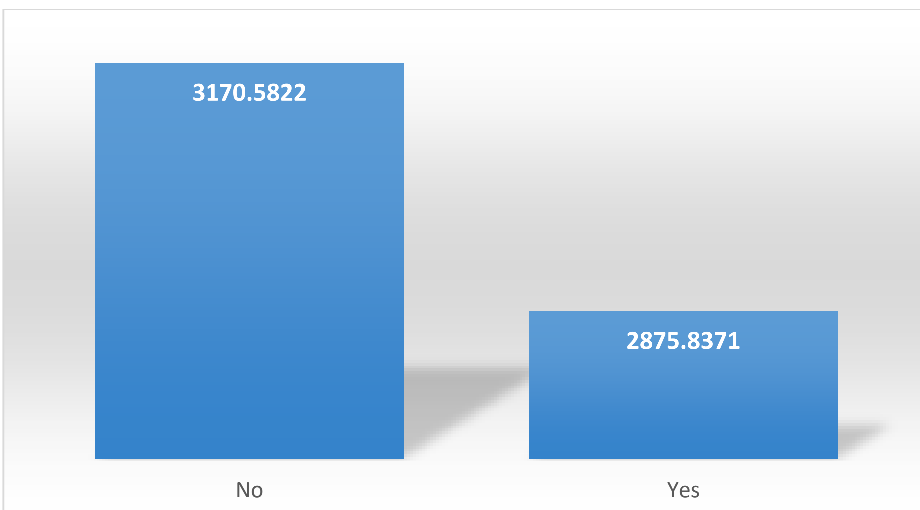

Figure (22): Bar chart for average wage by disability status

- The average wage for people with no disability is 3170.5822 EGP, while for people with disability is 2875.8371, and this difference is significant such that t-test: 7.650, and p-value associated with the t-test=0.000<0.05

- **Wage and age**

- There is a significant relationship between wage and age as p-value of Pearson correlation coefficient= 0.000 < 0.01.
- **Pearson correlation coefficient= 0.073:** this indicates a positive weak relationship; on average as age increases by 1 year, wage tends to increase by 0.073 units of wage.

- **Wage and educational level**

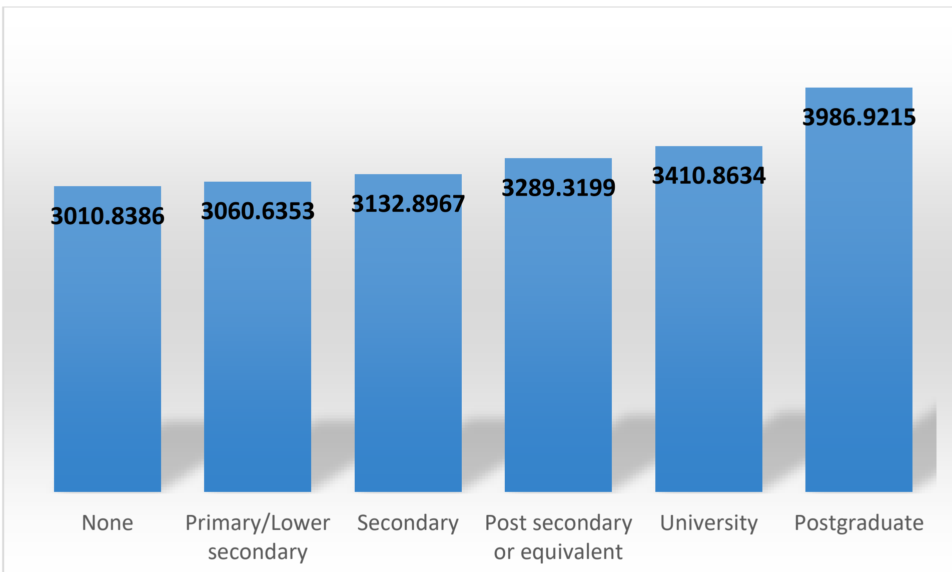

Figure (23): Bar chart for average wage by educational level

- The average wage for people not educated is 3010.8386 EGP, people from primary/lower secondary is 3060.6353, people from secondary is 3132.8967, people from post secondary or equivalent is 3289.199, people from university is 3410.8634 EGP, while people from postgraduate is 3986.9215 EGP, and this difference is significant as F-test computed from analysis of variance “ANOVA”: 74.587, and p-value associated with the F-test=0.000<0.05

- **Wage and marital status**

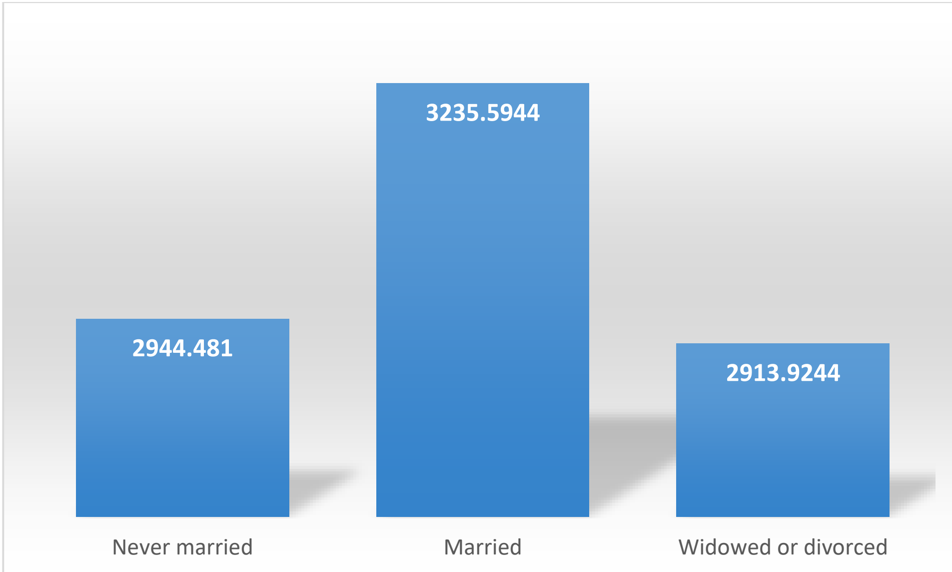

Figure (24): Bar chart for average wage by marital status

- The average wage for people who are never married is 2944.481 EGP, people who are married is 3235.5944 EGP, while people who are widowed or divorced is 2913.9244 EGP, and this difference is significant as F-test computed from analysis of variance "ANOVA": 137.400, and p-value associated with the F-test=0.000<0.05

- **Wage and employment stability in the main job**

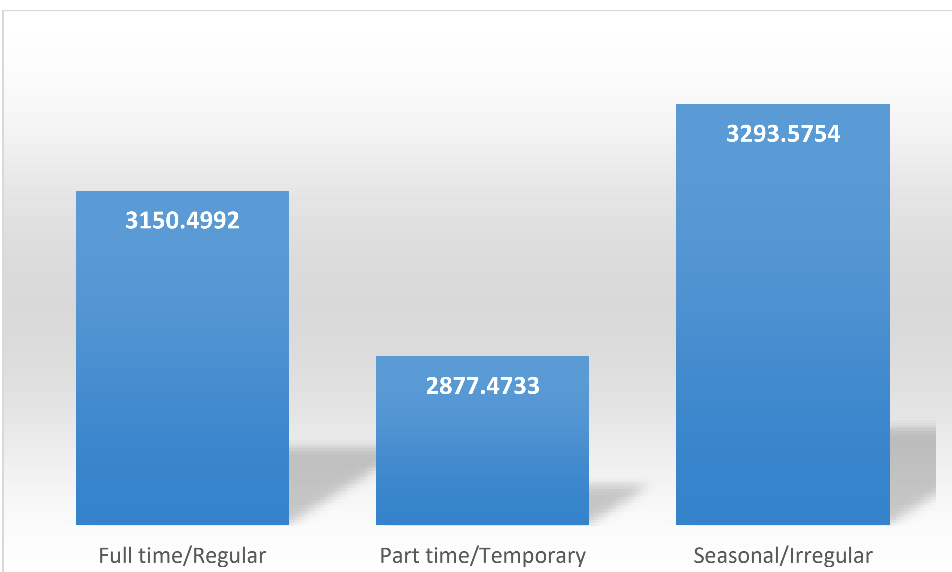

Figure (25): Bar chart for average wage by employment stability

- The average wage for people from full time/regular job is 3150.4992 EGP, people form part time/temporary job is 2877.4733 EGP, while from people from seasonal/irregular job is 3293.5754 EGP, and this difference is significant as f-test computed from analysis of variance “ANOVA”: 79.786. and p-value associated with F-test=0.000<0.05
-

- **Wage and total weekly working hours**

- There is a significant relationship between wage and total weekly working hours as p-value of Pearson correlation coefficient= 0.000 < 0.01.
- **Pearson correlation coefficient= 0.110:** this indicates a positive weak relationship; on average as total weekly working hours increase by 1 unit, wage tends to increase by 0.110 units of wage.

- **Wage and coronavirus caused a decrease in wage/income**

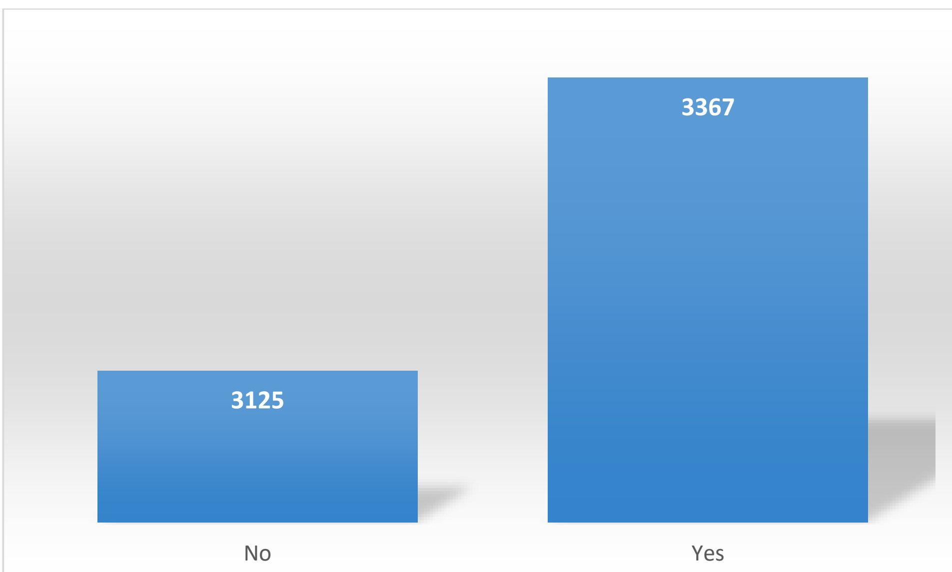

Figure (26): Bar chart for average wage by coronavirus caused a degree in wage/income.

- The average wage for people who did not face a decrease in wage/income from coronavirus is 3125 EGP, while for people who faced a decrease in wage/income from coronavirus is 3367 EGP, and this difference is significant such that t-test: -10.573, and p-value associated with t-test=0.000<0.05

Correlation matrix between continuous variables:

Table (9): Correlation matrix for continuous variables

		Age	Wage	Total weekly working hours in the main job
Age	Pearson Correlation	1	.073**	-.057**
	Sig. (2-tailed)		.000	.000
	N	45453	18735	18735
Wage	Pearson Correlation	.073**	1	.110**
	Sig. (2-tailed)	.000		.000
	N	18735	18735	18735
Total weekly working hours in the main job	Pearson Correlation	-.057**	.110**	1
	Sig. (2-tailed)	.000	.000	
	N	18735	18735	18735

Chapter 5

Model 1: Binary Logistic Regression Analysis of Employment Status

5. Model 1: Binary Logistic Regression Analysis of Employment Status

This chapter includes the analysis that investigates the factors influencing employment status within a given reference period using the data from the Labor Force Survey (LFS) 2022. By examining a comprehensive dataset, we aim to understand how gender, urban/rural residence, disability status, marital status, educational level, and age contribute to the likelihood of being employed or unemployed. This study employs logistic regression models to analyze the associations between these explanatory variables and employment status. Specifically, we develop two models: one considering only the main effects and another incorporating interaction terms to capture potential synergistic effects between variables.

Before, employing the binary logistic regression it is worth mentioning that we restricted the age range **of participants to 18 to 69 years**. This age restriction ensures that our analysis focuses on the working-age population, thereby providing a more accurate understanding of employment dynamics. These changes were made to enhance the clarity and relevance of our findings. (n=43,049) will be our sample size when employing the binary logistic regression model.

The 2 models will contain the employment status as the response variable with 2 categories, 0=unemployed and 1=employed. And these are the explanatory variables with their categories and reference category:

Table (10): Reference categories of independent variables of model 1.1

Independent Variable	**Categories**	**Reference Category**
Gender	Binary nominal: 1=Male 2=Female	Female
Urban/Rural Residence	Binary nominal: 0=Rural 1=Urban	Urban
Disability Status	Binary nominal: 0= No 1=Yes	Yes

Marital Status	Nominal: 1=Never married 2=Married 3= Divorced/Widowed	Divorced/Widowed
Educational Level	ordinal: 1=None 2=Primary/Lower secondary 3=Secondary 4=Post secondary or equivalent 5=University 6=Postgraduate	Postgraduate
Age	Continuous:18-69 years	

5.1. Logistic Regression Analysis of Employment: Main Effects Only (model1.1)

Research Questions Addressed in this model:

1. Gender-based Disparities: The analysis explores how gender influences employment status in the Egyptian labor force. It examines whether significant disparities exist between males and females, considering employment rates.
2. Urban vs. Rural Disparities: By comparing urban and rural areas in Egypt, the study assesses how residency location affects labor force participation. It seeks to identify differences in employment opportunities.
3. Impact of Disability: The analysis investigates the extent to which disability affects labor force outcomes in Egypt. It examines whether individuals with disabilities face barriers in accessing employment opportunities compared to their non-disabled counterparts.

Model Overview: This logistic regression model without interaction terms evaluates the odds ratios and statistical significance of independent variables—gender, marital status, disability status, educational level, age, and urban/rural residence—in predicting employment status during the reference period. The dataset comprises 43,049 observations from the Labor Force Survey (LFS) 2022, providing a robust foundation for analyzing these relationships.

The logistic regression equation for Model 1.1 is:

$$\log\left(\frac{p(employed)}{1-p(employed)}\right) = \beta 0 + \beta 1(gender) + \beta 2(\ residence) + \beta 3(disability\ status) + \beta 4(marital\ status) + \beta 5(educational\ level) + \beta 6(age)$$

Before running the model, the absence of multicollinearity must be checked first:

Table (11): Multicollinearity check for model 1.1

Model		Collinearity Statistics	
		Tolerance	VIF
	(Constant)		
	Marital status	.650	1.537
	Gender	.915	1.093
	Age	.658	1.521
	Urban/Rural residence	.956	1.047
	Educational level	.897	1.115
	Disability status	.945	1.058

-All the VIF values are comfortably below 10, indicating a low risk of multicollinearity. This is a positive sign, suggesting that the predictor variables (marital status, rural/urban residence, educational level, disability status, sex, and age) are relatively independent of each other.

Testing prediction power of the model:

To assess the predictive power of the model, several metrics are utilized. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) serves as a comprehensive measure of the model's ability to discriminate between employed and unemployed individuals based on the selected predictors. Additionally, sensitivity and specificity metrics are employed to evaluate the model's ability to correctly classify true positive (employed) and true negative (unemployed) cases, respectively. Furthermore, overall accuracy provides an aggregate measure of the model's predictive performance across all classifications. These metrics collectively enable an evaluation of the model's reliability in predicting labor force outcomes based on gender, urban/rural residence, disability status, and other key variables pertinent to the Egyptian context.

- Roc curve and AUC:

We first wanted to know the optimal cutoff point using sensitivity/specificity vs cutoff points graph as follows:

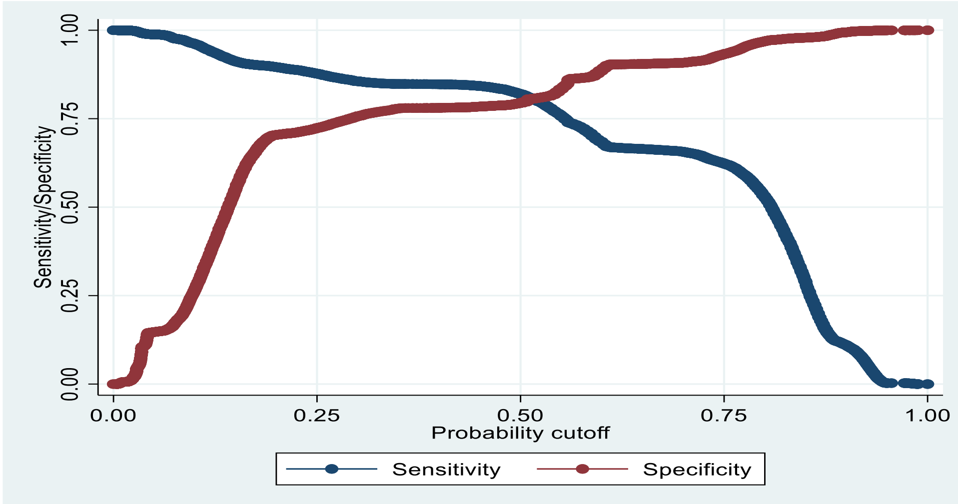

Figure (27): Plot fitting sensitivity & specificity against cutoff point for model 1.1

From figure (27), we can determine the best cutoff point is 0.526.

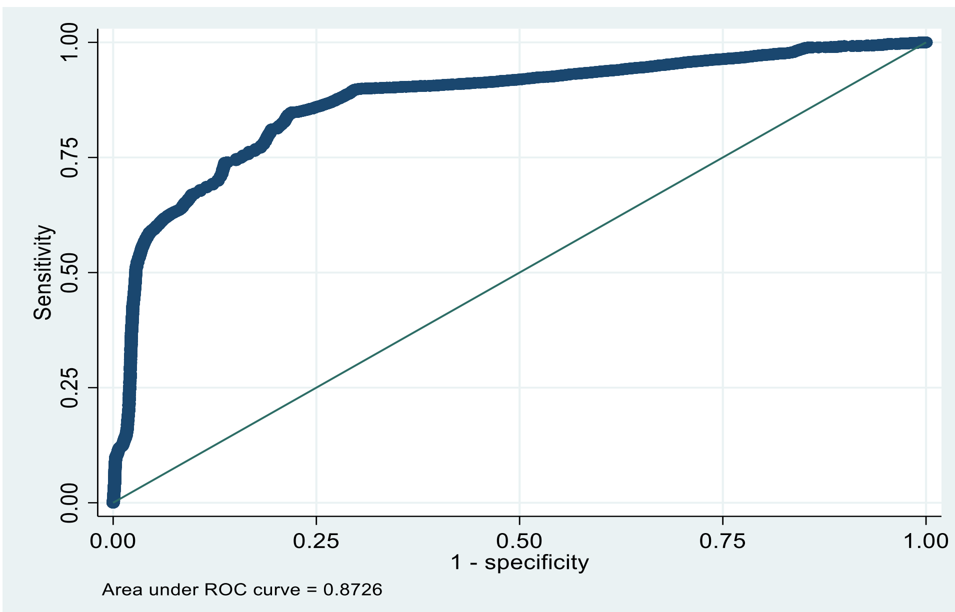

Figure (28): ROC Curve for model 1.1.

A ROC curve is drawn and from the graph we can notice that it is close to the upper left corner which means the model has high overall accuracy. We can also deduce that the model has a good predicative power, as the AUC is equal to 0.8726 (87.26%) which indicates that the model (the predictions) has high true positive rate.

- Classification table (overall correct classification, sensitivity& specificity):

Observed		Predicted		
		Employment Status		Percentage Correct
		Unemployed	Employed	
Employment Status	**Unemployed**	**19810**	**4639**	**81.0**
	Employed	**3816**	**14784**	**79.5**
Overall Percentage				**80.4**

Table (12): Classification table for model 1.1

Sensitivity =0.7948 Specificity=0.8103.

The sensitivity is 79.48 % which is the probability of correctly specifying employed people while the specificity is 81.03% which is the probability of correctly specifying unemployed people. The percentage of correct classification is 80.4%. since the sensitivity, specificity & percentage of correct classification are all greater than 60%, therefore the model has acceptable predictive power.

Based on the comprehensive evaluation of the logistic regression model's predictive performance, including its high AUC-ROC of 0.8726 (87.26%), sensitivity of 79.48%, specificity of 81.03%, and overall correct classification rate of 80.4%, it is evident that the model is well-suited for analyzing labor force outcomes in the Egyptian context. These metrics indicate strong discriminatory power and accuracy in distinguishing between employed and unemployed individuals based on gender, urban/rural residence, disability status, and educational attainment.

Fitting the model:

Omnibus Tests of Model Coefficients:

Chi square value= 21344.999 df = 11 p value= .000

- The high LR chi2 value (21344.999) and p-value<0.05 indicate that the model significantly improves the fit compared to a null model, suggesting that the predictors collectively contribute significantly to the model.

Table (13): Model 1.1 coefficients

Variable	Odds Ratio	Std. Error	z-Value	P>z	95% Confidence Interval	
Sex= Male	35.21373	1.137812	110.22	0.000	33.05281	37.51593
Rururb= Rural	1.220707	.0333547	7.30	0.000	1.157053	1.287863
Disable= No	3.642911	.2065373	22.80	0.000	3.259786	4.071065
mart2= Never Married	.1796876	.0135222	-22.81	0.000	.1550464	.2082449
mart2= Married	1.010445	.0596163	0.18	0.860	.9001019	1.134315
Educ= None	.0814018	.0206883	9.87	0.000	.0494652	.1339577
Educ= Primary/Lower Secondary	.0681486	.0173748	-10.54	0.000	.0413465	.1123246
Educ= Secondary	.0838998	.0212861	-9.77	0.000	.0510274	.137949
Educ=Post Secondary or Equivalent	.1704232	.0447231	-6.74	0.000	.1018951	.2850389
Educ= University	.202438	.0515198	-6.28	0.000	.1229317	.33336
Age	.9822666	.0013242	-13.27	0.000	.9796746	.9848655
_cons	.9088787	.2493797	-0.35	0.728	.5308255	1.556181

Interpreting parameters:

1. Gender-based Disparities

How does gender influence employment status in the Egyptian labor force?

Model Findings: The odds ratio for gender (being male) is 35.21373, which is statistically significant ($p < 0.05$). This indicates that males have significantly higher odds of being employed compared to females. Specifically, the estimated odds of being employed for males are about 35.21 times higher than for females, holding other variables constant, and with confidence 95%

2. Urban vs. Rural Disparities

How does residency location (urban vs. rural) affect labor force participation?

Model Findings: The odds ratio for rural residence is 1.220707, which is statistically significant ($p < 0.05$). This suggests that individuals living in rural areas have higher odds (about 22.07% higher) of being employed compared to those living in urban areas, holding other variables constant. and with confidence 95%

3. Impact of Disability

How does disability status affect labor force outcomes in Egypt?

Model Findings: The odds ratio for disability status (being non-disabled) is 3.642911, which is statistically significant ($p < 0.05$). This indicates that individuals without disabilities have significantly higher odds of being employed compared to individuals with disabilities. Specifically, the estimated odds of being employed for non-disabled individuals are about 3.64 times higher than for disabled individuals, holding other variables constant, and with confidence 95%

<u>Additional Variables:</u>

4.Marital Status: Being never married significantly decreases the odds of being employed by 82.03% compared to being widowed/divorced (OR = 0.1796876, $p < 0.05$), while being married does not show a significant difference compared to being widowed/divorced, and with confidence 95%

5.Educational Level: Higher education levels (none, primary/lower secondary, secondary, post-secondary or equivalent, university) significantly decrease the odds of being employed compared to postgraduate education, indicating that higher education levels correlate with better employment prospects, and with confidence 95%

6.Age: Each additional year of age slightly decreases the odds of being employed by 1.77% (OR = 0.9822666, $p < 0.05$), showing that younger individuals have slightly higher employment odds, and with confidence 95%.

-We can see that all dependent variables except married people are statistically insignificant at the 0.05 level.

Summary of model1.1:

The main findings of this logistic regression model (Model 1.1) are as follows:

-Gender-based Disparities: The analysis shows that gender significantly influences employment status in the Egyptian labor force. Males have significantly higher odds of being employed compared to females. The estimated odds of being employed for males are about 35.21 times higher than for females, holding other variables constant.

-Urban vs. Rural Disparities: Residency location (urban vs. rural) affects labor force participation. Individuals living in rural areas have higher odds (about 22.07% higher) of being employed compared to those living in urban areas, holding other variables constant.

-Impact of Disability: Disability status significantly affects labor force outcomes. Individuals without disabilities have significantly higher odds of being employed compared to individuals with disabilities. The estimated odds of being employed for non-disabled individuals are about 3.64 times higher than for disabled individuals, holding other variables constant.

5.2. Logistic Regression Analysis of Employment: Main Effects and interaction term (model1.2)

This section presents the logistic regression model with interaction terms, which serves to answer research questions 4 and 5. These questions focus on understanding the intersectional impacts of gender, urban/rural residency, and disability status on labor force disparities in Egypt. Specifically, we aim to:

4. Explore how gender, urban/rural residency, and disability status intersect to contribute to labor force disparities, identifying specific patterns and dynamics associated with these intersections.
5. Identify subgroups, such as disabled women in rural areas, that experience the greatest labor market disparities when considering the intersection of gender, residence, and disability status.

Model overview: The primary objective of this model is to investigate the combined effects of gender, residence, and disability status on employment outcomes. By including interaction terms, the model provides a nuanced understanding of how these factors jointly influence labor market disparities within the Egyptian context.

The <u>interaction terms</u> in this model are:

- Gender and urban/rural residence
- Gender and disability status
- Disability status and urban/rural residence
- Gender, urban/rural residence, and disability status

The logistic regression equation for Model 1.2 is:

$$\log\left(\frac{p(employed)}{1-p(employed)}\right) = \beta 0 + \beta 1(gender) + \beta 2(urban\backslash rural\ residence) + \beta 3(disability\ status) + \beta 4(marital\ status) + \beta 5(educational\ level) + \beta 6(age) + \sum \beta ij(interaction\ terms)$$

-We have previously checked the absence of multicollinearity (look at table ()), so there is no need to recheck it again.

Testing prediction power of the model:

The model's performance is assessed using metrics such as the Area Under the Receiver Operating Characteristic Curve (AUC-ROC), sensitivity, specificity, and overall accuracy. These metrics help evaluate the model's predictive power and reliability in classifying employment status based on the selected predictors and their interactions.

- Roc curve and AUC.

We will first want to know the optimal cutoff point using sensitivity/specificity vs cutoff points graph like the previous model as follows:

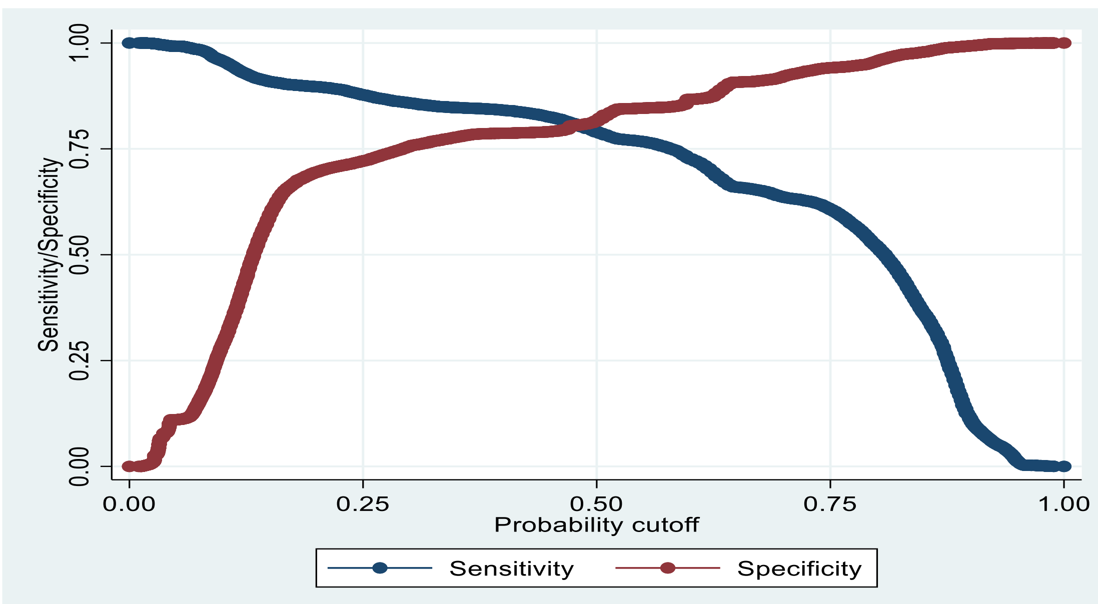

Figure (29): plot fitting sensitivity & specificity against cutoff point for model 1.2

From figure (29), we can determine the best cutoff point is 0.484.

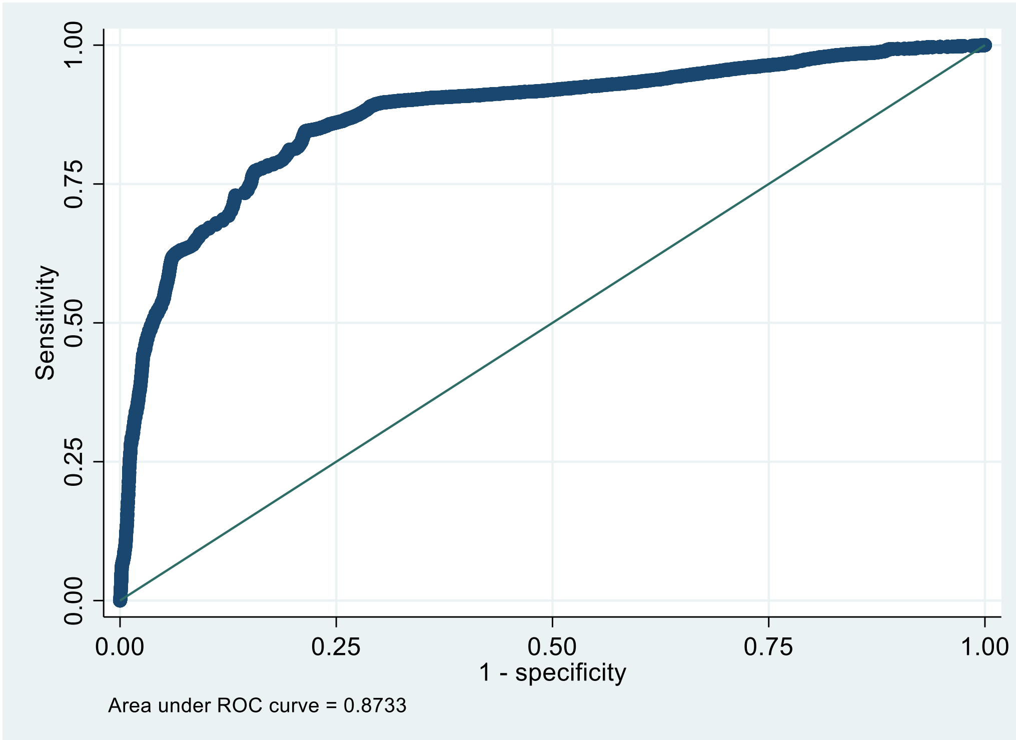

Figure (30): ROC Curve for model 1.2.

A ROC curve is drawn and from the graph we can notice that it is close to the upper left corner which means the model has high overall accuracy. We can also deduce that the model has a good predicative power, as the AUC is equal to 0.8733 (87.33%) which indicates that the model (the predictions) has a high true positive rate.

- Classification table (overall correct classification, sensitivity& specificity):

Observed		**Predicted**		
		Employment Status		**Percentage Correct**
		Unemployed	**Employed**	
Employment Status	**Unemployed**	**19735**	**4714**	**80.7**
	Employed	**3674**	**14926**	**80.2**
Overall Percentage				**80.5**

Table (14): Classification table for model 1.2

Sensitivity =0.8025 Specificity=0.8072.

The sensitivity is 80.25 % which is the probability of correctly specifying employed people while the specificity is 80.72% which is the probability of correctly specifying unemployed people. The percentage of correct classification is 80.5%. since the sensitivity, specificity & percentage of correct classification are all greater than 60%, therefore the model has acceptable predictive power.

Based on the comprehensive evaluation of the logistic regression model & predictive performance, including its high AUC-ROC, sensitivity, specificity and overall correct classification rate it is evident that the model is well-suited for analyzing labor force outcomes in the Egyptian context.

<u>Fitting the model:</u>

Omnibus Tests of Model Coefficients:

Chi square value= 21670.892 df = 15 p value= .000

- The high LR chi2 value (21670.892) and p-value<0.05 indicate that the model significantly improves the fit compared to a null model, suggesting that the predictors collectively contribute significantly to the model.

Table (15): Model 1.2 coefficients

Variable	Odds Ratio	Std. Error	z-Value	P>z	95% Confidence Interval
Rururb= Rural	1.090176	.2343298	0.40	0.688	.7153744 - 1.661344
Sex= Male	7.213053	1.400965	10.17	0.000	4.929392 - 10.55468
Rururb*sex= Rural*Male	1.541468	.377727	1.77	0.077	.9535708 - 2.491817
Disable= No	1.45198	.2498873	2.17	0.030	1.036258 - 2.03448
Sex*disabl= Male*No	3.268699	.648201	5.97	0.000	2.216037 - 4.821397
Rururb*disabl= Rural*No	.6697071	.1465955	-1.83	0.067	.436075 - 1.02851
mart2= Never Married	.1797718	.0134475	-22.94	0.000	.1552563 - .2081585
mart2= Married	1.072069	.0625978	1.19	0.233	.9561395 - 1.202054
Educ= None	.0932133	.0232418	-9.52	0.000	.0571794 - .1519556
Educ= Primary/Lower Secondary	.0761386	.0190441	-10.30	0.000	.0466335 - .1243114
Educ= Secondary	.0933562	.0232293	-9.53	0.000	.0573249 - .1520348
Educ=Post Secondary or Equivalent	.1918206	.0494072	-6.41	0.000	.1157847 - .3177891
Educ= University	.2235925	.0558076	-6.00	0.000	.1370886 - .3646807
Age	.9807708	.0013316	-14.30	0.000	.9781644 - .9833841
Sex*disable*rururb= Male*No*Rural	1.475843	.3706759	1.55	0.121	.9020913 - 2.414514
_cons	2.687335	.8314981	3.19	0.001	1.465374 - 4.928278

Interpreting interaction terms:

How do gender, urban/rural residency, and disability status intersect to contribute to labor force disparities?

- Gender and Urban/Rural Residence (Gender*Urban/Rural Residence):

The interaction term between gender and urban/rural residence (rururb*sex) has an odds ratio of 1.541 and a p-value of 0.077. This suggests that the result is not statistically significant at the 0.05 level, it indicates that the effect of Urban/rural residence on the odds of being employed doesn't significantly differs among males and females. and with confidence 95%

- Gender and Disability Status (Gender*Disability Status):

The interaction term between gender and disability status (sex*disabl) has an odds ratio of 3.269 and a p-value of < 0.05. This significant result indicates that the effect of disability status on the odds of being employed significantly differs among males and females. and with confidence 95%

This highlights a substantial disparity where disabled individuals, particularly women, face significantly lower employment odds, whereas non-disabled males are much more likely to be employed.

- Disability Status and Urban/Rural Residence (Disability Status*Urban/Rural Residence):

The interaction term between disability status and urban/rural residence (rururb*disabl) has an odds ratio of 0.670 and a p-value of 0.067. This suggests that the result is not statistically significant at the 0.05 level, it indicates that the effect of Urban/rural residence on the odds of being employed doesn't significantly differs among disabled and not disabled people. and with confidence 95%

- Gender, Urban/Rural Residence, and Disability Status (Gender*Urban/Rural Residence*Disability Status):

The three-way interaction term (sex*disabl*rururb) has an odds ratio of 1.476 and a p-value of 0.121. This result suggests that non-disabled males in rural areas have 1.476 times higher odds of being employed compared to the baseline group. and with confidence 95%

Although statistically significant, it suggests that the combination of being male, non-disabled, and living in rural areas might improve employment odds compared to other groups, but the disparity is less pronounced than in the two-way interactions.

Summary of model 1.2:

Which subgroups experience the greatest labor market disparities?

- Disabled Women in Rural Areas:

The individual and interaction terms suggest that disabled women, particularly in rural areas, face the greatest employment challenges.

For instance, the significant interaction term for gender and disability status (sex*disabl) underscores the compounded disadvantage for disabled females compared to non-disabled males.

- Rural Disabled Individuals:

The interaction term between disability status and urban/rural residence (rururb*disabl) implies that rural residency exacerbates the employment challenges for disabled individuals, though this effect is not statistically robust (p = 0.067).

- Non-Disabled Males:

The consistent high odds ratios for non-disabled males across various interactions (sex*disabl and sex*disabl*rururb) indicate that this group experiences significantly better employment outcomes compared to other subgroups.

This highlights the intersectional advantage where being male and non-disabled leads to markedly better employment prospects

Chapter 6

Model 2: wage gap models using Blinder Oaxaca Decomposition

6. Model 2: wage gap models using Blinder Oaxaca Decomposition method

This chapter includes the analysis that investigates the wage gap by urban/rural residence, gender and disability status using Labor Force Survey (LFS) 2022. This method is meant to explain the difference in the means of the response variable between two groups, then this difference is decomposed into 2 parts:

The "twofold" decomposition,

$$R = Q + U$$

where the first component,

$$Q = \{E(X_A) - E(X_B)\}' \beta^*$$

is the part of the outcome differential that is explained by group differences in the predictors (the "quantity effect"), **and the second component,**

$$U = E(X_A)' (\beta_A - \beta^*) + E(X_B)' (\beta^* - \beta_B)$$

is the unexplained part. The latter is usually attributed to discrimination.

Explained part: The part of the difference that can be explained by differences in observable characteristics between the two groups; the explanatory variables in our model.

Unexplained part: The part of the difference that cannot be explained by differences in observable characteristics due to discrimination.

As mentioned before the sample size was restricted in the modeling part so our sample size in the wage gap models will be 18,600. Moreover, we will use the natural logarithm of wage to be our response variable (Ln wage) to reduce from outliers and skewness. And these are the explanatory variables we will use in our models with the reference category for the categorical variables.

Table (16): Explanatory variables with the reference category for categorical variables

Independent Variable	Categories	Reference Category
Gender	Binary nominal: 1=Male 2=Female	Female
Urban/Rural Residence	Binary nominal: 0=Rural 1=Urban	Urban
Disability Status	Binary nominal: 0= No 1=Yes	Yes
Marital Status	Nominal: 1=Never married 2=Married 3= Divorced/Widowed	Divorced/Widowed
Educational Level	Nominal: 1=None 2=Primary/Lower secondary 3=Secondary 4=Post-secondary or equivalent 5=University 6=Postgraduate	Postgraduate
Employment Stability in the main Job	Nominal: 1=Full time/regular 2=Part time/temporary 3=Seasonal/irregular	Seasonal/irregular
Coronavirus caused a decrease in wage	Binary nominal: 0=No 1=Yes	Yes
Age	Continuous:18-69 years	
Total weekly working hours	Continuous	

6.1. Model (2.1) Wage gap by urban/rural residence

In this model our response variable is Ln wage, our explanatory variables are: gender, age, marital status, educational level, employment stability, coronavirus caused a decrease in wage and total weekly working hours, grouped by urban/rural residence.

Research questions addressed in this model

RQ2: How do labor force disparities vary based on residence in Egypt, examining differences in wages gaps between urban and rural areas?

Before running the model, the absence of multicollinearity must be checked first:

Table (17): Multicollinearity check for model 2.1

Variable	VIF
Age	**1.59**
Marital Status	**1.54**
Employment Stability	**1.16**
Gender	**1.11**
Educational Level	**1.10**
Total Weekly Working Hours	**1.07**
Coronavirus Caused a Decrease in Wage	**1.02**

-All the VIF values are comfortably below 10, indicating a low risk of multicollinearity. This is a positive sign, suggesting that the predictor variables (age, marital status, employment stability, gender, educational level, coronavirus caused a decrease in wage and total weekly working hours) are relatively independent of each other.

Fitting the model:

Table (18): Fitting the model (2.1)

Ln Wage	Coefficient	std. err.	z	P>z	95% conf. interval	
Overall						
Rural	7.98797	.0032893	2428.48	**0.000**	7.981523	7.994417
Urban	8.010653	.004526	1769.93	**0.000**	8.001783	8.019524
Difference	-.0226835	.005595	-4.05	**0.000**	-.0336495	-.0117175
Explained	-.010825	.0024184	-4.48	**0.000**	-.0155651	-.006085
Unexplained	-.0118585	.0056256	-2.11	**0.035**	-.0228844	-.0008326
Explained						
Age	-.0020584	.0004662	-4.41	**0.000**	-.0029723	-.0011446
Total weekly working hours	-.012344	.0011884	-10.39	**0.000**	-.0146732	-.0100148
Gender=Male	.006343	.0008259	7.68	**0.000**	.0047243	.0079618
Marital Status=Never married	.0011078	.0008752	1.27	**0.206**	-.0006075	.0028231
Marital Status=Married	.0050917	.0012287	4.14	**0.000**	.0026835	.0074999
Educational Level=None	-.0351927	.0041289	-8.52	**0.000**	-.0432853	-.0271001
Educational Level=Primary/lower secondary	-.0048409	.0017911	-2.70	**0.007**	-.0083514	-.0013304
Educational Level=Secondary	-.0133266	.0026787	-4.98	**0.000**	-.0185767	-.0080766

Educational Level=Post-Secondary or equiv.	.0039269	.0009541	4.12	**0.000**	.0020568	.0057969
Educational Level=University	.0215368	.0048932	4.40	**0.000**	.0119463	.0311273
Employment Stability=Full time/regular	.0162903	.0011597	14.05	**0.000**	.0140174	.0185632
Employment Stability=Part time/temporary	.0047691	.0008693	5.49	**0.000**	.0030652	.0064729
Coronavirus caused a Decrease in wage=No	-.002128	.0004192	-5.08	**0.000**	-.0029496	-.0013064

With 95% confidence

Interpreting the overall:

- The average Ln wage for rural areas is 7.98797 while for urban areas is 8.010653, it is clear that rural areas to be paid less on average with difference equal to -0.0226835.
- It is clear that the explained and unexplained parts are significant (p-value less than 0.05) indicating that the wage gap is due to observed characteristics (the explanatory variables included in our model) and also due to discrimination between the 2 groups.
- -0.010825 of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, gender, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups. Each of the coefficients in the explained part represents the contribution of that variable to the wage gap, assuming that both groups (rural and urban) receive the same returns to that characteristic.
- The rest of the wage gap -0.0118585 is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination between urban and rural in which urban areas have higher wages than rural areas.

Interpreting the significant variables in the explained part:

- The coefficient of age is -0.0020584 and statistically significant as (p-value < 0.05) and it means that differences in age between the rural and urban groups contribute to a decrease in the wage gap by 0.0020584. In other words, if the rural group were to have the same age distribution as the urban group, the wage gap would decrease by 0.0020584, holding all other variables constant.
- The coefficient of total weekly working hours is -.012344 and statistically significant as (p-value < 0.05) and it means that differences in weekly working hours between the rural and urban groups contribute to a

decrease in the wage gap by 0.012344. In other words, if the rural group were to have the same distribution of weekly working hours as the urban group, the wage gap would decrease by 0.012344, holding all other variables constant.

- The coefficient of gender= male is 0.006343 and statistically significant as (p-value < 0.05) which means that differences in being male between the rural and urban groups contribute to an increase in the wage gap by 0.006343 compared to being a female. In other words, being males in rural group, this will increase the wage gap by 0.006343 compared to being a female, holding all other variables constant.
- The coefficient of marital status= married is 0.0050917 and statistically significant as (p-value < 0.05) which means that differences in being married between the rural and urban groups contribute to an increase in the wage gap by 0.0050917 compared to divorced/widowed. In other words, if the rural group were to have the same marital status distribution which is married as the urban group, the wage gap would increase by 0.0050917 compared to being divorced/widowed, holding all other variables constant.
- The coefficient of educational level=none is -0.0351927 and statistically significant as (p-value < 0.05) which means that differences in being not educated between the rural and urban groups contribute to a decrease in the wage gap by -0.0351927 compared to postgraduate education. In other words, if the rural group were to have the same educational level which is none as the urban group, the wage gap would decrease by -0.0351927 compared to postgraduate education, holding all other variables constant.

- The coefficient of educational level=primary/lower secondary is -0.0048409 statistically significant as (p-value < 0.05) which means that differences in being primary/lower secondary education between the rural and urban groups contribute to a decrease in the wage gap by 0.0048409 compared to postgraduate education. In other words, if the rural group were to have the same educational level which is primary/lower secondary as the urban group, the wage gap would decrease by 0.0048409 compared to postgraduate education, holding all other variables constant.

The coefficient of educational level=secondary is -0.0133266 and statistically significant as (p-value < 0.05) which means differences in being secondary education between the rural and urban groups contribute to a decrease in the wage gap by 0.0133266 compared to postgraduate education. In other words, if the rural group were to have the same educational level which is secondary as the urban group, the wage gap would decrease by 0.0133266 compared to postgraduate education, holding all other variables constant.

- The coefficient of educational level=post-secondary or equivalent is 0.0039269 and statistically significant as (p-value < 0.05) which means differences in being post-secondary or equivalent between the rural and urban groups contribute to an increase in the wage gap by 0.0039269 compared to postgraduate education. In other words, if the rural group were to have the same educational level which is post-secondary or equivalent as the urban group, the wage gap would increase by 0.0039269 compared to postgraduate education, holding all other variables constant.

- The coefficient of educational level=university is .0215368 statistically significant as (p-value < 0.05) which means that differences in being university education between the rural and urban groups contribute to an increase in the wage gap by 0.0215368 compared to postgraduate education. In other words, if the rural group were to have the same educational level which is university as the urban group, the wage gap would increase by 0.0215368 compared to postgraduate education, holding all other variables constant.
- The coefficient of employment stability=full time/regular is 0.0162903 statistically significant as (p-value < 0.05) which means that differences in being in full time/regular job between the rural and urban groups contribute to an increase in the wage gap by 0.0162903 compared to in seasonal/irregular job. In other words, if the rural group were to have the same employment stability which is full time as the urban group, the wage gap would increase by 0.0162903 compared to in seasonal/irregular job, holding all other variables constant.
- The coefficient of employment stability=part time/temporary is 0.047691 statistically significant as (p-value < 0.05) means that that differences in being in part time/temporary job between the rural and urban groups contribute to an increase in the wage gap by 0.047691 compared to in seasonal/irregular job. In other words, if the rural group were to have the same employment stability which is part time as the urban group, the wage gap would increase by 0.047691 compared to in seasonal/irregular job, holding all other variables constant.
- The coefficient of coronavirus caused a decrease in wage=no is -0.002128 and statistically significant as (p-value < 0.05) which means that differences in coronavirus did not cause a decrease in wage between the rural and urban groups contribute to a decrease in the wage gap by 0.002128 compared to coronavirus caused a decrease in wage. In other words, if the rural group were to be the same in coronavirus did not cause a decrease in wage as the urban group, the wage gap would decrease by 0.002128 compared to coronavirus caused a decrease in wage, holding all other variables constant.

Answering the research questions:

RQ2: How do labor force disparities vary based on residence in Egypt, examining differences in wages gaps between urban and rural areas?

The average Ln wage for rural areas is 7.98797 while for urban areas is 8.010653, it is clear that rural areas to be paid less with difference equal to -0.0226835. -0.010825 of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, gender, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups. Each of the coefficients in the explained part represents the contribution of that variable to the wage gap, assuming that both groups (rural and urban) receive the same returns to that characteristic, while the rest of the wage gap - 0.0118585 is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination.

Summary for model 2.1 wage gap by urban/rural residence

The average Ln wage for rural areas is 7.98797 while for urban areas is 8.010653, it is clear that rural areas to be paid less on average with difference equal to -0.0226835. 47.72% of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, gender, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups, while 52.28% is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination between urban and rural in which urban areas have higher wages than rural areas. From table (19) we can examine the variables that increase or decrease the wage gap

Table (19): The effect of the variables on the wage gap between urban and rural areas

Variable	**Effect on wage gap between urban and rural areas**
Age	Decrease wage gap between urban and rural areas
Total weekly working hours	Decrease wage gap between urban and rural areas
Gender=Male	Increase wage gap between urban and rural areas
Marital Status=Never married	No significant effect
Marital Status=Married	Increase wage gap between urban and rural areas
Educational Level=None	Decrease wage gap between urban and rural areas

Educational Level=Primary/lower secondary	Decrease wage gap between urban and rural areas
Educational Level=Secondary	Decrease wage gap between urban and rural areas
Educational Level=Post-Secondary or equiv.	Increase wage gap between urban and rural areas
Educational Level=University	Increase wage gap between urban and rural areas
Employment Stability=Full time/regular	Increase wage gap between urban and rural areas
Employment Stability=Part time/temporary	Increase wage gap between urban and rural areas
Coronavirus caused a Decrease in wage=No	Decrease wage gap between urban and rural areas

6.2. Model (2.2) Wage gap by gender

In this model our response variable is Ln wage, our explanatory variables are: urban/rural residence, age, marital status, educational level, employment stability, coronavirus caused a decrease in wage and total weekly working hours, grouped by gender.

Research questions addressed in this model:

RQ1: What is the gender-based disparities within the Egyptian labor force, considering wage gap?

Before running the model, the absence of multicollinearity must be checked first:

Table (20): Multicollinearity check for model 2.2

Variable	VIF
Age	**1.60**
Marital Status	**1.52**
Employment Stability	**1.15**
Urban/rural residence	**1.09**
Educational Level	**1.11**
Total Weekly Working Hours	**1.06**
Coronavirus Caused a Decrease in Wage	**1.02**

-All the VIF values are comfortably below 10, indicating a low risk of multicollinearity. This is a positive sign, suggesting that the predictor variables (age, marital status, employment stability, urban/rural residence, educational level, coronavirus caused a decrease in wage and total weekly working hours) are relatively independent of each other.

Fitting the model:

Table (21): Fitting the model (2.2)

Ln wage	Coefficient	std. err.	z	P>z	95% conf. interval	
Overall						
Male	8.015083	.0028068	2855.59	**0.000**	8.009581	8.020584
Female	7.894378	.0078284	1008.43	**0.000**	7.879035	7.909722
Difference	.1207042	.0083163	14.51	**0.000**	.1044045	.1370039
Explained	-.0096424	.0036959	-2.61	**0.009**	-.0168862	-.0023986
Unexplained	.1303466	.0085036	15.33	**0.000**	.1136798	.1470134
Explained						
Age	-.0040309	.0007798	-5.17	**0.000**	-.0055593	-.0025026
Total weekly working hours	.01385	.0013886	9.97	**0.000**	.0111284	.0165716
Urban/rural residence=Rural	-.0010841	.0005277	-2.05	**0.040**	-.0021184	-.0000499
Marital status=Never married	-.0018644	.0014632	-1.27	**0.203**	-.0047322	.0010035
Marital status=Married	.0030559	.0009564	3.20	**0.001**	.0011814	.0049303
Educational level=none	-.0057662	.0030491	-1.89	**0.059**	-.0117423	.0002098
Educational level=primary/lower secondary	-.0311187	.0038774	-8.03	**0.000**	-.0387182	-.0235191
Educational level=secondary	-.0405735	.0056469	-7.19	**0.000**	-.0516413	-.0295058
Educational level=post-secondary or equiv.	.0056404	.0014311	3.94	**0.000**	.0028354	.0084454
Educational level=university	.0340264	.0077293	4.40	**0.000**	.0188772	.0491756
Employment stability=full time/regular	.0259847	.0015919	16.32	**0.000**	.0228647	.0291047
Employment stability=part time/temporary	-.0054146	.001016	-5.33	**0.000**	-.0074059	-.0034234
Coronavirus caused a decrease in wage=no	-.0023473	.0005635	-4.17	**0.000**	-.0034518	-.0012429

With 95% confidence

Interpreting the overall:

- The average ln wage for males is 8.015083, while for females it is 7.894378. The difference in average ln wage is 0.1207042, indicating that females are paid less on average.
- It is clear that the explained and unexplained parts are significant (p-value less than 0.05) indicating that the wage gap is due to observed characteristics (the explanatory variables included in our model) and also due to discrimination between the 2 groups.

- -0.0096424 of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups. Each of the coefficients in the explained part represents the contribution of that variable to the wage gap, assuming that both groups (males and females) receive the same returns to that characteristic.
- The rest of the wage gap 0.1303466 is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination between males and females as males tend to be paid more.

Interpreting the significant variables in the explained part:

- The coefficient of age is -0.0040309 and statistically significant as (p-value < 0.05) and it means that differences in age between males and females contribute to a decrease in the wage gap by 0.0040309. In other words, if females had the same age distribution as males, the wage gap would decrease by 0.0040309, holding all other variables constant.
- The coefficient of total weekly working hours is 0.01385 and statistically significant as (p-value < 0.05) and it means differences in weekly working hours between males and females contribute to an increase in the wage gap by 0.01385. In other words, if females had the same distribution of weekly working hours as males, the wage gap would increase by 0.01385, holding all other variables constant.
- The coefficient of urban/rural= rural is -0.0010841 and statistically significant as (p-value < 0.05) which means differences in being rural between the males and females contribute to a decrease in the wage gap by -0.0010841 compared to urban. In other words, if the female group were to have the same urban/rural residence distribution which is rural as the male group, the wage gap would decrease by -0.0010841 compared to urban, holding all other variables constant.
- The coefficient of marital status= married is 0.0030559 and statistically significant as (p-value < 0.05) which means that in being married between the males and females contribute to an increase in the wage gap by 0.0030559 compared to divorced/widowed. In other words, if the female group were to have the same marital status distribution which is married as the male group, the wage gap would increase by 0.0030559 compared to divorced/widowed, holding all other variables constant.
- The coefficient of educational level=primary/lower secondary is -0.0311187 statistically significant as (p-value < 0.05) which means that differences in having primary/lower secondary education between males and females contribute to a decrease in the wage gap by 0.0311187 compared to from postgraduates. In

other words, if the female group were to have the same distribution of having primary/lower secondary education as the male group, the wage gap would decrease by 0.0311187 compared to from postgraduates, holding all other variables constant.

- The coefficient of educational level=secondary is -0.0405735 and statistically significant as (p-value < 0.05) which means that differences in having secondary education between males and females contribute to a decrease in the wage gap by 0.0405735 compared to from postgraduates. In other words, if the female group were to have the same distribution of having secondary education as the male group, the wage gap would decrease by 0.0405735 compared to from postgraduates, holding all other variables constant.
- The coefficient of educational level=post-secondary or equivalent is 0.0056404 and statistically significant as (p-value < 0.05) which means differences in having post-secondary or equivalent education between males and females contribute to an increase in the wage gap by 0.0056404 compared to from postgraduates. In other words, if the female group were to have the same distribution of having post-secondary or equivalent education as the male group, the wage gap would increase by 0.0056404 compared to from postgraduates, holding all other variables constant.
- The coefficient of educational level=university is 0.0340264 statistically significant as (p-value < 0.05) which means that differences in having university education between males and females contribute to an increase in the wage gap by 0.0340264 compared to in postgraduates. In other words, if the female group were to have the same distribution of having university education as the male group, the wage gap would increase by 0.0340264 compared to from postgraduates, holding all other variables constant.
- The coefficient of employment stability=full time/regular is 0.0259847 statistically significant as (p-value < 0.05) which means that differences in being full time/regular employees between males and females contribute to an increase in the wage gap by 0.0259847 compared to in seasonal/irregular job. In other words, if the female group were to have the same distribution of being full time/regular employees as the male group, the wage gap would increase by 0.0259847 compared to in seasonal/irregular job, holding all other variables constant.
- The coefficient of employment stability=part time/temporary is -0.0054146 statistically significant as (p-value < 0.05) means that that differences in being part time/temporary employees between males and females contribute to a decrease in the wage gap by -0.0054146 compared to in seasonal/irregular job. In other words, if the female group were to have the same distribution of being part time/temporary employees

as the male group, the wage gap would decrease by -0.0054146 compared to in seasonal/irregular job, holding all other variables constant.

- The coefficient of coronavirus caused a decrease in wage=no is -0.0023473 and statistically significant as (p-value < 0.05) which means that differences in coronavirus did not cause a decrease in wage between the males and females contribute to a decrease in the wage gap by 0.0023473 compared to coronavirus caused a decrease in wage. In other words, if the rural group were to be the same in coronavirus did not cause a decrease in wage as the urban group, the wage gap would decrease by 0.0023473 compared to coronavirus caused a decrease in wage, holding all other variables constant.

Answering the research questions:

RQ1: What is the gender-based disparities within the Egyptian labor force, considering wage gap and employment status?

The average ln wage for males is 8.015083, while for females it is 7.894378. The difference in average ln wage is 0.1207042, indicating that females are paid less on average. -0.0096424 of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups. Each of the coefficients in the explained part represents the contribution of that variable to the wage gap, assuming that both groups (males and females) receive the same returns to that characteristic, while the rest of the wage gap 0.1303466 is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination.

Summary for model 2.1 wage gap by gender

The average ln wage for males is 8.015083, while for females it is 7.894378. The difference in average ln wage is 0.1207042, indicating that females are paid less on average. -7.99% of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups, while 107.99% is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination between males and females in which males have higher wages than females. From table (22) we can examine the variables that increase or decrease the wage gap.

Table (22): The effect of the variables on the wage gap between males and females

Variable	Effect of the wage gap between males and females
Age	Decrease wage gap between males and females
Total weekly working hours	Increase wage gap between males and females
Urban/rural residence=Rural	Decrease wage gap between males and females
Marital status=Never married	No significant effect on wage gap between males and females
Marital status=Married	Increase wage gap between males and females
Educational level=none	No significant effect on wage gap between males and females
Educational level=primary/lower secondary	Decrease wage gap between males and females
Educational level=secondary	Decrease wage gap between males and females
Educational level=post-secondary or equiv.	Increase wage gap between males and females
Educational level=university	Increase wage gap between males and females
Employment stability=full time/regular	Increase wage gap between males and females
Employment stability=part time/temporary	Decrease wage gap between males and females
Coronavirus caused a decrease in wage=no	Decrease wage gap between males and females

6.3. Model (2.3) Wage gap by disability status

In this model our response variable is Ln wage, our explanatory variables are: urban/rural residence, gender, age, marital status, educational level, employment stability, coronavirus caused a decrease in wage and total weekly working hours, grouped by disability status.

Research questions addressed in this model

RQ3: To what extent does disability impact labor force outcomes in Egypt, including wage gaps?

Before running the model, the absence of multicollinearity must be checked first:

Table (23): Multicollinearity check for model 2.3

Variable	VIF
Age	**1.61**
Marital Status	**1.55**
Employment Stability	**1.17**
Urban/rural residence	**1.10**
Gender	**1.11**
Educational Level	**1.13**
Total Weekly Working Hours	**1.10**
Coronavirus Caused a Decrease in Wage	**1.02**

-All the VIF values are comfortably below 10, indicating a low risk of multicollinearity. This is a positive sign, suggesting that the predictor variables (age, marital status, employment stability, gender, urban/rural residence, educational level, coronavirus caused a decrease in wage and total weekly working hours) are relatively independent of each other.

Fitting the model:

Table (24): Fitting the model (2.3)

Ln wage	Coefficient	std. err.	z	P>z	95% conf. interval	
overall						
Without disability	8.000963	.0027214	2940.01	0.000	7.995629	8.006297
With disability	7.902131	.0142323	555.23	0.000	7.874236	7.930025
Difference	.0988327	.0144901	6.82	0.000	.0704326	.1272328
Explained	-.0044504	.0043861	-1.01	0.310	-.013047	.0041461
Unexplained	.1032831	.0142751	7.24	0.000	.0753045	.1312618
Explained						
Age	-.0151517	.0023968	-6.32	0.000	-.0198494	-.010454
Total weekly working hours	.0036946	.0012627	2.93	0.003	.0012198	.0061694
Gender=Male	-.0018827	.0016877	-1.12	0.265	-.0051906	.0014251
Urban/rural residence=Rural	.0010127	.0005788	1.75	0.080	-.0001218	.0021472
Marital Status=Never married	-.0029138	.002124	-1.37	0.170	-.0070768	.0012492
Marital Status=Married	-.0055132	.0016292	-3.38	0.001	-.0087063	-.0023201
Educational Level=None	.063551	.0090587	7.02	0.000	.0457962	.0813057
Educational Level=Primary/lower secondary	-.0013658	.0041953	-0.33	0.745	-.0095884	.0068567
Educational Level=Secondary	-.0173466	.0055237	-3.14	0.002	-.0281728	-.0065203
Educational Level=Post-secondary or equiv	-.0027032	.0014069	-1.92	0.055	-.0054606	.0000542
Educational Level=University	-.015663	.0039212	-3.99	0.000	-.0233484	-.0079775
Employment Stability=Full time/regular	.0013702	.0020222	0.68	0.498	-.0025933	.0053337
Employment Stability=Part time/temporary	-.0034966	.0018036	-1.94	0.053	-.0070316	.0000385
Coronavirus caused a decrease in wage=no	-.0080423	.0013957	-5.76	0.000	-.0107779	-.0053067

With 95% confidence

Interpreting the overall:

- The average ln wage for individuals without disabilities is 8.000963, while for with disabilities it is 7.902131. The difference in average ln wage is 0.0988327, indicating that individuals with disabilities are paid less on average.
- It is clear that only the unexplained parts are significant (p-value less than 0.05) indicating that the wage gap is only due to discrimination between the 2 groups.

- -0.0044504 is the wage gap due to the explained components but not significant (p-value>0.05) the rest of the wage gap 0.1032831 is the unexplained part which cannot be explained by (age, total weekly working hours, gender, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage)and is attributed to discrimination between individuals with disability and without in which individuals without disability have higher wages than individuals with.

Answering the research questions:

RQ3: To what extent does disability impact labor force outcomes in Egypt, including wage gaps?

The average ln wage for individuals without disabilities is 8.000963, while for with disabilities it is 7.902131. The difference in average ln wage is 0.0988327, indicating that individuals with disabilities are paid less on average. Wage gap is only due to discrimination between the 2 groups equal to 0.1032831.

Summary for model 2.3 wage gap by disability status

The average ln wage for individuals without disabilities is 8.000963, while for with disabilities it is 7.902131. The difference in average ln wage is 0.0988327, indicating that individuals with disabilities are paid less on average. The explained part is not significant. Hence, we can conclude that the wage gap in this case is only due to discrimination between the 2 groups. 104.40% of the wage gap is the unexplained part which cannot be explained by (age, total weekly working hours, gender, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage) and is attributed to discrimination between individuals with disability and without in which individuals without disability have higher wages than individuals with.

Chapter 7: Main Findings

7.1 Conclusion

Here are the results of our analysis on disparities in the Egyptian labor force:

Urban vs. Rural Disparities: Our findings indicate that individuals living in rural areas have 22.07% higher odds of being employed compared to those in urban areas. However, the average Ln wage for rural areas is 7.98797 while for urban areas is 8.010653, it is clear that rural areas to be paid less on average with difference equal to -0.0226835. 47.72% of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, gender, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups, while 52.28% is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination between urban and rural in which urban areas have higher wages than rural areas. From table (19) we can examine the variables that increase or decrease the wage gap.

Gender-based Disparities: The analysis shows a significant gender disparity in the Egyptian labor force. Males have 35.21 times higher odds of being employed than females. Furthermore, the average ln wage for males is 8.015083, while for females it is 7.894378. The difference in average ln wage is 0.1207042, indicating that females are paid less on average. -7.99% of the wage gap can be explained by differences in the characteristics between the two groups, the explained part of the decomposition is the portion of the wage gap that can be explained by differences in the characteristics (age, total weekly working hours, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage) between the two groups, while 107.99% is the unexplained part which cannot be explained by these characteristics and is attributed to discrimination between males and females in which males have higher wages than females. From table (22) we can examine the variables that increase or decrease the wage gap.

Disability Status Disparities: Disability status significantly impacts labor force outcomes. Non-disabled individuals have 3.64 times higher odds of being employed than disabled individuals. Moreover, the average ln wage for individuals without disabilities is 8.000963, while for with disabilities it is 7.902131. The difference in average ln wage is 0.0988327, indicating that individuals with disabilities are paid less on average. The explained part is not significant. Hence, we can conclude that the wage gap in this case is only due to discrimination between the 2 groups. 104.40% of the wage gap is the unexplained part which cannot be explained by (age, total weekly working hours, gender, urban/rural residence, marital status, educational level, employment stability and coronavirus caused a decrease in wage) and is attributed to discrimination between individuals with disability and without in which individuals without disability have higher wages than individuals with.

Gender and Urban/Rural Residence: The effect of urban/rural residence on employment odds does not significantly differ between males and females.

Gender and Disability Status: Disabled individuals, particularly women, face significantly lower employment odds compared to non-disabled individuals of the same gender.

Disability Status and Urban/Rural Residence: The effect of urban/rural residence on employment odds does not significantly differ between disabled and non-disabled individuals.

Gender, Urban/Rural Residence, and Disability Status: Non-disabled males in rural areas have higher odds of employment compared to the baseline group. However, this advantage is less pronounced than in other interactions.

Disabled Women in Rural Areas: Disabled females, especially in rural areas, face the greatest employment challenges.

Rural Disabled Individuals: Rural residency exacerbates employment challenges for disabled individuals, although this effect is not statistically robust.

Non-Disabled Males: Non-disabled males consistently experience significantly better employment outcomes compared to other subgroups.

In conclusion, the labor force disparities in Egypt are significantly influenced by urban/rural residence, gender, and disability status. These disparities are further exacerbated by factors such as discrimination. Therefore, it is crucial to consider these intersections when developing policies and interventions aimed at promoting labor force equality in Egypt. This will ensure that all individuals, regardless of their residence, gender, or disability status have equal opportunities in the labor market.

7.2 Recommendations

Addressing Urban-Rural Disparities: Implement policies that aim to reduce the wage gap between urban and rural areas. This could include investing in rural infrastructure, providing incentives for businesses to operate in rural areas, and implementing rural employment programs.

Promoting Gender Equality: Implement policies and programs that promote gender equality in the workforce. This could include enforcing equal pay legislation, promoting women's participation in traditionally male-dominated industries, and providing support for women to balance work and family responsibilities.

Supporting Disabled Individuals: Implement policies and programs that support the employment of disabled individuals. This could include enforcing anti-discrimination laws, providing reasonable accommodations in the workplace, and offering vocational training programs for disabled individuals.

Reducing the Impact of the Coronavirus Pandemic: Implement policies and programs that mitigate the impact of the coronavirus pandemic on wage gaps. This could include providing financial support for businesses affected by the pandemic and providing retraining programs for individuals who have lost their jobs due to the pandemic.

Addressing Discrimination: Implement policies and programs that address discrimination in the workforce. This could include enforcing anti-discrimination laws, promoting diversity and inclusion in the workplace, and providing diversity training for employers.

Data Collection and Monitoring: Regularly collect and analyze data on labor force disparities in order to monitor the effectiveness of policies and interventions, and to identify areas where further action is needed.

7.3 Limitations

Data Source: The study relies solely on the Labor Force Survey (LFS) 2022 dataset. Although valuable, this dataset may not capture all aspects of the labor market, particularly the informal sectors or hidden unemployment.

Disability Status Data: The reliance on wave 3 data for disability status limits the analysis, preventing a longitudinal perspective on how disability status impacts labor force participation and wages over time. This also restricts the ability to compare disability-related disparities across different survey waves.

Blinder-Oaxaca Decomposition: This method helps understand wage gaps but does not account for potential interactions between factors such as gender, residence, and disability status.

Lack of Qualitative Data: The study relies solely on quantitative data, which may not capture the lived experiences and perspectives of individuals facing labor market disparities.

Focus on Specific Disparities: The study focuses on gender, urban/rural residence, and disability status, without exploring other potential disparities such as those based on ethnicity, religion, or socioeconomic background.

REFRENCES:

1. Assaad, R. (2000). Structural Adjustment, Demography, and the Egyptian Labor Market in the 1990s. World Bank Group. http://documents.worldbank.org [Internet] 1st July 2010 http://documents.worldbank.org/curated/en/795451468744333367/Structural-adjustment-demography-and-the-Egyptian-labor-market-in-the-1990s [Accessed February 2024]
2. Biltagy, M. (2018). Gender wage disparities in Egypt: Evidence from ELMPS 2006 and 2012.The Quarterly Review of Economics and Finance. https://doi.org [Internet] 28th November 2018 https://doi.org/10.1016/j.qref.2018.11.010 [Accessed February 2024]
3. Helmy, Heba. (2011). Egypt's rural-urban income disparity: Is the gap diverging?. Contemporary Arab Affairs. 4. 148-173. The Research Gate. https://www.researchgate.net [Internet] April 2011 https://www.researchgate.net/publication/233092721_Egypt's_ruralurban_income_disparity_Is_the_gap_diverging/citation/download [Accessed February 2024]
4. Hendy, R. (2015). Women's Participation in the Egyptian Labor Market: 1998 – 2012. The Economic Research Forum. https://erf.org.eg [Internet] May 2015 https://erf.org.eg/publications/womens-participation-in-the-egyptian-labor-market-1998-2012/?tab=undefined&c=undefined [Accessed February 2024]
5. Jann, Ben (2008). The Blinder-Oaxaca decomposition for linear regression models. The Stata Journal 8(4): 453-479. https://journals.sagepub.com/home/STJ [Internet] May 2008 https://doi.org/10.1177/1536867X0800800401 [Accessed June 2024]
6. Krafft, C., Assaad, R., and Keo, C. (2019). The Evolution of Labor Supply in Egypt from 1988-2018: A Gendered Analysis. The Economic Research Forum. https://erf.org.eg [Internet] October 2019 https://erf.org.eg/publications/the-evolution-of-labor-supply-in-egypt-from-1988-2018-a-gendered-analysis/ [Accessed February 2024]
7. Nazier, H. and Ramadan, R. (2018) Ever married women's participation in labor market in Egypt: constraints and opportunities, Middle East Development Journal, 10:119-151, DOI: 10.1080/17938120.2018.1443605 [Internet] 28th February 2018 https://doi.org/10.1080/17938120.2018.1443605 [Accessed February 2024]

8. Said, M. , Galal, R. , Joekes, S. , and Sami, M. (2018). Gender Diversity, Productivity, and Wages in Egyptian Firms. The Economic Research Forum. https://erf.org.eg [Internet] June 2018 https://erf.org.eg/publications/gender-diversity-productivity-wages-egyptian-firms/ [Accessed February 2024]
9. Shalkamy, M. (2021). Disability, Care and Women's Labor Force Participation in Egypt. Baseera Foundation. https://en.enow.gov.eg [Internet] March 2021 https://en.enow.gov.eg/PapersDetail/23 [Accessed February 2024]
10. Zeitoun, N. ,Elassiouty, S. ,Rahman, T. ,Osman, M. , Girgis, H. , Qarout, D. ,and Elkhorazaty, N. Social Norms and Female Labor Force Participation in Egypt. World Bank Group. http://documents.worldbank.org [Internet] 20th June 2023 https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099061823220039155/p17968608f6dd00350b350081904da12d91 [Accessed February 2024]
11. Zaghloul, W. (2022). Revisiting Gender Wage Gap In Egypt: Empirical Evidence Using 2018 ELMPS Data [Master's Thesis, the American University in Cairo]. AUC Knowledge Fountain. https://Aucegypt.edu [Internet] 25th June 2022 https://fount.aucegypt.edu/etds/1914 [Accessed February 2024]