Title: Determinants of Employment Status in the UK Labour Market: An Analysis Using the 2015 Quarterly Labour Force Survey

Table Of Contents

1.0 Introduction

For UK country economics the UK labor market is a very important component because it represents key factors like economies, employment, economies based social factors. There is a sudden deflection in the UK labour market due to changes like technological, policy making wise. To do a comprehensive research and study one can get the QLFS january- March 2015 data from Uk data service website. The dataset is important for policymakers, researchers and for academic purposes because it provides the socio-economics factor of the UK labour market trend.

1.1 Research Aim

The aim of this study is to investigate the factors influencing employment status in the UK labour market, with a focus on educational qualifications, socio-economic classification, and regional disparities.

1.2 Research Objectives

1. To examine the factors affecting employment status in the UK labour market.

2. To analyze the impact of educational qualifications on employment status.

3. To investigate the relationship between socio-economic classification (NS-SEC) and employment status.

4. To assess the effect of regional differences on employment status.

1.3 Research Questions

1. What factors influence employment status in the UK labour market?

2. How do educational qualifications affect employment status?

3. What role does socio-economic classification play in shaping employment status?

4. How do regional disparities impact employment status?

2.0 Literature Review

Uk has changed rapidly, especially for work because of technology and wages (Acemoglu, D 2011). There are numerous policies introduced in the last decade for the UK labour market. In the Uk for low -skilled workers there is a low wage and especially for working males.

During the great-recession of 2008-09 there was a great employment among the youth and for getting the employment the youth there were skilled programmed in youth so that they could have high probability(Bell, D.N et.al, 2011) in getting Job. During 2013 it is seen that minimisation in incomes arises the more labour supply and workers can (Blundell, R et.al, 2014) work at any type of job low skilled also this arises a strong competition for the Job in Uk.

Regional disparity is one of the challenges in Uk because one region is detached from the rest of the UK in scale of prosperity and growth of economics. There is a spatial-imbalance in the economics of Uk from the front period of 1988-2013. Beginning 1980, there was a lack of demand in the Northern part of (Martin, R., Pike et.al, 2015) Uk because of old industries facing direct competitors. To make the growth of the depressed northern britain region provide growth of economy and innovations and infrastructure.
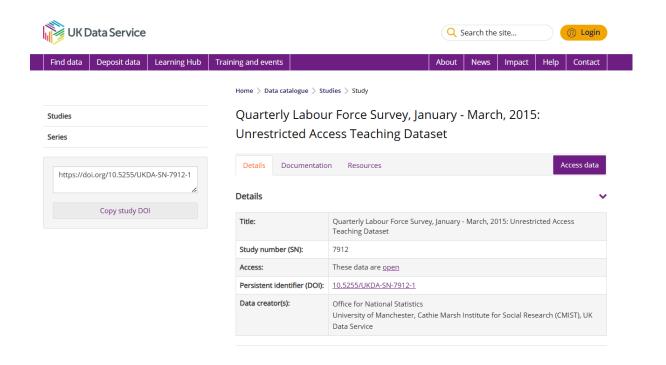
3.0 Methodology and Design

3.1 Data Collection and Sampling

The Quality labour force survey dataset is obtained from the UK Data Archive with appropriate permissions which is from the time period of January - March, 2015. It is a quarterly dataset and contains 22,428 cases with 13 variables.

The "open" dataset ensures timely access and representativeness of the UK labor market.

Dataset: [Uk data service](#)

3.1.1 Introduction to the Dataset:

Dataset website	[Uk data service](#)
Dataset category	Open dataset
Dataset name	Quarterly Labour Force Survey, January - March, 2015: Unrestricted Access Teaching Dataset
Study number (SN)	7912
Time period	January 2015 - March 2015
Country	**United Kingdom**
Number of Records	22,428 cases
Number of variables	13
Dataset Link	QLFS Dataset

Figure 1: QLFS dataset from uk data service site

3.1.2 Variables Mapping:

Based on the Data Dictionary, the variables in the dataset are as follows:

Variable name	Variable label	Variable type
CASENEW	New random ID number	Scalar
PWT14	Person weight	Scalar
SEX	Sex of respondent	Nominal
AGEEULR	Age bands in 5-year intervals	Ordinal
MARSTA3R	Marital status	Nominal
HIQUL15D	Highest qualification	Nominal
ETHUK7R	Ethnicity	Nominal
ILODEFR	Economic activity	Nominal
STAT3R	Employment status	Nominal
FTPTWK	Full-time or part-time in main job	Nominal
TOTHRS	Total hours worked in reference week	Scalar
NSECMJ3R	NS-SEC 3 class (main job)	Ordinal
GOVTOF2	Government Office Region	Nominal

3.1.2.1 Independent Variables Table:

	Type of Variable	Variable Name	Variable Definition	Data Values	Data Type
1.	**Outcome Variable**	Employment Status (STAT3R)	Employment category of the respondent	1 = Employee, 2 = Self-employed, 3 = Government scheme/unpaid worker	Nominal
2.	**Key Explanatory Variables**	Highest Qualification (HIQUL15D	Highest level of education completed	1 = Degree, 2 = Higher education, 3	Nominal

				= A-Level, 4 = GCSE, 5 = Other, 6 = No qualification	
3.		Socio-economic Classification (NSECMJ3R)	Socio-economic class based on occupation	1 = Higher managerial, 2 = Intermediate, 3 = Routine/manual, 4 = Never worked/unemployed	Ordinal
4.		Government Office Region (GOVTOF2)	Region of the UK where the respondent resides	1 = London, 2 = Scotland, 3 = North West, etc. (12 regions)	Nominal
5.	**Control Variables:** Demographic Controls	Gender (SEX)	Respondent's gender	1 = Male, 2 = Female	Nominal
6.		Age (AGEEULR)	Age group of the respondent in 5-year intervals	1 = 15-19, 2 = 20-24, ..., 12 = 70+	Ordinal
7.		Marital Status (MARSTA3R)	Current marital status	1 = Single, 2 = Married/Cohabiting, 3 = Divorced/Widowed	Nominal
8.		Ethnicity (ETHUK7R)	Respondent's ethnic group	1 = White, 2 = Indian, 3 = Black, etc. (7 categories)	Nominal
9.	**Control Variables:** Labour Market	Economic Activity (ILODEFR)	Employment status as per International Labour	1 = Employed, 2 = Unemployed	Nominal

		Controls		Organization (ILO) classification	, 3 = Inactive	
10.			Full-time or Part-time (FTPTWK)	Whether the respondent works full-time or part-time	1 = Full-time, 2 = Part-time	Nominal
11.			Total Hours Worked (TOTHRS)	Total hours worked in the reference week	Continuous range: 0 - 97	Scalar
12.		**Technical Variables**	Person Weight (PWT14)	Statistical weight assigned to respondent	Adjusts for survey sample representation	Scalar
13.			Person Weight Rescaled (PWT14R)	Rescaled statistical weight	Adjusts for survey sample representation	Scalar

3.1.2.2 Dependent Variables Table:

Dependent Variable Name	*Variable label*	*Variable type*	*Variable Mapping*
stat3r	Employment status	nominal	Employee, Self-employed, Government scheme or unpaid family worker

- <mark>employment status (STAT3R) as the dependent variable and it will focus on the employment status of the employee which type of employment is done by the employee.</mark>

3.2 Data Analysis and Methods

In this study the research works with data would be quantitative analysis. There is a need to check if there are any missing values. If a small amount of missing values then remove the missing values or either replace missing values with **mean** for continuous variables and **mode** for the missing values for categorical data. (Alwateer, M., Atlam et.al, 2024) However there are many packages in R so one can use **Mice** for imputing missing values would be a great choice. In this study (Prasada, Abikesh et.al, 2020) it is important to detect outliers using z-scores and to handle outliers need to either remove or keep depending on data outliers proportion. There is a need to check asymmetry in the dataset for this to check skewness and if data is negatively skewed if it has skewness level < - 0.05. (Maillie, David 2019) To handle skewness you have to apply log transformation using R-packages. To get the most statistical values and frequency analysis from data then have to use descriptive statistics from this for each numerical variable will get mean, median, mode, minimum , (Ghashim, E et.al, 2017) maximum, standard deviation, skewness and other statistical values. To check the linear relationship between two quantitative variables one can use techniques like Pearson, Spearman. In this study will perform a chi-square test of independence and measure of associations (Alberti, Gianmarco 2022) such as odds-ratio, p-value, Cramer's V, Cohen's w, Cohen's k.

Logistic regression and Multiple Regression to analyze the impact of independent variables (e.g., education, socio-economic classification, regional disparities) on employment status from these significance levels and odds-ratio will be used to (Panda NR, Kumar PJ et.al, 2022) get the idea of important predictors for dependent variables.

In this study anova test will be conducted to get the difference between the two groups with these methods will be employed like(Baker, D.Het.al, 2022) F-test, one-way anova, Tukey HSD test. To get the insights from the data and understand the data visually one has to build some visualisations for numerically building boxplot and histogram and for categorically dataset design (Pant, Anjali et.al, 2019) bar and pie chart and can design some other plots according to dataset nature.For implementing all these data analysis techniques, visualisations and machine learning will use **R programming language**, packages using **R-studio** software.

Regional Disparity Analysis: Region dummies (GOVTOF2) will be used as moderators in regression models to examine how regional differences influence employment status (STAT3R). Additionally, **subsample regressions** will be conducted for each region to compare the impact of factors such as educational qualifications (HIQUL15D) and socio-economic classification (NSECMJ3R) on employment status. This approach will provide insights into whether the effects of education and socio-economic classification vary across different regions of the UK.

4.0 Justification for Using 2015 Data

The study uses the QLFS January-March 2015 dataset because it provides a comprehensive snapshot of the UK labour market during a specific period of interest, particularly in the context of post-recovery from the 2008 financial crisis. This dataset is well-documented and widely used in academic research, making it easier to compare findings with existing studies. While more recent datasets are available, the 2015 dataset offers a robust foundation for analyzing labour market trends. Additionally, the 2015 dataset is unrestricted and accessible for teaching purposes, whereas more recent datasets may require safeguarded access, which can limit their usability for this study.

References

Acemoglu, D. and Autor, D. (2011) 'Skills, tasks and technologies: Implications for employment and earnings', in Card, D. and Ashenfelter, O. (eds.) *Handbook of labor economics*. Volume 4, Part B. Elsevier, pp. 1043–1171. Available at: https://doi.org/10.1016/S0169-7218(11)02410-5.

Alberti, G. (2022) ''chisquare': R package for chi-square test of independence'. Available at: https://www.researchgate.net/publication/359107636_'chisquare'_R_package_for_Chi-Square_Test_of_Independence.

Alwateer, M., Atlam, E.-S., Abd El-Raouf, M.M., Ghoneim, O.A. and Gad, I. (2024) 'Missing data imputation: A comprehensive review', *Journal of Computer and Communications*, 12(11), pp. XX-XX. Available at: https://doi.org/10.4236/jcc.2024.1211004.

Baker, D.H. (2022) *Research methods using R: Advanced data analysis in the behavioural and biological sciences*. Oxford University Press. Available at: https://eprints.whiterose.ac.uk/181926/.

Bell, D.N. and Blanchflower, D.G. (2011) 'Young people and the Great Recession', *Oxford Review of Economic Policy*, 27(2), pp. 241-267. Available at: https://docs.iza.org/dp5674.pdf.

Blundell, R., Crawford, C. and Jin, W. (2014) 'What can wages and employment tell us about the UK's productivity puzzle?', *The Economic Journal*, 124(576), pp.

377-407. Available at: https://www.ucl.ac.uk/~uctp39a/Blundell_Crawford_Jin_IFSwp201311.pdf.

Florisson, R. (2022) *The insecure work index: Two decades of insecurity*. Work Foundation, Lancaster University. Available at: https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/lums/work-foundation/UKInsecureWorkIndex.pdf.

Ghashim, E. and Boily, P. (2017) *Basics of R for data analysis*. Centre for Quantitative Analysis and Decision Support, Carleton University, Ottawa. Available at: https://www.data-action-lab.com/wp-content/uploads/2018/10/DSRS_BR.pdf.

Maillie, D. (2019) 'Dealing with skewness on any dataset'. Available at: https://www.researchgate.net/publication/335566169_Dealing_with_Skewness_on_any_Dataset.

Martin, R., Pike, A., Tyler, P. and Gardiner, B. (2015) 'Spatially rebalancing the UK economy: Towards a new policy model?', *Regional Studies*, 50(2), pp. 342–357. Available at: https://doi.org/10.1080/00343404.2015.1118450.

Pant, A. and Rajput, R. (2019) 'Introduction to research data and its visualization using R'. Available at: https://www.researchgate.net/publication/336982016_Introduction_To_Research_Data_And_Its_Visualization_Using_R.

Panda, N.R., Kumar, P.J., Mohanty, J.N. and Bhuyan, R. (2022) 'A review on logistic regression in medical research', *National Journal of Community Medicine*, 13(4), pp. 265-270. Available at: https://doi.org/10.55489/njcm.134202222.

Prasada, A., Mahapatra, A., Nanda, A., Mohapatra, B., Padhy, A. and Padhy, I. (2020) 'Concept of outlier study: The management of outlier handling with significance in inclusive education setting', *Asian Research Journal of Mathematics*, 16, pp. 7-25. Available at: https://doi.org/10.9734/ARJOM/2020/v16i1030228.

The 8th Annual International Conference (AIC) 2018 on Science and Engineering (2019) *IOP Conference Series: Materials Science and Engineering*, 523, p.012070. Available at: https://doi.org/10.1088/1757-899X/523/1/012070.