

**ECM104 Quantitative Research Methods  
Assessed Exercise 2020-2021**

**The Dynamic Relationship Between Unemployment in London and Other  
Regions of the United Kingdom**

This exercise is used to evaluate the ECM104 module. It counts for 100% of the mark for the module. There are no other exercises and no examination.

You must use Stata to complete the work. Do not use directly copied and pasted results tables from Stata in the main text of your answers. Any Stata output you wish to present should appear in clearly labelled appendices placed at the end of your write-up. These should be appropriately referenced in the main part of your solutions so material you refer to is easy to locate. In your main text present only the key results required in as succinct a way as possible. Certainly do not provide additional output that does not address the question.<sup>1</sup>

**Warning**

**This work must be conducted on an individual basis. It is expected and required that each person's work should be clearly their own and not be identical to any part of any other person's work, including that of other students.**

**Deadline for Completion of this Assessed Exercise**

**The deadline for the submission of this assignment is  
12.00 noon Friday January 22<sup>nd</sup> 2021**

---

<sup>1</sup> You must summarize the results from Stata in your answers. The complete output tables may include a lot of information that is not needed for the answer, in which case it should not appear. The complete tables *should* appear in an appendix to show that you have used Stata to complete the exercise.

## Introduction

Unemployment, and the impact of economic crises, tend to vary across space as well as time. It is therefore of interest to explore how the unemployment rates of different regions are related. In particular, concerning the UK, London and the south-east of England display persistently lower unemployment rates than other regions of the country. In this exercise you will explore the relationship between the unemployment rate in London and another UK region.

You will apply the time series techniques considered in ECM104 lectures and classes, including unit root tests, diagnostic testing of models, cointegration analysis, and error correction modelling. Koop (2013) discusses these methods (see especially chapters 9, 10, and 11).<sup>2</sup>

## Data

The data is contained in the Excel file called '*Regional Unemployment Data (2020) (2)*' available in the Assignments section of the Blackboard site. It contains monthly data for London and eleven other UK regions. These are given in table 1 below.

**Table 1: Regions and Variable Names**

Region	Variable Name
London	L
Northern Ireland	NI
North East	NE
North West	NW
Scotland	S
Wales	W
South West	SW
East	E

The data also includes a variable ***Date*** which provides the year and month of each observation.

The data are in percentage terms, and seasonally adjusted, for the period April 1992 to July 2020 inclusive (340 observations). The last few observations coincide with the onset of the COVID-19 pandemic. The data were obtained from the website of the UK Office for National Statistics (Office for National Statistics, 2020).

For this question, you are required to consider two series, that for London and one for another UK region. The other regional series you must use is determined by the allocation rule given in table 2 below.

Import the data into Stata and identify the series you will use. Set up a monthly time variable based on the variable *Date*. Use the *gen* and *tsset* commands introduced in Stata Exercise 5.<sup>3</sup>

---

<sup>2</sup> Useful texts including sections on unit root testing and cointegration are Brooks (2003) and Patterson (2000).

<sup>3</sup> The commands required are:

```
gen Month=monthly(Date, "YM")
```

```
tsset Month, monthly
```

**Table 2: Series by Student Number**

<b>If the last two digits of your student number are</b>	<b>You must use the variables:</b>
00 to 13	L and NI
14 to 27	L and NE
28 to 41	L and W
42 to 55	L and S
56 to 69	L and NW
70 to 84	L and SW
85 to 99	L and E

### **Questions**

#### **1. Ocular Econometrics**

- Provide time series line plots of each of your two unemployment series. (10 marks)
- Describe the evolution of the series over time commenting especially on any co-movements or changes in behaviour over time and the stationarity or otherwise of the series. (10 marks)

#### **2. Non-Stationarity of Unemployment**

Use augmented Dickey-Fuller (ADF) tests with an intercept to determine the order of integration of your two series. Recall that you will have to select an appropriate augmentation order for the tests. This is best done using an information criterion. (8 marks)

#### **3. The Long Run Relationship Between Unemployment in London and the Other Region**

- Determine if your two unemployment series are cointegrated using the Engle-Granger procedure. (See Stata Exercise 5 for how to do this). *Use London as the explanatory variable, not the dependent variable.* (8 marks)
- Comment on the value of the slope coefficient in the cointegrating regression. (8 marks)
- Save the residuals from the cointegrating regression as *RES1* and provide a time series plot of them. Comment any particular characteristics you observe in the graph. (8 marks)

#### **4. The Dynamic Relationship Between Unemployment Rates: Error Correction Models**

- Estimate two error correction models for the relationship between your unemployment series. One should have no additional lags of the differences (equation (1)), the other should include 8 lags of the differences (model (2)).<sup>4</sup> Let *L* represent the unemployment rate in London and *OTHER* that

---

<sup>4</sup> Remember an error correction model uses the differences of the variables of interest and the lagged residuals from the cointegrating regression.

of your other region. Then the models you should estimate are set out in equations (1) and (2):

$$\Delta OTHER_t = \beta_o + \beta_1 \Delta L_t + \beta_2 RES_{t-1} + u_t, \quad (1)$$

$$\Delta OTHER_t = \beta_o + \beta_1 \Delta L_t + \beta_2 RES_{t-1} + \sum_{i=1}^8 \alpha_i \Delta L_{t-i} + \sum_{i=1}^8 \delta_i \Delta OTHER_{t-i} + u_t \quad (2)$$

where *OTHER* is the unemployment rate of your other region,  $\Delta$  is the differencing operator such that  $\Delta X_t = X_t - X_{t-1}$ ,  $u_t$  is the regression disturbance, and  $\beta_i, i = 0, 1, 2$ ,  $\alpha_i, i = 1, 2, \dots, 8$ , and  $\delta_i, i = 1, 2, \dots, 8$  are coefficients.<sup>5</sup> (8 marks)

b) Use diagnostic tests<sup>6</sup> of the model errors and information criteria to determine which of these models is superior. (8 marks)

c) Use your preferred model to consider the following.

i) How, ***in the long run***, does unemployment in your other region adjust to that in London? Is the adjustment significant? (8 marks)

ii) How statistically reliable is the model on which you base these conclusions? (8 marks)

iii) Does your estimated equation include any insignificant coefficients? (8 marks)

iv) What impact, if any, has the COVID-19 pandemic had on your results. Explain. (8 marks)

## References

Brooks, C. (2002) *Introductory Econometrics for Finance*. Cambridge. Cambridge University Press.

Koop, G. (2013) *Analysis of Economics Data* (fourth edition). Chichester: Wiley

Office for National Statistics (2020)

<https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/datasets/labourmarketstatistics> Accessed 5/11/2020.

Patterson, K.D. (2000) *An Introduction to Applied Econometrics: a Time Series Approach*. New York: St. Martin's Press.

Simon P. Burke  
November 5<sup>th</sup> 2020

<sup>5</sup> Note that in equations (1) and (2), terms in  $L$  appear on the right hand side only. This is consistent with the long run relationship estimated in question 3, part a), which must have  $L$  on the right hand side as well (i.e. as the explanatory variable).

<sup>6</sup> Diagnostic tests are the tests of serial correlation, heteroscedasticity and non-normality of the model errors.