U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

Unemployment Hysteresis in OECD Countries – The impact of sampling on unit root tests

Pedro Guilherme Santos Martins

Dissertation

Master in Economics

Supervised by
Prof Doutora Maria Eduarda da Rocha Pinto Augusto da Silva

2022

Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof Doutora Maria Eduarda da Rocha Pinto Augusto da Silva and to PhD student Eduardo André Moura Martins Costa for the invaluable support, patience, feedback and guidance. They provided me with the knowledge and expertise to develop my dissertation with remarkable quality. Without them, I would not be able to flourish as a dissertation developer and writer.

Words cannot express how grateful I am for my parents' emotional and financial support in all my academic, personal and professional endeavours. They always provided me with all the necessary things for me to thrive, and there are no statements to describe how meaningful they are in my life. They are my life's basis and my idols. Being next to them symbolises my safest place and I could not feel prouder for being their son.

To all my friends and family that accompanied me in my emotional growth journey and encouraged me to pursue my dreams and never give up, I am extremely grateful for having you in my life. A special thanks must be given to Ana, Beatriz and my grandmother for being one of my biggest supporters.

Abstract

Unemployment has considerable economic and social adverse impacts as it is strongly correlated to inefficiencies and thus it is imperative to understand the unemployment dynamics. The literature on unemployment presents three main theories about unemployment dynamics: the NAIRU, the Hysteresis and the Structuralist hypotheses which are linked to the presence of unit roots in the unemployment dynamics. Both NAIRU and the structuralism hypotheses propose that unemployment is a mean-reverting process and thus incompatible with the presence of unit roots in its dynamics. On the contrary, if unit roots are present, one may conclude in favour of Hysteresis, implying that unemployment does not tend to return to equilibrium since any "natural" rate deviations will be permanent. The empirical literature presents conflicting conclusions on the presence/absence of unit roots in unemployment dynamics. The aim of this work is to assess the impact of sampling choices, namely the time frame and the frequency of the data, on unit-root tests applied to time series of unemployment rates. For that purpose, we study the unemployment dynamics of 38 OECD countries, from the first observation available on the OECD database to 2020, using univariate and panel unit root tests, thus extending results presented in the literature to more extensive samples. To assess the sensitivity of the unit-root tests to sampling characteristics - time frame and time frequency - an exhaustive analysis of the tests results is performed, increasing sequentially the sample and considering different univariate and panel unit-root tests. The results show that the conclusions of most tests are highly sensitive to sampling choices, suggesting that selecting samples to test for unit root presence in unemployment rates must be made thoughtfully since the results may influence macroeconomic guidance on the pursuance of costly stabilization policies and tools.

JEL codes: C12; C52; C80; E24; J64

Keywords: Unemployment; Unit roots; Sampling characteristics; OECD

Resumo

É imperativo compreender a dinâmica do desemprego, uma vez que este está fortemente correlacionado com ineficiências, provocando adversos impactos sociais e econômicos. A literatura apresenta três teorias sobre o desemprego: a NAIRU, a Histerese e o Estruturalismo, relacionadas com a presença de raízes unitárias na sua dinâmica. Tanto a NAIRU como o Estruturalismo propõem que o desemprego é um processo que retorna ao equilíbrio, sendo incompatível com a presença de raízes unitárias na sua dinâmica. Contrariamente, se raízes unitárias estão presentes, pode-se concluir a favor da Histerese, implicando que qualquer desvio da taxa de "equilíbrio" se perpetuará. A literatura empírica apresenta conclusões conflituantes sobre a presença/ausência de raízes unitárias na dinâmica do desemprego. O objetivo da presente dissertação é avaliar o impacto da escolha da amostra, nomeadamente o intervalo de tempo e a frequência dos dados, nos testes de raízes unitárias aplicados a séries temporais de taxas de desemprego. Para tal, avaliámos a dinâmica do desemprego de 38 países da OCDE, desde a primeira observação disponível na base de dados da OCDE até 2020, usando testes de raiz unitária em contexto univariado e de dados em painel, estendendo resultados apresentados na literatura a amostras mais extensas. Para avaliar a sensibilidade dos testes às características da amostra – intervalo e frequência temporal – realizámos uma análise exaustiva dos resultados dos testes, aumentado sequencialmente a amostra e considerando os diferentes testes de raízes unitárias. Os resultados demonstram que as conclusões da maioria dos testes apresentam forte sensibilidade às escolhas de amostragem, sugerindo que a seleção das amostras para testar a presença de raiz unitária nas séries do desemprego deve ser feita cuidadosamente, uma vez que os resultados podem influenciar a tomada de decisão de dispendiosas políticas macroeconômicas de estabilização do desemprego.

Códigos JEL: C12; C52; C80; E24; J64

Palavras-chave: Desemprego, Raízes Unitárias, Caracteristicas da amostra, OCDE

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Unemployment has considerable economic and social adverse impacts as it is strongly correlated to inefficiencies. It jeopardises the long-run economic growth and increases redistributive pressures and distortions at the economic level, rising poverty and inequality, and limiting labour mobility. At the social level, unemployment can impact the mental distress of the unemployed individuals, resulting in the loss of the workers' skills, leading to the lack of human capital or, ultimately, it can cause civil unrest. Accordingly, it becomes imperative to understand the unemployment dynamics, so that policies that counteract this undesirable phenomenon can be implemented and insights for well-designed macroeconomic policy-making can be provided.

The literature on unemployment presents three main theories about unemployment dynamics: the NAIRU, the Hysteresis and the Structuralist hypotheses.

Authors such as Friedman and Phelps (Friedman, 1968; Phelps, 1967, 1972) argue that macroeconomic shocks affect the unemployment rate temporarily, implying that unemployment follows a stationary, mean-reverting process. Thus, a shock, such as technological development or monetary policy changes, does not affect the natural rate of unemployment associated with an equilibrated labour market. Additionally, the concept of persistence in the unemployment rate, also known as the NAIRU (Non-Accelerating Inflation Rate of Unemployment) hypothesis, states that the usually slow speed of adjustment observed in the unemployment rate towards the long-run equilibrium can be explained by real wage rigidities and union behaviour.

On the other hand, Blanchard and Summers (1986, 1987) introduce the concept of hysteresis under which transitory shocks permanently affect the level of unemployment, i.e., any deviations of unemployment from its "natural" rate will be perpetual. According to them, this definition justifies the rationales such as the insider-outsider theory and the adverse impact of unemployment on human and physical capital stocks, mostly during high and persistent unemployment periods. Blanchard and Summers (1986) argue that current employees have all the bargaining power in the wage-setting, setting the wages above the market-clearing equilibrium, leaving no opportunities for the unemployed in the labour market. Phelps (1972) and Pissarides (1992) emphasise that the unemployed population's skills may depreciate since they find obstacles in adjusting workers' skills to those demanded by the available jobs. Long-term unemployment gives a wrong signal to employers about unemployed people, undermining future employment, and might increase the irrational behaviour of unemployed citizens, like the decrease in job searching, resulting in the deterioration of the human capital stock.

The structuralist theory of Perron (1989); Phelps (1994) presents an alternative to the hysteresis hypothesis. This theory implies that some occasional shocks, such as technological development, monetary and fiscal policy adjustments, lead to permanent changes in the equilibrium rate of unemployment, even though most of these shocks are temporary, i.e., unemployment is described by a stationary process subject to occasional but permanent structural changes.

The dynamics of the unemployment data are usually assessed via unit root tests. Both NAIRU and the structuralism hypotheses propose that unemployment is a mean-reverting process and thus incompatible with the presence of unit roots in its dynamics. On the contrary, if unit roots are present, one may conclude in favour of Hysteresis, implying that unemployment does not tend to return to equilibrium since any "natural" rate deviations will be permanent.

Even though it may appear simple to conclude on the unemployment dynamics, researchers have no consensus on which theory suits better unemployment dynamics. In fact, some studies provide empirical evidence for the natural rate hypothesis, meaning that the unemployment rate follows a mean-reverting process, i.e., stationary series, such as in Bolat, Tiwari, and Erdayi (2014); Camarero and Tamarit (2004); Furuoka (2017); Khraief, Shahbaz, Heshmati, and Azam (2020). Other authors, however, provide empirical evidence in favour of hysteresis in unemployment dynamics, i.e., nonstationary series, such as Bakas and Makhlouf (2019); Chang (2011); Chang, Lee, Nieh, and Wei (2005); Ener and Arica (2011); H.-Y. Lee, Wu, and Lin (2010). Furthermore, some empirical investigations conclude in favour of the structuralist hypothesis, like Ayala, Cuñado, and Gil-Alana (2012); Franchi and Ordóñez (2008); C.-C. Lee and Chang (2008). Meanwhile, it is crucial to note that these studies rely on either different unit root testing methodologies or different samples, which provides the rationale to explain the discrepancies in the conclusions on the unemployment dynamics present in the literature.

The first unit root test was introduced in 1979 by Dickey and Fuller (Dickey & Fuller, 1979). Since then, numerous alternatives have been developed and proposed in the literature to improve the performance of the tests and account for more complex dynamics in the data-generating processes.

The present research categorises the unit root testing framework into univariate and panel environments. The univariate tests use the information of an individual (country) to conclude individually for unit roots in a sample. The panel framework analysis corresponds to using a series of multiple individuals simultaneously to conduct the tests, where the conclusions are dependent on all individuals in the sample.

The standard unit roots tests under the univariate framework include the Dickey-Fuller test (Dickey & Fuller, 1979) and its augmented version (ADF) (Said & Dickey, 1984), developed to deal with the serial correlation of the error by including lagged terms of the dependent variable in the regression. These tests and those developed by Phillips and Perron (1988) - PP test - and Elliott, Rothenberg, and Stock (1996) - DF-GLS test - state in the null hypothesis presence of the unit root (non-stationarity) and its absence (stationarity) in the alternative. Kwiatkowski, Phillips, Schmidt, and Shin (1992) propose the KPSS test considering stationarity as the null hypothesis and non-stationarity in the alternative. Kwiatkowski et al. (1992) and Choi (1994) suggest combining tests with distinct null hypotheses to confirm the conclusion about unit-root presence.

The research developed by Perron (1989) is the milestone to accept that the standard unit root tests are biased toward the non-rejection of the unit root null hypothesis when the presence of a structural break characterises the sample; thus, other researchers also developed unit root tests considering structural breaks to handle the biases. The test developed by Bai and Perron (2003) is the most widely used used test to control for time series structural breaks. Perron (1990, 1997), Zivot and Andrews (1992), and J. Lee and Strazicich (2003, 2004) present unit root tests that account for the changes in parameters caused by structural breaks.

Panel unit root tests were developed to tackle the issue of the low power observed in the univariate framework analysis. Two generations of panel unit root tests may be defined. Tests of the first generation were developed under the assumption of cross-sectional independence between the individuals in the sample. While some of these tests have non-stationarity as the null hypothesis – Maddala and Wu (1999), Choi (2001), Levin, Lin, and James Chu (2002) and Im, Pesaran, and Shin (2003) - others have stationarity as the null hypothesis - Hadri (2002). The second generation of tests allows for cross-sectional dependency between the individuals in a sample, such as the tests developed by Moon and Perron (2004), Bai and Ng (2004), Choi (2006) and Pesaran (2007). Neither of the previously mentioned tests takes into account structural breaks. Therefore, some authors developed panel unit root tests that consider possible structural breaks and cross-sectional dependency - Im, Lee, and Tieslau (2005), Westerlund and Edgerton (2008), Banerjee and Carrion-i Silvestre (2015), and more recently, Arsova and Karaman Örsal (2021).

Since no empirical investigations assess the effects of the time span and time-frequency choices in the results of the unit-root tests, the present dissertation focuses on these issues by examining unemployment dynamics in the 38 OECD countries. The investigation contributes to the existing literature by (i) applying a vast battery of unit root tests to more extensive samples than those used in previous empirical studies; (ii) detecting how the sampling frequency and time spans might produce distinct unit-root tests conclusions.

The results of this work provide insights to future research on unit root analysis, which is essential for macroeconomic guidance about pursuing costly macroeconomic stabilisation policies and tools.

The dissertation is organised as follows: Section 2 introduces the fundamental concepts and reviews the literature regarding unit root tests on unemployment dynamics analysis; Section 3 displays the data used for the investigation, the unit root tests under analysis and applied econometric methodologies; Section 4 discusses the empirical results. Concluding remarks are given in Section 5.

Chapter 2

Literature Review

2.1 Fundamental concepts

Before proceeding to the literature review, a brief introduction of five fundamental concepts is relevant: unemployment rate, stationarity, unit root processes, structural breaks and cross-sectional dependence.

2.1.1 Unemployment Rate

Unemployment rate, as defined by Eurostat[1], is the ratio between the number of unemployed persons and the labour force - represented by the sum of the employed and unemployed citizens.

According to Eurostat, unemployed persons are "all persons 15 to 74 years of age (16 to 74 years in ES, IT and the UK) who were not employed during the reference week, had actively sought work during the past four weeks and were available to begin working immediately or within two weeks." Alternatively, a person is considered to be employed if has "worked at least one hour for pay or profit during the reference week or were temporarily absent from such work" (Eurostat). The labour force corresponds to the sum of unemployed and employed citizens.

Thus, unemployment rate in a certain time, u_t, is formulated as follows:

$$u_t = \frac{\text{Unemployed Citizens}}{\text{Labour Force}}$$

2.1.2 Stationarity

Stationarity is a fundamental notion for studying time series, as it qualifies robust analytic procedures by providing statistical stability or equilibrium to the series. Different forms of

[1] https://ec.europa.eu/eurostat/cache/metadata/EN/lfsi_esms.htm#unit_measure1645711702136

stationarity are defined, depending either on the statistical properties that are restricted or the applied procedures to achieve it.

A time series is considered strictly stationary if the joint distribution of $X_{t_1}, ..., X_{t_n}$ is the same as the joint distribution of $X_{t_1+\tau}, ..., X_{t_n+\tau}$, for all $t_1, ..., t_n$, and τ. Therefore, shifting time by a random value does not change the process's distribution, i.e., the probabilistic behaviour is the same across time. Therefore, the parameters that characterize the distribution of the process depends on the lag (τ) and do not depend on t, Maddala and Kim (1999).

Being defined in terms of the distribution function, stationarity is difficult to verify, thus the concept defined in terms of moments is favoured. Weak, covariance or even second-order stationarity is characterized as follows:

$$\begin{cases} E[X_t] = E[X_{t+\tau}] = u \\ V[X_t^2] = V[X_{t+\tau}^2] = \sigma^2 \\ cov(X_{t1}, X_{t2}) = cov(X_{t_1+\tau}, X_{t_2+\tau}) = \gamma_{t_1,t_2} = \gamma_\tau \end{cases}$$

The covariance-stationarity restricts the mean and variance of X_t to be finite and invariant in time, and the covariance to be a function of the time distance, γ_{t_1,t_2}, between two distinct observations $\tau = t_1 - t_2$, i.e., the autocorrelation function depends only on the lag between the observations. This form of stationarity is preferable in modelling time series since it is equivalent to strict-stationarity when the series follows the normal distribution and imposes fewer restrictions while guaranteeing statistical results viability.

Furthermore, a trend-stationary process materialises when an underlying trend is a predictable deterministic component displaying linear or quadratic dynamics that can be removed, resulting in a stationary process (Maddala & Kim, 1999). In the presence of a shock, a considerable and quick shift to the value of the series, a trend-stationary process is mean-reverting, i.e., over time, the series will converge to the mean, which remains unaffected by the shock.

2.1.3 Unit root processes

A random walk represents the prototype for a unit root process (Pfaff, 2008). A unit root time-series process displays an unpredictable systematic pattern, leading to distorted analysis and forecasts. A pure random walk without drift is defined as follows:

$$y_t = y_{t-1} + \varepsilon_t = y_0 + \sum_{s=1}^{t} \varepsilon_t$$

Where ε_t is a white noise (set of independent and identically distributed, zero mean random variables). If it is assumed that $y_0 = 0$, and the white noise process is independent of y_0, then $E[y_t] = 0$ and $var(y_t) = t\sigma^2$. Since the variance grows with time, a random walk is a non-stationary time series. Note that $\Delta y_t = y_t - y_{t-1} = \varepsilon_t$ and so the first difference of a random walk is stationary. Thus, by differencing the time series, stationarity is achieved, with the mean and variance of the process being constant over time. A non-stationary process transformed into a weakly-stationary process by differencing is said to be a difference-stationary process. A process that has to be differenced d times in order to achieve weak stationarity is said to be integrated of order d, denoted by $I(d)$ (Engle & Granger, 1987). To assess the value of d and therefore stationarity, researchers rely on unit root tests.

2.1.4 Structural breaks

Unexpected shifts in the process parameters that rule the time-series dynamics lead to effects known in the literature as structural breaks. External events are the leading cause of the changes, provoking significant alterations in the time-series pattern. A structural break can be represented by dummy variables formulating the changes in parameters as follows (Zivot & Andrews, 1992):

$$y_t = \alpha + \beta t + \theta DU_t + \gamma DT_t + \varepsilon_t$$

where DU_t is a dummy variable capturing a shift in the intercept, $DU_t = 1$ if $t > TB$, and zero otherwise. TB corresponds to the break moment. DT_t is a dummy variable capturing the shift in the trend, where $DT_t = t - TB$ if $t > TB$, and zero otherwise.

2.1.5 Cross-sectional Dependence

When conducting investigations that resort to panel data (time series of various individuals), the concept of cross-sectional dependence, denominated as CSD, is crucial. The individuals of the panel can be correlated since some common factors, such as oil prices, interest rates' movements and warfare states, affects all of them in diverse levels.

Whereas some tests do not account for this characteristic, denominated as first-generation panel unit-root tests, others make this consideration when analysing time series, designated by second-generation unit-root tests.

2.2 Empirical studies of unemployment dynamics

Numerous empirical studies have investigated unemployment dynamics in OECD countries by scrutinising unit root properties of the unemployment rate. Availability of data and the significant international impact of the organisation explains the large number of OECD based studies. While some authors conclude in favour of the hysteresis hypothesis, others reject it, and a third group supports the structuralist theory by controlling for the presence of structural breaks. Table 2.1 displays a summary of OECD based studies

The table evidences that even identical univariate unit root tests lead to different conclusions. For example, Mitchell (1993) finds unit roots in all time series except in Italy, concluding for the hysteresis hypothesis for almost all countries under analysis. Røed (1996) and Røed (2002) also conclude that the hysteresis thesis better describes unemployment dynamics except for the USA. Røed reduces the number of analysed countries from the former to the latter study because the 2002 research also includes the assessment of unemployment and vacancy rates effects on inflation dynamics, and a limitation of data is present, which requires a reduction in the countries for the analysis. Fève, Hénin, and Jolivaldt (2003) performs the same methodology as in Røed (2002) and adds eleven countries for unit root presence examination. The investigation achieves a contradictory conclusion for Canada in respect to Røed's study, but overall also favours the hysteresis hypothesis for most countries. It is worth highlighting that the main differences between these studies are the samples' period and frequency characteristics, reinforcing the importance of detecting whether and how these characteristics impact the tests' conclusions.

When controlling for structural breaks in univariate analysis, Papell, Murray, and Ghiblawi (2000) find evidence favouring the structuralist thesis for ten countries, where an unit root is not identified and thus hysteresis is rejected. C.-C. Lee and Chang (2008) performs a similar methodology and selects higher sampling spans to avoid seasonal adjustments and short sample periods shortcomings. For the limitation of data availability, there is a reduction in the number of countries under analysis. The investigation finds that a unit root is not present in all time series, favouring the structuralist thesis for all countries. Contradictory conclusions are present for all countries where Papell et al. (2000) find unit root presence, favouring the hysteresis hypothesis. The sample span is considerably different, and the number of considered breaks is distinct: the first study considers one break while the second examines between two and four. These differences explain the divergent conclusions and reinforce the necessity of studying the sample span impact on unit-root tests results.

Resorting to first generation panel unit root tests as central methodology, Song and Wu (1998) firmly reject the hysteresis in unemployment. Camarero and Tamarit (2004) engage in a similar strategy and reach the same conclusion. Regardless, the authors apply panel unit root test allowing individual analysis and find unit root presence for seven countries thus favouring the hysteresis hypothesis. Camarero, Carrion-i Silvestre, and Tamarit (2006) employ univariate and second generation panel unit root tests, simultaneously considering structural breaks in the data-generating process. The results evidences for unit root absence both in the panel and the country-by-country analysis, favouring the natural rate hypothesis and the structuralist thesis when accounting for structural breaks. J.-D. Lee, Lee, and Chang (2009) achieve the same conclusion of hysteresis rejection, while applying similar methodology. The results depict the invariability of the tests outcomes to different approaches, and the variability regarding the different assumptions, such as the presence of structural breaks and cross-sectional dependency or independence.

The sample selection depends on the data available at the commencement of the study and the requirements of the tests, with some panel unit-root tests demanding balanced data, i.e., equal number of observations without ruptures in the sample. The prolonged study period and the longstanding interval between the study termination and its revision and official publication explain the disparity between the last year of analysis and the publication moment.

Using similar samples (time spans and frequencies), various authors can reach different conclusions concerning unemployment dynamics in OECD countries, explained by different assumptions about the number of structural breaks, the employed unit root tests, and the characteristics of the panel data (cross-sectional dependence or independence).

Table 2.1: A summary of OECD based studies

Author(s)	Countries	Years of Analysis	Frequency	Methodology	Hysteresis conclusion
Mitchell (1993)	15	1960 (earliest) - 1991	Quarterly	ADF + PP tests	Confirmed except for Italy
Røed (1996)	16	1970 - 1994	Quarterly	ADF + Exact maximum likelihood stationarity test	Confirmed except for USA; Ambiguous results in four countries
Song and Wu (1998)	15	1970 - 1992	Quarterly	Panel Unit root test	Rejected
Papell et al. (2000)	16	1955 - 1997	Yearly	ZA type	Rejected in ten countries
Røed (2002)	10	1960 (earliest) - 1995	Yearly	ADF + KPSS	Confirmed except for USA
Fève et al. (2003)	21	1966 (earliest) - 1999	Quarterly	ADF + KPSS	Confirmed except for Canada, Netherlands and USA
Camarero and Tamarit (2004)	19	1956 - 2001	Yearly	MADF (joint panel ADF) + SURADF	MADF - Rejected; SURADF - confirmed in seven countries
Camarero et al. (2006)	19	1956 (earliest) - 2001	Yearly	Univariate and panel tests with CSD and structural breaks	Rejected in panel and country-by-country analysis
C.-C. Lee and Chang (2008)	14	1885 (earliest) - 2004	Yearly	Univariate unit root tests with structural breaks	Rejected
J.-D. Lee et al. (2009)	19	1960 - 2004	Yearly	Univariate and panel unit root tests with breaks	Rejected

Chapter 3

Data and Methodology

This chapter describes the data to evaluate unemployment dynamics in OECD countries, the unit root tests applied to the time series, and the methodology employed to conduct the study and achieve the conclusions.

To denote the unemployment rate time series throughout the study we use u_t as notation.

3.1 Data

We study unemployment dynamics in thirty-eight OECD countries by examining their unemployment time series. These experiments are performed using data collected from the OECD database on 16/11/2021. Table 3.1 displays the countries under consideration and the first observation available. The last observation under examination corresponds to the last observation of the year 2020, and there is no missing observations in the time frame for all countries.

For fourteen countries there is not enough yearly data observations to apply all the methods, which hinders part of the analysis. The fourteen countries are: Colombia, Costa Rica, Estonia, Greece, Hungary, Iceland, Israel, Latvia, Lithuania, Poland, Slovakia, Slovenia, Switzerland and Turkey.

For New Zealand and Switzerland, there is no monthly data.

Country	**Start**	**Country**	**Start**	**Country**	**Start**
Australia	1979	Greece	1999	Norway	1989
Austria	1993	Hungary	1996	Poland	1997
Belgium	1983	Iceland	2003	Portugal	1983
Canada	1955	Ireland	1983	Slovakia	1998
Chile	1986	Israel	2012	Slovenia	1996
Colombia	2007	Italy	1983	South Korea	1990

Country	Start	Country	Start	Country	Start
Costa Rica	2011	Japan	1955	Spain	1987
Czechia	1993	Latvia	1999	Sweden	1983
Denmark	1983	Lithuania	1999	Switzerland	2010
Estonia	1997	Luxembourg	1983	Turkey	2005
Finland	1988	Mexico	1987	United Kingdom	1983
France	1983	Netherlands	1983	United States	1955
Germany	1991	New Zealand	1986		

Table 3.1: OECD countries and first observation available

3.2 Unit root tests under study

This section introduces the unit root tests, both univariate and panel, considered in this study. The tests considered in this study are chosen for two main reasons: they are implemented credibly in software such as R and STATA, and represent the primordial unit root tests in the literature, having served as the basis for developing more complex tests.

3.2.1 Univariate Environment

Dickey-Fuller and Augmented version of the test

Dickey and Fuller (1979) develop the first test for assessing the presence of a unit root in a time series. Considering the model:

$$u_t = DT_t + \gamma u_{t-1} + \varepsilon_t$$

where u_t is the unemployment rate, DT_t is a deterministic component (zero, drift, drift plus trend), and ε_t is the error term described as a white noise process with mean 0 and variance σ^2, the hypotheses of interest are the unit root presence $H_0 : \gamma = 1$ against the stationary alternative $H_1 : \gamma < 1$.

In order to facilitate the test, u_{t-1} is subtracted to both equation sides:

$$\Delta u_t = u_t - u_{t-1} = DT_t + \rho u_{t-1} + \varepsilon$$

The test hypotheses become $H_0 : \rho = 0$ equivalent to $H_0 : \gamma = 1$ against the alternative hypothesis $H_1 : \rho < 0$ equivalent to $H_1 : \gamma < 1$

The Dickey-Fuller test assumes error homogeneity, an assumption that is commonly violated. To overcome this issue, Said and Dickey (1984) augment the test with higher-order differences (lags) of the dependent variable for residual structure correlation correction. Thus, the augmented Dickey-Fuller test exploits the following model specification:

$$\Delta u_t = DT_t + \rho u_{t-1} + \sum_{j=1}^{K} \beta_j \Delta u_{t-j} + \varepsilon_t$$

where K is the lag-augmentation with $\varepsilon_t \sim \textit{idd}N\left(0, \sigma^2\right)$, a set of independent and identically distributed white noise process with a normal distribution of mean 0 and variance σ^2. If no correction is made to account for the serial correlation of the errors, $K = 0$, one obtains the standard Dickey-Fuller test.

The model can be constructed under three deterministic component assumptions:

$$\Delta u_t = \alpha + \beta t + \rho u_{t-1} + \sum_{j=1}^{K} \beta_j \Delta u_{t-j} + \varepsilon_t \tag{3.1}$$

$$\Delta u_t = \alpha + \rho u_{t-1} + \sum_{j=1}^{K} \beta_j \Delta u_{t-j} + \varepsilon_t \tag{3.2}$$

$$\Delta u_t = \rho u_{t-1} + \sum_{j=1}^{K} \beta_j \Delta u_{t-j} + \varepsilon_t \tag{3.3}$$

In 3.1, a time trend plus a drift formulate the deterministic component while in 3.2 only a drift formulates the deterministic component and in 3.3 there is no deterministic component.

There are, in the literature, different testing procedures concerning the employment of the DF and ADF test. While some suggest starting with the less-restrictive model, equation (3.1), others recommend starting with the most restrictive model, equation (3.3). Applying the procedure proposed by Pfaff (2008) and displayed in Figure 3.1, one may conclude that either the series is stationary around a zero or non-zero mean or stationary around a linear trend or the series contains a unit root with a zero or a non-zero drift.

The unit root null hypothesis is tested using a conventional t-ratio for ρ. However, the test does not follow a conventional distribution, and one needs to compare the test statistic with critical values derived in Dickey and Fuller (1981).

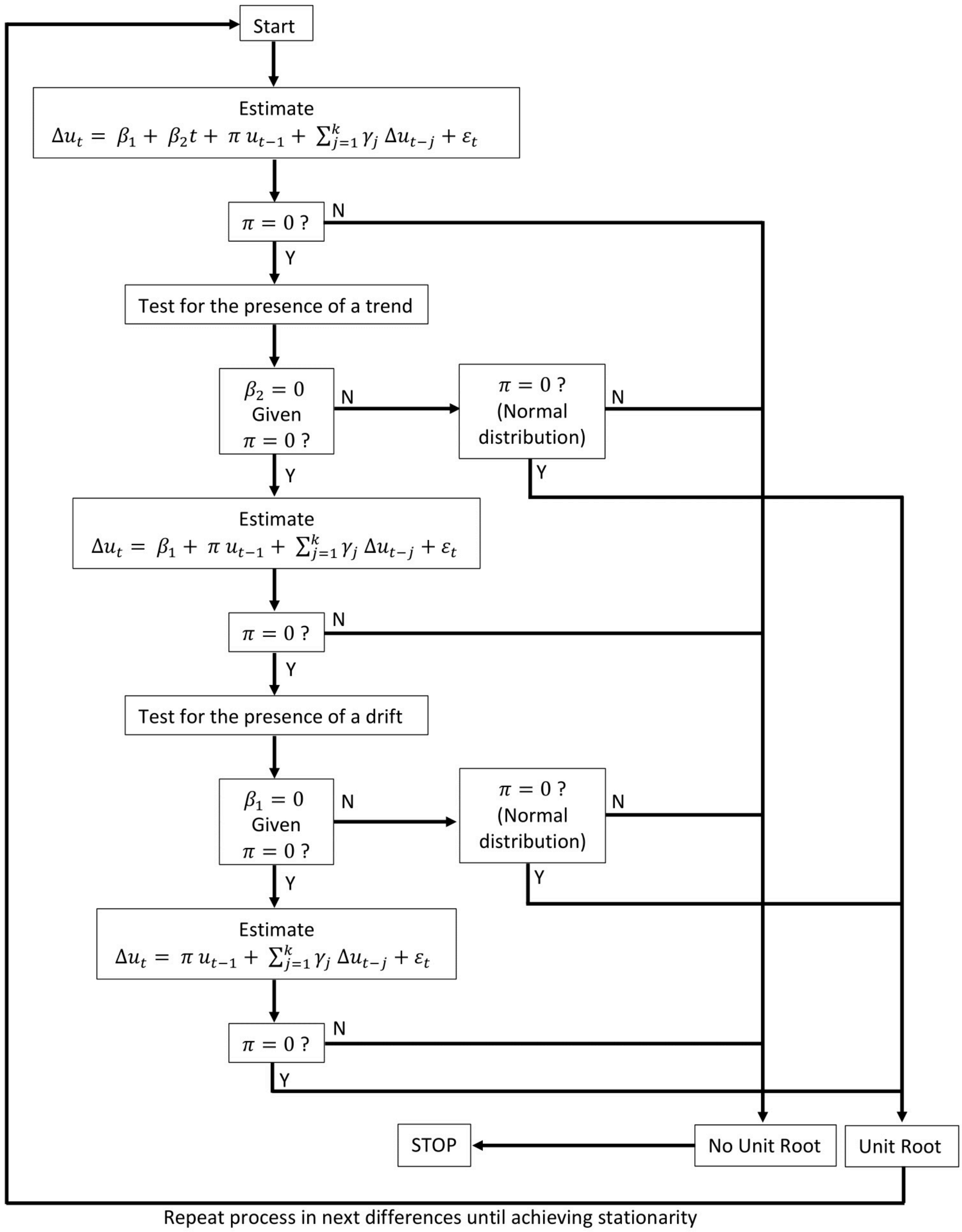

Figure 3.1: Pfaff (2008) ADF procedure

KPSS test - Kwiatkowski, Phillips, Schmidt and Shin

Kwiatkowski et al. (1992) propose a Lagrange Multiplier (LM) test for assessing either trend or level stationarity. The null hypothesis here is that of stationarity, opposing the ADF test unit root null hypothesis. The KPSS test considers the following model formulation:

$$u_t = \xi t + r_t + \varepsilon_t$$
$$r_t = r_{t-1} + \mu_t$$

where ε_t is stationary and r_t is a random walk with an error term μ_t that is assumed to be *iid* $\sim (0, \sigma^2_\mu)$. The initial value of r_t is fixed to r_0, representing the intercept (level).

The null hypothesis states that the error term variance is constant, $H_0 : \sigma^2_\mu = 0$, leading to stationarity of the time series. If $\xi = 0$, the model frames a constant as the deterministic regressor; otherwise, it formulates the deterministic component with a constant plus a trend. Under the null hypothesis, and since ε_t is stationary, u_t is either level-stationary ($\xi = 0$) or trend-stationary ($\xi \neq 0$).

To calculate the test statistic, we first need to regress u_t as above, and calculate its residuals partial sums $\hat{\varepsilon}$ as follows:

$$S_t = \sum_{i=1}^{t} \hat{\varepsilon}_i, t = 1, ..., T$$

The test statistic is defined as:

$$LM = \frac{\sum_{t=1}^{T} S_t^2}{\hat{\sigma}^2_\varepsilon}$$

with $\hat{\sigma}^2_\varepsilon$ being an estimate of the error variance from the first step.

Then we compare the LM test statistic with tabulated critical values, concluding on the stationarity of the series.

The general rule about statistical testing states that it is impossible to prove a null hypothesis, only affirm it. However, if the test rejects the null hypothesis, it is conceivable to guarantee that it is not valid. Thus, taking the null of stationarity represents a conservative strategy by placing the hypothesis of interest in the alternative. Pfaff (2008) argues that if the results of the ADF test indicate the presence of a unit root, but the KPSS result contradicts it, the researcher should be cautious and decide on the KPSS test conclusion.

Elliott-Rothenberg-Stock - DF-GLS test

Elliott et al. (1996) propose an alternative to solve the issue of the ADF and KPSS tests low power when the data-generating process has an autoregressive parameter close to the unity. To that end, they propose to first detrend the data using a generalised least squares (GLS) procedure and then applying the ADF test.

The null hypothesis is that u_t is a random walk, possibly with drift, and two alternative hypotheses arise: u_t is stationary around a linear time trend, or u_t is stationary with a possible non-zero mean. Under the first alternative hypothesis, a GLS regression estimates the intercept and trend by generating new variables, $\tilde{u}_t$, x_t and z_t defined as follows:

$$\begin{cases} \tilde{u}_1 = & u_1 \\ \tilde{u}_t = & u_t - \alpha^* u_{t-1},\ t = 2, \ldots, T \end{cases}$$

$$x_1 = 1,$$

$$x_t = 1 - \alpha^*, \quad t = 2, ..., T,$$

$$z_1 = 1,$$

$$z_t = t - \alpha^*(t-1),$$

$$\alpha^* = 1 - (13.5/T),$$

Then, an OLS regression estimates the equation: $\tilde{u}_t = \delta_0 x_t + \delta_1 z_t + \varepsilon_t$, where the OLS estimators $\hat{\delta}_0$ and $\hat{\delta}_1$ remove the trend from u_t by generating:

$$u_t^* = u_t - (\hat{\delta}_0 + \hat{\delta}_1 t) \tag{3.4}$$

Finally, an ADF test is applied to the transformed variable, u_t^*. This procedure is denominated as GLS detrending.

Under the second alternative, define $\alpha^* = 1 - (7/T)$, and eliminate z from the GLS regression (3.4), computing $u_t^* = u_t - \hat{\delta}_0$. Afterwards, one needs to fit the augmented Dickey–Fuller regression to the above variable, and perform the test $H_0 : \beta = 0$, comparing the test statistic with the tabulated critical values to conclude on the unit root presence. This procedure is denominated as GLS demeaning.

Zivot and Andrews - ZA - Unit root test

Zivot and Andrews (1992) design a framework to assess the presence of a unit root in a time series when one structural break is present in the DGP. The procedure consists in endogenously determining the time break (from now on TB) by using a data-dependent algorithm to dictate the position of the break point. The TB is the one that minimises the one-sided t-statistic of the null hypothesis: $\rho = 1$ in models **A**, **B** and **C** in equation (3.5. Therefore, the break occurs in the least favourable point for the unit-root presence null hypothesis.

To test for the presence of a unit root against the alternative of stationarity with a one-time structural break in the data-generating process, Zivot and Andrews use the following hypotheses:

$$
\begin{array}{ll}
H_0: & u_t = \alpha + u_{t-1} + e_t \\
H_1: & \\
\textbf{Model A:} & \Delta u_t = \alpha + \beta t + \theta DU1_t + \rho u_{t-1} + \sum_{j=1}^{k} c_i \Delta u_{t-j} \; + \varepsilon_t \\
\textbf{Model B:} & \Delta u_t = \alpha + \beta t + \gamma DT1_t + \rho u_{t-1} + \sum_{j=1}^{k} c_i \Delta u_{t-j} + \varepsilon_t \\
\textbf{Model C:} & \Delta u_t = \alpha + \beta t + \theta DU1_t + \gamma DT1_t + \rho u_{t-1} + \sum_{j=1}^{k} c_i \Delta u_{t-j} + \varepsilon_t
\end{array}
\tag{3.5}
$$

where DU_t and DT_t are dummy variables that capture intercept and trend breaks, respectively, formally represented in 3.6 and 3.7. Thus, while model **A** allows for a one-time shift in the intercept, model **B** only entitles a one-time change in the slope of the trend function. Furthermore, model **C** qualifies a shift in the intercept and a broken trend.

$$
DU_t = \begin{cases} 1, & \text{if } t > TB \text{ ;} \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}
$$

$$
DT_t = \begin{cases} t - TB, & \text{if } t > TB \text{ ;} \\ 0, & \text{otherwise.} \end{cases} \tag{3.7}
$$

While the models' null hypothesis states $\rho = 0$, implying unit root series with drift without any structural break, i.e., a random walk with drift and without breaks, the alternative hypothesis states $\rho < 0$, indicating a trend-stationary with one break series. The break can be in the intercept, trend or both deterministic components and the null hypothesis rejection occurs when ρ is statistically significant. Thus, this model allows the confrontation between the hysteresis and structuralist hypotheses.

3.2.2 Panel Environment

Levin Lin and Chu - LLC test

Levin et al. (2002) propose the LLC test that assesses the unit root presence in panel data by using the information of various individuals of the sample to execute the test. As a first-generation panel unit root test, cross-sectional independence is the main characteristic, furthermore considers homogeneous lagged dependent variable coefficients across all selected individuals, thus basing the analysis on the pooled estimator of the autoregressive parameter, $\hat{\rho}$. Regarding the unemployment rate observed on T periods, for N countries, the model can have the following formulations:

$$
\Delta u_{i,t} = \rho u_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta u_{i,t-z} + \varepsilon_{i,t} \tag{3.8}
$$

$$\Delta u_{i,t} = \alpha_i + \rho u_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta u_{i,t-z} + \varepsilon_{i,t} \tag{3.9}$$

$$\Delta u_{i,t} = \alpha_i + \delta_i t + \rho u_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta u_{i,t-z} + \varepsilon_{i,t} \tag{3.10}$$

The difference between the models relies on the individual exogenous variables that are under consideration. Whereas model 3.8 attributes no individual exogenous variables, model 3.9 assigns individuals intercepts, and model 3.10 ascribes individuals intercepts and time trends as exogenous variables.

For all individuals and observations the errors $\varepsilon_{i,t} \sim$ *i.i.d.*$(0, \sigma^2_{\varepsilon_{i,t}})$ are assumed to be independent across all the units of the sample. The hypotheses of interest are (respectively for models 3.8, 3.9 and 3.10):

$$H_0 : \rho = 0$$

$$\begin{cases} H_0 : \rho = 0 \wedge \alpha_i = 0, \forall i \\ H_1 : \rho < 0 \wedge \alpha_i \in R \end{cases}$$

$$\begin{cases} H_0 : \rho = 0 \cap \delta_i = 0, \forall i \\ H_1 : \rho < 0 \wedge \delta_i \in R \end{cases}$$

The restrictiveness occurring in both hypotheses implies identical autoregressive parameters (for instance, convergence/divergence of growth models happening at the same rate) between the individuals in the panel, a condition usually not satisfied. This restricts all individuals in the sample as unit-root processes, favouring the hysteresis hypothesis, or as stationary processes, supporting the NAIRU hypothesis.

Im, Pesaran and Shin - IPS test

The IPS test of Im et al. (2003) also belongs to the first-generation panel unit root tests, characterised by cross-sectional independence. Similarly to the LLC test, and considering individual effects with no time trend, and individual effects with time trends as exogenous variables, the IPS test is based on the following formulations (respectively):

$$\Delta u_{i,t} = \alpha_i + \rho_i u_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta u_{i,t-z} + \varepsilon_{i,t} \tag{3.11}$$

$$\Delta u_{i,t} = \alpha_i + \delta_i t + \rho_i u_{i,t-1} + \sum_{z=1}^{p_i} \beta_{i,z} \Delta u_{i,t-z} + \varepsilon_{i,t} \tag{3.12}$$

Another common feature between the LLC and the IPS tests is that both are panel extensions of the ADF test, based on t-ratios. However, while the LLC pools the equations and computes a pooled test statistic resulting in a homogeneous alternative, the IPS averages the individual test statistics for N cross-section units, obtaining group-mean tests, allowing for a heterogeneous alternative hypothesis that qualifies unit-root presence for some individuals in the panel.

Thus, the hypotheses of interest are:

$$\begin{cases} H_0 : \rho_i = 0, \forall i \\ H_1 : \rho_i < 0, i = 1, ..., N_1 \wedge \rho_i = 0, i = N_1 + 1, ..., N \end{cases}$$

The null hypothesis also implies auxiliary assumptions about the individuals' effects, particularly $\alpha_i = 0$ in model 3.11 and $\delta_i = 0$ in model 3.12 for all individuals in the sample, such as the LLC formulation. This restrictiveness is not present under the alternative hypothesis.

Under the null hypothesis, all individuals have unemployment time series characterized by unit-root processes, and under the alternative, at least one individual in the panel has in its DGP a unit-root while the remaining are time series are stationary processes, allowing for hysteresis conclusion for some countries in the panel and the NAIRU conclusion for the remaining individuals.

Fisher's type tests: Maddala and Wu and Choi

Maddala and Wu (1999) and Choi (2001) present techniques that use combinations of observed significance levels from individual tests (p-value based approach), arguing that the p-values from independent ADF tests can be combined obtaining a test on the joint hypothesis concerning all individuals.

The models for the fisher's type tests can be defined under the three assumptions about the individuals exogenous variables (no variables, intercepts and intercepts plus time trends), such as derived in models 3.8, 3.9 and 3.10 and they all assume cross-sectional independence.

Maddala and Wu (1999) propose a test statistic defined as follows:

$$P_{MW} = -2 \sum_{i=1}^{N} log(p_i)$$

which has a χ^2 distribution with $2N$ degrees of freedom, when T tends to infinity and N is fixed.

Banerjee (1999) notices the simplicity of the test and its robustness to the statistic choices, lag length and sample size and asserts the attractiveness of the test.

Choi (2001) proposes in his paper three test statistics, namely the modified P test, the inverse normal and the logit tests. The test statistic are respectively defined as follows:

$$P_m = -\frac{1}{\sqrt{N}} \sum_{i=1}^{N} (ln(\rho_i) + 1)$$

$$Z = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Phi^{-1}(\rho_i)$$

where $\Phi(\bullet)$ is the standard normal cumulative distribution function;

$$L = \sum_{i=1}^{N} ln(\frac{\rho_i}{1-\rho_i})$$

Considering the heterogeneous models, the hypothesis of interest are:

$$\begin{cases} H_0 : \rho_i = 0, \forall i \\ H_1 : \rho_i < 0, i = 1, ..., N_1 \wedge \rho_i = 0, i = N_1 + 1, ..., N \end{cases}$$

similarly to the IPS formulation.

Hadri test

Hadri (2002) suggests a test representing an extension of the Kwiatkowski et al. (1992) test, opposing the above-mentioned tests by taking the null of stationarity. Hadri proposes a residual-based LM test where the null hypothesis supports that the individual series $u_{i,t}$ (for $i = 1, ..., N$, and $t = 1, ..., T$) are stationary, either around a deterministic level or a deterministic trend (depending on how the composition of the deterministic component is employed), against the alternative of at least unit root in the panel data.

Hadri (2002) considers two models:

$$u_{i,t} = r_{i,t} + \varepsilon_{i,t} \tag{3.13}$$

$$u_{i,t} = r_{i,t} + \beta_i t + \varepsilon_{i,t} \tag{3.14}$$

Where $r_{i,t}$ is a random walk: $r_{i,t} = r_{i,t-1} + \mu_{i,t}$, with $\mu_{i,t}$ being i.i.d. (0,σ^2_μ), and $\mu_{i,t}$ and $\varepsilon_{i,t}$ being independent. The null hypothesis is $\sigma^2_\mu = 0$. Since $\varepsilon_{i,t}$ are assumed to be i.i.d., under the null, $u_{i,t}$ is stationary around a deterministic level in model 3.13 and stationary around a deterministic trend in model 3.14.

Analogously to all panel unit root tests displayed under section (3.1.2), the Hadri test belongs to the first-generation panel unit root tests, with cross-sectional independence assumption prevailing. And similarly to the IPS and fisher type tests, it allows for an heterogeneous alternative hypothesis, with the hypotheses of interest being formulated as follows:

$$\begin{cases} H_0: & \text{all individuals have stationary processes} \\ H_1: & \text{at least one individual has a unit-root process} \end{cases}$$

3.2.3 Lag-length selection

Most unit root tests require a choice of parameters that may influence the results. Particularly important is the choice of a modelling correction that accounts for the serial correlation of the error by including lagged terms of the dependent variable. Thus, it is vital to determine the optimal lag length, which represents a trade-off between having better size properties (related to size distortions) and better test power. Therefore, two problems may arise: if the lag is too small, the results are biased, and if it is too large, the power of the test suffers a reduction.

The literature provides different methods for the lag length selection. The current investigation follows the methodology suggested by Schwert (1989), where the maximum lag length is determined using the short and long approaches, formulated respectively:

$$l_4 = int[4(*(\frac{T}{100})^{\frac{1}{4}}]$$
$$l_{12} = int[12(*(\frac{T}{100})^{\frac{1}{4}}]$$

where T corresponds to the sample size.

Under the univariate environment both approaches are under investigation. Under the panel environment we only consider the short approach since the long, by selecting the number of lags considering the number observations of the full panel, achieves a value that does not bare econometric reasoning to apply for some individuals in the panel, where the number of observations is not extensive enough. The software recognizes this limitation and does not allow the application of this approach.

After determining the maximum number of lags, the Schwarz (1978) information criteria selects the optimal lag length, which accounts for the achieved maximum lag length and sample characteristics. This information criteria is applied for the ADF and the panel unit root tests that are formulated similarly to the ADF regression (IPS, LLC and Fisher's type tests). For the remaining unit-root tests there is no information criteria selecting the best lag due to software limitations, thus the number of lags considered are the ones derived under the short approach for the panel unit root tests and both approaches for the univariate unit root tests.

3.3 Design of the experimental study

This research is performed using R Software (R Core Team, 2021) and resorts to the package *urca* developed by Pfaff (2008) to apply the univariate unit root tests, and the package *plm* designed by Croissant and Millo (2018) to operate the panel unit root tests.

3.3.1 Univariate Environment

Initially, we assess the unit root presence for the countries under analysis for the largest period available, intending to extend previous OECD based studies to a larger number of countries and years of analysis. This method is applied to the three observation's frequencies (monthly, quarterly and annual).

To assess the impact of sample choices, time span and sample frequency, in the tests for unit roots in unemployment rates time series, the tests are performed sequentially in samples of increasing size (one unit at a time), starting from a minimum admissible sample size that guarantees satisfactory properties of the tests, good power and low size distortions (Afriyie, Twumasi-Ankrah, Gyamfi, Arthur, & Pels, 2020; Arltová & Fedorová, 2016; Shiller & Perron, 1985), corresponding to twenty five observations.

We inspect if the tests' conclusions are robust to the sample span, denoted as R_S, by assessing if the tests's conclusions are the same for different sample spans. We inspect if the tests' conclusions are robust to the sampling frequency, denoted as R_F, by evaluating if the tests provide the same conclusion at the various frequencies.

This methodology is sequentially applied to the two lag structures for the lag length selection, short and long approaches, referred above.

3.3.2 Panel Environment

Whereas the Fisher's type tests are applied to unbalanced panels, the remaining tests require balanced panels. Thus, the procedures applied to perform the analysis are slightly different for these two types of tests.

For the tests that consider unbalanced panels, we start by assessing the unit root presence considering all observations available for all individuals. We apply the test to that respective panel at the three observations' frequencies, allowing the extension of previous investigations to a larger number of individuals and number of observations that constitute the panel. We then compare the results at the different frequencies to determine if the results are robust to the sampling frequency, R_F.

To assess if the tests' results are robust to the sampling span, R_S, we construct panels that possess twenty-five observations (Hlouskova & Wagner, 2005) for each individual, starting from the first observation, and run the tests. Then, we increase the number of observations of each individual by unity, until the panel is includes all data available, and we apply the tests in each iteration. This allows monitoring the results' behaviour to the different sample spans. This procedure is performed to the three sampling frequencies, and entitles a judgment concerning R_S.

For the tests that require balanced data panels, we perform an examination considering six cases regarding the sample spans. For the first case, we construct panels that include all countries from whom we have sufficient data (more than twenty-five observations), and we run the tests by following the unity increase methodology (from twenty-five observations until the full data is included) to further analysis of the test results robustness to the sample span, R_S, and observation frequency, R_F, and conclusion of the test for the full panel.

Then, we sequentially increase the sample span by eliminating individuals from the panels, to perform the same results scrutiny by following the respective procedure. Table 3.2 reveals the individuals and the beginning of the sampling span that constitutes the panel. We further assess if the tests' results are robust to the sample span, R_S, and we thoroughly examine the robustness of the tests to the sampling frequency, R_F, by comparing results at different frequencies for a given sample span.

Case	Individuals in the panel	Sample Span		
		Annual	Quarterly	Monthly
1	All countries under analysis	1993	2012	2012
2	Austria, Czechia and Germany excluded		1990	
3	Only includes: Australia, Canada, Japan and USA		1980	
4	Only includes: Canada, Japan and USA		1970	
5	Only includes: Canada, Japan and USA		1960	
6	Only includes: Canada, Japan and USA		1955	

Table 3.2: Balanced data panels for unit root tests

Chapter 4

Results

This chapter presents the results of applying the methodology described in Section 3.3. One of the goals is to provide insights on the presence of unit roots in unemployment dynamics in OECD countries by applying a vast battery of tests to a large number of countries and years of analysis. The other is to scrutinise the tests' results to determine whether robustness to the sampling span, R_S, and sampling frequency, R_F, exists. Such analyses allow understanding the impact of sampling decisions when running unit-root tests, explaining the disparity of conclusions presented in the literature when resorting to similar unit-root tests. This chapter begins by providing two illustrative examples, one considering the ADF univariate test and another considering the IPS under a panel data environment, to guide the readers on how data were analysed to achieve the conclusions in each unit-root test. Then, it displays the results in a more general way.

4.1 Illustrative examples

4.1.1 ADF test - USA case

Firstly, ADF unit roots test conclusions are examined considering the largest span available for the three frequencies of USA data. Then, the tests are applied sequentially to samples of increasing size to assess how the tests' conclusions behave with such increments at the different frequencies.

Figures 4.1 and 4.2 display the evolution of the USA's unemployment rate time series, u_t, from the first observation available to 2020 at the three sampling frequencies for short and long lag versions of the test. Ensuring a minimum of the first twenty-five observations and increasing the sample size in a one-time unit, the plots represent the ADF test conclusion, highlighted with different colours, for different sample spans. The grey colour represents the minimum number of observations, the red indicates unit root presence, supporting the hysteresis hypothesis, and the green denotes unit root absence, favouring the NAIRU thesis.

Considering the test conclusion for the complete series, observed at the largest span available in Figures 4.1 and 4.2, the ADF test favours the NAIRU hypothesis for all frequencies and lag choice versions once that at the maximum observation of plots lacks the unit root conclusion. Also, from such figures, it is possible to infer that test conclusions are not robust to the sample span and the sampling frequency. Indeed, conclusions for both short and long lag approaches depend on the selected sample span, and the test validates different conclusions at different frequencies for a given sample span.

The same analysis is performed for the remaining univariate unit root tests, exposed in Section 3.2.1. The Annex A contains the results.

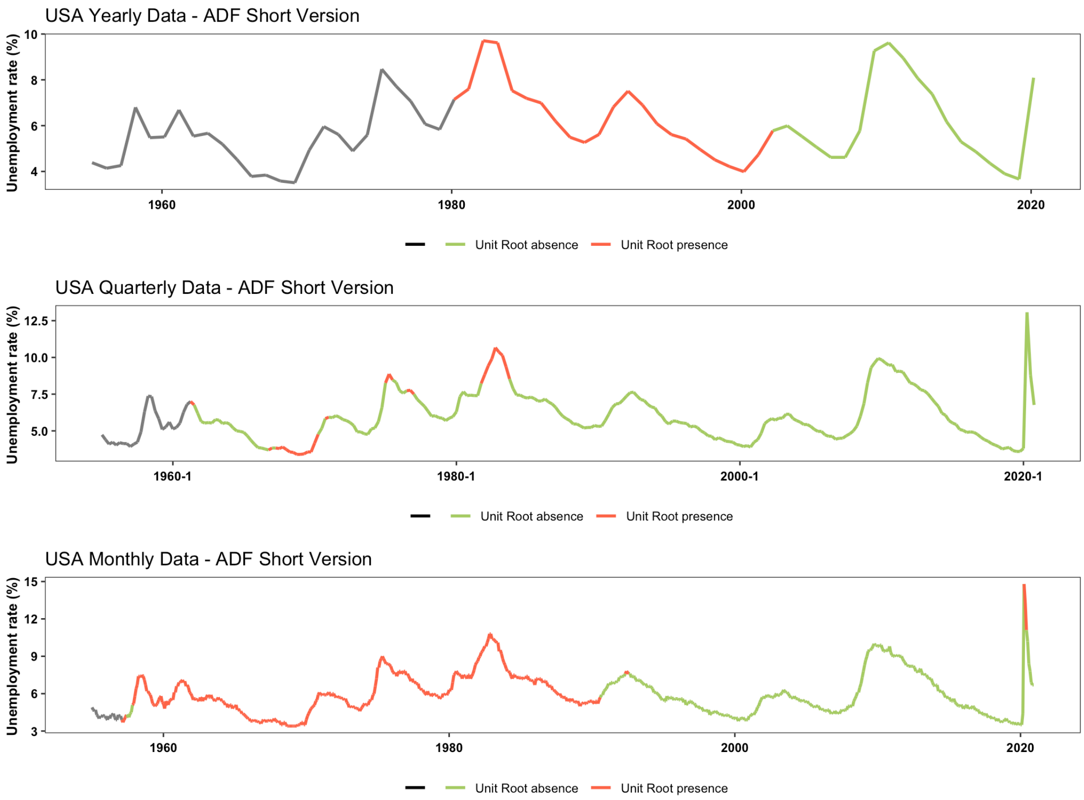

Figure 4.1: ADF test, short lag's version - USA

4.1.2 IPS test

The IPS test illustrated considers the case 2, referred to in Table 3.2, a panel data from 1990 to 2020, including the countries: Australia, Belgium, Canada, Chile, Denmark, Finland, France, Ireland, Italy, Japan, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Portugal, South Korea, Spain, Sweden, UK and USA. The minimum number of observations for each country is 25 to perform the test under the increasing time units framework. Ad-

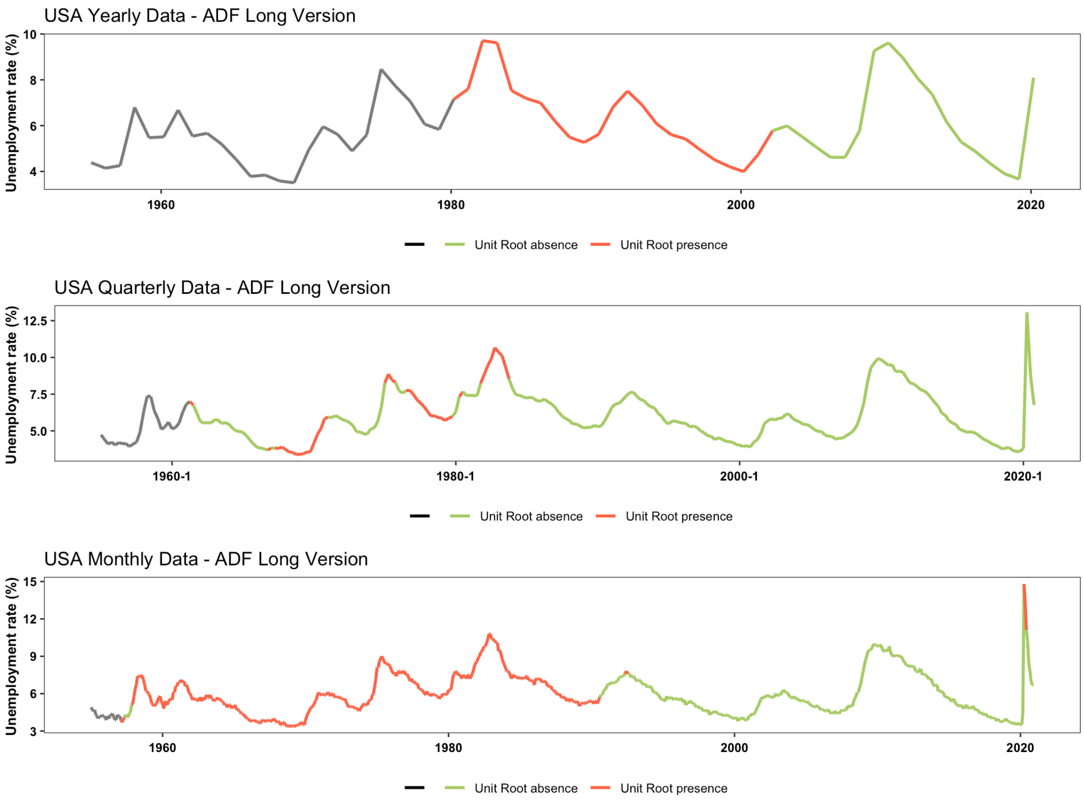

Figure 4.2: ADF test, long lag's version - USA

ditionally, this test analysis considers the possibility of two individual exogenous variables assumptions - individual intercepts or individual intercepts plus time trends.

Figures 4.3, 4.4 and 4.5 display the test results for the annual, quarterly and monthly cases, respectively. In such figures, the x-axis represents the cumulative analysis span in the panel, i.e., each date on this axis indicates the test conclusion under a panel containing historical data for all countries up to a particular axis date. The y-axis labels 0 to test results supporting the null hypothesis of all individuals presenting unit-root processes in the DGP and 1 favouring the alternative hypothesis that at least one country features a stationary process. In the plots, red lines denote the assumption of individual intercepts, and purple lines represent individual intercepts plus time trends.

From Figure 4.3, for both exogenous variables assumptions cases, when the panel contains all data from 1990 to 2020, the test results favour that at least one country displays unemployment time series characterized by a stationary process. Also, the test results are robust to the sampling span, R_S, since the results are invariant to the considered sample span.

Results for the quarterly sampled case, depicted in Figure 4.4, also favour at least one country's unemployment series featuring stationarity with complete data. However, in this

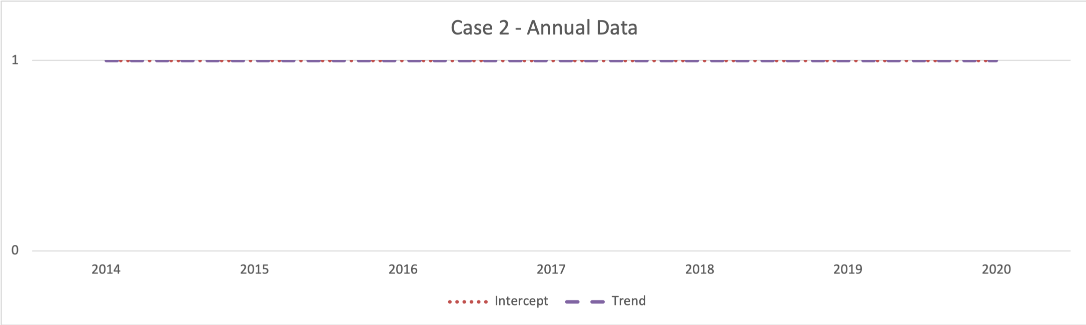

Figure 4.3: IPS test - Annual Data

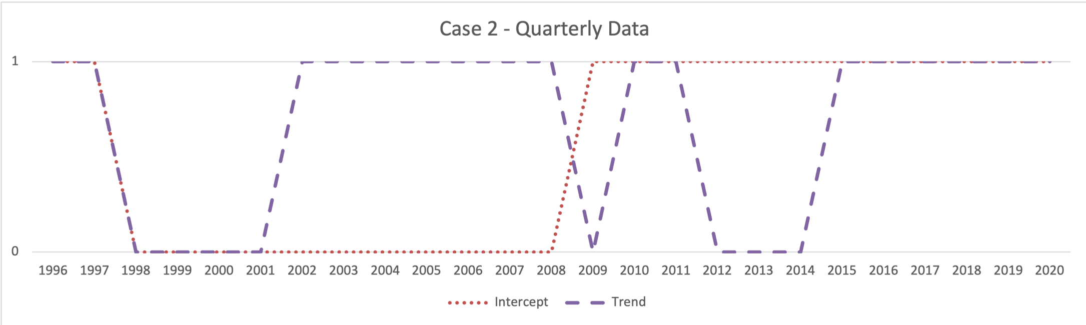

Figure 4.4: IPS test - Quarterly Data

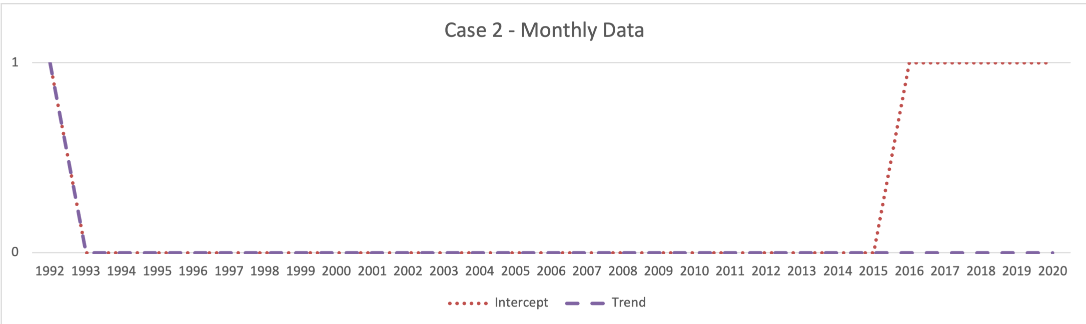

Figure 4.5: IPS test - Monthly Data

case, conclusions are not robust to the sample span, R_S. The monthly sampled case converges in such a conclusion regarding R_S; however, as Figure 4.5 depicts, when applying the test to the entire sample span, results contradict the previous ones when considering exogenous variables assumptions. Monthly panel data results favour the null hypothesis that all countries contain a unit-root process in their unemployment dynamics.

Regarding R_F, findings point to a lack of robustness to sampling frequency since the results are inconsistent among the three considered frequencies in specific sample spans. For instance, when the panels are constituted with sample spans from 1990 to 2015, whereas the annual and quarterly cases favour the alternative hypothesis, the monthly case favours the null hypothesis.

4.2 Univariate Environment

4.2.1 ADF test

Resorting to the methodology defined in Chapter 3, we start by applying the test to the largest time span available. Figure 4.6 displays in blue the percentage of countries where the test fails to reject the hysteresis null hypothesis for the two lag structures. The results reveal the unemployment dynamics as unit root processes for at least 54% of 38 countries under analysis, regardless of the sampling frequency and the lag choice. Additionally, it is possible to note the impact of the data frequency on the test's results; the higher the sampling frequency, the higher the tendency of the test to conclude in favour of hysteresis. Whereas annual frequency series point from 54% to 58% of countries' unemployment dynamics as unit-roots, quarterly series indicate from 71% to 74%, and monthly series from 86% to 89%.

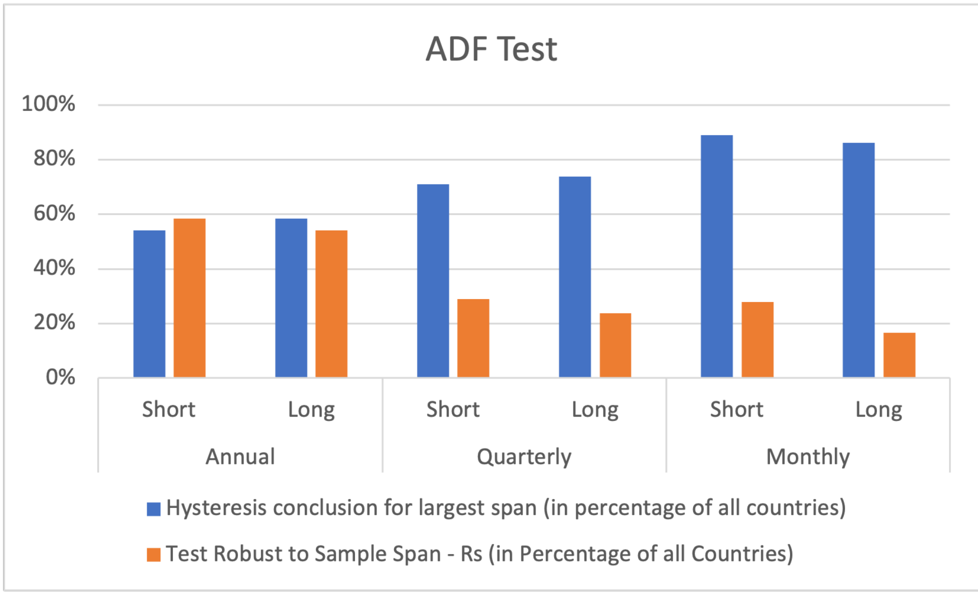

Figure 4.6: ADF Test

Figure 4.6 also presents, in orange, the percentage of countries where the ADF test delivers the same conclusion for different sample spans at the diverse frequencies, being robust to the sample span, R_S. Our results show that the smallest percentage of countries where the test is robust to the sample span is 17% out of the 38 countries under analysis, regardless of the frequency and lag structure. Furthermore, there is a pattern of decreasing percentage of countries where the test is robust to the sample span, R_S, as we apply the test to higher frequencies. Whereas annual frequency series point from 54% to 56% of the countries presenting sensible results to the sampling span, quarterly series exhibit from 24% to 29%, and monthly series from 17% to 28%.

For a given sample span, our results reveal that around one-third of the countries deliver the same conclusion for the different frequencies, being the test robust to the sampling frequency, R_F, as illustrated in Table 4.1. There is one country (Switzerland) where only monthly observations are available, thus it is not possible to assess if the test is robust to the sampling frequency in that country, illustrated by the 2% in the table. We can conclude that the majority of the countries exhibit outcome variation related to the observation's frequency, reinforcing the importance of making consideration of this sampling characteristic when comparing different empirical studies.

Not Robust	65.8%
Robust, R_F	31.6%
Not possible to compare	2.6%

Table 4.1: ADF Test - Robust Frequency Analysis (in percentage of countries)

4.2.2 KPSS test

For this test, we consider two possibilities for the deterministic component, namely, a constant and a constant plus a trend, as suggested in the literature. Figures 4.7 and 4.8 reveal, in blue, the percentage of countries where the test finds a unit root regarding the largest span available case and, in orange, the percentage of countries where the test is robust to the sample span.

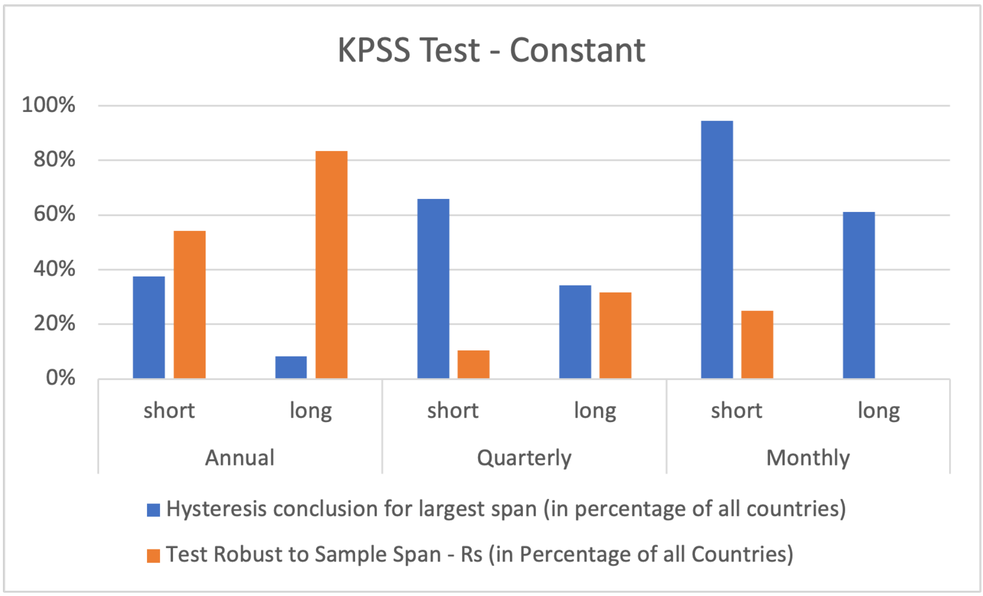

Figure 4.7: KPSS Test - Constant as deterministic component

Similarly to the ADF test's results, a pattern of increased propensity to find a unit root in the unemployment dynamics emerges as we employ the test to higher frequency data, regardless of the deterministic component. When only a constant characterizes the deterministic component the test characterizes unemployment as unit root processes from 8% to 38% of the 38 countries in annual series, from 34% to 66% in quarterly series and from 61% to 91% in monthly series. When adding a trend as a deterministic component, our results reveal unemployment as unit root processes from 13% to 21% in annual series, 29% to 87% in quarterly series, and 78% to 100% in monthly series.

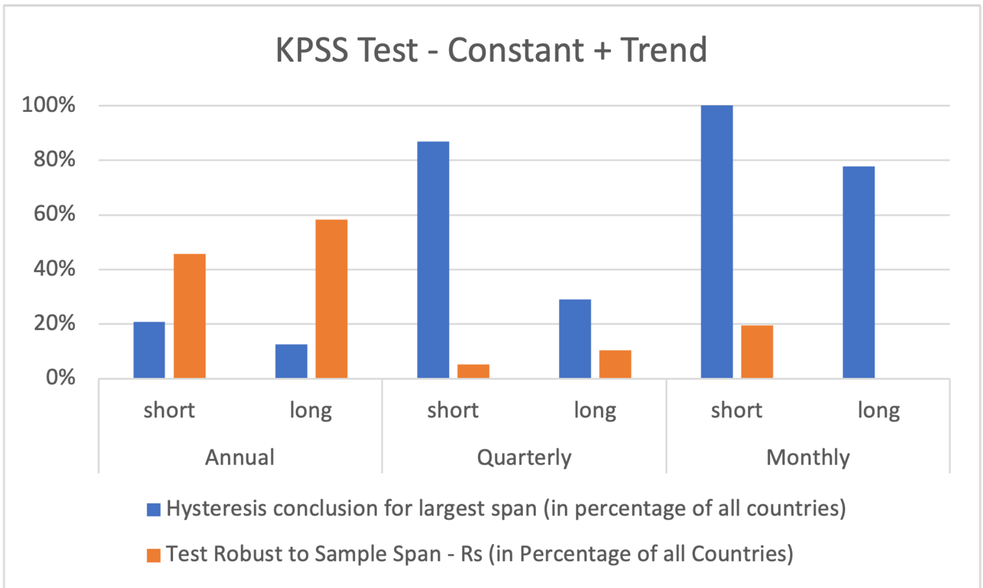

Figure 4.8: KPSS Test - Constant + Trend as deterministic components

Simultaneously, there is a decreasing percentage of countries where the test deliver conclusions that are robust to the sample span, R_S, as we use higher frequency data. However, as shown in Figure 4.7 and Figure 4.8, for the short version of lags, the test presents a lower percentage of countries with results that are robust to the sample span when using quarterly data compared to the monthly data case. Our results indicate that, for all countries, the test lacks sampling span robustness for the long version of lag in all countries, when operating with monthly data.

	Constant	Constant + Trend
Not Robust	73.7%	84.2%
Not Robust in some lag versions	21.1%	10.6%
Robust, R_F	2.6%	2.6%
Not possible to compare	2.6%	2.6%

Table 4.2: KPSS Test - Robust Frequency Analysis (in percentage of countries)

For the same sample span, our study shows that the test delivers different conclusions at different frequencies for around 74% of the 38 countries when only a constant characterizes the deterministic component, and for around 84% of the countries when a trend is also considered as deterministic component, meaning that the test's conclusions are not robust to the sampling frequency. Furthermore, only around 1 country deliver the same conclusion for the different frequencies, being robust to the sampling frequency, R_F, as shown in Table 4.2, illustrating a considerable decline when analogised to the ADF test results. Regardless, for some countries, the robustness of the test to the sampling frequency depends on the lag structure.

4.2.3 DF-GLS test

As suggested in Pfaff (2008), there are two methods, demeaning and detrending, to define the dynamics that reasonably describe unemployment. In both cases, our results indicate that when applied to the largest span available, the test's results favour the hysteresis hypothesis for the majority of the countries. As can be observed in Figures 4.9 and 4.10, the lowest percentage of countries where the test finds a unit root in the DGP is 54%.

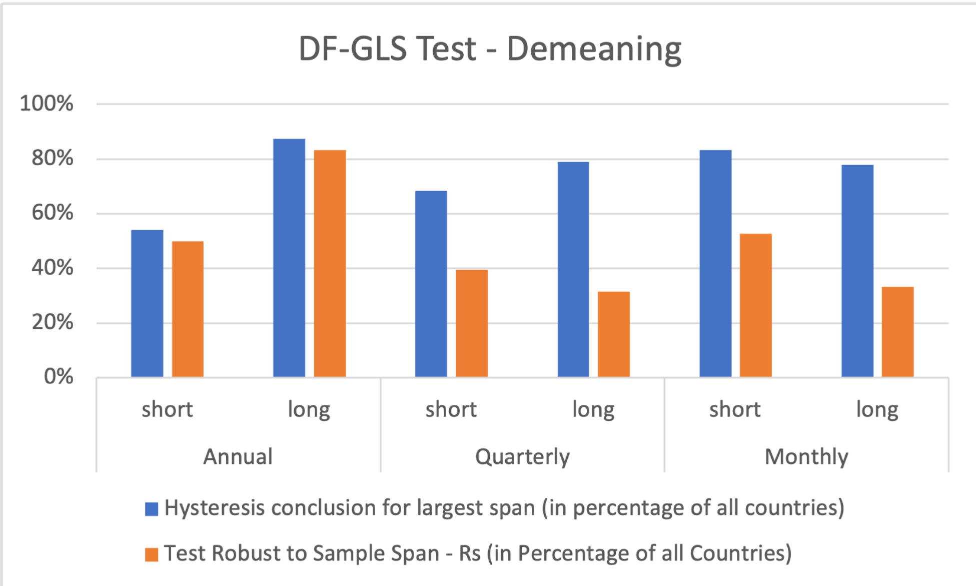

Figure 4.9: DF-GLS Test - Demeaning

For the demeaning method, our results show an increasing predisposition to locate a unit root in the DGP for the short version of lags while operating with higher frequency data, from 54% in the annual series to 68% in the quarterly series to 83% in the monthly series. However, for the long version of lags, there is a decreasing tendency to find a unit root in the DGP, from 88% to 79%, and then to 78%, operating with annual, quarterly and monthly series, respectively.

In the detrending method, a decreasing pattern is present when considering the long version of lags, where the test provides hysteresis hypothesis support for 96% of the 38 countries with annual series, 92% with quarterly series and 86% with monthly series. For the short version of lags, the test finds unit-root presence for more than 90% when operating with annual and monthly series, and 79% when applying the test to quarterly series.

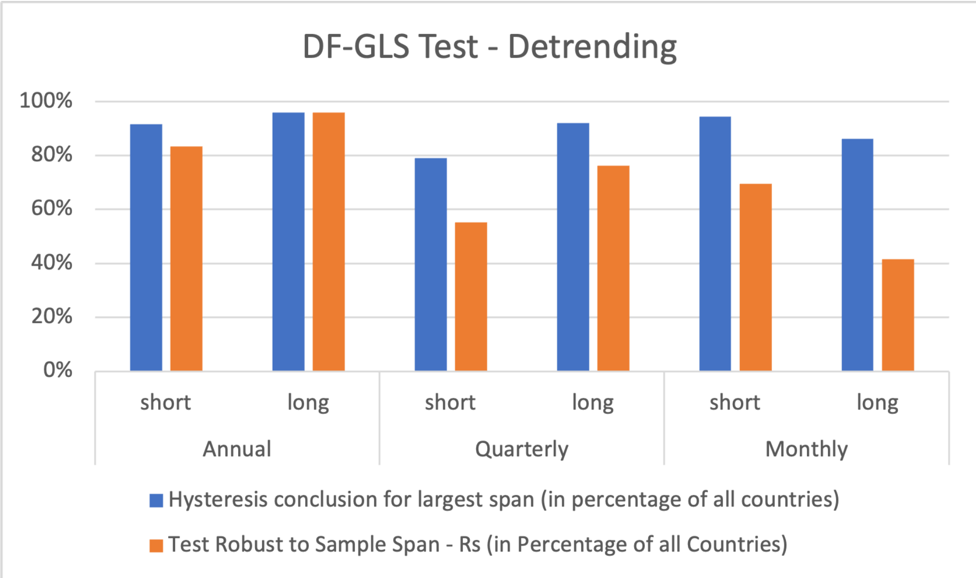

Figure 4.10: DF-GLS Test - Detrending

Similar patterns of decreasing percentage of countries where the test provides robust results to the sampling span, R_S, are present when operating with higher sample frequencies, considering the long version of lags in the detrending procedure. However, regarding the remaining cases, there is a contraction in the percentage of countries that present this robustness when moving from annual to quarterly series, followed by an enlargement in that percentage shifting from quarterly to monthly series. For instance, regarding the demeaning procedure and the short version of lags, while 50% of the countries show the same result for different sample spans in annual series, this ratio is decreased to 39% in quarterly series and increased to 53% operating with monthly series.

	Demeaning	Detrending
Not Robust	42.1%	26.3%
Not Robust in some lag versions	23.7%	13.2%
Robust, R_F	31.6%	57.9%
Not possible to compare	2.6%	2.6%

Table 4.3: DF-GLS Test - Robust Frequency Analysis (in percentage of countries)

The DF-GLS test bears the highest percentage of countries where the test delivers conclusions that are robust to the sampling frequency, R_F, when compared to the remaining univariate framework tests. As displayed in Table 4.3, in the detrending procedure, our analysis indicates that more than half of the countries deliver the same conclusion when comparing the different observations' frequencies, being robust to this sampling characteristic.

4.2.4 Zivot-Andrews test

This test allows confrontation between the hysteresis hypothesis and the structuralist thesis. Our research establishes that most countries are illustrated as unit root processes, favouring the hysteresis hypothesis, when applied to the most extensive sample spans available, as portrayed in Figures 4.11, 4.12 and 4.13.

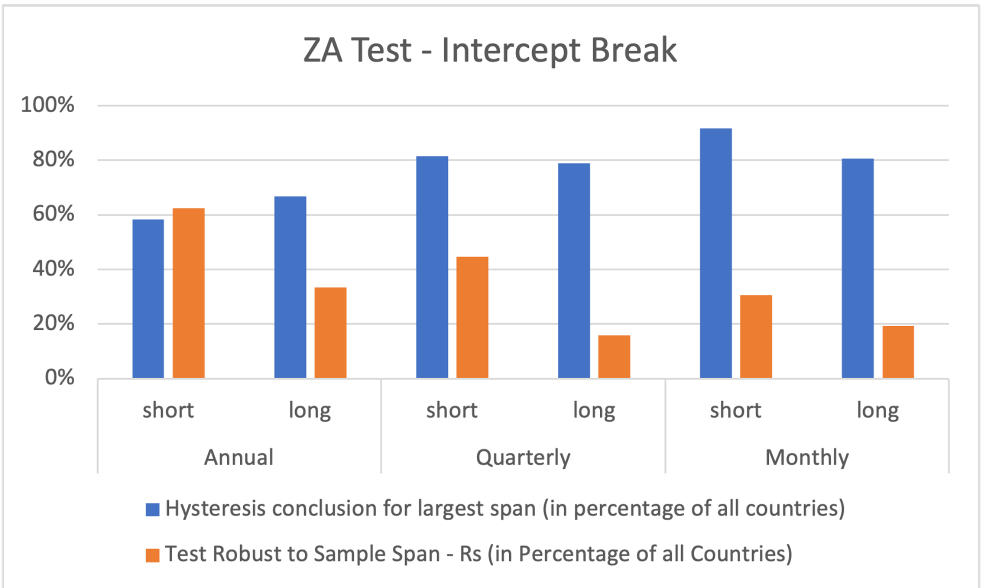

Figure 4.11: ZA Test - Intercept Break

Apart from the long version of lags, when considering a break in the intercept and trend, we can observe an increasing tendency to describe the unemployment rate as a unit root process as the test is applied to higher frequency series, as depicted in the figures. For the case described in the beginning of the paragraph, we can observe a contraction in the percentage of countries where the test's conclusion favours the hysteresis hypothesis when moving from yearly to quarterly data, from 83% to 76%, and an enlargement of that percentage when shifting from quarterly to monthly data, 76% to 86%, as displayed in Figure 4.13.

When controlling for a trend break, almost all countries have unemployment dynamic characterised by a unit root, and thus no break in trend in the DGP, describing the time series as unit-root processes in a minimum of 88% out of the 38 countries, meaning that the structuralist thesis is not appropriate to describe unemployment dynamics for this case.

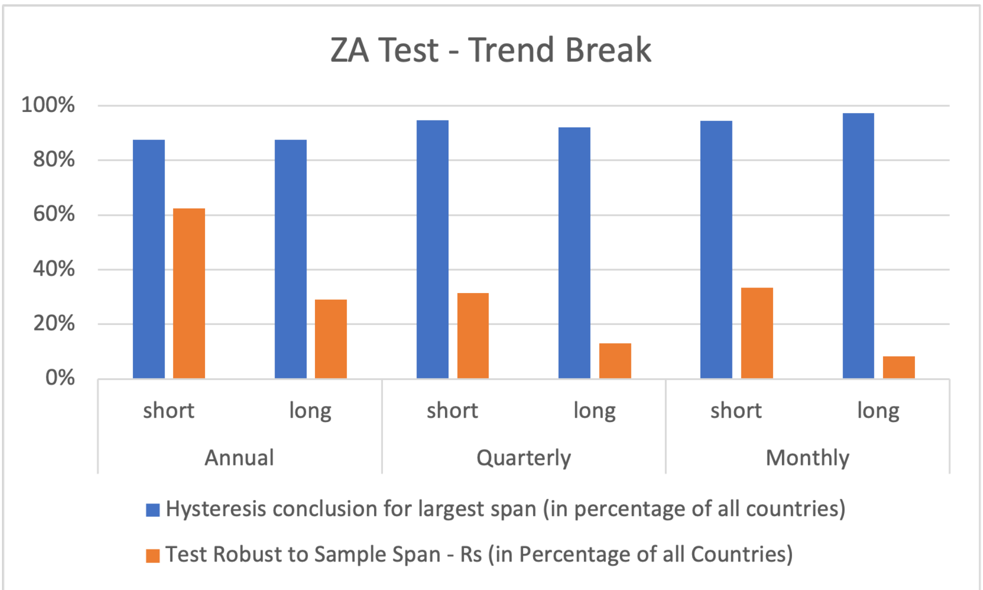

Figure 4.12: ZA Test - Trend Break

Analysing whether the tests' conclusions are robust to the sample span, R_S, our results exhibit a decrease in the percentage of countries where the tests provide robust results as we apply them to higher frequencies in the short lag's version of the intercept break and both the intercept and trend breaks cases, and in the long lag's version of the trend break case. For the remaining cases, our results show a contraction in the percentage of countries where the test is robust to the sample span, R_S, as we operate from yearly to quarterly data, followed by an increase in that percentage shifting from quarterly to monthly data.

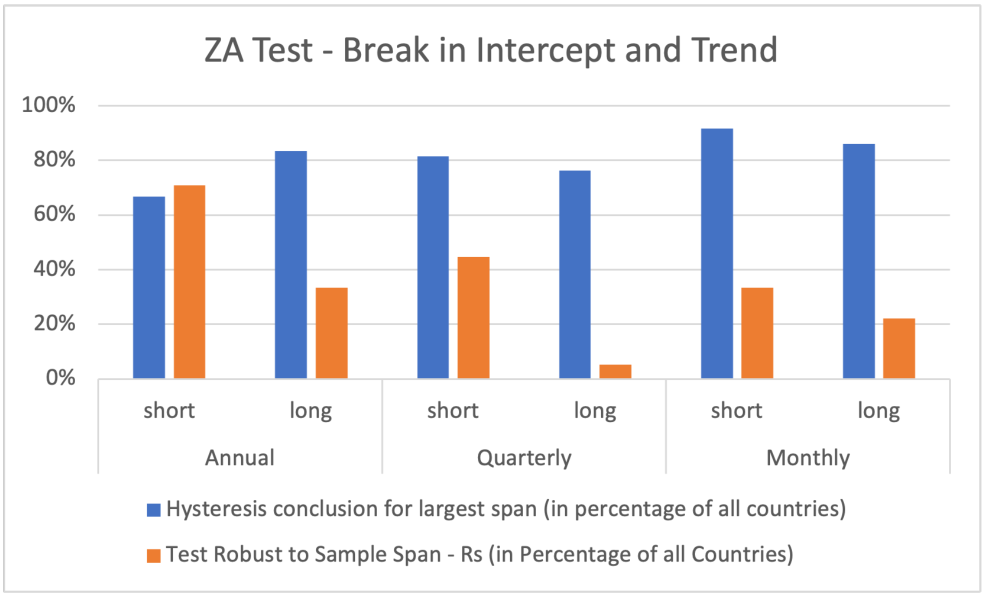

Figure 4.13: ZA Test - Break in Trend and Intercept

Regardless, and besides the short lag version operating with yearly data where the percentage of countries with robust results to the sample span is above 60%, our results depict a considerable impact of the sampling span in the results of the test, with the highest percentage of countries with robust outcomes for this sampling characteristic being 45%, expressing the necessity of regarding this sampling feature when conducting and comparing empirical investigations.

	Intercept	Trend	Intercept + Trend
Not Robust	57.9%	50%	57.9%
Not Robust in some lag versions	34.2%	34.2%	31.6%
Robust, R_F	5.3%	13.2%	7.9%
Not possible to compare	2.6%	2.6%	2.6%

Table 4.4: ZA Test - Robust Frequency Analysis (in percentage of countries)

By analysing the tests' results robustness to the sampling frequency, our results depict a significant impact of the data frequency on the test's results. As Table 4.4 displays, controlling for an intercept break in the DGP, around 58% of the countries lack results that are robust to the sample frequency, and the same result materialises when controlling for a break in both the intercept and trend. When handling for a break in the trend component, 50% of countries present this sensitivity to the sampling frequency.

4.3 Panel Environment

4.3.1 LLC Test

We start by examining the test's results for the full panel at the three observation frequencies for the different cases, displayed in Table 3.2. Table 4.5 exhibits the hypothesis supported by the tests' results, where H_0, displayed in bold, represents the null hypothesis, sustaining the hysteresis hypothesis for all individuals in the panel, and H_a characterizes the alternative hypothesis, sustaining the stationarity thesis for the unemployment dynamics to all countries.

Case	Annual			Quarterly			Monthly		
	None	Intercept	Trend	None	Intercept	Trend	None	Intercept	Trend
1	H_a	H_a	H_a	H_a	H_a	$\boldsymbol{H_0}$	H_a	H_a	$\boldsymbol{H_0}$
2	H_a	H_a	H_a	H_a	H_a	H_a	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
3	$\boldsymbol{H_0}$	H_a	H_a	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
4	$\boldsymbol{H_0}$	H_a	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
5	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
6	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	H_a	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$

Table 4.5: LLC tests' results for full panel

As can be inferred from Table 4.5, there are numerous situations where the test's results support the same hypothesis at different frequencies, providing robust results to the sampling frequency, R_F, when considering the full samples. In cases 1 and 6, we have consistent results at the different frequencies when considering no individual exogenous variables and when intercept components are incorporated, favouring the stationarity hypothesis for all individuals in the panel in case 1, and the hysteresis thesis for all countries in case 6. In case 2, we find our results not to be R_F, since we achieve different conclusions at the three frequencies, for the three exogenous variables' assumptions. In case 3, only when considering no exogenous variables do we have results that are R_F, favouring the hysteresis null hypothesis. In case 4, we find results that are R_F when no variables and when an intercept plus a time trend are considered as individual exogenous variables, where the results favour the hysteresis null hypothesis for all countries. Finally, in case 5, our results conclude in favour of the hysteresis thesis for all countries, for all observation frequencies and for all exogenous variables' assumptions.

Next we increase the time by one unit and analyse robustness of the test to the sampling span and frequency. Table 4.6 displays our conclusion regarding R_S for the different cases and exogenous variables assumptions. Table 4.7 provides a thorough insight on R_S by providing the percentage of results where the hysteresis null hypothesis is supported by the test, on the full amount of results that arise from the unity increase methodology. In that table it is highlighted in red the values where the variation of the test's results to the sample span are higher than 15%.

Case	Annual			Quarterly			Monthly		
	None	Intercept	Trend	None	Intercept	Trend	None	Intercept	Trend
1	Yes	Yes	Yes	Yes	No	No	No	No	No
2	No	Yes	Yes	No	No	No	No	No	No
3	No	No	Yes	Yes	No	No	Yes	Yes	Yes
4	Yes	No	No	Yes	No	No	Yes	No	No
5	Yes	No	No	No	No	No	No	No	No
6	Yes	No	No	No	No	No	No	No	No

Table 4.6: LLC test - R_S analysis

In annual data we have nine situations where the test's results are R_S, being this frequency the one with more results that are R_S. However, when scrutinizing this result with the ones provided in Table 4.7, we can conclude that, even though in the monthly data case we only have four situations where the test's results are R_S, this is the frequency where the test's results are more invariant to the considered sample span since the highest variation is around 30% (while in 70% of the cases, that arise from the unity increase methodology, the test favours the hysteresis null hypothesis, for the remaining it favours the stationarity alternative) whereas for annual frequency is around 50% and around 40% for the quarterly case.

Nevertheless, we should bear in mind that, even though the quarterly case has a highest variation that is lower than the annual frequency, the quarterly has more results that present a considerable divergence of results when considering different sampling spans (highlighted in red in Table 4.7).

Case	Annual			Quarterly			Monthly		
	None	Intcpt.	Trend	None	Intcpt.	Trend	None	Intcpt.	Trend
1	0%	0%	0%	0%	41.7%	25.0%	1.2%	78.6%	13.1%
2	14.3%	0%	0%	67.0%	39.0%	19.0%	79.9%	91.1%	70.1%
3	88.2%	11.8%	0%	100%	99.3%	85.7%	100%	100%	100%
4	100%	40.7%	37.0%	100%	96.7%	86.1%	100%	94.4%	94.2%
5	100%	89.2%	45.9%	91.8%	94.5%	60.5%	93.1%	99.9%	88.7%
6	100%	90.5%	52.4%	96.3%	69.6%	28.3%	99.6%	99.7%	89.8%.

Table 4.7: LLC test - R_S analysis in percentage

Table 4.8 displays our results regarding R_F. As can be inferred from the table, it is only when no variables are considered as individual exogenous variables that we achieve the same results at the different frequencies for the different sampling spans, being the results robust to the sampling frequency, R_F.

Case	None	Intercept	Trend
1	Yes	No	No
2	No	No	No
3	No	No	No
4	Yes	No	No
5	Yes	No	No
6	Yes	No	No

Table 4.8: LLC Test - R_F Analysis

4.3.2 IPS Test

Analysing the test's results for the full panel in each case (1 to 6) presented in Table 3.2, our results show that, when applied to annual data, the test provides results supporting the alternative hypothesis, where at least one country has unemployment dynamics characterized by a stationary process. Apart from case 1, and as displayed in Table 4.9, we can observe results that are R_F when an intercept is included as individual exogenous variable, and results that are not R_F when a trend is added as exogenous component (whereas for annual data our results support the alternative hypothesis, for the remaining frequencies it supports the null hypothesis - unemployment dynamics characterized by unit-root processes for all individual in the panel).

Case	Annual		Quarterly		Monthly	
	Intercept	Trend	Intercept	Trend	Intercept	Trend
1	H_a	H_a	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
2	H_a	H_a	H_a	H_a	H_a	$\boldsymbol{H_0}$
3	H_a	H_a	H_a	$\boldsymbol{H_0}$	H_a	$\boldsymbol{H_0}$
4	H_a	H_a	H_a	$\boldsymbol{H_0}$	H_a	$\boldsymbol{H_0}$
5	H_a	H_a	H_a	$\boldsymbol{H_0}$	H_a	$\boldsymbol{H_0}$
6	H_a	H_a	H_a	$\boldsymbol{H_0}$	H_a	$\boldsymbol{H_0}$

Table 4.9: IPS Test's results for full panel

Applying the unity increase methodology we can analyse the robustness of the test's results to the sampling span. Table 4.10 exhibits our conclusion regarding R_S for the different cases, exogenous variables' assumptions and observation frequencies. As can be inferred from the table, when applied to annual data the test provides the highest number of situations where the results are R_S, followed by the monthly data case and the quarterly one.

Case	Annual		Quarterly		Monthly	
	Intercept	Trend	Intercept	Trend	Intercept	Trend
1	Yes	Yes	No	No	No	No
2	Yes	Yes	No	No	No	No
3	No	Yes	No	No	No	Yes
4	No	No	No	No	No	No
5	No	No	No	No	No	No
6	No	No	No	No	No	Yes

Table 4.10: IPS Test - R_S Analysis

Nevertheless, when we perform a more thoroughly R_S analysis, an interesting result arises. Table 4.11 presents the percentage of results where the tests favour the hysteresis null hypothesis on the full amount of results that arise from the unity increase methodology, for each case and exogenous variables' assumptions. Even though a unique case presents R_S results when applying the test to monthly data, this is the frequency that presents the highest results stability regarding the sampling span, since the highest variation of the results is around 25% - whereas for around 78% of the results support the hysteresis null hypothesis, for the remaining it supports the alternative one. When applying the test to annual data, our results present a high sensitivity to the sampling span when they are not R_S, since all of them present variation that is higher than 15% (highlighted in red in Table 4.11).

Case	Annual		Quarterly		Monthly	
	Intercept	Trend	Intercept	Trend	Intercept	Trend
1	0%	0%	33.3%	33.3%	95.2%	13.1%
2	0%	0%	44.0%	28.0%	77.9%	98.0%
3	17.6%	0%	77.9%	86.4%	98.9%	100%
4	44.4%	37.0%	81.1%	93.3%	87.6%	95.1%
5	73.0%	32.4%	88.6%	77.7%	95.6%	86.9%
6	66.7%	57.1%	63.3%	59.2%	91.9%	100%

Table 4.11: IPS Test, R_S analysis in percentage

Table 4.12 exhibits our results regarding R_F when applying the unity increase method. As can be inferred from the table, in all cases and exogenous variables assumptions, our results are not robust to the sampling frequency R_F since for at least a sample span our results deliver different conclusions at different frequencies.

Case	Intercept	Trend
1	No	No
2	No	No
3	No	No
4	No	No
5	No	No
6	No	No

Table 4.12: IPS Test - R_F analysis

4.3.3 Fisher's Type Tests

We start by applying the Fisher's type tests to the full panel, considering the different individuals exogenous variables specifications and the three observations' frequencies. Table 4.13 exhibits the hypothesis supported by the tests' results, where H_0, exhibited in bold, represents the null hypothesis, sustaining hysteresis hypothesis for all individuals in the panel, and H_a characterizes the alternative hypothesis where at least one individual in the panel has a unit-root process in the DGP.

Frequency	Exogenous Variables	Test			
		MW	Choi - P	Choi - Inv Normal	Choi - Logit
Annual	None	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
	Intercept	H_a	H_a	H_a	H_a
	Intercept + Trend	H_a	H_a	H_a	H_a
Quarterly	None	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
	Intercept	H_a	H_a	H_a	H_a
	Intercept + Trend	H_a	H_a	H_a	H_a
Monthly	None	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$
	Intercept	H_a	H_a	H_a	H_a
	Intercept + Trend	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$	$\boldsymbol{H_0}$

Table 4.13: Fisher's Type tests results for full panel

With this method, two results arise. Firstly, the different tests provide the same conclusion when considering both the different frequencies and individual exogenous variables that are included in the ADF regressions while running the tests. This depicts results' consistency between the different Fisher's type tests. Secondly, we achieve results that are mostly robust to the frequency of observation, R_F, since for a given formulation of the exogenous variables, we obtain congruous results at the different frequencies. However, when considering individual intercepts and time trends as exogenous variables, our results find contradictory conclusions when comparing the results provided by both the annual and quarterly frequencies with the monthly frequency case (whereas using annual and quarterly data our outcomes favour the alternative hypothesis, applying the tests to monthly data our results favour the null hypothesis).

Employing the unity increase methodology to these panel unit root tests, we can assess if the outcomes are robust to the sample span, R_S. Table 4.14 evidences the percentage of cases where the tests' results favour the null hypothesis, considering the different panels that arise from increasing the span of each individual per iteration. As Table 4.14 depicts, when utilising annual data to implement the tests, we consistently conclude for the hysteresis thesis for all individuals in the panel when no exogenous variables are included in the ADF regressions, and we achieve invariable results favouring the alternative hypothesis for the remaining exogenous variables' options. Therefore, we can infer that out results are robust to the sample span, R_S, when operating with annual data, independently of the test and individuals exogenous variables we consider.

Frequency	Exogenous Variables	Test			
		MW	P test	Inv. Normal	Logit
Annual	None	100%	100%	100%	100%
	Intercept	0%	0%	0%	0%
	Intercept + Trend	0%	0%	0%	0%
Quarterly	None	94.2%	94.2%	93.3%	94.1%
	Intercept	0.4%	0.4%	5.4%	4.6%
	Intercept + Trend	7.9%	7.9%	20.0%	31.4%
Monthly	None	90.4%	90.2%	99.1%	99.1%
	Intercept	63.0%	55.9%	53.5%	55.1%
	Intercept + Trend	97.1%	96.9%	99.9%	99.9%

Table 4.14: Fisher's Type tests - R_S Analysis

Performing the same analysis, but utilising quarterly and monthly data, we can conclude that our results are not robust to the sample span, R_S. However, when considering the quarterly data case, we can infer that, apart from the inverse normal and logit tests when considering an intercept plus a time trend as individuals exogenous variables, our results are to a certain extent consistent to the different sample spans that constitute the panels, since more than 90% of the results favour the null hypothesis when no exogenous variables are considered, and more than 90% favours the alternative hypothesis for the remaining cases. When employing the tests to monthly data we achieve similar conclusions since our results favour the null hypothesis for more than 90% of the cases when considering no exogenous variables and an intercept plus a time trend as individuals exogenous components. When we consider an intercept as individuals exogenous variables, our results depict the highest variability with more than 50% of the tests results providing evidence favouring the null hypothesis when applied to the different sample spans that constitute the panels.

4.3.4 Hadri Test

Opposing the remaining panel unit root tests under analysis, the Hadri test takes the stationarity as null hypothesis, where all individuals in the panel present stationary processes in the DGP, and the hysteresis as the alternative one, where at least one individual in the panel has unemployment dynamics characterized by unit-root process.

The test requires balanced data, thus the methodology applied to perform the analysis is the one that corresponds to this characteristic. It considers two possibilities for the individual exogenous variables - intercept and intercept plus a time trend.

By applying our methodology, we consistently achieve the same result, regardless of the different cases (Table 3.2) and exogenous variables assumptions, rejecting the stationarity null hypothesis for all cases, even when assessing the unity increase methodology.

Therefore, we can conclude that the Hadri test is the only test under analysis that provides results that are consistently robust to the sampling span, R_S, and sampling frequency, R_F. For this reason, we have decided not to include any visual or graphical visualisation of the results. However, one can find part of the R's Software output for case 1, where the test is applied to monthly data and only an intercept is considered as individual exogenous variable, in Annex B.

Chapter 5

Conclusions

One of the key points in the study of the dynamics of an economic time series is the presence of units roots assessed by unit root tests. In the case of unemployment the presence of unit roots is linked to three different hyothesis concerning its dynamics. The first is hysteresis which states that the dynamics of unemployment are driven by the presence of a unit root and therefore external shocks will endure. Another hypothesis, NAIRU, states that unemployment is a stationary process, conveying it is mean-reverting and thus shocks only temporarily affect unemployment. Finally the structuralism hypothesis, controls for structural breaks and thus if unit-roots are absent, unemployment exhibits permanent changes in the equilibrium rate in the presence of occasional but permanent structural changes.

Empirical investigations concerning the presence of unit roots in unemployment time series achieve conflicting conclusions when resorting to identical unit-root procedures applied to data with distinct sampling characteristics namely, span and observation frequency. Thus, our research provides insights on the impact of these sampling characteristics in the unit root tests by studying the unemployment dynamics of thirty-eight OECD countries, considering increasing sample spans and three sampling frequencies both in univariate and panel contexts.

In the univariate context, we start by applying the univariate unit-root tests to the largest span available to each country. Our results show that the tests are not robust to the sampling frequency. KPSS is the test whose results vary the most with sampling frequency. On the other hand, the ZA test with a break in the trend component is the most robust to sampling frequency.

The impact of the sampling span on the tests is assessed by considering rolling samples, i.e., samples that increase the size by one time unit. We can observe a disparity in the percentage of countries for which the tests' results are robust to the sampling span at the different observation frequencies. We note that the annual frequency is the one with the highest number of countries with results robust to the sampling span for the majority of the tests.

When comparing the tests' results at different frequencies for a given sampling span, we can assess whether the results are robust to the sampling frequency. The KPSS is the test with the highest percentage of countries whose results that are not robust to the sampling frequency, followed by the ADF, the ZA and the DF-GLS test.

Therefore, for all the univariate unit-root tests, we have countries where the tests display sensitive results to the sampling characteristics, explaining the conflicting results of empirical investigations in the literature, and reinforcing the need to consider these factors when studying the hysteresis hypothesis in unemployment analyses.

Under the panel environment, we employ different methods for panel unit-root tests that require balanced and unbalanced data.

The Fisher's type tests require unbalanced data. Therefore we resort to a different methodology than the remaining panel unit-root tests. The tests provide invariant conclusions considering identical sampling frequencies and exogenous variables' assumptions being consistent between them. Apart from incorporating an intercept and a time trend as exogenous variables, the tests deliver robust results to the sampling frequency.

Applying the unity increase methodology and operating with annual data, the Fisher's type tests deliver robust results to the sampling span. The majority of our tests' results provide the same conclusion with this approach working with quarterly data, the inverse normal and logit tests being the outliers. With monthly data, considering an intercept as the exogenous variable is the only case with results that deliver a considerable variation.

Therefore, we can conclude that by performing unemployment dynamics analysis using these panel unit-root tests, we can expect that the sampling characteristics have a lower impact compared with the univariate environment. However, it facilitates an explanation of contradictory conclusions present in the literature and reinforces the influence of sampling features in the analysis.

Operating with the panel tests requiring balanced data panels, we perform a more thorough analysis. The LLC and IPS tests deliver results that highly depend on the panels' spans and data frequency. No pattern emerges since our conclusions on sampling and frequency robustness depend on the chosen panel spans, reinforcing the impact of the sampling span in conducting unit-root analysis.

It must be noted that the Hadri test delivers results that are invariant to sampling features.

Our research contributes to the literature by employing unit-root tests to a larger dataset than empirical investigations in the literature. Our central focus is on understanding the influence of sampling features in conducting unemployment dynamics analysis.

Even though we thoroughly investigate the impact of sampling features in unit-root tests conclusions, we deal with software limitations. The software does not allow for certain tests' specifications and lacks credible implementations of more advanced unit-root tests.

We consider that similar methods should be employed to analyse sampling characteristics' influence in more recent unit-root tests, to assess their conclusions' sensitivity to those features.

Bibliography

Afriyie, J., Twumasi-Ankrah, S., Gyamfi, K., Arthur, D., & Pels, W. (2020, 12). Evaluating the performance of unit root tests in single time series processes. *Mathematics and Statistics*, *8*, 656-664.

Arltová, M., & Fedorová, D. (2016, 09). Selection of unit root test on the basis of length of the time series and value of ar(1) parameter. *Statistika - Statistics and Economics Journal*, *96*, 47-64.

Arsova, A., & Karaman Örsal, D. D. (2021). A panel cointegrating rank test with structural breaks and cross-sectional dependence. *Econometrics and Statistics*, *17*, 107-129.

Ayala, A., Cuñado, J., & Gil-Alana, L. A. (2012, November). Unemployment hysteresis: empirical evidence for Latin America. *Journal of Applied Economics*, *15*, 213-233.

Bai, J., & Ng, S. (2004). A panic attack on unit roots and cointegration. *Econometrica*, *72*(4), 1127-1177.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, *18*(1), 1-22.

Bakas, D., & Makhlouf, Y. (2019, 02). Can the insider–outsider theory explain unemployment hysteresis in OECD countries? *Oxford Economic Papers*, *72*(1), 149-163.

Banerjee, A. (1999). Panel data unit roots and cointegration: An overview. *Oxford Bulletin of Economics and Statistics*, *61*(S1), 607-629.

Banerjee, A., & Carrion-i Silvestre, J. L. (2015). Cointegration in panel data with structural breaks and cross-section dependence. *Journal of Applied Econometrics*, *30*(1), 1-23.

Blanchard, O., & Summers, L. (1986). Hysteresis and the european unemployment problem. *NBER Macroeconomics Annual*, 15-78.

Blanchard, O., & Summers, L. (1987). Hysteresis in unemployment. *European Economic Review*, *31*(1-2), 288-295.

Bolat, S., Tiwari, A. K., & Erdayi, A. U. (2014). Unemployment hysteresis in the eurozone area: evidences from nonlinear heterogeneous panel unit root test. *Applied Economics Letters*, *21*(8), 536-540.

Camarero, M., Carrion-i Silvestre, J. L., & Tamarit, C. (2006). Testing for hysteresis in unemployment in oecd countries: New evidence using stationarity panel tests with breaks*.

Oxford Bulletin of Economics and Statistics, *68*(2), 167-182.

Camarero, M., & Tamarit, C. (2004). Hysteresis vs. natural rate of unemployment: new evidence for {OECD} countries. *Economics Letters*, *84*(3), 413-417.

Chang, T. (2011). Hysteresis in unemployment for 17 oecd countries: Stationary test with a fourier function. *Economic Modelling*, *28*(5), 2208-2214.

Chang, T., Lee, K.-C., Nieh, C.-C., & Wei, C.-C. (2005). An empirical note on testing hysteresis in unemployment for ten european countries: panel suradf approach. *Applied Economics Letters*, *12*(14), 881-886.

Choi, I. (1994). Residual-based tests for the null of stationarity with applications to u.s. macroeconomic time series. *Econometric Theory*, *10*(3/4), 720–746.

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance*, *20*(2), 249-272.

Choi, I. (2006). *Combination unit root tests for cross-sectionally correlated panels*. Cambridge University Press.

Croissant, Y., & Millo, G. (2018). *Panel data econometrics with R*. Wiley.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*(366), 427–431.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, *49*(4), 1057–1072.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, *64*(4), 813–836.

Ener, M., & Arica, F. (2011). Is there hysteresis in unemployment in oecd countries? evidence from panel unit root test with structural breaks. *Chinese Business Review*, *10*(4), 294-304.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276.

Franchi, M., & Ordóñez, J. (2008). Common smooth transition trend-stationarity in european unemployment. *Economics Letters*, *101*(2), 106-109.

Friedman, M. (1968). The role of monetary policy. *The American Economic Review*, *58*(1), 1–17.

Furuoka, F. (2017, 11). A new approach to testing unemployment hysteresis. *Empirical Economics*, *53*.

Fève, P., Hénin, P., & Jolivaldt, P. (2003). Testing for hysteresis: Unemployment persistence and wage adjustment. *Empirical Economics*, *28*(3), 535-552.

Hadri, K. (2002, 03). Testing for stationarity in heterogeneous panel data. *The Econometrics Journal*, *3*(2), 148-161.

Hlouskova, J., & Wagner, M. (2005, 02). The performance of panel unit root and stationarity tests: Results from a large scale simulation study. *Econometric Reviews*, *25*.

Im, K.-S., Lee, J., & Tieslau, M. (2005). Panel lm unit-root tests with level shifts*. *Oxford*

Bulletin of Economics and Statistics, *67*(3), 393-419.

Im, K. S., Pesaran, M., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, *115*(1), 53-74.

Khraief, N., Shahbaz, M., Heshmati, A., & Azam, M. (2020). Are unemployment rates in oecd countries stationary? evidence from univariate and panel unit root tests. *The North American Journal of Economics and Finance*, *51*, 100838.

Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, *54*(1-3), 159-178.

Lee, C.-C., & Chang, C.-P. (2008). Unemployment hysteresis in oecd countries: Centurial time series evidence with structural breaks. *Economic Modelling*, *25*(2), 312-325.

Lee, H.-Y., Wu, J.-L., & Lin, C.-H. (2010). Hysteresis in east asian unemployment. *Applied Economics*, *42*(7), 887-898.

Lee, J., & Strazicich, M. (2003). Minimum lagrange multiplier unit root test with two structural breaks. *The Review of Economics and Statistics*, *85*(4), 1082-1089.

Lee, J., & Strazicich, M. (2004, 02). Minimum lm unit root test with one structural break. *Economics Bulletin*, *33*.

Lee, J.-D., Lee, C.-C., & Chang, C.-P. (2009). Hysteresis in unemployment revisited: Evidence from panel lm unit root tests with heterogeneous structural breaks. *Bulletin of Economic Research*, *61*(4), 325-334.

Levin, A., Lin, C.-F., & James Chu, C.-S. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, *108*(1), 1-24.

Maddala, G. S., & Kim, I.-M. (1999). *Unit Roots, Cointegration, and Structural Change* (No. 9780521587822). Cambridge University Press.

Maddala, G. S., & Wu, S. (1999). A comparative study of unit root tests with panel data and a new simple test. *Oxford Bulletin of Economics and Statistics*, *61*(S1), 631-652.

Mitchell, W. F. (1993). Testing for unit roots and persistence in oecd unemployment rates. *Applied Economics*, *25*(12), 1489-1501.

Moon, H. R., & Perron, B. (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics*, *122*(1), 81-126.

Papell, D., Murray, C., & Ghiblawi, H. (2000). The structure of unemployment. *Review of Economics and Statistics*, *82*(2), 309-315.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, *57*(6), 1361-1401.

Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business & Economic Statistics*, *8*(2), 153–162.

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables.

Journal of Econometrics, *80*(2), 355-385.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265-312.

Pfaff, B. (2008). *Analysis of integrated and cointegrated time series with r* (Second ed.). New York: Springer.

Phelps, E. S. (1967). Phillips curves, expectations of inflation and optimal unemployment over time. *Economica*, *34*(135), 254–281.

Phelps, E. S. (1972). *Inflation policy and unemployment theory; the cost-benefit approach to monetary planning* [Book]. Norton New York.

Phelps, E. S. (1994). *Structural slumps : the modern equilibrium theory of unemployment, interest, and assets* [Book]. Cambridge, Mass.: Harvard University Press.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, *75*(2), 335–346.

Pissarides, C. (1992). Loss of skill during unemployment and the persistence of employment shocks. *The Quarterly Journal of Economics*, *107*(4), 1371-1391.

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Røed, K. (1996). Unemployment hysteresis - macro evidence from 16 oecd countries. *Empirical Economics*, *21*(4), 589-600.

Røed, K. (2002). Unemployment hysteresis and the natural rate of vacancies. *Empirical Economics*, *27*(4), 687-704.

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*(3), 599–607.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464.

Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. *Journal of Business & Economic Statistics*, *7*(2), 147–159.

Shiller, R. J., & Perron, P. (1985). Testing the random walk hypothesis: Power versus frequency of observation. *Economics Letters*, *18*(4), 381-386.

Song, F. M., & Wu, Y. (1998). Hysteresis in unemployment: Evidence from oecd countries. *The Quarterly Review of Economics and Finance*, *38*(2), 181-192.

Westerlund, J., & Edgerton, D. L. (2008). A simple test for cointegration in dependent panels with structural breaks*. *Oxford Bulletin of Economics and Statistics*, *70*(5), 665-704.

Zivot, E., & Andrews, D. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, *10*(3), 251-70.

Annex

A USA - Illustrative example

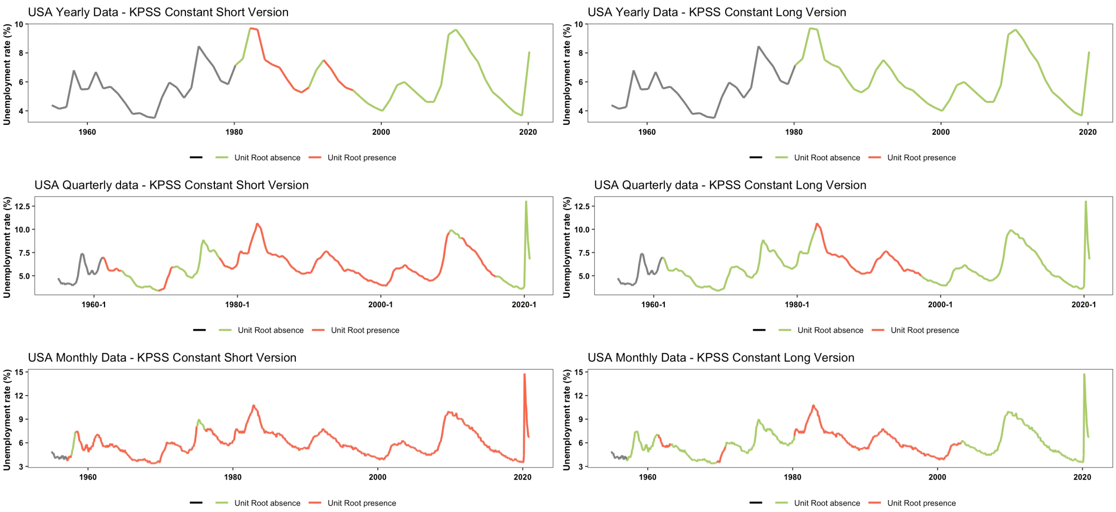

Figure 5.1: KPSS Test: Constant - USA

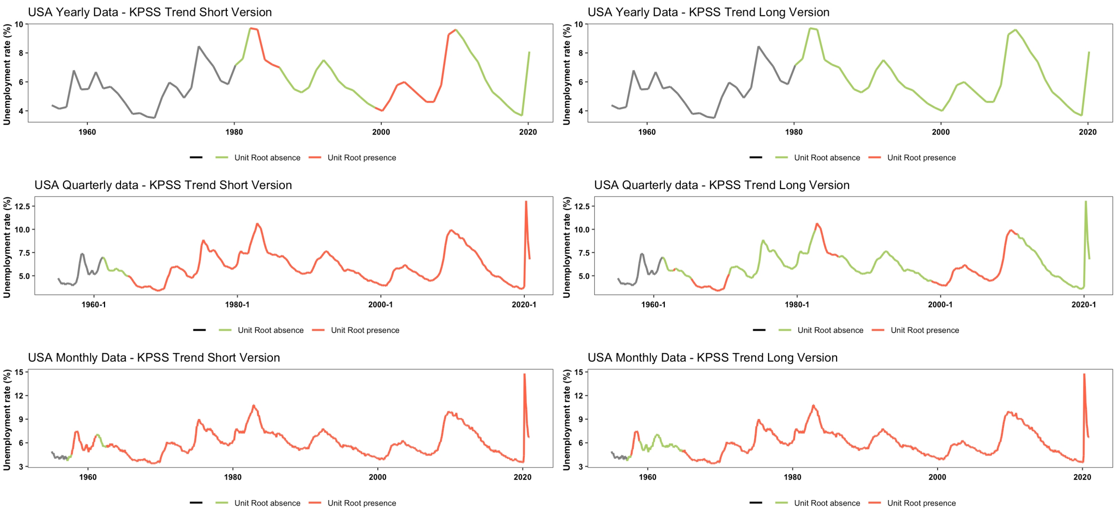

Figure 5.2: KPSS Test: Constant + Trend - USA

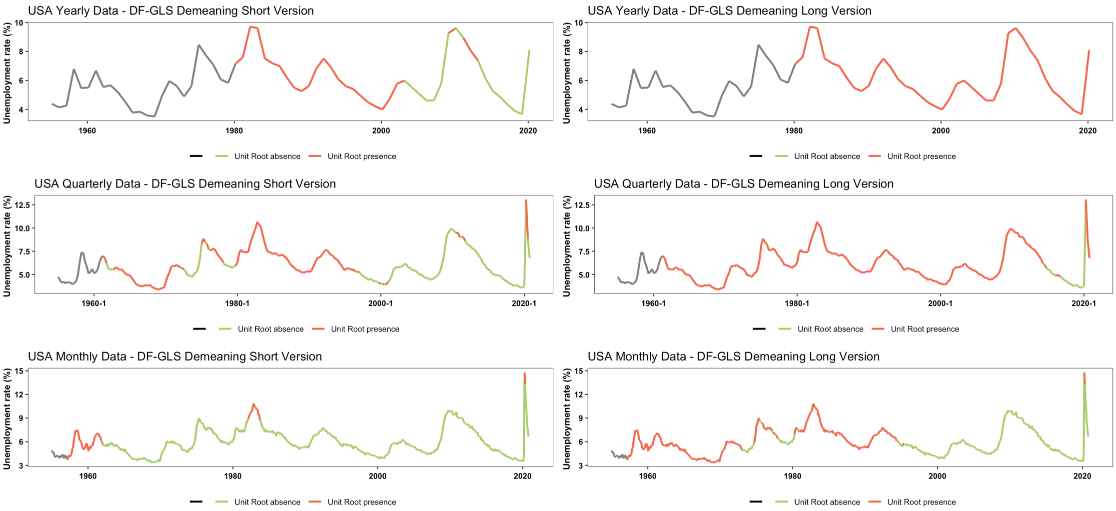

Figure 5.3: DF-GLS Test: Demeaning - USA

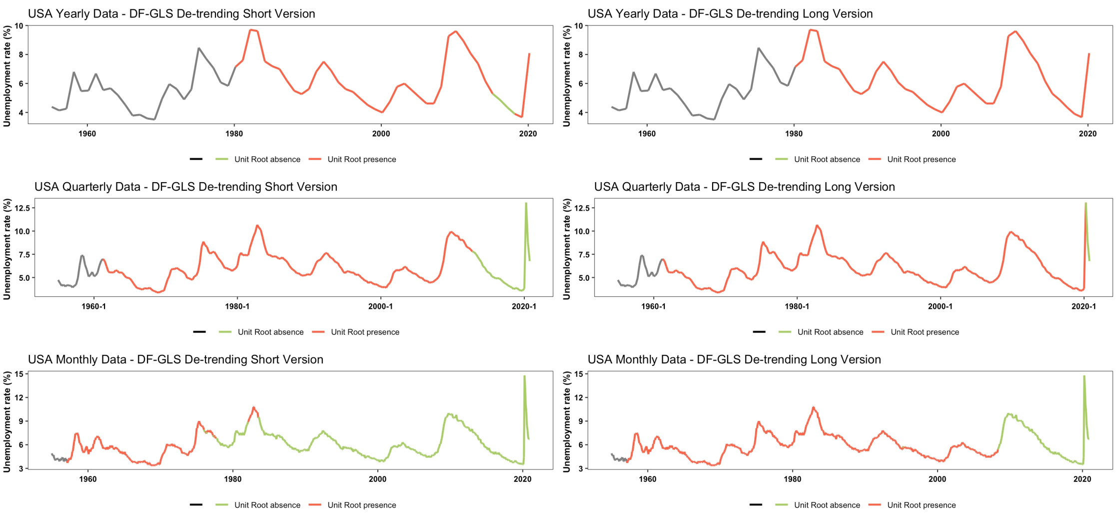

Figure 5.4: DF-GLS Test: Detrending - USA

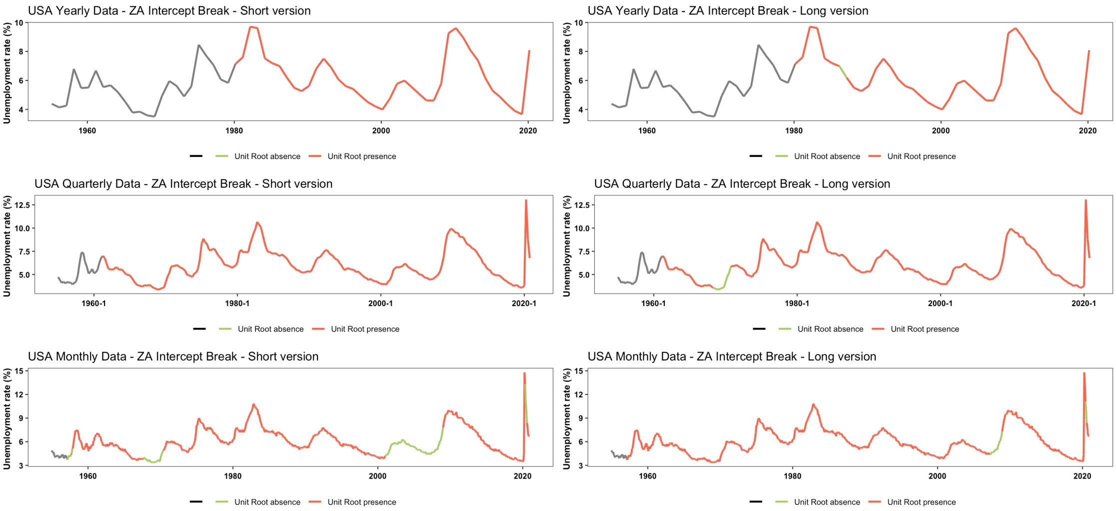

Figure 5.5: ZA Test: Break in intercept - USA

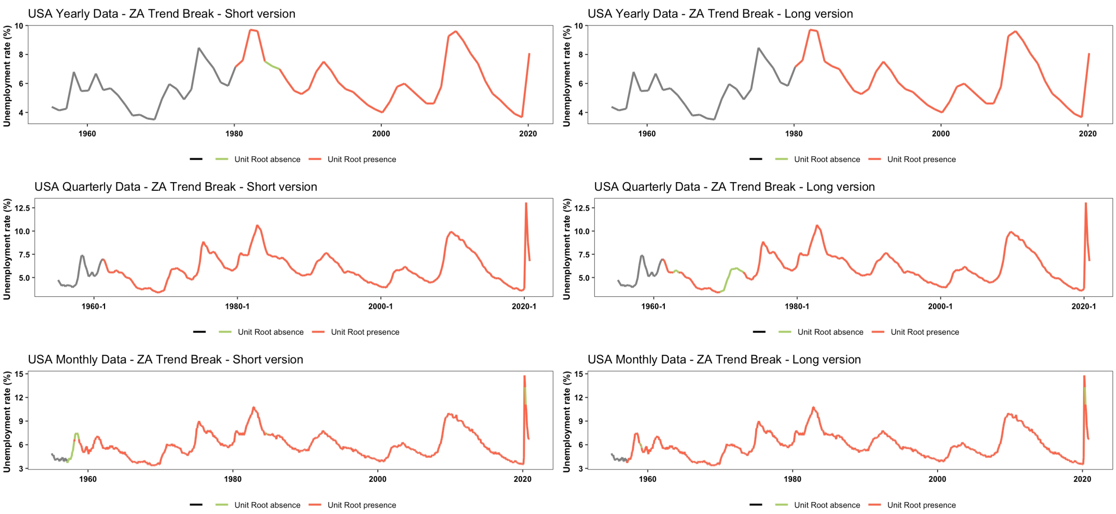

Figure 5.6: ZA Test: Break in trend - USA

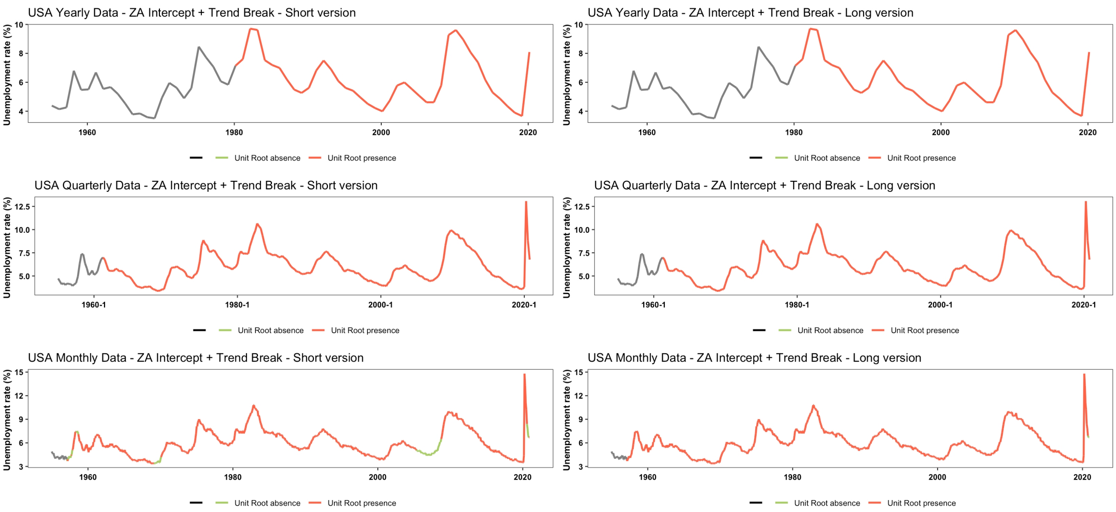

Figure 5.7: ZA Test: Break in intercept and trend - USA

B Hadri Test - R's Output

Max.Size	Test	Statistic	p.value	alternative
25	Hadri Test (ex. var.: Individual Intercepts)	73.7268	4.00E-04	at least one series has a unit root
26	Hadri Test (ex. var.: Individual Intercepts)	76.0917	6.00E-04	at least one series has a unit root
27	Hadri Test (ex. var.: Individual Intercepts)	78.4957	8.00E-04	at least one series has a unit root
28	Hadri Test (ex. var.: Individual Intercepts)	80.9488	9.00E-04	at least one series has a unit root
29	Hadri Test (ex. var.: Individual Intercepts)	83.6102	7.00E-04	at least one series has a unit root
30	Hadri Test (ex. var.: Individual Intercepts)	85.8232	4.00E-04	at least one series has a unit root
31	Hadri Test (ex. var.: Individual Intercepts)	88.1101	8.00E-04	at least one series has a unit root
32	Hadri Test (ex. var.: Individual Intercepts)	90.4919	7.00E-04	at least one series has a unit root
33	Hadri Test (ex. var.: Individual Intercepts)	93.1356	5.00E-04	at least one series has a unit root
34	Hadri Test (ex. var.: Individual Intercepts)	96.3808	2.00E-04	at least one series has a unit root
35	Hadri Test (ex. var.: Individual Intercepts)	99.8853	5.00E-04	at least one series has a unit root
36	Hadri Test (ex. var.: Individual Intercepts)	103.5237	9.00E-04	at least one series has a unit root
37	Hadri Test (ex. var.: Individual Intercepts)	107.2191	1.00E-04	at least one series has a unit root
38	Hadri Test (ex. var.: Individual Intercepts)	111.0256	5.00E-04	at least one series has a unit root
39	Hadri Test (ex. var.: Individual Intercepts)	115.0175	9.00E-04	at least one series has a unit root
40	Hadri Test (ex. var.: Individual Intercepts)	118.8174	9.00E-04	at least one series has a unit root
41	Hadri Test (ex. var.: Individual Intercepts)	122.5239	2.00E-04	at least one series has a unit root
42	Hadri Test (ex. var.: Individual Intercepts)	126.4254	1.00E-04	at least one series has a unit root

Table 5.1: Hadri test - example