# IDENTIFICATION OF FACTORS AFFECTING THE UNEMPLOYMENT OF EDUCATED YOUTH IN SRI LANKA 

Lakmini Uthpala Mallawarachchi (168838U)

Degree of Master of Science in Business Statistics

Department of Mathematics

Faculty of Engineering

University of Moratuwa

Sri Lanka

December 2019

# IDENTIFICATION OF FACTORS AFFECTING THE UNEMPLOYMENT OF EDUCATED YOUTH IN SRI LANKA 

Lakmini Uthpala Mallawarachchi

(168838U)

Dissertation submitted in partial fulfillment of the requirements for the Degree of Master of Science in Business Statistics

Department of Mathematics

Faculty of Engineering
University of Moratuwa

Sri Lanka

December 2019

## DECLARATION OF THE CANDIDATE AND THE SUPERVISOR

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).
Signature:
Date:
L. U. Mallawarachchi

The above candidate has carried out research for the Masters Dissertation under my supervision.

Signature:
Date:

Prof. T. S.G. Peiris

## ACKNOWLEDGEMENT

During the period of the preparation of my research project, I worked with a large number of people, contributions of whom in various ways to the project and the making of the research deserved exceptional mention. It's my pleasure to express my sincere gratitude to all of them in my humble acknowledgement.

Firstly, I would like to express my gratitude to Prof. T. S. G. Peiris, Senior Professor in Applied Statistics, and the Course Coordinator of M. Sc. / Postgraduate Diploma in Business Statistics for his supervision, guidance and support from the first day of the programme. His true intuition has made him as a constant sanctuary of ideas, thoughts and passions in business statistics which exceptionally inspire and enrich my growth as a postgraduate student. Sir, I'm so much grateful to you in every possible way and hope to keep our collaboration in the future as well.

I sincerely thank all the lecturers involved in this course for their valuable guidance, assistance and comments to complete this programme. Further, special thanks go to the Director General, Department of Census and Statistics and all the officers for their support and guidance in providing the labour force data for this research study.

Next, I would like to give my sincerest gratitude to my loving parents. Thank you for your support, encouragement and for always being there. In addition, I would like to extend my special thanks to my relatives and cousins, for their continuous support, care and for helping me to survive all the stress throughout this period and not letting me give up. Thank you for always being with me.

Finally, I would like to express my deepest appreciation to my beloved colleagues Tharindi Abeywickrama, Jayani De Costa, Keshika De Soysa and Dinithi Kande Arachchi for their encouragement and efforts in helping me to finalize my research project.


#### Abstract

As the unemployment of the educated youth has been a common problem over the past years in Sri Lanka, this study was initiated to determine the factors affecting the educated unemployed youth (EUY). Data from the Annual Labour Force Survey in 2016 conducted by the Department of Census and Statistics (DCS) was used for this study. Data has been collected using the two stage stratified sampling technique. This analysis is based on 4002 individuals whose age category of (15-24) and qualified with the G.C.E (O/L) and G.C.E (A/L) qualifications. For the analysis, seven variables used were gender, race, religion, marital status, education attainment, literacy in English and the residential sector. When each variable was considered separately, it was found that gender, religion, education attainment, literacy level in English and residential sector have significant (p<0.05) impact on EUY. The educated female unemployed youth is significantly higher than that of males. The highest percentage of EUY was found among Buddhists while the lowest percentage was found among Muslims. The rate of EUY having G.C.E (A/L) qualification is significantly less than that of G.C.E. ( $\mathrm{O} / \mathrm{L}$ ) qualification. Of the residential sector, the highest unemployed rate ( $78 \%$ ) was found in the estate sector. When all the seven variables were considered simultaneously using Binary logistic model, only gender, religion, marital status, education attainment and residential sector were found significant on EUY. The overall correct classification rate of the Binary Logistic model is $76.2 \%$. When only two way interactions were considered, the model was found to be significant based on Hosmer and Lemeshow statistic ( $\mathrm{p}=0.460$ ) and the overall correct classification was improved to $77.1 \%$. The significant interaction was found between gender and education attainment, religion, marital status and English literacy. The percentage of married male EUY is significantly lower than that of unmarried male, but the percentage of EUY is almost same for females irrespective of marital status. The lowest percentage of EUY was noted for males having G.C.E (A/L). The percentage of male EUY who are unable to read and write is almost the same than that of males who are able to read and write, percentage of female EUY is much higher among those unable to read and write than that of able to read and write. The inferences derived in this study can effectively be utilized to reduce the EUY in Sri Lanka.


Keywords: Educated Youth, Labour Force, Logistic Regression, Two Stage Stratified Sampling,
Unemployment.

## CHAPTER 1

## INTRODUCTION

### 1.1 Background

In the global context, the problem of unemployment has been a deeply rooted problem. This is due to various factors such as discrimination and inequalities on the basis of social class and status, ethnicity, gender etc (UNDP, 2015). Women are most likely to expose to the issue of gender inequality, which reduces their active participation in labour force. It has been estimated that all the Asian countries have incurred a loss of (42-47) USD billion per year with the restrictions on women's engagement in employment opportunities (UNESCAP, 2016).

Sri Lanka is considered to be a Lower Middle-Income country with a GDP per capita of USD 3,835 in 2016 with an entire population of 21.2 million people (Central Bank Report, 2017). Further, the country has been ranked $73^{\text {rd }}$ out of 187 countries with regard to the Human Development Index (HDI) in 2013 with an HDI of 0.750 and an inequality-adjusted HDI of 0.643. In the Gender Inequality Index (GII), Sri Lanka was placed at $75^{\text {th }}$ position with a GII of 0.383 (UNDP, 2014).

At the end of the 30 years of civil war in 2009, Sri Lankan economy has been grown with an average of $6.2 \%$ over the period from 2010 to 2016 and the services sector has contributed for 60\% of the GDP in 2016 (Central Bank Report, 2017). Over the past three decades, youth unemployment has been recognized as a serious problem in Sri Lanka, which has negatively influenced on the overall productivity and the economic development of the country (Central Bank Report, 2017).

### 1.2 Youth Unemployment in Sri Lanka

According to the Department of Census and Statistics, unemployment has been defined as those people who are available and or looking for work, and who did not work and have tried out to find a job within the last four weeks and waiting to accept a job within next two weeks. Sri Lankans have been struggling with the problem of youth unemployment since 1960s. The problem of increasing the youth unemployment rates have been further confirmed by the
following labour force statistics obtained from the Department of Census and Statistics as shown in Table 1.1.

Table 1.1: Youth unemployment rates in Sri Lanka (2011-2016)

| Year | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Males | 12.8 | 14.0 | 15.6 | 16.1 | 16.6 | 17.1 |
| Females | 25.0 | 23.5 | 23.5 | 27.5 | 27.4 | 29.2 |
| Total | 17.2 | 17.2 | 19.2 | 20.3 | 20.8 | 21.6 |

Source: Department of Census and Statistics (2011-2016)

As indicated in the above Table 1.1, among the youth population, unemployment rates of males have increased from $12.8 \%$ (in 2011) to $17.1 \%$ (in 2016), while the female unemployment rates have also increased from $25 \%$ (in 2011) to $29.2 \%$ (in 2016). Thus, it is clear that female unemployment rates are higher than that of males in all the years.

### 1.3 Distribution of Youth Unemployment by Provinces

The variation of the rate of youth unemployment between provinces is generally higher than the variation between years within the province (Table 1.2).

Table 1.2: Youth unemployment rates by provinces (2011-2016)

| Province | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Western | 14.3 | 14.6 | 16.9 | 14.3 | 15.9 | 14.7 |
| Central | 22.4 | 19.4 | 25.4 | 23.9 | 25.5 | 26.0 |
| Southern | 22.5 | 24.4 | 25.6 | 28.4 | 27.9 | 30.3 |
| Northern | 15.4 | 18.9 | 17.7 | 19.1 | 18.8 | 24.7 |
| Eastern | 20.4 | 20.8 | 14.8 | 16.4 | 21.7 | 20.1 |
| North Western | 14.1 | 17.8 | 13.9 | 21.3 | 15.2 | 16.0 |
| North Central | 10.0 | 12.8 | 15.2 | 13.6 | 16.3 | 19.7 |
| Uva | 13.1 | 12.6 | 14.9 | 21.1 | 22.0 | 24.8 |
| Sabaragamuwa | 20.8 | 17.1 | 27.3 | 30.9 | 29.6 | 30.0 |

Source: Department of Census and Statistics (2011-2016)

Further, it indicates that Southern province has the highest youth unemployment rate in all the years (Table 1.2).

### 1.4 Distribution of G.C.E. (O/L) \& G.C.E. (A/L) Passed Percentages

The results in Table 1.3 indicates the educational qualifications obtained by the Sri Lankan youth during the period of 2011 to 2016. These percentages were computed with respect to the total number of candidates sat for the examination.

Table 1.3: Percentage of candidates who passed G.C.E. (O/L) and G.C.E. (A/L) Examinations (2011-2016)

| Year | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| :--- | ---: | ---: | ---: | ---: | ---: | :---: |
| Passed G.C.E. O/L Examination | 60.8 | 64.74 | 66.67 | 69.02 | 69.33 | 69.94 |
| Passed G.C.E. A/L Examination |  | 63.15 | 58.56 | 61.25 | 62.35 | 63.36 |

Source: Department of Examinations (2011-2016)

These results very clearly indicate that more than $60 \%$ was able to successfully get through these examinations. Although the majority of the Sri Lankan youth is educated, they do not have enough employment opportunities in the job market and this has been further confirmed by the following labour force statistics obtained from the Department of Census and Statistics in 2016 (Table 1.3).

Table 1.4: Youth unemployment rates with respect to their level of education (2011-2016)

| Year | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| G.C.E. O/L | 18.3 | 20.5 | 21.9 | 22.8 | 23.9 | 22.4 |
| G.C.E. A/L and above | 33.9 | 30.6 | 33.5 | 31.3 | 34.0 | 32.5 |

Source: Department of Census and Statistics (2011-2016)

According to the Table 1.4, youth unemployment rates are higher among the educated group with G.C.E. A/L and above educational qualifications than those who obtain G.C.E. O/L qualification. Therefore, this study focuses on identifying the factors affecting the unemployment of educated youth with G.C.E O/L and G.C.E. A/L qualifications.

### 1.5 Ratio of the Youth Unemployment Rates to Adult Unemployment

Below Table 1.5 illustrates the ratio of the youth unemployment rates to adult unemployment rates in Sri Lanka over the period 2011 to 2016.

Table 1.5: Ratio of the youth unemployment rates to the adult unemployment rates in Sri Lanka (2011-2016)

| Year | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Males | 10.8 | 11.2 | 11.1 | 10.8 | 12.8 | 15.3 |
| Females | 6.0 | 6.4 | 6.5 | 7.5 | 5.7 | 7.2 |

Source: Department of Census and Statistics (2011-2016)
As shown in Table 1.5, the ratio of the youth unemployment rate to the adult unemployment rate of males ranges from 10.8 to 15.3 while range of females is in between 6.0 to 7.2. In addition, the ratio above one indicates that youth unemployment is significantly higher than the adult unemployment. The corresponding rates by provinces are shown in Table 1.6.

Table 1.6: Youth unemployment as a proportion of total unemployment by Province (2016)

| Provinces | Percentage |
| :--- | :---: |
| Sabaragamuwa | 64.1 |
| Uva | 59.7 |
| Nothern | 59.6 |
| North Central | 58.7 |
| Southern | 56.2 |
| North Western | 53.4 |
| Western | 53.0 |
| Central | 51.8 |
| Eastern | 50.9 |
| Sri Lanka | 55.9 |

Source: Department of Census and Statistics (2016)
Figures in Table 1.6 clearly indicates that the corresponding proportions in Sabaragamuwa, Uva, Northern, North Central and Southern provinces are higher than the national proportion of $55.9 \%$, however in all the provinces this ratio is higher than $50 \%$.

### 1.6 NEET Rate

The Table 1.7 indicates the gender composition of the youth NEET (Rate of youth not in employment, education and training) Rate during the period 2011 to 2016.

Table 1.7: NEET rate as a percentage to total youth by gender (2011-2016)

| Year | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Males | 13.0 | 12.7 | 17.0 | 17.6 | 16.3 | 17.0 |
| Females | 33.4 | 34.7 | 35.5 | 37.4 | 34.6 | 34.5 |

Source: Department of Census and Statistics (2011-2016)

Accordingly, the percentage of the females who are not in employment, education or training is comparatively higher than that of males over the period from 2011 to 2016.

Table 1.8: NEET rate as a percentage to total youth by educational qualifications (2016)

| Educational Qualifications | Males | Females |
| :--- | :---: | :---: |
| G.C.E. O/L | 13.2 | 26.7 |
| G.C.E. A/L and above | 22.8 | 34.7 |

Source: Department of Census and Statistics (2016)

Results in Table 1.8 indicates that female NEET rate is higher from that of male for both the education groups. The lowest NEET rate is reported for the group with G.C.E. (O/L) level of education for both males and females.

All the above results highlighted that although the majority of the Sri Lankan youth is educated, they do not get enough opportunities in the labour market. Therefore, this research study is aimed in identifying the factors affecting the unemployment of the educated youth with G.C.E. O/L and G.C.E. A/L qualifications in Sri Lanka.

### 1.7 Research Objectives

On view of the above explanation, the objectives of the study are to:
> Identify the significant factors that influence on the educated unemployed youth with G.C.E. O/L qualifications and above.
> Develop a predictive model for the educated unemployed youth.

### 1.8 Outline of the Dissertation

The dissertation will be organized with six chapters as follows. The Chapter 1 includes the introduction of the study; the Chapter 2 consists with both national and international studies related to the research topic carried out previously. The Chapter 3 explains information about the data and the statistical methodologies used for the study. The Chapter 4 gives the basic statistical analysis of the variables under each scenarios separately. The development of the predictive model using the four methods are discussed in Chapter 5. The Chapter 6 provides the conclusion from the data analysis along with the appropriate recommendations.

## CHAPTER 2

## LITERATURE REVIEW

This chapter consists of detailed investigation of the previous studies related to the unemployment of educated youth in Sri Lanka and other countries.

### 2.1 Factors Influencing on Women Unemployment in Sri Lanka

In the Sri Lankan labour force, the participation of females are less in comparison to males. As stated in the World Bank report in 2017, less representation of females in the labour force was mainly due to socio cultural factors, mismatching skills and gender discrimination practices. These factors were elaborated further in the following sections.

### 2.1.1 Impact of Socio-cultural factors on Unemployment

Firstly the household roles and responsibilities of women especially when they get married at young ages, they are less likely to engage in the labour force. According to 2015 statistics, marriage persons indicated a lower percentage (4.4\%) in the odds of Female Labour Force Participation (FLFP), while men's participation was comparatively high (11\%). In the early stages before 2010, married women with small children showed a less chance of becoming a paid employee and also there earnings were less with compared to the men (World Bank, 2017).

Based on the study conducted using a sample size of 150 by Gunatilaka (2013), due to certain cultural beliefs and norms, females are supposed to engage in in-house activities. This may cause for the reasons for gender gap in LFP by women. This study claimed that $70 \%$ of the married women in Sri Lanka are having at least 1 child under the age of 5 and are less likely to engage in work rather than spending their time at homes. This rate is comparatively higher $(75 \%)$ with women in the urban areas than in the rural areas ( $71 \%$ ). The study also revealed that the married females who are the heads of households have certain cultural constraints such as status-related perceptions, attitudes regarding their roles as married women and the gender division of household and care labour within the family unit. But the problem is, with
all these constraints they are encouraged to seek for employment, although there were some restrictions imposed by the private sector on the nature and type of work that women were able to take up (Gunatilaka, 2013).

A study conducted based on 200 households who are having at least one pre-school child in urban areas, by Gunathilaka (2010) suggested that non market work of women is not depending on their husband's income, but rather on the proportion of household activities and child care of their adults in the household. If these women can share their house work and child caring services with others, they are most likely to engage in the labour force. There are several factors influencing on the decision of getting formal child care such as the age of their children, household income, cost of the day care centers, quality of child care etc. Therefore, it's not simply the availability of those service providers but the quality and the affordability really affects the participation of the females in the labour force.

### 2.1.2 Impact of Education on Unemployment

Mismatching skills with respect to the availability of jobs is another reason for the increase in unemployment. Qualitative studies revealed that majority of the Sri Lankan women are preferred to continue their higher studies in the subject areas like humanities and arts but in the job market, there are more demand and job opportunities available for the people who continue their studies in the subject areas like Information Technology and Management (World Bank, 2017). Furthermore, it confirms with the following Table 2.1 that indicates the number of candidates who passed from G.C.E A/L examination and qualified to enter into any higher educational institute with respect to different subject streams.

Table 2.1: Percentage of candidates admitted for the higher educational institutions with respect to different subject streams (2014-2016)

| Subject <br> Stream | 2014 |  |  | 2015 |  |  | 2016 |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | No. <br> Qualified | No. <br> Admitted | $\%$ <br> Admitted | No. <br> Qualified | No. <br> Admitted | $\%$ <br> Admitted | No. <br> Qualified | No. <br> Admitted | $\%$ <br> Admitted |
| Arts | 68,565 | 8,617 | 12.57 | 65,511 | 9,891 | 15.1 | 64,218 | 9,418 | 14.67 |
| Commerce | 40,602 | 5,299 | 13.05 | 40,918 | 5,441 | 13.3 | 38,768 | 5,793 | 14.94 |
| Physical <br> Science | 15,634 | 5,287 | 33.82 | 15,809 | 5,374 | 33.99 | 18,211 | 6,023 | 33.07 |
| Biological <br> Science | 23,419 | 6,316 | 26.97 | 24,877 | 6,288 | 25.28 | 25,876 | 7,032 | 27.18 |
| Other | 1,352 | 124 | 9.17 | 1,582 | 210 | 13.27 | 2,020 | 324 | 16.04 |
| Total | 149,572 | 25,643 | 17.14 | 155,550 | 29,055 | 18.68 | 160,517 | 30,662 | 19.1 |

Source: University Grants Commission (2018)

According to the Table 2.1, it can be seen that the highest number of persons qualified are from Arts followed by Commerce irrespective of years. However, the percentage of admitted persons is high among Physical Science stream followed by Biological Science stream.

Table 2.2: Distribution of top 20 occupations on total recruitments (2015-2017)

| Rank | Occupation | 2015 |  | 2016 |  | 2017 |  |
| :---: | :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | No | $\%$ | No | $\%$ | No | $\%$ |
| 1 | Sewing Machine Operators | 297,627 | 32.7 | 240,688 | 27.3 | 188,117 | 28.4 |
| 2 | Security Guards | 179,284 | 19.7 | 171,135 | 19.4 | 114,370 | 17.3 |
| 3 | Other Manufacturing Labourers | 100,282 | 11.0 | 107,594 | 12.2 | 84,629 | 12.8 |
| 4 | Commercial and Sales <br> Representatives | 78,634 | 8.6 | 79,348 | 9.0 | 52,344 | 7.9 |
| 5 | Shop Sales Assistants | 33,844 | 3.7 | 45,158 | 5.1 | 30,206 | 4.6 |
| 6 | Cleaners and Helpers in <br> Offices, Hotels and Other <br> Establishments | 24,783 | 2.7 | 26,008 | 2.9 | 28,024 | 4.2 |
| 7 | Manufacturing Supervisors | 29,185 | 3.2 | 32,861 | 3.7 | 26,073 | 3.9 |
| 8 | General Office Clerks | 20,442 | 2.2 | 23,117 | 2.6 | 20,410 | 3.1 |
| 9 | Tailors, Dressmakers, Furriers <br> and Hatters | 19,276 | 2.1 | 22,528 | 2.6 | 16,617 | 2.5 |
| 10 | Stall and Market Salespersons | 11,217 | 1.2 | 16,378 | 1.9 | 13,878 | 2.1 |
| 11 | Hand Packers | 14,745 | 1.6 | 14,444 | 1.6 | 12,916 | 1.9 |
| 12 | Mechanical Engineering <br> Technicians | 16,933 | 1.9 | 14,267 | 1.6 | 9,646 | 1.5 |
| 13 | Secondary Education Teachers | 7,588 | 0.8 | 9,438 | 1.1 | 9,558 | 1.4 |
| 14 | Stock Clerks | 9,764 | 1.1 | 10,511 | 1.2 | 8,902 | 1.3 |
| 15 | Sewers, Embroiders and <br> Related Workers | 12,269 | 1.3 | 13,768 | 1.6 | 8,249 | 1.2 |
| 16 | Information and <br> Communication Technology <br> Operation Technicians | 8,664 | 1.0 | 9,406 | 1.1 | 8,128 | 1.2 |
| 17 | Textile, Leather and Related <br> Pattern Makers and Cutters | 12,422 | 1.4 | 10,502 | 1.2 | 8,039 | 1.2 |
| 18 | 0.7 | 10,437 | 1.2 | 7,848 | 1.2 |  |  |

Table 2.2: Distribution of top 20 occupations on total recruitments (2015-2017) (Continued)

| Rank | Occupation |  | 2015 |  | 2016 |  | 2017 |  |
| :---: | :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | No | $\%$ | No | $\%$ | No | $\%$ |  |
| 19 | Accounting and Book Keeping <br> Clerks | 13,101 | 1.4 | 13,087 | 1.5 | 7,815 | 1.2 |  |
| 20 | Civil Engineering Labourers | 13,408 | 1.5 | 11,026 | 1.3 | 7,005 | 1.1 |  |
|  | Total (Top 20 Occupations <br> with respect to 2017) | 910,069 | 100.0 | 881,701 | 100.0 | 662,774 | 100.0 |  |

Source: Department of Census and Statistics (2015-2017)
According to the Annual Labour Demand Survey (2017) carried out by the Department of Census and Statistics, majority of the workers are engaged as sewing machine operators followed by security guards over the period of 2015 to 2017. Therefore, it's clear that highly demanded occupations were in contrast with most of the students selected subject streams.

### 2.1.3 Impact of Training Courses on Unemployment

Although, the technical education and vocational training are considered to be important in improving the employability of job seekers, Gunatilaka (2008) found that most of these programs are not being able to provide the expected results. One reason she claimed is that these training sessions are conducted based on the Western province, where the unemployment rate is considerably less. Other reasons are: (i) these training programs are not updated with the expected level of standards (ii) inadequate teachers/ resource personnel and teaching aids and insufficient practical exposure.

Another important fact is that although most of these training programs are designed for those people who have completed their G.C.E.O/L examination, a considerable amount of participants were from those who are qualified with G.C.E. A/L examination. This confirms that their formal schooling had not prepared their students for the job market. That is one of the main reason for the existence of mismatching skills with respect to the availability in jobs (Gunatilaka, 2008).

### 2.1.4 Impact of Gender discrimination on Unemployment

Gender may be discriminated based on the nature of the job. Especially females are given less priority in the jobs related to the field of construction and IT based on its nature. At times in the hiring and recruitment process, less priority is giving for those who do not have contacts with the existing employees at the work place (World Bank, 2017).

A study conducted by Gunatilaka (2008) had shown that the majority of the Sri Lankan employed persons are engaged in the informal work arrangements and among them only about $6 \%$ are having permanent engagements. In comparison to females, males comprises a larger portion of the employed in the informal economy. The share of informal work is highest for both males and females in the age category of 15-19 age group. Nevertheless she claimed that in order to reduce the unemployment, informal job creation can make a direct impact by creating more employment opportunities than in the formal job creation.

### 2.2 Distribution of Unemployment in Asia-Pacific Countries

In the global context, several studies were carried out in relation to the factors influencing educated unemployment (Asian Development Bank, 2015). Most of such studies indicate that although the gaps between males and females on education and health have been reducing in the Asia Pacific countries, still there are wage gaps in the labour market due to the less participation of females because of the cultural and social norms. This confirms by the percentage of values of the labour force participation rates of different countries with respect to their gender (Figure 2.1).


Figure 2.1: Labour Force Participation in Different Countries

## Source: World Bank (2017)

Figure 2.1, further indicates that the labour force participation of males is greater than the females who are in the age category of 15 years and above in all eight countries.

### 2.3 Distribution of Female Unemployment in Specific Countries



Figure 2.2: Female Labour Force Participation by Selected Country, Economic Status and Region (1993-2016)

According to Figure 2.2, Nepal, Sub Saharan Africa, Upper middle income countries, middle income countries and Malaysia indicate a higher rate in the Female labour force participation in comparison to Sri Lankan females. The Human Development Report (UNDP, 2015) indicates that women accounted for $52 \%$ of the world's work, which constitutes $52 \%$ of the work in the world and it's the total of $21 \%$ of the paid work and $31 \%$ of unpaid work. In contrast, men's total work consists of $38 \%$ of paid work, and only $10 \%$ of unpaid work. It is concluded that the women's share of the unpaid work is nearly three times that of men's unpaid work. The report of MGI (2015) states that if women were given equal chance as men in the labour market $26 \%$ could be added to global annual GDP by 2025.

### 2.4 Distribution of Unemployment in World Scenarios

Following Table 2.3 indicates the statistics of the world unemployment during the year 2016 and predicted values for the years 2017 and 2018.

Table 2.3: Unemployment Trends and Projections (2016-2018)

| Country Grouping | Unemployment Rates (2016- <br> $2018)$ |  |  | Unemployment in Millions <br> $(2016-2018)$ |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 2016 <br> Actual | 2017 <br> Predicted | 2018 <br> Predicted | 2016 <br> Actual | 2017 <br> Predicted | 2018 <br> Predicted |
|  | 5.7 | 5.8 | 5.8 | 197.7 | 201.1 | 203.8 |
| Developed Countries | 6.3 | 6.2 | 6.2 | 38.6 | 37.9 | 38.0 |
| Emerging Countries | 5.6 | 5.7 | 5.7 | 143.4 | 147.0 | 149.2 |
| Developing Countries | 5.6 | 5.5 | 5.5 | 15.7 | 16.1 | 16.6 |

Source: International Labour Organization (2017)

As indicated in Table 2.3, unemployment rate was $6.3 \%$ in the developed countries, $5.6 \%$ in emerging countries and developing countries in 2016. The overall global unemployment rate is $5.7 \%$ and it's expected to rise by $0.1 \%$ in 2018 . The details of the unemployment rates in Asia and the Pacific Region during the year 2016 and predicted values for the year 2017 and 2018 are shown in Table 2.4.

Table 2.4: Unemployment Trends and Projections in Asia and the Pacific Region (20162018)

| Country/Region | Unemployment Rates <br> $(2016-2018)$ |  |  | Unemployment in Millions <br> $(2016-2018)$ |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 2016 | 2017 | 2018 | 2016 | 2017 | 2018 |
| Asia and the Pacific | 4.2 | 4.2 | 4.3 | 84.4 | 85.4 | 86.5 |
| Eastern Asia | 4.5 | 4.5 | 4.5 | 41.6 | 41.9 | 42.4 |
| China | 4.6 | 4.6 | 4.7 | 37.3 | 37.6 | 37.9 |
| Japan | 3.1 | 3.0 | 3.0 | 2.0 | 1.9 | 2.0 |
| Korea, Republic of | 3.7 | 3.6 | 3.7 | 1.0 | 1.0 | 1.0 |
| South-Eastern Asia |  |  |  |  |  |  |
| and the Pacific | 3.8 | 3.8 | 3.9 | 13.3 | 13.7 | 14.0 |
| Australia | 5.7 | 5.5 | 5.3 | 0.7 | 0.7 | 0.7 |
| Indonesia | 5.6 | 5.8 | 5.9 | 7.1 | 7.4 | 7.7 |
| Southern Asia | 4.1 | 4.1 | 4.1 | 29.5 | 29.8 | 30.2 |
| India | 3.5 | 3.4 | 3.4 | 17.7 | 7.8 | 18.0 |

Source: International Labour Organization (2017)

According to Table 2.4, it is proved that in 2016 the overall unemployment rate in the Asian and the Pacific Region was $4.2 \%$. In the Eastern Asia the unemployment rate was $4.5 \%$, in the South-Eastern Asia and the Pacific the rate was $3.8 \%$ and in Southern Asia the unemployment rate was $4.1 \%$. Psacharopoulos and Tzannatos (1989) found that factors such as age, fertility and religion affect the Labour Force Participation Rate (LFPR) irrespective of the country that is considered. Additionally, Uwakwe (2004) stated that in Nigerian family responsibilities, pregnancy, and physical factors such as nutrition, water and health services are significantly influential factors on LFPR.

In Turkey, State and Planning Organization and the World Bank (2010) found that the FLFP is due to both socioeconomic and cultural factors such as household responsibilities, childcare/eldercare, urbanization and marital status. According to Faridi et al. (2009), factors like close relatives' educational status, household assets, spouse participation in economic activities, number of children, age of children and husband salary influence the female's decision on whether to participate or not participate in the labor market.

According to a study carried out in Pakistan by Khadim and Akram (2013) broadly identified three categories of factors that directly influences on the labour force participation in economic activity such as individual and demographic factors as age, education, marital status, socio economic condition factors as per capita income of the household, number of dependents, household type, geographic location factors as whether in urban and rural residences. According to the past trends, the labour force participation of women within the age group of 15-64 over the past two decades have declined from $57 \%$ to $55 \%$ worldwide. Participation of women in the Middle East and North African countries are lesser than $25 \%$. As estimated by Gallup, within the South Asian region men are having twice of the opportunity than women to engage in full time jobs (World Bank, 2017).

Due to certain gender specific constraints in comparison to male farmers, female farmers are less active in commercial farming and more likely to have lower output per unit of land. In the area of Central Highlands of Ethiopia, their output value of per hectare of the female headed households have been estimated $35 \%$ lesser in comparison to the male headed households According to a research findings of ILO, women who are in paid work earn less than $10 \%-30 \%$ on average with compared to men in 83 countries. These gaps are more in Middle East, North African and in OECD countries (ILO, 2017).

As stated by Amin and Alam (2008) in Malaysia the Muslim women whether they are married or not were significantly less likely to work for any payment rather than Buddhists and Hindus. As the single women living in the urban areas were not interested in engage in the work force, there was no significant difference in urban married women of different religions are working for payment.

Developing and middle income countries like Mexico, Columbia, Argentina, Brazil, states of India use subsidies to or public provision of child care and they are willing to bear the cost of women when they are involved in economic activities in their households. Further, several studies have found that subsidized childcare and free kindergarten facilities, combined with the transformation of public kindergarten from part-time to full-time, are having a significant influence on the labour force participation decision of mothers with preschool children (Baker, Glyn, \& Howell, 2005).

### 2.5 Summary of the Chapter 2

In the local context, household duties and responsibilities, certain cultural beliefs and social norms, mismatching skills with the availability of jobs, gender discrimination in job search, hiring and promotion process have been identified as the primary reasons for the labour force participation decision resulted in the increase of unemployment. Similarly, in the global context, age and fertility, religion, physical factors like nutrition and health conditions, cultural factors like household responsibilities, childcare, eldercare, urbanization, marital status, number of children, age of children, husbands' salary influences on the labour force participation decision of especially in women. However, no specific studies was found for the reasons of unemployment of youth having G.C.E. (O/L) \& G.C.E. (A/L) educational qualifications.

## CHAPTER 3

## MATERIALS AND METHODS

### 3.1 Materials

### 3.1.1 Secondary Data

For this study, data were obtained from Sri Lanka Annual Labour Force Survey in 2016 conducted by the Department of Census and Statistics, Sri lanka. The survey has been carried out from January to December in 2016, using a sample of 25,750 housing units which includes 85,082 individuals. Further, this provides the estimates of national, provincial and district level during the reference period of 2016. This survey has collected information on demographic characteristics of the usual residents of the household economically active and inactive conditions, employment and unemployment characteristics, underemployment and informal sector employment information.

### 3.1.2 Sampling Methodology of the LFS

The annual labour force survey is conducted by using the two stage stratified sampling technique and selected a sample of 25,750 housing units. It indicates the persons who are only living in housing units and excludes the institutional population. For the sampling frame, Census of Population and Housing in 2012 has been used. Here, the census blocks prepared for the Census of Population and Housing in 2012 have been selected as the primary sampling units and the secondary sampling units are the housing units selected from the 2575 primary sampling units. By using the method of systematic random sampling, from each of the selected primary sampling unit, 10 housing units (SSU) are selected for the survey.

By using the Neymann allocation method, Primary Sampling Units (PSUs) of 2575 were allocated to each district and to each sectors (Urban, Rural and Estate). Then the sample which is allocated for each district is equally distributed among 12 months (DCS, 2016).

Following Table 3.1 indicates the distributions of the sample by district for the year 2016.

Table 3.1: Sample allocation by District (2016)

| District | No of Housing Units | District | No of Housing Units |
| :--- | :---: | :--- | :---: |
| Colombo | 2850 | Kilinochchi | 400 |
| Gampaha | 2400 | Batticaloa | 820 |
| Kalutara | 1450 | Ampara | 900 |
| Kandy | 1500 | Trincomalee | 600 |
| Matale | 700 | Kurunegala | 1750 |
| Nuwara Eliya | 900 | Puttalam | 800 |
| Galle | 1400 | Anuradhapura | 850 |
| Matara | 1250 | Polonnaruwa | 650 |
| Hambantota | 900 | Badulla | 850 |
| Jaffna | 750 | Moneragala | 650 |
| Mannar | 370 | Ratnapura | 1150 |
| Vavunia | 400 | Kegalle | 1100 |
| Mullaitivu | 360 | Total | 25750 |

Source: Department of Census and Statistics (2016)

Table 3.2, illustrates an outline of the Labour Force Survey Data in 2016 obtained from the Department of Census and Statistics.

Table 3.2: Descriptive statistics extracted from Sri Lanka Annual Labour Force Survey (2016)

| Labour Force Characteristics | Estimate | $\%$ | Standard <br> Error | Coefficient <br> of Variation |
| :--- | :---: | :---: | :---: | :---: |
| Population (15 years \& above) | $15,448,679$ | 100 | 98,235 | 0.64 |
| Labour Force | $8,310,682$ | 100 | 66,023 | 0.79 |
| Inactive | $7,137,997$ | 100 | 57,995 | 0.81 |
| Employed | $7,947,683$ | 100 | 64,616 | 0.13 |
| Unemployed | 362,999 | 100 | 11,186 | 3.08 |

[^0]According to the Table 3.2, the estimated total population in 2016 was $15,448,679$ and estimated labour force of the country was $8,310,682$ and further these results indicate that the estimated employed total population was $7,137,997$ while the unemployed population amounting to 362,999 .

Out of the entire population, 85,082 of individuals were selected for the Annual Labour Force Survey in 2016 and this analysis is carried out based on 4,002 individuals those who are among the youth within the age category of (15-24) and qualified with the G.C.E. O/L and G.C.E. A/L Academic qualifications.

### 3.1.3 Variables Used for the Analysis

Following Table 3.3 indicates a list of variables selected for the analysis.

Table 3.3: List of Variables used for the analysis

| Variables | Code |  |
| :--- | :--- | :---: |
|  | Male | 1 |
|  | Female | 2 |
| Race - X2 | Sinhala | 1 |
|  | Tamil | 2 |
|  | Malay | 3 |
|  | Other | 9 |
| Marital Status - X4 | Buddhist | 1 |
|  | Hindu | 2 |
|  | Muslim | 3 |
|  | Other | 9 |
| Education Attainment - X5 | Single | 1 |
|  | Married | 2 |
|  | Other | 3 |
|  | Passed G.C.E. O/L | 1 |
| Literacy in English - X6 | Ability to read and write | 2 |
|  | Unable to read and write | 2 |
|  | Urban | 1 |
|  | Rural | 2 |
|  | Estate | 3 |

### 3.1.4 Concepts and Definitions

The concepts and definitions used by the Department of Census and Statistics to disseminate labour force information will be used for this study. The recommended statistical definitions are as follows.
> Labour force: The labour force comprises of the economically active population with the age of 10 and over during the reference period. (Usually one week, the inquiry relates to activity or status is referred in the preceding week of the survey week.) Officially the Department of Census and Statistics used the lower age limit as 10 up to 2012 but a very small number of cases are reported in 10-14 age groups therefore from 2013 lower age limit as age 15 and above is considered as the working age population and upper age limit is not defined.
> Economically Active Population: These are the persons who are/were employed or unemployed in the survey reference period.
> Employed: These are the persons who worked as paid employees, employers, own account workers (self-employed), or unpaid family workers are said to be employed in the reference period. These indicate persons with a job but not at work during the reference period.
> Unemployed: These are the persons who are seeking and available for work, but had no employment during the reference period.
> Currently Economically Active: These persons were employed or unemployed during the current reference period are considered as currently economically active.
> Not in the Labour Force (not economically active): These are the persons who are categorized as "not in the labour force" are neither working not available/ looking for work. Persons who are not in the labour force are due to causes such as full time care of the household, full time students, retired or old age, infirmed or disabled, or are not willing to work for any of the reasons (DCS, 2016).

### 3.2 Statistical Techniques Used

### 3.2.1 Analysis of 2-way Freq. Tables

In typical two factors (A \& B) having 2 levels can be illustrated as shown below.
Table 3.4: Two way Frequency Table

| $A$ | B |  | Total |
| :--- | :---: | :---: | :---: |
|  | $\mathrm{B}_{1}$ | $\mathrm{~B}_{2}$ |  |
| $\mathrm{~A}_{1}$ | $\mathrm{f}_{11}$ | $\mathrm{f}_{12}$ | $\mathrm{f}_{1 .}$ |
| $\mathrm{A}_{2}$ | $\mathrm{f}_{21}$ | $\mathrm{f}_{22}$ | $\mathrm{f}_{2 .}$ |
| Total | $\mathrm{f}_{.1}$ | $\mathrm{f}_{.2}$ | $\mathrm{f} .$. |

Let $\left\{\mathrm{f}_{\mathrm{ij}}\right\}=$ Observed frequency of the row category $=\mathrm{i}$ and column category $=\mathrm{j}$

## Hypotheses

$\mathrm{H}_{0}$ : Factor A is independent of factor B or there is no significant association between the two factors A and B.
$\mathrm{H}_{1}$ : There is a significant association between the two variables.

### 3.2.2 Tests used for the analysis of 2-D table

The three common statistics used to test the above hypothesis are based on 2 way frequency table.
$>$ Pearson's Chi-Square Test (exact $)=\mathcal{X}^{2}=\sum \frac{(\text { Observed-Expected })^{2}}{\text { Expected }}$
$>$ Yates Correction Chi-Square Test $=\sum \frac{(\text { Observed }- \text { Expected }-0.5)^{2}}{\text { Expected }}$
$>$ Likelihood Ratio $=\sum$ Observed $* \log \left(\frac{\text { observed }}{\text { Expected }}\right)$
Each statistic is distributed $x^{2}(\mathrm{r}-1)(\mathrm{c}-1) \quad$ where r and c are the number of rows and columns. Source: Peiris (2018)

### 3.2.3 Binary Logistic Regression Analysis

The goal of logistic regression is to predict the dichotomous characteristic of dependent variable based on the continuous and/ or categorical independents and to identify the variability of the dependent variable on the basis of the independents, to rank the relative importance of independents, to assess the interaction effects, and to measure the impact of covariate control variables. Logistic regression comes under the family of generalized linear models. The impact of predictor variables is usually explained in terms of odd ratios.

## > Odd Ratios

The odds are simply the ratios of the proportions for the two possible outcomes of the binary logistic regression. If p is the proportion of the event outcome then $(1-\mathrm{p})$ is the proportion for the next outcome and the odds of an event is defined as,

$$
\text { Odds }=\frac{\mathrm{p}}{1-\mathrm{p}}
$$

The $\log$ transformation of $p$ is also called as the logit of $p$ or logit $(p)$ and thus it is defined as,

$$
\operatorname{Logit}(p)=\log \left(\frac{p}{1-p}\right)
$$

Logistic regression uses maximum likelihood estimation by changing the dependent in to logit variable. Estimation in logistic regression chooses parameters that maximize the likelihood function. Continuous variables are not taken as dependent variables in logistic regression and also only one dependent variable can exists.

Binary logistic regression is a branch of logistic regression which is used when the dependent is a dichotomous. Assume $\mathrm{X}_{1 \mathrm{i}}, \mathrm{X}_{2 \mathrm{i}} \ldots \ldots . \mathrm{X}_{\mathrm{ki}}$ are the explanatory variables for the $\mathrm{i}^{\text {th }}$ individual. The response is a dichotomous variable having two outcomes. Then the binary logistic model gives the relationship between the response and explanatory variables as follows.

$$
\log \left(\frac{\mathrm{p}_{\mathrm{i}}}{1-\mathrm{p}_{\mathrm{i}}}\right)=\beta_{0}+\beta_{1} x_{1 \mathrm{i}}++\beta_{2} x_{2 \mathrm{i}}+\ldots \ldots \ldots \ldots+\beta_{\mathrm{k}} x_{\mathrm{ki}}
$$

Where $\mathrm{i}=1,2, \ldots . . \mathrm{n} \quad X_{i 0}=1$ for all $\mathrm{i}=1,2, \ldots . . n$
$p_{i}$ is the probabilities of occurrence of the response of the interest of the $\mathrm{i}^{\text {th }}$ individual.
The model can alternatively be expressed in the form of,

$$
\mathrm{p}_{\mathrm{i}}=\frac{\exp \left(\beta_{0}+\beta_{1} x_{1 i}+\beta_{2} x_{2 \mathrm{i}} \ldots \ldots . . . . . .+\beta_{\mathrm{kx}} x_{\mathrm{ki}}\right)}{1+\exp \left(\beta_{0}+\beta_{1} x_{1 i}+\beta_{2} x_{2 \mathrm{i}} \ldots \ldots \ldots \ldots . .+\beta_{\mathrm{kx}} x_{\mathrm{ki}}\right)}
$$

Where $\mathrm{i}=1,2, \ldots . \mathrm{n} \quad$ (Peiris, 2018)

### 3.2.4 Variable Selection Methods in Binary Logistic Model

The four main methods that are used for the variable selection are as follows.
> Forward Selection (Likelihood Ratio) - It's a stepwise selection method which uses the entry testing on the basis of the significance of the score statistic, and removal testing is done on the basis of the probability of a likelihood-ratio statistic which uses the maximum partial likelihood estimates.
> Forward Selection (Wald) - It's a stepwise selection method which uses the entry testing on the basis of the significance of the score statistic, and removal testing is done on the basis of the probability of the Wald statistic.
> Backward Elimination (Likelihood Ratio) - It's Backward stepwise selection method which uses the removal testing on the basis of the probability of likelihood ratio statistic and uses the maximum partial likelihood estimates.
> Backward Elimination (Wald) - It's a backward stepwise selection method which uses the removal testing on the basis of the probability of the Wald statistic (Peiris, 2018).

### 3.2.5 Model Selection Procedure

Forward selection method (Likelihood Ratio) is used to select the most suitable model. This method starts with the null model (simplest model only with the intercept). Then the most significant variable (main effect) is added to the model. The variable with the lowest p value (at a given significant level) is considered to choose the most significant variable. In this way by adding one variable at a time to the each new model, the finalized model is defined when there is no further improvement. Similarly, higher order interactions are added thereafter.

In order to assess the model fit,
$\mathrm{H}_{0}$ : The model fits the data vs $\mathrm{H}_{1}$ : Model does not fits the data
In order to test the significance of logistic regression model, following tests can be used.

## - Deviance Test

In each of these competing models, both the null and the alternative models are separately fitted and at each step, log-likelihood statistic is computed as shown below. The test statistic is indicated by $D$-Deviance which is twice as the difference in these log-likelihoods.

$$
D=-2 \log \left(\frac{\text { likelihood for null model }}{\text { likelihood for alternative model }}\right)=2 \log (L)
$$

Under $\mathrm{H}_{0}: \mathrm{D} \sim \mathrm{X}^{2}{ }_{1}$
D statistic is the difference of log likelihood between two models.

## - Hosmer and Lemeshow Test

This statistical test is used to find out the goodness of fit for the logistic regression models. Most commonly it is used in the risk prediction model. The test identifies whether the observed data can be matched with the expected data under hypothetical model.

The test statistic H is given by,

$$
\mathrm{H}=\sum_{g=1}^{n} \frac{\left(o_{g}-E_{g}\right)^{2}}{N_{g} p_{g}\left(1-p_{g}\right)}
$$

$O_{g}$ - Number of observed cases in $\mathrm{g}^{\text {th }}$ group
$E_{g}-$ Number of expected cases $\mathrm{g}^{\text {th }}$ group under the fitted model
$g$ - Number of groups

## Hypotheses

$\mathrm{H}_{0}$ : Model is significant vs $\mathrm{H}_{1}$ : Model is not significant

Under $\mathrm{H}_{0}$, the test statistic is asymptotically follows Chi-square g-2 df (Peiris, 2018).

## CHAPTER 4

## INFLUENCE OF THE SELECTED EXPLANATORY VARIABLES ON EDUCATED UNEMPLOYED YOUTH (EUY)

This section investigates the influence of each selected explanatory variable separately on EUY. The results of the statistical analysis and the interpretations are given in this chapter.

### 4.1 Influence of Gender on EUY

Following Table 4.1, shows the results of the analysis of 2-way frequency table between gender and EUY.

Table 4.1: Influence of Gender on EUY

|  |  | Labour_Force |  | Total |  |
| :--- | :--- | :--- | ---: | ---: | ---: |
|  |  | Yes | No |  |  |
| Gender | Male | Count | 515 | 1143 | 1658 |
|  |  | \% within Gender | $31.1 \%$ | $68.9 \%$ | $100.0 \%$ |
|  | Female | Count | 446 | 1898 | 2344 |
|  |  | \% within Gender | $19.0 \%$ | $81.0 \%$ | $100.0 \%$ |

Chi Square test statistic $-X_{(1)}^{2}=77.075(\mathrm{p}=0.000)$

As the chi square test statistic (77.075) is highly significant ( $\mathrm{p}=0.000$ ), it can be concluded that there is a significant influence of gender on EUY. Among the educated female youth, $81.0 \%$ of females are not in the labour force (unemployed) and among the educated male youth, $68.9 \%$ of males are not in the labour force (Table 4.1). Nevertheless, as chi square is significant, it can be concluded that the percentage of educated female unemployed youth is significantly higher than the percentage of educated male unemployed youth in Sri Lanka.

### 4.2 Influence of Race on EUY

The results of 2- way frequency table between race and EUY are shown in Table 4.2.

Table 4.2: Influence of Race on EUY


As the results of the Chi Square statistic (7.331) is not significant ( $\mathrm{p}>0.05$ ), it can be concluded that there is no significant influence of race on EUY. Thus, it can be concluded that the percentage of educated unemployed youth is significantly different among races. The lowest percentage $(74.8 \%$ ) can be seen among Sinhalese while the highest percentage ( $81.8 \%$ ) is among different races other than Sinhala, Tamil or Malay.

### 4.3 Influence of Religion on EUY

Below Table 4.3, shows the results of the analysis of 2-way frequency table between religion and EUY.

Table 4.3: Influence of Religion on EUY

|  |  |  | Total |  |  |
| :--- | :--- | :--- | ---: | ---: | ---: |
|  |  | Yes |  | No |  |
| Religion | Buddhist | Count | 649 | 2000 | 2649 |
|  |  | \% within Religion | $24.5 \%$ | $75.5 \%$ | $100.0 \%$ |
|  |  | Hindu | Count | 105 | 463 |
|  | \% within Religion | $18.5 \%$ | $81.5 \%$ | $100.0 \%$ |  |
|  |  | Muslim | Count | 92 | 342 |
|  |  | \% within Religion | $21.2 \%$ | $78.8 \%$ | $100.0 \%$ |
|  | Other | Count | 115 | 236 | 351 |
|  |  | \% within Religion | $32.8 \%$ | $67.2 \%$ | $100.0 \%$ |
|  |  |  |  |  |  |

Chi Square test statistic - $x^{2}{ }_{(3)}=26.468(\mathrm{p}=0.000)$

The result of the Chi Square statistics $(26.468, \mathrm{p}=0.000)$ confirms that there is a significant influence of religion on EUY. Thus, it can be concluded that the percentage of educated unemployed youth is significantly different among religions. The corresponding percentages among Buddhists, Hindus, Malays and others are $75.5 \%, 81.5 \%, 78.8 \%$ and $67.2 \%$ respectively. (Table 4.3).

### 4.4 Influence of Marital status on EUY

Following Table 4.4, shows the results of the analysis of 2-way frequency table between marital status and EUY.

Table 4.4: Influence of Marital status on EUY

|  |  |  | Labour |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Yes | No |  |
|  |  | Count | 836 | 2684 | 3520 |
|  |  | \% within Marital_Status | 23.8\% | 76.3\% | 100.0\% |
| Marital St | Married | Count | 124 | 351 | 475 |
| Marital_Status | Married | \% within Marital_Status | 26.1\% | 73.9\% | 100.0\% |
|  |  | Count | 1 | 6 | 7 |
|  |  | \% within Marital_Status | 14.3\% | 85.7\% | 100.0\% |
|  | hi Squa | test statistic - $x^{2}$ | 673 (p= |  |  |

The results of the chi square statistics $(11.673, p=0.441)$ confirms that there is no significant influence of the marital status on EUY. Among the educated youth of singles, $76.3 \%$ of the majority are not in the labour force and among the married educated youth $73.9 \%$ are not in the labour force. As chi square is not significant, it can be concluded that marital status is not significantly influence on the educated youth unemployment irrespective of the gender (Table 4.4).

### 4.5 Influence of Educational Attainment on EUY

The results of 2-way frequency table between education attainment and EUY are shown in Table 4.5.

Table 4.5: Influence of Educational Attainment on EUY

|  |  |  | Labour_Force |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Yes | No |  |
| Education_Attainment |  | Count | 368 | 1678 | 2046 |
|  | Passed G.C.E. (O/L) | \% within | 18.0\% | 82.0\% | 100.0\% |
|  |  | Count | 593 | 1363 | 1956 |
|  | Passed G.C.E. (A/L) | \% within <br> Education Attainment | 30.3\% | 69.7\% | 100.0\% |

Chi Square test statistic - $\mathcal{X}^{2}{ }_{(5)}=83.327(\mathrm{p}=0.000)$

The Chi Square test statistic is significant ( $\mathrm{p}=0.000$ ). Thus, it can be concluded that there is a significant influence of educational attainment on EUY. The percentage of educated youth unemployed among those who passed G.C.E. (O/L) examination (82.0\%) significantly higher than that of those who passed G.C.E. (A/L) examination (69.7\%) (Table 4.5).

### 4.6 Influence of Literacy in English on EUY

Below Table 4.6, shows the results of the analysis of 2-way frequency table between the literacy in English and EUY.

Table 4.6: Influence of Literacy in English on EUY


Chi Square test statistic - $x^{2}{ }_{(6)}=4.412(\mathrm{p}=0.036)$
The results of the chi square test statistics (4.412, $\mathrm{p}=0.036$ ) indicate that there is a significant influence of the level of English literacy on EUY. According to the Table 4.6, among the educated youth who have the ability to read and write, $74.3 \%$ are not in the labour force and among the youth who are unable to read and write, $77.2 \%$ are not in the labour force. As chi
square is significant, it can be concluded that among the English literate persons, the percentage of educated unemployed youth is significantly lower than the percentage of educated employed youth among non-English literate persons.

### 4.7 Influence of Residential Sector on EUY

The results of 2- way frequency table between residential sector and EUY is shown in Table 4.7.

Table 4.7: Influence of Residential Sector on EUY

|  |  |  | Labour_Force |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Yes | No |  |
| Residential _ Sector | Urban | Count | 237 | 564 | 801 |
|  |  | \% within Sector | 29.6\% | 70.4\% | 100.0\% |
|  | Rural | Count | 704 | 2406 | 3110 |
|  | Rural | \% within Sector | 22.6\% | 77.4\% | 100.0\% |
|  | Estate | Count | 20 | 71 | 91 |
|  | Estate | \% within Sector | 22.0\% | 78.0\% | 100.0\% |

Chi Square test statistic - $x^{2}{ }_{(7)}=17.079(\mathrm{p}=0.000)$

Results of the Chi Square statistics $(17.079, p=0.000)$ confirms that there is a significant influence of the living area on EUY. In the urban sector $70.4 \%$, in the rural sector $77.4 \%$ and in the estate sector $78.0 \%$ are not in the labour force. It can be concluded that among the sectors, the percentage of educated unemployed youth in both rural and estate sector is significantly higher than that of in urban sector when each variable is considered separately.

### 4.8 Summary of the Chapter 4

The results of the Chi Square analysis confirmed that out of all the selected variables namely, gender, religion, education attainment, literacy level in English and residential sector have significant impact while race and marital status are not significantly impact on the unemployment of the educated youth in Sri Lanka. Furthermore, for each level of the above seven factors, the percentage of educated unemployed youth is higher than that of educated employed youth.

## CHAPTER 5

## MODELLING THE EDUCATED UNEMPLOYED YOUTH (EUY): LOGISTIC REGRESSION APPROACH

In this section, the effects of each of the selected seven variables have been analyzed simultaneously by using the binary logistic regression method. The seven variables are shown in Table 5.1.

Table 5.1: List of variables used in the model

| Variable | Variable Label |
| :--- | :---: |
| Gender | X1 |
| Race | $\mathbf{X 2}$ |
| Religion | X3 |
| Marital Status | X4 |
| Education Attainment | $\mathbf{X 5}$ |
| Literacy in English | $\mathbf{X 6}$ |
| Residential Sector | X7 |

The dichotomous variable is the labour force participation which indicates whether the youth is in the labour force or not.

Labour force participation $\left\{\begin{array}{l}1 \text { If educated youth is not in the labour force (unemployed) } \\ \mathbf{0} \text { If educated youth is in the labour force (employed) }\end{array}\right.$

In order to find out the significant variables, when all the variables considered simultaneously the following four approaches were carried out.
> Forward Selection (Likelihood) Method
> Forward Selection (Wald) Method
> Backward Elimination (Likelihood Ratio) Method
> Backward Elimination (Wald) Method

As similar results were obtained in all the four methods, only the results obtained in Forward Selection (Likelihood) Method is illustrated below.

### 5.1 Use Forward Selection (Likelihood) Method

The FS method is based on the likelihood ratio and the model starts only with the constant model. Then most influential variable is added to the model. This continues until there is no further improvements for the model. The reference groups used for the analysis is indicated in Table 5.2.

Table 5.2: Reference Categories used for the model

| Variables | Reference Category |
| :--- | :---: |
| Gender (X1) | Female |
| Race (X2) | Other |
| Religion (X3) | Other |
| Marital Status (X4) | Other |
| Education Attainment (X5) | Passed G.C.E. (A/L) |
| Literacy in English (X6) | Unable to read and write |
| Residential Sector (X7) | Estate |

### 5.1.1 Goodness of Fit of the Fitted Model for the Main Effects

The goodness of fit for the overall model can be discussed by using model Chi Square test.
Table 5.3 and Table 5.4 show the results for the overall significance of the model.

Table 5.3: Model Summary

| Step | -2 Log likelihood | Cox \& Snell $R$ <br> Square | Nagelkerke R <br> Square |
| :--- | ---: | ---: | ---: |
| 1 | $4328.207^{\mathrm{a}}$ | .021 | .031 |
| 2 | $4230.149^{\mathrm{a}}$ | .044 | .067 |
| 3 | $4209.651^{\mathrm{a}}$ | .049 | .074 |
| 4 | $4195.495^{\mathrm{a}}$ | .053 | .079 |
| 5 | $4181.065^{\mathrm{b}}$ | .056 | .084 |

These indicators sometimes known as Pseudo $\mathrm{R}^{2}$. It indicates the explained variation in the dependent variable based on the final model varies from $5.6 \%$ to $8.4 \%$ depending on Cox \& Snell $\mathrm{R}^{2}$ and Nagelkerke $\mathrm{R}^{2}$ respectively (Table 5.3).

Table 5.4: Results of Hosmer and Lemeshow Test

| Step | Chi-square | df | Sig. |
| :--- | ---: | ---: | ---: |
| 1 | .000 | 0 |  |
| 2 | 23.802 | 2 | .000 |
| 3 | 22.069 | 6 | .001 |
| 4 | 24.573 | 7 | .001 |
| 5 | 27.309 | 8 | .001 |

According to Table 5.4, the value given in the fourth column is the probability of the chi square statistic used to test the null hypothesis ( $\mathrm{H}_{0}$ : Model is significant) (Section 3.25, 25 pg ). In other words, this is the probability of obtaining this chi-square statistic (27.309) for goodness of fit of the model. In this case, the models in each steps are not significant as the corresponding p values are less than 0.05 . This implies that the model can be further improved by including the interaction terms as discussed in section 5.2.

The SPSS output of the final model is shown in Table 5.5.

Table 5.5: Final results of the Logistic Regression model via Forward Selection (LR) Method

|  | B | S.E. | Wald | df | Sig. | $\operatorname{Exp}(\mathrm{B})$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Gender (X1) |  |  |  |  |  |  |
| Male | -. 824 | . 080 | 106.965 | 1 | . 000 | . 439 |
| Religion (X3) |  |  | 20.605 | 3 | . 000 |  |
| Buddhist | . 343 | . 128 | 7.187 | 1 | . 007 | 1.409 |
| Hindu | . 679 | . 163 | 17.334 | 1 | . 000 | 1.973 |
| Muslim | . 573 | . 169 | 11.517 | 1 | . 001 | 1.774 |
| Marital_Status (X4) |  |  | 14.992 | 2 | . 001 |  |
| Step $5^{\text {e }} \quad$ Single | -. 561 | 1.129 | . 247 | 1 | . 619 | . 571 |
| Married | -1.016 | 1.134 | . 802 | 1 | . 370 | . 362 |
| Education_Attainment (X5) Passed G.C.E. (O/L) | . 780 | . 079 | 98.674 | 1 | . 000 | 2.182 |
| Residential_Sector (X7) |  |  | 15.354 | 2 | . 000 |  |
| Urban | -. 166 | . 281 | . 349 | 1 | . 555 | . 847 |
| Rural | . 200 | . 275 | . 529 | 1 | . 467 | 1.221 |
| Constant | 1.286 | 1.157 | 1.234 | 1 | . 267 | 3.618 |

According to Table 5.5, the fitted model can be written as follows;

$$
\begin{gathered}
\log \left(\frac{\mathrm{P}}{1-\mathrm{P}}\right)=1.286-0.824\left(x_{1=1}\right)+0.343\left(x_{3=1}\right)+0.679\left(x_{3=2}\right)+0.573\left(x_{3=3}\right)-0.561\left(x_{4=1}\right)- \\
1.016\left(x_{4=2}\right)+0.780\left(x_{5=1}\right)-0.166\left(x_{7=1}\right)+2.00\left(x_{7=2}\right)
\end{gathered}
$$

According to the results in Table 5.5, it can be interpreted that there is a significant odds of educated males ( 0.439 ) who are unemployed compared to females. In comparison to the other category of religion, the most significant religion is Hindu (1.973) followed by Muslims (1.774) and Buddhists (1.409) who are among the unemployed educated youth. Odds of unemployed single youth is significantly higher (0.571) than those who are married (0.362). The results of the model indicates that the significant odds of unemployed youth who passed G.C.E. (O/L) examination (2.182) compared to those unemployed who passed the G.C.E. (A/L) examination. Further, the odds of educated unemployed youth representing the rural sector increases by 1.221 compared to those who are in the urban sector.

In order to check the goodness of fit of the model, predicted values were obtained at the critical level probability of 0.5 (Table 5.6).

Table 5.6: Observed and Predicted results of the EUY

|  |  |  | Predict | oup | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Yes | No |  |
|  |  | Count | 21 | 940 | 961 |
| Labour Force |  | \% within Labour_Force | 2.2\% | 97.8\% | 100.0\% |
| Labour_Force | No | Count | 14 | 3027 | 3041 |
|  |  | \% within Labour_Force | 0.5\% | 99.5\% | 100.0\% |
| Total |  | Count | 35 | 3967 | 4002 |
| Total |  | \% within Labour_Force | 0.9\% | 99.1\% | 100.0\% |

According to Table 5.6, 3027 are predicted correctly as unemployed (99.5\%) and from 961 educated youth employed, 21 are predicted correctly as employed ( $2.2 \%$ ). The overall correct classification by the model is $\frac{21+3027}{4002}=76.2 \%$

Based on the results obtained, it can be concluded that out of the selected seven variables, five variables are significantly influence on the educated youth unemployment and in order to study more details, 2-way interactions were tested by using the Forward LR method (Table 5.7).

### 5.2 Model with Two Way Interactions

In order to improve the model, the importance of the inclusion of interaction terms were pointed out in Section 5.1. However, the inclusion of all interactions make the model complicated. Therefore, higher order interactions were not considered. Furthermore among the ten possible 2 -way interactions, only four 2 -way interactions were identified by the model.

### 5.2.1 Goodness of Fit of the Fitted Model with Two Way Interactions

Table 5.7 and Table 5.8 indicate the results for the overall significance of the model.

Table 5.7: Model Summary for Two way Interactions

| Step | -2 Log likelihood | Cox \& Snell $R$ <br> Square | Nagelkerke $R$ <br> Square |
| :--- | ---: | ---: | ---: |
| 1 | $4251.936^{\mathrm{a}}$ | .039 | .059 |
| 2 | $4137.147^{\mathrm{a}}$ | .066 | .099 |
| 3 | $4117.842^{\mathrm{b}}$ | .071 | .106 |
| 4 | $4097.599^{\mathrm{b}}$ | .076 | .113 |
| 5 | $4080.992^{\mathrm{b}}$ | .079 | .119 |
| 6 | $4067.485^{\mathrm{b}}$ | .082 | .124 |
| 7 | $4063.532^{\mathrm{b}}$ | .083 | .125 |

According to Table 5.7, the explained variation in the dependent variable based on the final model varies from $8.3 \%$ to $12.5 \%$ depending on Cox \& Snell $\mathrm{R}^{2}$ and Nagelkerke $\mathrm{R}^{2}$ respectively.

Table 5.8: Results of Hosmer and Lemeshow Test

| Step | Chi-square | df | Sig. |
| :--- | ---: | ---: | ---: |
| 1 | .000 | 1 | 1.000 |
| 2 | 17.645 | 3 | .001 |
| 3 | .039 | 3 | .998 |
| 4 | 6.793 | 7 | .451 |
| 5 | 6.166 | 7 | .520 |
| 6 | 6.775 | 7 | .453 |
| 7 | 6.704 | 7 | .460 |

Table 5.8 indicates the probability of obtaining this chi-square statistic (6.704) for goodness of fit of the model. In this case, the model is statistically significant as $\mathrm{p}=0.460$ ( $\mathrm{p}>0.05$ ).

The SPSS output of the final model with 2-way interactions is shown in Table 5.9.

Table 5.9: Final results of the Two way interactions via Forward Selection (LR) Method

|  | B | S.E. | Wald | df | Sig. | $\operatorname{Exp}(\mathrm{B})$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Religion (X3) |  |  | 33.712 | 3 | . 000 |  |
| Buddhist | . 412 | . 171 | 5.777 | 1 | . 016 | 1.509 |
| Hindu | 1.109 | . 234 | 22.426 | 1 | . 000 | 3.032 |
| $7^{\text {Step }}$ g Muslim | 1.115 | . 253 | 19.349 | 1 | . 000 | 3.048 |
| Education_Attainment (X5) Passed G.C.E.(O/L) | 1.152 | . 120 | 91.562 | 1 | . 000 | 3.165 |
| Residential_Sector (X7) |  |  | 15.831 | 2 | . 000 |  |
| Urban | -. 158 | . 288 | . 300 | 1 | . 584 | . 854 |
| Rural | . 224 | . 281 | . 633 | 1 | . 426 | 1.251 |
| $\begin{aligned} & \text { Gender (X1) * Education_Attainment (X5) } \\ & \text { Male * Passed G.C.E. (O/L) } \end{aligned}$ | -. 585 | . 168 | 12.130 | 1 | . 000 | . 557 |
| Gender (X1) * Religion (X3) |  |  | 16.434 | 3 | . 001 |  |
| Male * Buddhist | -. 206 | . 257 | . 642 | 1 | . 423 | . 814 |
| Male * Hindu | -. 853 | . 333 | 6.566 | 1 | . 010 | . 426 |
| Male * Muslim | -1.046 | . 352 | 8.814 | 1 | . 003 | . 351 |
| Gender (X1) * Marital_Status (X4) |  |  | 62.352 | 2 | . 000 |  |
| Male * Single | -. 181 | . 257 | . 500 | 1 | . 480 | . 834 |
| Male * Married | -2.650 | . 401 | 43.641 | 1 | . 000 | . 071 |
| Gender(X1) * English_Literacy(X6) <br> Male * Able to read and write | . 239 | . 121 | 3.919 | 1 | . 048 | 1.270 |
| Constant | . 366 | . 310 | 1.396 | 1 | . 237 | 1.442 |

The fitted model with the two way interactions can be written as follows;

$$
\begin{gathered}
\log \left(\frac{\mathrm{P}}{1-\mathrm{P}}\right)=0.366+0.412\left(x_{3=1}\right)+1.109\left(x_{3=2}\right)+1.115\left(x_{3=3}\right)+1.152\left(x_{5=1}\right)-0.158\left(x_{7=1}\right)+ \\
0.224\left(x_{7=2}\right)-0.585\left(x_{1=1^{*}} x_{5=1}\right)-0.206\left(x_{1=1^{*}} x_{3=1}\right)-0.853\left(x_{1=1 *} x_{3=2}\right)-1.046\left(x_{1=1 *} x_{3=3}\right)- \\
1.181\left(x_{1=1 *} x_{4=1}\right)-2.650\left(x_{1=1 *} x_{4=2}\right)+0.239\left(x_{1=1 *} x_{6=1}\right)
\end{gathered}
$$

According to Table 5.9, it can be interpreted that there is a significant odds of unemployed educated males who are Buddhists (0.814) followed by male Hindus (0.426) and male Muslims ( 0.351 ) when females with other category are controlled. Similarly, it can be seen that there are more single youth males (0.834) than the married males $(0.071)$ who are educated and not in the labour force. In addition, there is a significant amount of unemployed educated youth males in the rural sector (1.251) than the urban sector (0.854). Regarding the
education attainment, there's a significant amount of unemployed youth males who passed G.C.E. (O/L) examinations (0.557) compared to unemployed youth males who passed G.C.E. (A/L) examination. Similarly, more unemployed males are English literate (1.270) compared to females who are illiterate in English.

In order to check the goodness of fit of the model, predicted values were obtained at the critical level probability of 0.5 (Table 5.10).

Table 5.10: Observed and Predicted results of the EUY


According to Table 5.10, 3028 are predicted correctly as unemployed (99.6\%) and from 961 educated youth employed, 59 are predicted correctly as employed ( $6.1 \%$ ). The overall correct classification by the model is $\frac{59+3028}{4002}=77.1 \%$

### 5.3 Summary of 2-way Interactions

As four 2-way interactions were found significant, it would be more appropriate to derive percentages from those tables than percentages based on main effects only.

Table 5.11: Influence of Gender and Religion on EUY

| X3 $=$ Religion |  |  |  | $\mathrm{Y}=$ Labour force |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Yes | No |  |
| Buddhist | X1 | Male | Count | 334 | 780 | 1114 |
|  |  |  | \% within X1 | 30.0\% | 70.0\% | 100.0\% |
|  |  | Female | Count | 315 | 1220 | 1535 |
|  |  |  | \% within X1 | 20.5\% | 79.5\% | 100.0\% |
|  | Total |  | Count | 649 | 2000 | 2649 |
|  |  |  | \% within X1 | 24.5\% | 75.5\% | 100.0\% |
| Hindu | X1 | Male | Count | 66 | 153 | 219 |
|  |  |  | \% within X1 | 30.1\% | 69.9\% | 100.0\% |
|  |  | Female | Count | 39 | 310 | 349 |
|  |  |  | \% within X1 | 11.2\% | 88.8\% | 100.0\% |
|  | Total |  | Count | 105 | 463 | 568 |
|  |  |  | \% within X1 | 18.5\% | 81.5\% | 100.0\% |
| Muslim | X1 | Male | Count | 62 | 119 | 181 |
|  |  |  | \% within X1 | 34.3\% | 65.7\% | 100.0\% |
|  |  | Female | Count | 30 | 223 | 253 |
|  |  |  | \% within X1 | 11.9\% | 88.1\% | 100.0\% |
|  | Total |  | Count | 92 | 342 | 434 |
|  |  |  | \% within X1 | 21.2\% | 78.8\% | 100.0\% |
| Other | X1 | Male | Count | 53 | 91 | 144 |
|  |  |  | \% within X1 | 36.8\% | 63.2\% | 100.0\% |
|  |  | Female | Count | 62 | 145 | 207 |
|  |  |  | \% within X1 | 30.0\% | 70.0\% | 100.0\% |
|  | Total |  | Count | 115 | 236 | 351 |
|  |  |  | \% within X1 | 32.8\% | 67.2\% | 100.0\% |
| Total | X1 | Male | Count | 515 | 1143 | 1658 |
|  |  |  | \% within X1 | 31.1\% | 68.9\% | 100.0\% |
|  |  | Female | Count | 446 | 1898 | 2344 |
|  |  |  | \% within X1 | 19.0\% | 81.0\% | 100.0\% |
|  | Total |  | Count | 961 | 3041 | 4002 |
|  |  |  | \% within X1 | 24.0\% | 76.0\% | 100.0\% |

As indicated in Table 5.11, the percentage of male EUY rate is almost same irrespective of religion. However the percentage of female EUY rate among Hindu (88.8\%) and Muslim ( $88.1 \%$ ) religions are significantly higher than that of Buddhists ( $75.5 \%$ ).

Table 5.12: Influence of Gender and Marital status on EUY

| X4 = Marital status |  |  |  | $\mathrm{Y}=$ Labour force |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Yes | No |  |
| Single | X1 | Male | Count | 456 | 1129 | 1585 |
|  |  |  | \% within X1 | 28.8\% | 71.2\% | 100.0\% |
|  |  | Female | Count | 380 | 1555 | 1935 |
|  |  |  | \% within X1 | 19.6\% | 80.4\% | 100.0\% |
|  | Total |  | Count | 836 | 2684 | 3520 |
|  |  |  | \% within X1 | 23.8\% | 76.3\% | 100.0\% |
| Married | X1 | Male | Count | 59 | 13 | 72 |
|  |  |  | \% within X1 | 81.9\% | 18.1\% | 100.0\% |
|  |  | Female | Count | 65 | 338 | 403 |
|  |  |  | \% within X1 | 16.1\% | 83.9\% | 100.0\% |
|  | Total |  | Count | 124 | 351 | 475 |
|  |  |  | \% within X1 | 26.1\% | 73.9\% | 100.0\% |
| Other | X1 | Male | Count | 0 | 1 | 1 |
|  |  |  | \% within X1 | 0.0\% | 100.0\% | 100.0\% |
|  |  | Female | Count | 1 | 5 | 6 |
|  |  |  | \% within X1 | 16.7\% | 83.3\% | 100.0\% |
|  | Total |  | Count | 1 | 6 | 7 |
|  |  |  | \% within X1 | 14.3\% | 85.7\% | 100.0\% |
| Total | X1 | Male | Count | 515 | 1143 | 1658 |
|  |  |  | \% within X1 | 31.1\% | 68.9\% | 100.0\% |
|  |  | Female | Count | 446 | 1898 | 2344 |
|  |  |  | \% within X1 | 19.0\% | 81.0\% | 100.0\% |
|  | Total |  | Count | 961 | 3041 | 4002 |
|  |  |  | \% within X1 | 24.0\% | 76.0\% | 100.0\% |

According to Table 5.12, the percentage of female EUY is almost the same irrespective of marital status. However the percentage of married male EUY (18.1\%) is significantly lower than the percentage of unmarried male EUY (71.2\%).

Table 5.13: Influence of Gender and Education Attainment on EUY

| X5 = Education Attainment |  |  |  | $\mathrm{Y}=$ Labour force |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Yes | No |  |
| Passed G.C.E. (O/L) | X1 | Male | Count | 259 | 694 | 953 |
|  |  |  | \% within X1 | 27.2\% | 72.8\% | 100.0\% |
|  |  | Female | Count | 109 | 984 | 1093 |
|  |  |  | \% within X1 | 10.0\% | 90.0\% | 100.0\% |
|  | Total |  | Count | 368 | 1678 | 2046 |
|  |  |  | \% within X1 | 18.0\% | 82.0\% | 100.0\% |
| Passed G.C.E. (A/L) | X1 | Male | Count | 256 | 449 | 705 |
|  |  |  | \% within X1 | 36.3\% | 63.7\% | 100.0\% |
|  |  | Female | Count | 337 | 914 | 1251 |
|  |  |  | \% within X1 | 26.9\% | 73.1\% | 100.0\% |
|  | Total |  | Count | 593 | 1363 | 1956 |
|  |  |  | \% within X1 | 30.3\% | 69.7\% | 100.0\% |
| Total | X1 | Male | Count | 515 | 1143 | 1658 |
|  |  |  | \% within X1 | 31.1\% | 68.9\% | 100.0\% |
|  |  | Female | Count | 446 | 1898 | 2344 |
|  |  |  | \% within X1 | 19.0\% | 81.0\% | 100.0\% |
|  | Total |  | Count | 961 | 3041 | 4002 |
|  |  |  | \% within X1 | 24.0\% | 76.0\% | 100.0\% |

As illustrated in Table 5.13, the percentage of EUY who passed G.C.E. (O/L) examination is comparatively higher than the EUY who passed G.C.E. (A/L) examination irrespective of the gender. The highest percentage of EUY (63.7\%) was noted for the males having G.C.E. (A/L).

Table 5.14: Influence of Gender and English Literacy on EUY

| X6 = English Literacy |  |  |  | $\mathrm{Y}=$ Labour force |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | Yes | No |  |
| Ability to read and write | X1 | Male | Count | 193 | 445 | 638 |
|  |  |  | \% within X1 | 30.3\% | 69.7\% | 100.0\% |
|  |  | Female | Count | 228 | 771 | 999 |
|  |  |  | \% within X1 | 22.8\% | 77.2\% | 100.0\% |
|  | Total |  | Count | 421 | 1216 | 1637 |
|  |  |  | \% within X1 | 25.7\% | 74.3\% | 100.0\% |
| Unable to read and write | X1 | Male | Count | 322 | 698 | 1020 |
|  |  |  | \% within X1 | 31.6\% | 68.4\% | 100.0\% |
|  |  | Female | Count | 218 | 1127 | 1345 |
|  |  |  | \% within X1 | 16.2\% | 83.8\% | 100.0\% |
|  | Total |  | Count | 540 | 1825 | 2365 |
|  |  |  | \% within X1 | 22.8\% | 77.2\% | 100.0\% |
| Total | X1 | Male | Count | 515 | 1143 | 1658 |
|  |  |  | \% within X1 | 31.1\% | 68.9\% | 100.0\% |
|  |  | Female | Count | 446 | 1898 | 2344 |
|  |  |  | \% within X1 | 19.0\% | 81.0\% | 100.0\% |
|  | Total |  | Count | 961 | 3041 | 4002 |
|  |  |  | \% within X1 | 24.0\% | 76.0\% | 100.0\% |

According to Table 5.14, the percentage of male EUY who are unable to read and write $(68.4 \%)$ is almost the same that of males who are able to read and write $(69.7 \%)$. Of the females, the percentage of EUY is higher among those who are unable to read and write $(83.8 \%)$ than among those who are able to read and write ( $77.2 \%$ ).

### 5.4 Summary of Chapter 5

In summary, it is proved that the best fitted logistic method of selection is invariant by the method of the identification of significant variables as results were same for all the forward and backward methods under LR and Wald criteria. Gender (X1), religion (X3), marital status (X4), education attainment (X5) and residential sector (X7) were found to be the most significant variables. In the two way interactions, Gender (X1) and Religion (X3), Gender (X1) and Marital status (X4), Gender (X1) and Education attainment (X5), Gender (X1) and English Literacy (X6) are found to be significant on the educated youth unemployment. The model with main effects only was improved by considering 2-way interaction terms and by the model with five main effects (gender (X1), religion (X3), marital status (X4), education attainment (X5) and residential sector (X7)) and four 2-way interaction terms were found significant based on Hosmer and Lemeshow Test (H-L statistic $=6.704, \mathrm{p}=0.460$ ). The overall correct classification rate was $77.1 \%$.

## CHAPTER 6

## CONCLUSIONS AND RECOMMENDATIONS

This study has been carried out to identify the factors affecting the unemployment of the educated youth (EUY) with G.C.E. O/L and G.C.E. A/L qualifications in Sri Lanka. A corresponding sample of size 4002 was obtained from Labour Force Survey 2016 conducted by Department of Census and Statistics with a sample of 25,750 . Based on the analysis carried out in Chapter 4 and Chapter 5, following conclusions and recommendations can be given.

### 6.1 Conclusions

- This analysis pointed out that females among the youth are more unemployed than males. In addition, among the unemployed youth, majority are Sinhalese and Buddhist. In terms of their marital status, majority of the unemployed youth are single.
- With respect to their educational qualifications, majority who are not in the labour force have obtained G.C.E. O/L qualification. In addition, most of these unemployed educated youth are in the estate sector.
- According to the analysis of 2-way frequency table, it shows that out of the selected seven variables gender, religion, literacy in English, education attainment and the sector do have significant influence while race and the marital status do not have significant influence with the educated youth unemployment in Sri Lanka.
- Binary logistic model with five variables (gender, religion, marital status, education attainment, sector residential) and four 2-way interaction terms (gender and Religion, gender and marital status, gender and education attainment, gender and English literacy) were found to be significant to explain the variation of EUY. The overall correct classification is $77.1 \%$.


### 6.2 Recommendations

- The youth age group of (15-24) should be considered as a prime age group to be involved in the labour force irrespective of their gender.
- Among the youth population, as the proportion of educated females are more likely to be unemployed, it is necessary to create more chances for them in the job market as in some industries, gender discrimination takes place in hiring and recruitment process.
- According to the job market requirements, the secondary level and the tertiary level education system needs to be updated and the skills needs to be developed so that it would not be difficult for the females to find suitable jobs and contribute for the labour force.
- Further, in order to increase the women involvement in labour force, it's essential to initiate flexible working hours in companies either on shift basis or as part time work, so that it would be a support for the females in balancing their personal life and work life simultaneously.
- With the use of technological advancements, it is high time to create job opportunities for women to work from home which will increase the involvement of females in the labour force as in developed countries.
- In addition, the government can promote the concept of 'Entrepreneurship' as it would be more useful for both males and females to balance their life while earning a reasonable income and creating more employment opportunities for others as well.


## REFERENCES

Asian Development Bank. (2015). Women in the Workforce. Manila: ADB.

Baker, D.A., Glyn, D., \& Howell, J. (2005). Labor Market Institutions and Unemployment: A Critical Assessment of the Cross-Country Evidence In Fighting Unemployment: The Limits of Free Market Orthodoxy, ed. David R. Howell. Oxford, U.K.: Oxford University Press.

Central Bank of Sri Lanka. (2017). Annual Report of Central Bank of Sri Lanka 2017.

DCS (2011-2017). Sri Lanka Labour Force Survey Data. Department of Census and Statistics.

DCS (2016). Sri Lanka Labour Force Survey Annual Report. Department of Census and Statistics.

Faridi, M. Z., Chaudhry, I. S., \& Anwar, M. (2009). The socio-economic and demographic determinants of women work participation in Pakistan: evidence from Bahawalpur District.

Gunatilaka, R. (2008). Informal Employment in Sri Lanka: Nature, Probability of Employment and Determinants of Wages, ILO - Asia Pacific Working Paper Series.

Gunatilaka, R. (2013). Women's participation in Sri Lanka's Labour Force: Trends, Drivers and Constraints. Colombo: ILO.

Gunatilaka, R. (2016). Women's Activity Outcomes, Preferences and Time Use in Western Sri Lanka. Colombo: ILO. Retrieved from https://www.researchgate.net/publication/319535440

International Labour Organization. (2017). World Employment Social Outlook: Trends 2017. Geneva.

Khadim, Z., \& Akram, W. (2013). Female Labor Force Participation in Formal Sector: An Empirical Evidence from PSLM (2007-08). Middle-East Journal of Scientific Research, 14 (11), 1480-1488.

McKinsey Global Institute. (2015). The Power of Parity: How Advancing Women's Equality Can Add $\$ 12$ Trillion to Global Growth. McKinsey Global Institute. http://www.mckinsey.com/ global-themes/employment-and-growth/howadvancing-womens-equality-can-add-12-trillionto-global-growth, Accessed 10 October 2018.

Psacharopoulous, G., \& Tzannatos, Z. (1989). Female Labor Force Participation: An International Perspective. The International Bank for Reconstruction and Development The World Bank.

Pieris T.S.G. (2018). - Lecture Notes, Analysis of Categorical Data in Business.

State Planning Organization of the Republic of Turkey and World Bank. (2010). Determinants of and trends in labor force participation of women in Turkey.

Sunday Leader. (2016) Women's Labour Force Participation in Sri Lanka. Retrieved from http://www.thesundayleader.lk/2017/03/12/womens-labour-force-participation-in-sri-lanka/

UNDP Human Development Report. (2014). Gender Inequality Index. Retrieved from http://hdr.undp.org/en/content/gender-inequality-index-gii

UNDP. (2015). Human Development Report 2015. New York: United Nations Development Programme.

UNESCAP. (2016). Boosting women's workforce participation in South and South West Asia. MPFD Policy Briefs (No. 36). Retrieved from https://www.unescap.org/sites/default/files/MPFD\ Policy\ Brief-SSWA-Women\'sparticipation.pdf

University Grants Commission. (2018). "University Admissions". Retrieved from http://www.ugc.ac.lk/en/university-admissions.html

Uwakwe, M. (2004). Factors Affecting Women's Participation In The Labour Force In Nigeria. Journal of Agriculture and Social Research (JASR), 4 (2), 43-53.

World Bank. (2010). The Challenge of Youth Employment in Sri Lanka. Washington D.C: The World Bank.

World Bank. (2013). Gender at Work, A companion to the World Development Report on Jobs. Washington, DC: World Bank.

World Bank. (2017). Sri Lanka Development Update: Creating Opportunities and Managing Risks for Sustained Growth. Washington D.C: The World Bank.

World Bank. (2017). Getting to Work: Unlocking Women's Potential in Sri Lanka's Labor Force. Washington D.C: The World Bank.

## APPENDICES

## APPENDIX A: SPSS Output of Logistic Regression using the selected explanatory

 variablesTable A1: Case Processing Summary

| Unweighted Cases $^{\mathrm{a}}$ | N | Percent |  |
| :--- | :--- | ---: | ---: |
| Included in Analysis |  |  | 4002 |
| Selected Cases | Missing Cases | 0 | 100.0 |
|  | Total | 0 | .0 |
| Unselected Cases | 4002 | 100.0 |  |
| Total | 0 | .0 |  |

Table A2: Dependent Variable Encoding

| Original Value | Internal Value |
| :--- | ---: |
| Yes | 0 |
| No | 1 |

Table A3: Categorical Variables Codings

|  |  | Frequency | Parameter coding |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | (1) | (2) | (3) |
| X2 | Sinhala |  | 2814 | 1.000 | . 000 | . 000 |
|  | Tamil | 738 | . 000 | 1.000 | . 000 |
|  | Malay | 439 | . 000 | . 000 | 1.000 |
|  | Other | 11 | . 000 | . 000 | . 000 |
| X3 | Buddhist | 2649 | 1.000 | . 000 | . 000 |
|  | Hindu | 568 | . 000 | 1.000 | . 000 |
|  | Muslim | 434 | . 000 | . 000 | 1.000 |
|  | Other | 351 | . 000 | . 000 | . 000 |
| X7 | Urban | 801 | 1.000 | . 000 |  |
|  | Rural | 3110 | . 000 | 1.000 |  |
|  | Estate | 91 | . 000 | . 000 |  |
| X4 | Single | 3520 | 1.000 | . 000 |  |
|  | Married | 475 | . 000 | 1.000 |  |
|  | Other | 7 | . 000 | . 000 |  |
| X5 | Passed G.C.E. (O/L) | 2046 | 1.000 |  |  |
|  | Passed G.C.E. (A/L) | 1956 | . 000 |  |  |
| X6 | Ability to read and write | 1637 | 1.000 |  |  |
|  | Unable to read and write | 2365 | . 000 |  |  |
| X1 | Male | 1658 | 1.000 |  |  |
|  | Female | 2344 | . 000 |  |  |

Block 0: Beginning Block

Table A4: Classification Table ${ }^{\text {a,b }}$

| Observed |  |  | Predicted |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Y |  | Percentage Correct |
|  |  |  | Yes | No |  |
| Step 0 |  | Yes | 0 | 961 | . 0 |
|  |  | No | 0 | 3041 | 100.0 |
|  |  | rcent |  |  | 76.0 |

a. Constant is included in the model.
b. The cut value is .500

Table A5: Variables in the Equation

|  |  | B | S.E. | Wald | df | Sig. | $\operatorname{Exp}(B)$ |
| :--- | :--- | :--- | ---: | :--- | :--- | :--- | ---: |
| Step 0 | Constant | 1.152 | .037 | 969.043 |  | 1 | .000 |
| 3.164 |  |  |  |  |  |  |  |

Table A6: Variables not in the Equation

|  |  |  | Score | df | Sig. |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Step 0 |  | Gender(1) | 77.075 | 1 | . 000 |
|  |  | Race | 7.331 | 3 | . 062 |
|  |  | Race(1) | 7.264 | 1 | . 007 |
|  |  | Race(2) | 4.098 | 1 | . 043 |
|  |  | Race(3) | 1.828 | 1 | . 176 |
|  |  | Religion | 26.468 | 3 | . 000 |
|  |  | Religion(1) | 1.018 | 1 | . 313 |
|  |  | Religion(2) | 11.082 | 1 | . 001 |
|  | Variables | Religion(3) | 2.114 | 1 | . 146 |
|  |  | Marital_Status | 1.636 | 2 | . 441 |
|  |  | Marital_Status(1) | 1.108 | 1 | . 293 |
|  |  | Marital_Status(2) | 1.293 | 1 | . 255 |
|  |  | Education_Attainment(1) | 83.327 | 1 | . 000 |
|  |  | English_Literacy(1) | 4.412 | 1 | . 036 |
|  |  | Residential_Sector | 17.079 | 2 | . 000 |
|  |  | Residential_Sector(1) | 17.058 | 1 | . 000 |
|  |  | Residential_Sector(2) | 14.486 | 1 | . 000 |
|  | Overall Statistics |  | 229.415 | 13 | . 000 |

Block 1: Method = Forward Stepwise (Likelihood Ratio)

Table A7: Omnibus Tests of Model Coefficients

|  |  | Chi-square | df | Sig. |
| :--- | :--- | ---: | ---: | ---: |
|  | Step | 83.835 | 1 | .000 |
| Step 1 | Block | 83.835 |  | 1 |
|  | Model | 83.835 | .000 |  |
|  | Step | 98.058 | 1 | .000 |
| Step 2 | Block | 181.893 | 1 | .000 |
|  | Model | 181.893 | 2 | .000 |
|  | Step | 20.497 | 2 | .000 |
| Step 3 | Block | 202.391 | 3 | .000 |
|  | Model | 202.391 | 5 | .000 |
|  | Step | 14.156 | 5 | .000 |
| Step 4 | Block | 216.547 | 2 | .001 |
|  | Model | 216.547 | 7 | .000 |
|  | Step | 14.430 | 7 | .000 |
| Step 5 | Block | 230.977 | 2 | .001 |
|  | Model | 230.977 | 9 | .000 |

Table A8: Model Summary

| Step | -2 Log likelihood | Cox \& Snell R <br> Square | Nagelkerke R <br> Square |
| :--- | ---: | ---: | ---: |
| 1 | $4328.207^{\mathrm{a}}$ | .021 | .031 |
| 2 | $4230.149^{\mathrm{a}}$ | .044 | .067 |
| 3 | $4209.651^{\mathrm{a}}$ | .049 | .074 |
| 4 | $4195.495^{\mathrm{a}}$ | .053 | .079 |
| 5 | $4181.065^{\mathrm{b}}$ | .056 | .084 |

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.
b. Estimation terminated at iteration number 5 because parameter estimates changed by less than . 001 .

Table A9: Hosmer and Lemeshow Test

| Step | Chi-square | df | Sig. |
| :--- | ---: | ---: | ---: |
| 1 | .000 | 0 | . |
| 2 | 23.802 | 2 | .000 |
| 3 | 22.069 | 6 | .001 |
| 4 | 24.573 | 7 | .001 |
| 5 | 27.309 | 8 | .001 |

Table A10: Contingency Table for Hosmer and Lemeshow Test

|  |  | $\mathrm{Y}=\mathrm{Yes}$ |  | $\mathrm{Y}=\mathrm{No}$ |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Observed | Expected | Observed | Expected |  |
| Step 1 | 1 | 593 | 593.000 | 1363 | 1363.000 | 1956 |
|  | 2 | 368 | 368.000 | 1678 | 1678.000 | 2046 |
|  | 1 | 256 | 287.395 | 449 | 417.605 | 705 |
| Step 2 | 2 | 337 | 305.605 | 914 | 945.395 | 1251 |
|  | 3 | 259 | 227.605 | 694 | 725.395 | 953 |
| Step 3 | 4 | 109 | 140.395 | 984 | 952.605 | 1093 |
|  | 1 | 203 | 234.938 | 355 | 323.062 | 558 |
|  | 2 | 132 | 115.927 | 210 | 226.073 | 342 |
|  | 3 | 238 | 208.386 | 612 | 641.614 | 850 |
|  | 4 | 160 | 151.360 | 465 | 473.640 | 625 |
|  | 5 | 85 | 85.679 | 320 | 319.321 | 405 |
| Step 4 | 6 | 41 | 40.672 | 175 | 175.328 | 216 |
|  | 7 | 77 | 89.491 | 608 | 595.509 | 685 |
|  | 8 | 25 | 34.548 | 296 | 286.452 | 321 |
|  | 1 | 218 | 252.688 | 385 | 350.312 | 603 |
|  | 2 | 152 | 135.014 | 266 | 282.986 | 418 |
|  | 3 | 82 | 76.625 | 202 | 207.375 | 284 |
|  | 4 | 188 | 164.018 | 515 | 538.982 | 703 |
|  | 5 | 136 | 124.158 | 397 | 408.842 | 533 |
| Step 5 | 6 | 79 | 76.139 | 321 | 323.861 | 400 |
|  | 7 | 25 | 34.087 | 192 | 182.913 | 217 |
|  | 8 | 63 | 74.065 | 532 | 520.935 | 595 |
|  | 9 | 18 | 24.207 | 231 | 224.793 | 249 |
|  | 1 | 101 | 104.739 | 120 | 116.261 | 221 |
|  | 2 | 130 | 155.201 | 267 | 241.799 | 397 |
|  | 3 | 136 | 120.332 | 227 | 242.668 | 363 |
|  | 4 | 126 | 112.644 | 269 | 282.356 | 395 |
|  | 5 | 127 | 131.608 | 448 | 443.392 | 575 |
|  | 6 | 172 | 143.769 | 481 | 509.231 | 653 |
|  | 7 | 69 | 73.399 | 338 | 333.601 | 407 |
|  | 8 | 31 | 47.553 | 286 | 269.447 | 317 |
|  | 9 | 54 | 53.479 | 414 | 414.521 | 468 |
|  | 10 | 15 | 18.277 | 191 | 187.723 | 206 |

Table A11: Classification Table ${ }^{\text {a }}$

| Observed |  | Predicted |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Y |  | PercentageCorrect |
|  |  | Yes | No |  |
| Step 1 | Y Yes | 0 | 961 | . 0 |
|  | No | 0 | 3041 | 100.0 |
|  | Overall Percentage |  |  | 76.0 |
| Step 2 | y Yes | 0 | 961 | . 0 |
|  | No | 0 | 3041 | 100.0 |
|  | Overall Percentage |  |  | 76.0 |
| Step 3 | y Yes | 29 | 932 | 3.0 |
|  | No | 40 | 3001 | 98.7 |
|  | Overall Percentage |  |  | 75.7 |
| Step 4 | Y Yes | 9 | 952 | . 9 |
|  | No | 9 | 3032 | 99.7 |
|  | Overall Percentage |  |  | 76.0 |
| Step 5 | Y Yes | 21 | 940 | 2.2 |
|  | No | 14 | 3027 | 99.5 |
|  | Overall Percentage |  |  | 76.2 |

a. The cut value is .500

Table A12: Model if Term Removed

| Variable |  | Model Log <br> Likelihood | Change in -2 <br> Log Likelihood | dig. of the <br> Change |  |
| :--- | :--- | ---: | ---: | ---: | ---: |
| Step 1 | Education_Attainment | -2206.021 | 83.835 | 1 | .000 |
| Step 2 | Gender | -2164.103 | 98.058 | 1 | .000 |
|  | Education_Attainment | -2167.918 | 105.686 | 1 | .000 |
|  | Gender | -2153.403 | 97.154 | 1 | .000 |
| Step 3 | Religion | -2115.074 | 20.497 | 3 | .000 |
|  | Education_Attainment | -2155.150 | 100.649 | 1 | .000 |
|  | Gender | -2146.660 | 97.825 | 1 | .000 |
|  | Religion | -2107.937 | 20.379 | 3 | .000 |
|  | Education_Attainment | -2146.525 | 97.556 | 1 | .000 |
|  | Residential_Sector | -2104.826 | 14.156 | 2 | .001 |
|  | Gender | -2145.122 | 109.178 | 1 | .000 |
|  | Religion | -2100.867 | 20.669 | 3 | .000 |
| Step 5 | Marital_Status | -2097.748 | 14.430 | 2 | .001 |
|  | Education_Attainment | -2141.446 | 101.827 | 1 | .000 |
|  | Residential_Sector | -2098.051 | 15.036 | 2 | .001 |

Table A13: Case Processing Summary

| Unweighted Cases $^{\text {a }}$ | N | Percent |  |
| :--- | :--- | ---: | ---: |
|  | Included in Analysis | 4002 | 100.0 |
| Selected Cases | Missing Cases | 0 | .0 |
|  | Total | 4002 | 100.0 |
| Unselected Cases |  | 0 | .0 |
| Total | 4002 | 100.0 |  |

Table A14: Dependent Variable Encoding

| Original Value | Internal Value |
| :--- | ---: |
| Yes | 0 |
| No | 1 |

Table A15: Categorical Variables Codings

|  |  | Frequency | Parameter coding |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | (1) | (2) | (3) |
| X2 | Sinhala |  | 2814 | 1.000 | . 000 | . 000 |
|  | Tamil | 738 | . 000 | 1.000 | . 000 |
|  | Malay | 439 | . 000 | . 000 | 1.000 |
|  | Other | 11 | . 000 | . 000 | . 000 |
|  | Buddhist | 2649 | 1.000 | . 000 | . 000 |
| X3 | Hindu | 568 | . 000 | 1.000 | . 000 |
|  | Muslim | 434 | . 000 | . 000 | 1.000 |
|  | Other | 351 | . 000 | . 000 | . 000 |
| X7 | Urban | 801 | 1.000 | . 000 |  |
|  | Rural | 3110 | . 000 | 1.000 |  |
|  | Estate | 91 | . 000 | . 000 |  |
| X4 | Single | 3520 | 1.000 | . 000 |  |
|  | Married | 475 | . 000 | 1.000 |  |
|  | Other | 7 | . 000 | . 000 |  |
| X5 | Passed G.C.E. (O/L) | 2046 | 1.000 |  |  |
|  | Passed G.C.E. (A/L) | 1956 | . 000 |  |  |
| X6 | Ability to read and write | 1637 | 1.000 |  |  |
|  | Unable to read and write | 2365 | . 000 |  |  |
| X1 | Male | 1658 | 1.000 |  |  |
|  | Female | 2344 | . 000 |  |  |

Table A16: Case Processing Summary

|  | Cases |  |  |  |  |  |
| :--- | :---: | ---: | ---: | ---: | ---: | ---: |
|  | Valid |  | Missing |  | Total |  |
|  | N | Percent | N | Percent | N | Percent |
|  | 4002 | $100.0 \%$ |  | 0 | $0.0 \%$ | 4002 |

Table A17: Labour_Force * Predicted group Crosstabulation


Table A18: Chi-Square Tests

|  | Value | df | Asymp. Sig. (2- <br> sided) | Exact Sig. (2- <br> sided) | Exact Sig. (1- <br> sided) |
| :--- | ---: | ---: | ---: | ---: | ---: |
| Pearson Chi-Square | $25.060^{\mathrm{a}}$ | 1 | .000 |  |  |
| Continuity Correction $^{\mathrm{b}}$ | 23.110 | 1 | .000 |  |  |
| Likelihood Ratio | 20.714 |  | 1 | .000 |  |
| Fisher's Exact Test |  |  |  |  |  |
| Linear-by-Linear Association | 25.054 |  | 1 | .000 |  |
| N of Valid Cases | 4002 |  |  |  |  |

a. 0 cells (.0\%) have expected count less than 5 . The minimum expected count is 8.40.
b. Computed only for a $2 \times 2$ table

Appendix B: SPSS Output of Logistic Regression of the selected explanatory variables with the use of interaction terms

## Block 0: Beginning Block

Table B1: Classification Table ${ }^{\text {a,b }}$

| Observed |  | Predicted |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Y |  | Percentage <br> Correct |
|  |  | Yes | No |  |
| Step 0 | Yes | 0 | 961 | . 0 |
|  | No | 0 | 3041 | 100.0 |
|  | Overall Percentage |  |  | 76.0 |

a. Constant is included in the model.
b. The cut value is .500

Table B2: Variables in the Equation

|  |  | B | S.E. | Wald | df | Sig. |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Step 0 | Constant | 1.152 | .037 | 969.043 |  | 1 |

Table B3: Variables not in the Equation

|  |  |  | Score | df | Sig. |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Gender(1) | 77.075 | 1 | . 000 |
|  |  | Race | 7.331 | 3 | . 062 |
|  |  | Race(1) | 7.264 | 1 | . 007 |
|  |  | Race(2) | 4.098 | 1 | . 043 |
|  |  | Race(3) | 1.828 | 1 | . 176 |
|  |  | Religion | 26.468 | 3 | . 000 |
|  |  | Religion(1) | 1.018 | 1 | . 313 |
|  |  | Religion(2) | 11.082 | 1 | . 001 |
| Step 0 | Variables | Religion(3) | 2.114 | 1 | . 146 |
|  |  | Marital_Status | 1.636 | 2 | . 441 |
|  |  | Marital_Status(1) | 1.108 | 1 | . 293 |
|  |  | Marital_Status(2) | 1.293 | 1 | . 255 |
|  |  | Education_Attainment(1) | 83.327 | 1 | . 000 |
|  |  | English_Literacy(1) | 4.412 | 1 | . 036 |
|  |  | Residential_Sector | 17.079 | 2 | . 000 |
|  |  | Residential_Sector(1) | 17.058 | 1 | . 000 |
|  |  | Residential_Sector(2) | 14.486 | 1 | . 000 |


| Gender * Race | 79.761 | 3 | .000 |
| :--- | ---: | ---: | ---: |
| Gender(1) by Race(1) | 35.008 | 1 | .000 |
| Gender(1) by Race(2) | 13.398 | 1 | .000 |
| Gender(1) by Race(3) | 11.053 | 1 | .001 |
| Gender *Religion | 58.004 | 3 | .000 |
| Gender(1) by Religion(1) | 30.143 | 1 | .000 |
| Gender(1) by Religion(2) | 4.762 | 1 | .029 |
| Gender(1) by Religion(3) | 10.897 | 1 | .001 |
| Gender *Marital_Status | 184.131 | 2 | .000 |
| Gender(1) by | 32.543 | 1 | .000 |
| Marital_Status(1) | 134.853 | 1 | .000 |
| Gender(1) by |  | 16.864 | 1 |

Block 1: Method = Forward Stepwise (Likelihood Ratio)

Table B4: Omnibus Tests of Model Coefficients

|  |  | Chi-square | df | Sig. |
| :--- | :--- | ---: | ---: | ---: |
|  | Step | 160.107 | 2 | .000 |
| Step 1 | Block | 160.107 | 2 | .000 |
|  | Model | 160.107 | 2 | .000 |
|  | Step | 114.788 | 1 | .000 |
| Step 2 | Block | 274.895 | 3 | .000 |
|  | Model | 274.895 | 3 | .000 |
|  | Step | 19.306 | 1 | .000 |
| Step 3 | Block | 294.200 | 4 | .000 |
|  | Model | 294.200 | 4 | .000 |
|  | Step | 20.243 | 3 | .000 |
| Step 4 | Block | 314.443 | 7 | .000 |
|  | Model | 314.443 | 7 | .000 |


|  | Step | 16.607 | 3 | .001 |
| :--- | :--- | ---: | ---: | ---: |
| Step 5 | Block | 331.050 | 10 | .000 |
|  | Model | 331.050 | 10 | .000 |
|  | Step | 13.507 | 2 | .001 |
| Step 6 | Block | 344.558 | 12 | .000 |
|  | Model | 344.558 | 12 | .000 |
|  | Step | 3.952 | 1 | .047 |
| Step 7 | Block | 348.510 | 13 | .000 |
|  | Model | 348.510 | 13 | .000 |

Table B5: Model Summary

| Step | -2 Log likelihood | Cox \& Snell R <br> Square | Nagelkerke R <br> Square |
| :--- | ---: | ---: | ---: |
| 1 | $4251.936^{\mathrm{a}}$ | .039 | .059 |
| 2 | $4137.147^{\mathrm{a}}$ | .066 | .099 |
| 3 | $4117.842^{\mathrm{b}}$ | .071 | .106 |
| 4 | $4097.599^{\mathrm{b}}$ | .076 | .113 |
| 5 | $4080.992^{\mathrm{b}}$ | .079 | .119 |
| 6 | $4067.485^{\mathrm{b}}$ | .082 | .124 |
| 7 | $4063.532^{\mathrm{b}}$ | .083 | .125 |

Table B6: Hosmer and Lemeshow Test

| Step | Chi-square | df | Sig. |
| :---: | :---: | :---: | :---: |
| 1 | .000 | 1 | 1.000 |
| 2 | 17.645 | 3 | .001 |
| 3 | .039 | 3 | .998 |
| 4 | 6.793 | 7 | .451 |
| 5 | 6.166 | 7 | .520 |
| 6 | 6.775 | 7 | .453 |
| 7 | 6.704 | 7 | .460 |

Table B7: Classification Table ${ }^{\text {a }}$

| Observed |  | Predicted |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Y |  | Percentage Correct |
|  |  | Yes | No |  |
| Step 1 | Yes | 59 | 902 | 6.1 |
|  | No | 13 | 3028 | 99.6 |
|  | Overall Percentage |  |  | 77.1 |
| Step 2 | Yes | 59 | 902 | 6.1 |
|  | No | 13 | 3028 | 99.6 |
|  | Overall Percentage |  |  | 77.1 |


a. The cut value is .500

Table B8: Contingency Table for Hosmer and Lemeshow Test


| Step 5 | 1 | 188 | 181.815 | 208 | 214.185 | 396 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 2 | 163 | 159.852 | 314 | 317.148 | 477 |
|  | 3 | 20 | 20.472 | 51 | 50.528 | 71 |
|  | 4 | 238 | 240.342 | 612 | 609.658 | 850 |
|  | 5 | 67 | 61.640 | 174 | 179.360 | 241 |
|  | 6 | 132 | 138.699 | 461 | 454.301 | 593 |
|  | 7 | 51 | 63.702 | 317 | 304.298 | 368 |
|  | 8 | 77 | 74.658 | 608 | 610.342 | 685 |
|  | 9 | 25 | 19.820 | 296 | 301.180 | 321 |
|  | 1 | 194 | 185.402 | 196 | 204.598 | 390 |
|  | 2 | 202 | 197.137 | 396 | 400.863 | 598 |
|  | 3 | 52 | 63.723 | 158 | 146.277 | 210 |
|  | 4 | 188 | 189.661 | 515 | 513.339 | 703 |
| Step 6 | 5 | 149 | 146.763 | 488 | 490.237 | 637 |
|  | 6 | 70 | 75.659 | 333 | 327.341 | 403 |
|  | 7 | 18 | 21.003 | 127 | 123.997 | 145 |
|  | 8 | 63 | 62.134 | 532 | 532.866 | 595 |
|  | 9 | 25 | 19.518 | 296 | 301.482 | 321 |
|  | 1 | 197 | 188.490 | 202 | 210.510 | 399 |
|  | 2 | 140 | 140.828 | 264 | 263.172 | 404 |
|  | 3 | 107 | 113.970 | 274 | 267.030 | 381 |
|  | 4 | 188 | 188.852 | 515 | 514.148 | 703 |
| Step 7 | 5 | 121 | 119.638 | 376 | 377.362 | 497 |
|  | 6 | 85 | 81.195 | 307 | 310.805 | 392 |
|  | 7 | 35 | 46.659 | 275 | 263.341 | 310 |
|  | 8 | 63 | 61.874 | 532 | 533.126 | 595 |
|  | 9 | 25 | 19.492 | 296 | 301.508 | 321 |

Table B9: Variables in the Equation

|  |  | B | S.E. | Wald | df | Sig. | Exp(B) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Step $1^{\text {a }}$ | Gender * Marital_Status |  |  | 127.003 | 2 | . 000 |  |
|  | Gender(1) by | -. 542 | . 076 | 50.269 | 1 | . 000 | . 581 |
|  | Marital_Status(1) |  |  |  |  |  |  |
|  | Gender(1) by | -2.961 | . 311 | 90.744 | 1 | . 000 | . 052 |
|  | Marital_Status(2) |  |  |  |  |  |  |
|  | Constant | 1.449 | . 053 | 758.075 | 1 | . 000 | 4.258 |
|  | Education_Attainment(1) | . 833 | . 079 | 110.388 | 1 | . 000 | 2.300 |
| Step $2^{\text {b }}$ | Gender * Marital_Status |  |  | 153.170 | 2 | . 000 |  |
|  | Gender(1) by | -. 647 | . 079 | 67.960 | 1 | . 000 | . 523 |
|  | Marital_Status(1) |  |  |  |  |  |  |



|  | Religion(1) | . 418 | . 171 | 5.974 | 1 | . 015 | 1.520 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Religion(2) | 1.111 | . 234 | 22.519 | 1 | . 000 | 3.038 |
|  | Religion(3) | 1.111 | . 253 | 19.257 | 1 | . 000 | 3.039 |
|  | Education_Attainment(1) | 1.153 | . 120 | 91.790 | 1 | . 000 | 3.169 |
|  | Residential_Sector |  |  | 13.786 | 2 | . 001 |  |
|  | Residential_Sector(1) | -. 150 | . 288 | . 272 | 1 | . 602 | . 860 |
|  | Residential_Sector(2) | . 201 | . 281 | . 512 | 1 | . 474 | 1.223 |
|  | Gender * Religion |  |  | 16.754 | 3 | . 001 |  |
|  | Gender(1) by Religion(1) | -. 199 | . 257 | . 603 | 1 | . 437 | . 819 |
|  | Gender(1) by Religion(2) | -. 872 | . 332 | 6.871 | 1 | . 009 | . 418 |
|  | Gender(1) by Religion(3) | -1.033 | . 352 | 8.624 | 1 | . 003 | . 356 |
|  | Gender * Marital_Status |  |  | 62.886 | 2 | . 000 |  |
|  | Gender(1) by | -. 052 | . 248 | . 043 | 1 | . 835 | . 950 |
|  | Marital_Status(1) |  |  |  |  |  |  |
|  | Gender(1) by | -2.540 | . 396 | 41.080 | 1 | . 000 | . 079 |
|  | Marital_Status(2) |  |  |  |  |  |  |
|  | Education_Attainment(1) by | -. 655 | . 164 | 15.991 | 1 | . 000 | . 519 |
|  | Gender(1) |  |  |  |  |  |  |
|  | Constant | . 376 | . 310 | 1.474 | 1 | . 225 | 1.456 |
|  | Religion |  |  | 33.712 | 3 | . 000 |  |
|  | Religion(1) | . 412 | . 171 | 5.777 | 1 | . 016 | 1.509 |
|  | Religion(2) | 1.109 | . 234 | 22.426 | 1 | . 000 | 3.032 |
|  | Religion(3) | 1.115 | . 253 | 19.349 | 1 | . 000 | 3.048 |
|  | Education_Attainment(1) | 1.152 | . 120 | 91.562 | 1 | . 000 | 3.165 |
|  | Residential_Sector |  |  | 15.831 | 2 | . 000 |  |
|  | Residential_Sector(1) | -. 158 | . 288 | . 300 | 1 | . 584 | . 854 |
|  | Residential_Sector(2) | . 224 | . 281 | . 633 | 1 | . 426 | 1.251 |
|  | Gender * Religion |  |  | 16.434 | 3 | . 001 |  |
|  | Gender(1) by Religion(1) | -. 206 | . 257 | . 642 | 1 | . 423 | . 814 |
| Step $7^{\text {g }}$ | Gender(1) by Religion(2) | -. 853 | . 333 | 6.566 | 1 | . 010 | . 426 |
|  | Gender(1) by Religion(3) | -1.046 | . 352 | 8.814 | 1 | . 003 | . 351 |
|  | Gender * Marital_Status |  |  | 62.352 | 2 | . 000 |  |
|  | Gender(1) by | -. 181 | . 257 | . 500 | 1 | . 480 | . 834 |
|  | Marital_Status(1) |  |  |  |  |  |  |
|  | Gender(1) by | -2.650 | . 401 | 43.641 | 1 | . 000 | . 071 |
|  | Marital_Status(2) |  |  |  |  |  |  |
|  | Education_Attainment(1) by | -. 585 | . 168 | 12.130 | 1 | . 000 | . 557 |
|  | Gender(1) |  |  |  |  |  |  |
|  | English_Literacy(1) by | . 239 | . 121 | 3.919 | 1 | . 048 | 1.270 |
|  | Gender(1) |  |  |  |  |  |  |
|  | Constant | . 366 | . 310 | 1.396 | 1 | . 237 | 1.442 |

Table B10: Model if Term Removed

| Variable |  | Model Log <br> Likelihood | Change in -2 <br> Log Likelihood | df | Sig. of the Change |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Step 1 <br> Step 2 | Gender * Marital_Status | -2206.021 | 160.107 | 2 | . 000 |
|  | Education_Attainment | -2125.968 | 114.788 | 1 | . 000 |
|  | Gender * Marital_Status | -2164.103 | 191.060 | 2 | . 000 |
|  | Education_Attainment | -2115.906 | 113.971 | 1 | . 000 |
| Step 3 | Gender * Marital_Status | -2112.232 | 106.623 | 2 | . 000 |
|  | Education_Attainment * | -2068.574 | 19.306 | 1 | . 000 |
|  | Gender |  |  |  |  |
|  | Religion | -2058.921 | 20.243 | 3 | . 000 |
|  | Education_Attainment | -2103.625 | 109.651 | 1 | . 000 |
| Step 4 | Gender * Marital_Status | -2102.071 | 106.544 | 2 | . 000 |
|  | Education_Attainment * | -2058.297 | 18.995 | 1 | . 000 |
|  | Gender |  |  |  |  |
|  | Religion | -2057.281 | 33.569 | 3 | . 000 |
|  | Education_Attainment | -2093.535 | 106.078 | 1 | . 000 |
| Step 5 | Gender * Religion | -2048.799 | 16.607 | 3 | . 001 |
|  | Gender * Marital_Status | -2084.029 | 87.066 | 2 | . 000 |
|  | Education_Attainment * | -2049.017 | 17.042 | 1 | . 000 |
|  | Gender |  |  |  |  |
|  | Religion | -2051.712 | 35.939 | 3 | . 000 |
|  | Education_Attainment | -2084.934 | 102.383 | 1 | . 000 |
|  | Residential_Sector | -2040.496 | 13.507 | 2 | . 001 |
| Step 6 | Gender * Religion | -2042.317 | 17.150 | 3 | . 001 |
|  | Gender * Marital_Status | -2077.201 | 86.917 | 2 | . 000 |
|  | Education_Attainment * | -2041.880 | 16.276 | 1 | . 000 |
|  | Gender |  |  |  |  |
|  | Religion | -2049.880 | 36.229 | 3 | . 000 |
|  | Education_Attainment | -2082.819 | 102.106 | 1 | . 000 |
|  | Residential_Sector | -2039.520 | 15.507 | 2 | . 000 |
|  | Gender * Religion | -2040.179 | 16.826 | 3 | . 001 |
| Step 7 | Gender * Marital_Status | -2074.759 | 85.987 | 2 | . 000 |
|  | Education_Attainment * | -2037.911 | 12.290 | 1 | . 000 |
|  | Gender |  |  |  |  |
|  | English_Literacy * Gender | -2033.742 | 3.952 | 1 | . 047 |

Table B11: Variables not in the Equation





|  |  | English_Literacy(1) by Gender(1) | 3.926 | 1 | . 048 |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Gender * Residential_Sector | 5.395 | 2 | . 067 |
|  |  | Gender(1) by | 4.308 | 1 | . 038 |
|  |  | Residential_Sector(1) |  |  |  |
|  |  | Gender(1) by | 2.565 | 1 | . 109 |
|  |  | Residential_Sector(2) |  |  |  |
|  | Overall St | tics | 14.142 | 13 | . 364 |
|  |  | Gender(1) | . 558 | 1 | . 455 |
|  |  | Race | 1.329 | 3 | . 722 |
|  |  | Race(1) | . 521 | 1 | . 470 |
|  |  | Race(2) | . 143 | 1 | . 705 |
|  |  | Race(3) | . 000 | 1 | . 991 |
|  |  | Marital_Status | . 711 | 2 | . 701 |
|  |  | Marital_Status(1) | . 622 | 1 | . 430 |
|  |  | Marital_Status(2) | . 543 | 1 | . 461 |
|  | Variable | English_Literacy(1) | . 001 | 1 | . 981 |
| Step 7 |  | Gender * Race | 1.432 | 3 | . 698 |
|  |  | Gender(1) by Race(1) | . 571 | 1 | . 450 |
|  |  | Gender(1) by Race(2) | 1.268 | 1 | . 260 |
|  |  | Gender(1) by Race(3) | . 234 | 1 | . 628 |
|  |  | Gender * Residential_Sector | 4.419 | 2 | . 110 |
|  |  | Gender(1) by | 3.181 | 1 | . 074 |
|  |  | Residential_Sector(1) |  |  |  |
|  |  | Gender(1) by | 1.678 | 1 | . 195 |
|  |  | Residential_Sector(2) |  |  |  |
|  | Overall St | tics | 10.229 | 12 | . 596 |

## Crosstabs

Table B12: Case Processing Summary

|  | Cases |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Valid |  | Missing |  | Total |  |
|  | N | Percent | N | Percent | N | Percent |
|  | 4002 | $100.0 \%$ |  | 0 | $0.0 \%$ | 4002 |

Table B13: Labour_Force * Predicted group Crosstabulation

|  |  |  | Predicted group |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | Yes | No |  |
| Labour_Force |  | Count | 59 | 902 | 961 |
|  |  | \% within Labour_Force | 6.1\% | 93.9\% | 100.0\% |
|  |  | Count | 13 | 3028 | 3041 |
|  |  | \% within Labour_Force | 0.4\% | 99.6\% | 100.0\% |
| Total |  | Count | 72 | 3930 | 4002 |
|  |  | \% within Labour_Force | 1.8\% | 98.2\% | 100.0\% |

Table B14: Chi-Square Tests

|  | Value | df | Asymp. Sig. (2sided) | Exact Sig. (2sided) | Exact Sig. (1sided) |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Pearson Chi-Square | $134.853^{\text {a }}$ | 1 | . 000 |  |  |
| Continuity Correction ${ }^{\text {b }}$ | 131.639 | 1 | . 000 |  |  |
| Likelihood Ratio | 109.925 | 1 | . 000 |  |  |
| Fisher's Exact Test |  |  |  | . 000 | . 000 |
| Linear-by-Linear Association | 134.819 | 1 | . 000 |  |  |
| $N$ of Valid Cases | 4002 |  |  |  |  |

a. 0 cells $(0.0 \%)$ have expected count less than 5 . The minimum expected count is 17.29.
b. Computed only for a $2 \times 2$ table


[^0]:    Source: Annual Labour Force Survey (2016)

