Lund University
Department of Statistics

Bachelor thesis, March 2012
Supervisor: Björn Holmquist

IS THERE HYSTERESIS IN UNEMPLOYMENT RATES?

AN ARFIMA APPROACH

AUTHOR: NICLAS LAVESSON

Abstract

This paper investigates if there is hysteresis in unemployment rates for 8 OECD countries. Univariate unit root tests (ADF and KPSS tests) are performed as an initial analysis of the unemployment rates. These tests suggest that the hysteresis hypothesis is true.

Since traditional unit root test have low power against fractionally integrated alternatives, the Sowell's (1992) exact maximum likelihood (EML) ARFIMA estimator is used in the paper as well. The main advantage of using this estimator is since a quantitative measure of the memory parameters is given. As a consequence, tests of the estimated memory parameters can be conducted and a more refined analysis is possible. When the EML estimator is used, the result found by the ADF and KPSS tests are strengthened; there is hysteresis in unemployment rates.

Table of contents

1. Introduction

Unemployment is a key macroeconomic variable and works as an economic indicator for governments and policymakers (Clemente et al., 2004). The design of unemployment insurances and wage negotiations between labor unions and employees are only some important issues that are affected by the unemployment rate (Meyer, 1990; Lemieux and MacLeod, 2000). Furthermore, there is a connection between low unemployment rates and the well-being of individuals in a country or region (Clark and Oswald, 1994). Since economic policy decisions and several other aspects are highly dependent on the characteristics of the unemployment series, there is a further need to investigate the time series properties of unemployment rates.

The aim of this paper is therefore to investigate if there is hysteresis in unemployment rates. The hysteresis hypothesis states that unemployment rates follow a random walk (i.e. at least one unit root is present in the series); (Blanchard and Summers, 1986). If unemployment rates follow a random walk or have unit roots, the series will not return to its initial mean value after a shock or disturbance in the long run. In other words, the long run equilibrium is affected once unemployment rates are subject to any distortions (such as recessions or other crises). Thus, in the long run, governments and policymakers need to take actions to adjust the unemployment rates to a "preferred" level. Altogether, the hysteresis hypothesis is concerned with the influence of historic unemployment rates on the long run equilibrium in unemployment (Gil-Alana, 2001).

The alternative to hysteresis is the natural rate of unemployment. According to this view, unemployment series do not follow a random walk or have unit roots. Unemployment rates are rather referred to as being mean-reverting since these series revert back to a certain mean value after a shock (see Baillie, 1996). The mean-reverting property also implies that there exists a definite equilibrium state of unemployment in the long run (Friedman, 1968; Phelps, 1967). Deviations, such as when the actual unemployment rate differs from the long run equilibrium state, are only temporary and fade away later in time (Friedman, 1968; Phelps, 1967); it is only a question of time and depends on the degree of persistence/integration in the series. Hence, unemployment series are unaffected of any shock or disturbance in the long run (Friedman, 1968; Phelps, 1967). Moreover, when the natural rate of unemployment applies, governments need not to act in order to lower the unemployment in the long run.

The distinction between the hysteresis hypothesis and the natural rate of unemployment is important from a practical point of view. For example, assume that unemployment rates incorrectly are expected to follow a natural rate of unemployment when the hysteresis hypothesis in fact is true. Since policymakers expect the series to return to its initial equilibrium, no policy actions are taken even though such are needed. As a consequence, a new equilibrium level is reached and possibly high unemployment rates become a problem in the long run. Thus, the negative effects of unemployment are permanent (Blanchard and Summers, 1986).

Some credibility to the hysteresis hypothesis is given in earlier literature. A vast amount of studies show that unemployment rates exhibit random walk (see e.g. Blanchard and Summers, 1986; Candelon et al., 2009; Gil-Alana, 1999; Holl and Kunst, 2011; Koustas and Veloce, 1996; León-Ledesma, 2002; Lin et al., 2008; Mitchell, 1993; Roed, 1996; Yilanci 2009). Gil-Alana (1999) investigates the time series characteristics of the unemployment rate in the United Kingdom. In his paper, the main result is that the UK unemployment exhibit random walk and some credibility is therefore given to the hysteresis hypothesis. Holl and Kunst (2011) find a similar result when investigating unemployment rates in 13 OECD countries.

There is also a profound body of literature that supports the idea of a natural rate of unemployment. Gustavsson and Österholm (2006) conclude that the natural rate of unemployment hypothesis is righteously favored over the hysteresis hypothesis. Camarero et al. (2006) find a similar result when investigating the unemployment rates in 19 OECD countries. Furthermore, numerous papers advocate that unemployment rates preferably are described by the natural rate of unemployment (see e.g. Lee, 2010; Lee and Chang, 2008).

Seemingly, the result differs between papers and there is no consensus whether there is hysteresis in unemployment rates. Some papers provide somewhat mixed results. Yilanci (2008) find that the hysteresis hypothesis is supported for ten out of 17 OECD countries. For the rest of the countries, unemployment is better described by the natural rate of unemployment hypothesis. In a comparative study, Koustas and Veloce (1996) find that the Canadian unemployment rate is more persistent than the unemployment rate in the United States. The main finding is however that there likely is hysteresis in the Canadian unemployment rate but not in the US unemployment rate. Moreover, Romero-Ávila and Usabiaga (2009) reveal that the occurrence of hysteresis is a European phenomenon. The hysteresis phenomenon is for instance not seen in the United States. In another study, Romero-Ávila and Usabiaga (2007) found that there is hysteresis in Spanish unemployment rates but not in the US unemployment rate.

In this paper, the main finding is that there likely is hysteresis in unemployment rates, hence supporting what several other studies have found. The hysteresis finding does not hold for all countries in the study; for instance, the analysis shows that the US unemployment rate likely not is characterized by hysteresis. This finding is valid at least before the turbulent 1990s. There are also some indications that the degree of persistence is higher after 1990 in most of the OECD countries in the study.

Since hysteresis in unemployment is consistent with a unit root series, traditional unit root tests are used when the hysteresis hypothesis is tested (see e.g. Roed, 1996). Several of the papers referred to above rely on traditional unit root test such as the test of Dickey and Fuller (1979) or Kwiatkowski-Phillips-Schmidt-Shin (1992); henceforth called the ADF- and KPSS-test, respectively. However, when such unit root tests are used, the time series is modeled either as a stationary process or as a non-stationary unit root process (i.e. random walk or a unit root process). This rigid definition implies that traditional unit root tests have low power against series that is close to a unit root process (i.e. fractionally integrated series); (see e.g. Blough, 1992; Cochrane, 1991; Diebold and Rudebusch, 1991; Hassler and Wolters, 1994; Leybourne and Newbold, 1999). Fractionally integrated series take any integration order and therefore a process can be a near unit root process but still being mean-reverting. Such series is often incorrectly specified as a random walk or a unit root process when traditional unit root tests are used. In the context of this paper, this implies that the hysteresis hypothesis incorrectly is favored over the natural rate of unemployment.

In order to investigate if there is hysteresis in unemployment rates, ADF- and KPSS-tests are initially conducted as a preliminary analysis. Since these tests have low power and often give unreliable conclusions, Sowell's (1992) exact maximum likelihood (EML) ARFIMA (autoregressive fractionally integrated moving average) estimator is used in the paper as well. One advantage of using an ARFIMA estimator is that the fractional integration order is estimated and hence a quantitative measure is given. Comparisons of the degree of integration in unemployment rates between countries are therefore possible. More importantly, the point estimates of the integration order in unemployment rates can be tested against different hypotheses. ARFIMA models are complicated to model and in order to choose the most suitable model the Akaike information criterion (AIC) is used. Before determining the

ARFIMA model, the residuals are tested for normality and several tests are conducted to assure that there is no autocorrelation in the residuals. The countries investigated in the paper are Australia, Belgium, Canada, France, Netherlands, Sweden, the United Kingdom and the United States. Since the finding that unemployment rates are best described by the hysteresis hypothesis is challenged by the fact that structural breaks may bias the result (see e.g. Arestis and Mariscal, 1999; Lee and Chang, 2008; Yilanci, 2009), some structural breaks are considered in the paper as well.

This paper extends the existing literature by studying several OECD countries when monthly unemployment data is used. Earlier studies use quite few observations and often rely on quarterly data (thus having fewer observations). Furthermore, the ARFIMA approach is used in the paper, both applied to a longer period of time but also for shorter time spans. Up to the author's knowledge, this is not a common approach taken in other papers, and the goal is that this approach provides further insights to the existing literature.

The remainder of this paper is organized as follows. In section 2, the methodology of the paper is described. Following this, in section 3, the monthly unemployment data used in the paper is presented. Several plots are provided and some descriptive statistics as well. In section 4, the analysis of the paper is found. The analysis begins with some traditional unit root tests (ADF and KPSS tests). Next, the ARFIMA estimates are provided and analyzed. The conclusions and some discussion about the findings in the paper are found in section 5.

2. Methodology

In this section, some methodology is provided. Fractional integration theory and Sowell's (1992) EML ARFIMA estimator are described in subsections 2.1 and 2.2, respectively. Lastly, in subsection 2.3, the model selection procedure is explained.

2.1 Fractionally integrated processes

A time series that can be expressed as an invertible autoregressive moving average (ARMA) process after being differenced d times is said to be integrated of order d (Brockwell and Davis, 2002). In traditional unit root literature, e.g. according to the integrated ARMA models

[ARIMA (p, d, q)] suggested by Box and Jenkins (1976), the integration parameter d is restricted to take integer values (e.g. 0,1,…); the parameters p and q are the number of AR and MA parts included in the model. According to fractional integration literature, the parameter d takes values containing a fractional part as well (e.g. 0.01, 0.02,…). Expressed in other words, the parameter d can take all real values in ARFIMA models. The fractional integration approach is therefore a generalization of the traditional view of time series. In the fractional integration literature, the parameter d is often referred to as the fractional integration parameter or (long) memory parameter (Baillie, 1996). Granger (1980), Granger and Joyeux (1980) and Hosking (1981) define a long memory ARFIMA (p,d,q) model (y_t) as

$$\Phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t \qquad t = 1,2,\dots T \tag{1}$$

Where d is the fractional integration parameter, L is the lag operator and ε_t is a white noise process (i.e. a process with independent and identically distributed observations with zero mean and the variance σ_ε^2). T is the total number of observations and $\Phi(L)$ and $\theta(L)$ represent the autoregressive (AR) operator and the moving average (MA) operator, respectively. The AR and MA operators have no common roots and all solutions of the characteristic polynomials lie outside the unit circle (Baillie, 1996). This is expressed as

$$\Phi(L) = 1 + \Phi_1 L^1 + \Phi_2 L^2 + \cdots + \Phi_p L^p \tag{2}$$

$$\theta(L) = 1 + \theta_1 L^1 + \theta_2 L^2 + \cdots + \theta_q L^q \tag{3}$$

Since the AR and MA parts have no influence on the long run behavior of the time series, these components are referred to as short memory components (Baillie, 1996; Funke, 1998). The long run properties of the time series are described by the memory parameter d (Baillie, 1996; Funke, 1998). Table 1 summarizes the characteristics associated with different values of d.

TABLE 1: Characteristics of the memory parameter

Memory	Interval (d)	Mean-reverting	Variance	Characteristics	Hypothesis supported
Short	$d = 0$	Yes	Finite	Covariance stationary	NRH
Long	$0 < d < 0.5$	Yes	Finite	Covariance stationary	NRH
Long	$0.5 \leq d < 1$	Yes	Infinite	Covariance non-stationary	NRH
Long	$d \geq 1$	No	Infinite	Covariance non-stationary	HH

Sources: Granger (1980), Granger and Joyeux (1980), Hosking (1981)
Remark: NRH = The natural rate of unemployment hypothesis. HH = The hysteresis hypothesis.

A process is said to have short memory when d is equal to zero ($d = 0$). Short memory time series are both stationary in mean and variance. This means that the series reverts back to its equilibrium level after a disturbance to the process (Baillie, 1996). The mean-reversion process is fairly rapid since there is no long memory in the time series when $d = 0$.

As Table 1 shows, there are several cases of when a process possesses long memory. When $0 < d < 0.5$ the series is more persistent than when $d = 0$. But since there is some long memory when $0 < d < 0.5$, it takes slightly longer time for the process to return its initial mean value than in the case when $d = 0$; thus, a greater value of the memory parameter implies more persistence in the series. Any process with a parameter value of d within the interval ($0.5 \leq d < 1$) is therefore even more persistent than when $0 \leq d < 0.5$. Table 1 reveals that the series is mean-reverting even though covariance non-stationarity is implied. Covariance non-stationarity occurs since the variance no longer is finite (Granger, 1980; Granger and Joyeux, 1980; Hosking, 1981). The series will however return to its mean value in the long run, it is only a matter of time.

The three cases covered so far give support to the idea of a natural unemployment rate (i.e. $d < 1$, the natural rate of unemployment hypothesis describes the process best); (see Table 1). In other words, as long as the series is mean-reverting, the unemployment equilibrium is unchanged in the long run and no policy measures are needed. The hysteresis hypothesis is supported when the series have long memory with the memory parameter being greater than or equal to one ($d \geq 1$). Since the series is neither mean-reverting nor covariance-stationary, the unemployment equilibrium will change in the long run as the series is subject to disturbances. From a statistical point of view, series with an integration order greater than or equal to one is problematic. Since the variance no longer is finite, and there is no mean value to return to, the series explodes as the number of observations increase (Greene, 2003). The series will therefore wander away in any direction if the process is left untreated. So if a shock occurs, the effect is neither predictable nor transitory. To achieve stationarity, the unit root series needs to be differenced at least one time when $d \geq 1$.

2.2 Sowell's (1992) Exact Maximum Likelihood (EML) estimator

Sowell's (1992) EML estimator assumes that a vector of observations $Y = (y_1, y_2, \dots, y_T)'$ generated by the fractionally integrated model described in Equation (1) follows a normal distribution with zero mean and the covariance matrix Σ. The maximum likelihood objective function (log likelihood function) of the ARFIMA process in Equation (1) is expressed as (see Sowell, 1992 and Franke et al., 2008)

$$l(\Phi, \theta, d; Y) = -\frac{T}{2}\log|\Sigma| - \frac{1}{2}Y'\Sigma^{-1}Y \quad (4)$$

The EML estimator by Sowell (1992) is given by maximizing the log likelihood function with respect to the parameters of interest (Φ, θ, d). Therefore, the EML estimate of the memory parameter is obtained through

$$\hat{d} = \arg\max_{\Phi,\theta,d}\left\{-\frac{T}{2}\log|\Sigma| - \frac{1}{2}Y'\Sigma^{-1}Y\right\} \quad (5)$$

An advantage with Sowell's (1992) EML estimator is that the short memory components (ARMA) are simultaneously calculated with the long memory component (Baillie, 1996). This means that the effects from each type of components are efficiently separated from each other. Moreover, a correct specification of the ARMA part in ARFIMA models is crucial in order to obtain an unbiased estimate of the memory parameter (Gil-Alana, 2001).

The MLE estimator in Equation (5) performs well in finite samples. Additionally, the EML estimator produces normally distributed estimates of the memory parameter when $d > 0$ (Dalhaus 1989; Yajima, 1985). Hence, theory based on normality assumptions applies. Dalhaus (1989) and Yajima (1985) also conclude that the MLE estimator is efficient (i.e. the estimator with the smallest variance among a set of unbiased estimators; Miller and Miller, 2003).

2.3 Model selection criterion

A commonly used approach when selecting the most appropriate ARFIMA model is first to determine the number of AR and MA components to be included in the model (see e.g. Box and Jenkins, 1976; Gil-Alana, 1999, 2001, 2002). Once this is done, some diagnostic tests are conducted to confirm that the model is correctly specified. In earlier papers the residuals are often tested for at least autocorrelation and normality (Gil-Alana, 1999, 2002; Kurita, 2010). Other papers test for ARCH errors as well (see e.g. Gil-Alana, 1999). Earlier research has however shown that testing for ARCH is unnecessary if the main interest is in the long run parameter; ARCH errors do not affect the memory estimate (d) (Hauser and Kunst, 1998).

Finally, some information criterion (IC) is used to choose the most suitable model. The Akaike information criterion (AIC) or some Bayesian variants (e.g. Schwarz IC) are often used (see e.g. Koustas and Veloce, 1996; Gil-Alana, 1999, 2001). The AIC tends to overfit models (include too many parameters) while Bayesian ICs sometimes underfits the model. Since the Bayesian ICs often lead to underfitted models when the sample is small, the AIC is used in this paper (Greene, 2003). In this paper, the following approach is taken when finding appropriate ARFIMA models

Step 1: Several ARFIMA models are fitted to the data material. An ARFIMA (3, *d*, 3) is modeled at first, next, a somewhat smaller model is estimated. This procedure continues for all combinations of AR and MA components until the smallest model is reached - an ARFIMA (0, *d*, 0); (see tables in the Appendix). Only models in which all AR and MA components are significantly different from zero at any reasonable level of significance (1%, 5% or 10%) continue to the next step in the procedure. Models with one or more insignificant variables are omitted from further analysis at this point.

Step 2: The models of interest (from step 1) are tested for autocorrelation in the residuals using the Lagrange Multiplier (LM) test by Breusch and Godfrey (BG test; see Breusch, 1978; Godfrey, 1978). If there is serial correlation between the residuals in any model, the model is likely misspecified. Models in which the null hypothesis of no autocorrelation is rejected at the 1% significance level are considered as invalid and are omitted from the analysis at this stage.

Step 3: Models that pass step 2 is next tested for normality in the residuals using the LM test of Jarque and Bera (1987); (JB test). The standard set in this paper is that models pass the JB test as long as the null hypothesis of normality is not rejected at the 1% significance level. The assumption of normality in the residuals is however not crucial for the findings in the paper. The samples in the paper are relatively large and hence tests based on the normality assumptions still applies.

Step 4: The last step is to choose the model that minimizes the value of AIC.

$$\text{AIC} = -2\, l(\Phi, \theta, d; Y) + 2K \qquad (6)$$

where *K* is the number of parameters (AR-, MA-components and the intercept) estimated in the model. Since the maximum of the log likelihood function provides the best model, a larger value of the log likelihood function leads to a smaller AIC according to Equation (6).

Furthermore, the AIC takes the number of parameters into account since the last part in Equation (6) penalizes for including more parameters in the model. Important to note is that the AIC is a measure that compares a set of models that are correctly specified (see e.g. Greene, 2003): in this paper, that are models that have passed the criterions stated above - significant AR and MA estimates, no serial correlation in the residuals and preferably (but not necessarily) normality in the residuals.

When comparing models with respect to the AIC values, models that pass both the test for autocorrelation as well as the test for normality are always preferred over a model that only pass the autocorrelation test. For instance; suppose that we have a correctly specified model (significant AR and MA parameter estimates, no autocorrelation and the residuals belong to a normal distribution) but with larger AIC value than a similar model that only passes the test for autocorrelation (but still have significant parameter estimates). In that case, the former model with the larger AIC value is preferred since that model passes both tests.

3. Data

The data used in the analysis is monthly observations of seasonally adjusted unemployment rates. The unemployment rates are measured as the number of unemployed persons as a percentage of the civilian labor force (OECD, 2011a). The number of observations differs for the studied countries. For Canada and the United States there are many observations. For these countries, unemployment data has been collected since January 1956 and January 1955, respectively, and data is accessible up to November 2011. Hence, in total, there are 671 and 683 observations for Canada and the United States. For the other countries in the study, the data material contains fewer observations. The data gathering often begins within the span 1970-1980 and proceeds until November 2011. Therefore, the number of observations varies between 407 and 503 depending on when the data gathering begins (see column 2 in Table 2 for further reference).

Table 2 provides some summary statistics of the countries' unemployment rates. Notice that individual comparisons of values between countries may not be completely legitimate since the number of observations differs between the countries and cover different time periods.

However, comparing summary statistics give some guideline of the characteristics of the data material, even though the result should be taken with some skepticism.

TABLE 2: Summary statistics

Country	Time period	Mean	Std.dev.	Median	Maximum	Minimum
Australia	1978M01 - 2011M11	7.10	1.85	6.68	11.21; (1992M12)	3.97; (2008M02)
Belgium	1970M01 - 2011M11	7.69	2.18	8.10	11.00; (1983M10)	2.16; (1970M09)
Canada	1956M01 - 2011M11	7.38	2.16	7.33	13.04; (1982M12)	2.76; (1956M10)
France	1978M01 - 2011M11	9.07	1.60	9.20	11.80; (1994M03)	4.39; (1978M01)
Netherlands	1970M01 - 2011M11	4.91	1.73	5.10	8.63; (1982M12)	0.91; (1970M06)
Sweden	1970M01 - 2011M11	4.78	2.80	3.70	10.50; (1997M06)	1.21; (1970M06)
United Kingdom	1971M01 - 2011M11	6.88	2.44	6.20	11.30; (1985M04)	2.26; (1973M12)
United States	1955M01 - 2011M11	5.95	1.60	5.70	10.80; (1982M11)	3.40; (1969M05)

Remark: The values in the table are the percentage of unemployed persons in the civilian labor force with respect to each measure (i.e. mean, std.dev, median, maximum and minimum).

In Table 2, it is seen that the mean unemployment rate is highest in Belgium and France; 7.69% and 9.07%. Historically, Canada has experienced high unemployment rates in average as well. Moreover, Netherlands and Sweden have had lower rates of unemployment than the other countries in the paper; see 4.91% and 4.78%. The differences in the mean unemployment rates between countries are notable. For instance, the difference between France and Sweden is 4.29 percentage points. Hence, the unemployment rate has been considerably higher in France than in Sweden; this becomes even clearer when the median is studied (compare 9.20% with 3.70%). The spread in unemployment rates has been largest in Sweden and the United Kingdom (2.80% and 2.44%), while the unemployment rates in France and the United States have shown less spread then other countries (both 1.60%).

Concerning the maximum rates of unemployment, Canada has experienced the highest unemployment rate (13.04%; December 1982). On the contrary, the lowest maximum occurred for Netherlands (8.63%; December 1982). The magnitude of the difference in unemployment between Canada and Netherlands is 4.41 percentage points and reveals that there are somewhat significant differences in maximum unemployment between the studied countries. The other countries have maximum values that typically lie with the span 10.50-11.80%. A similar pattern in the minimum unemployment is seen. In France, the minimum unemployment rate is 4.39% (January 1978) while Netherlands has experienced a level as low as 0.91% (June 1970). Once again, there is a substantial difference between the countries.

The maximum unemployment rates occurred during the 1980-1990s (Table 2). Figure 1-8 tell the same story. More or less every country had high unemployment during 1980-1990s. One reason to the high unemployment levels is the occurrence of the second oil crisis in 1979. It is also reasonable to believe that some aftereffects of the collapse of the Bretton Woods system in 1973, and the first oil crisis the same year, had a long lasting impact on future unemployment rates (Cohen, 2001).

Moreover, Table 2 tells that unemployment in general was low during 1970s. Additionally, Figure 1-8 reveal that unemployment was low during the late 1960s and early 1970s. As is seen from the figures, the unemployment rates began to rise at the middle of 1970s. The rise in unemployment can likely be explained by the collapse of the Bretton Woods system and the oil crisis in 1973 (Cohen, 2001).

The unemployment rates remain high after the peak period 1980-1990 according to Figure 1-8. If the time series for Belgium, Canada and Sweden are studied (Figure 2, 3 and 6) this becomes clear. Figure 6 illustrates the phenomena well. The unemployment rate is higher at the end of the data material than it is in comparison to the initial or lowest value in the beginning of the time series. This behavior is consistent with the hysteresis hypothesis. Australia works as an exception since unemployment is lower in 2008 than it was during 1970s and no tendencies of consistently higher unemployment rates are seen.

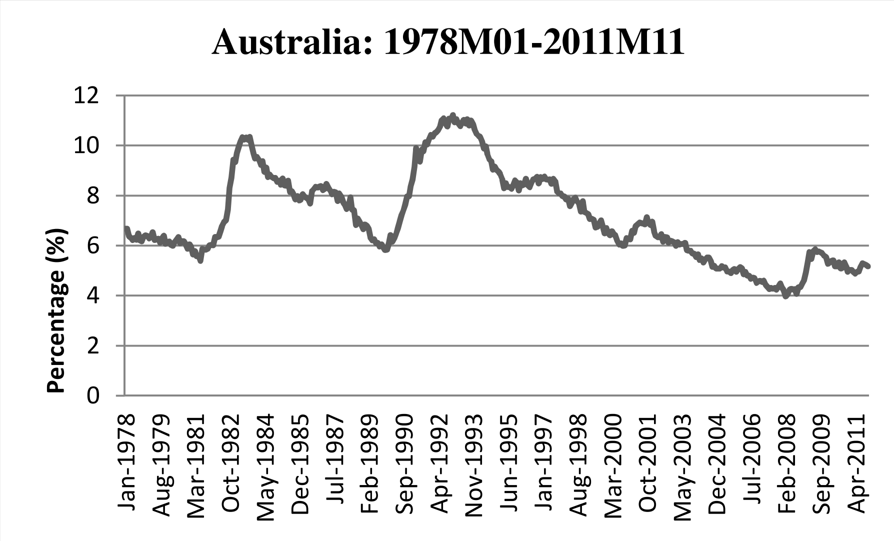

Figure 1: Time series graph, Australia

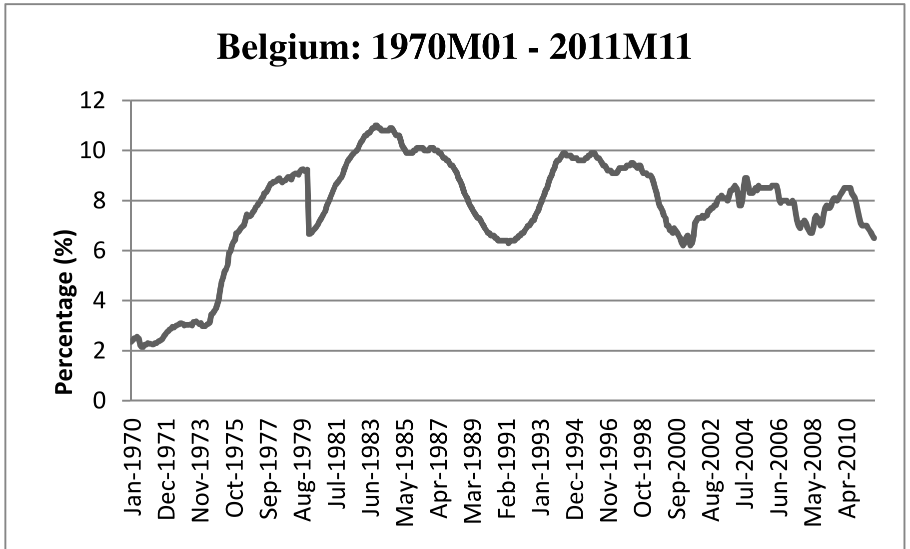

Figure 2: Time series graph, Belgium.

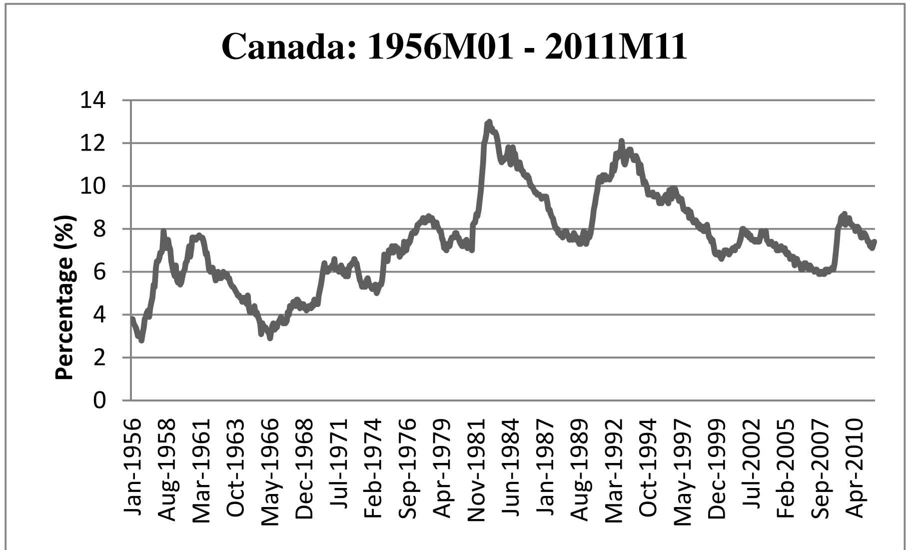

Figure 3: Time series graph, Canada

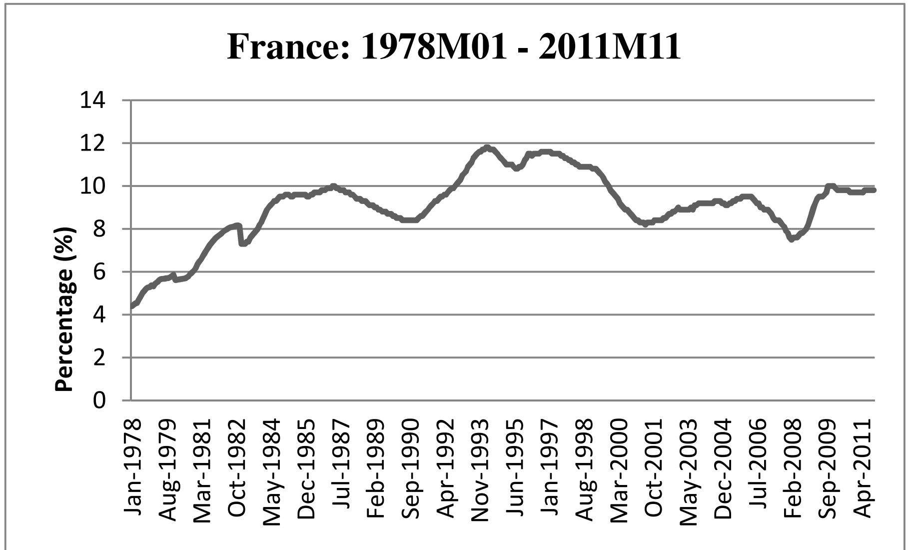

Figure 4: Time series graph, France

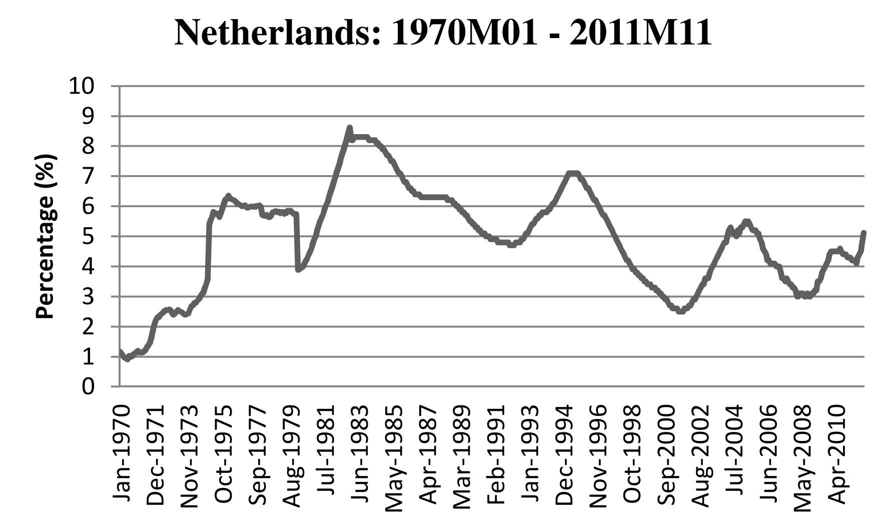

Figure 5: Time series graph, Netherlands

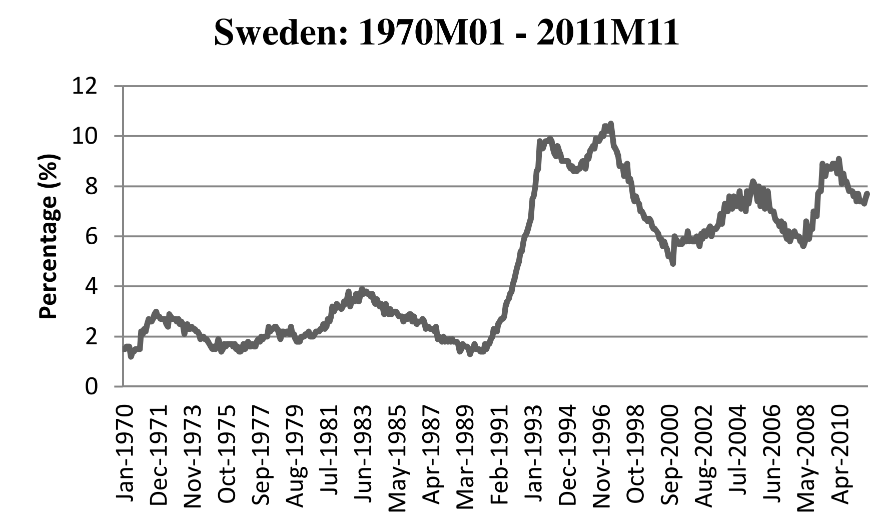

Figure 6: Time series graph, Sweden

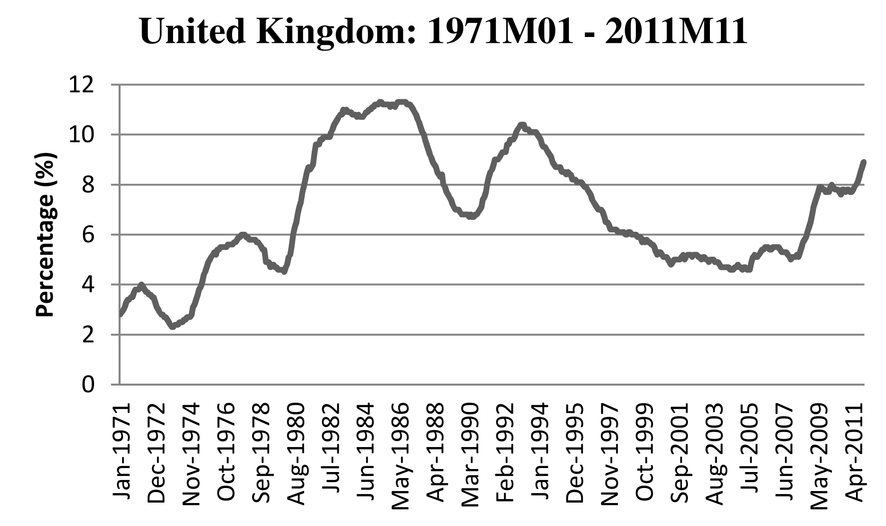

Figure 7: Time series graph, United Kingdom

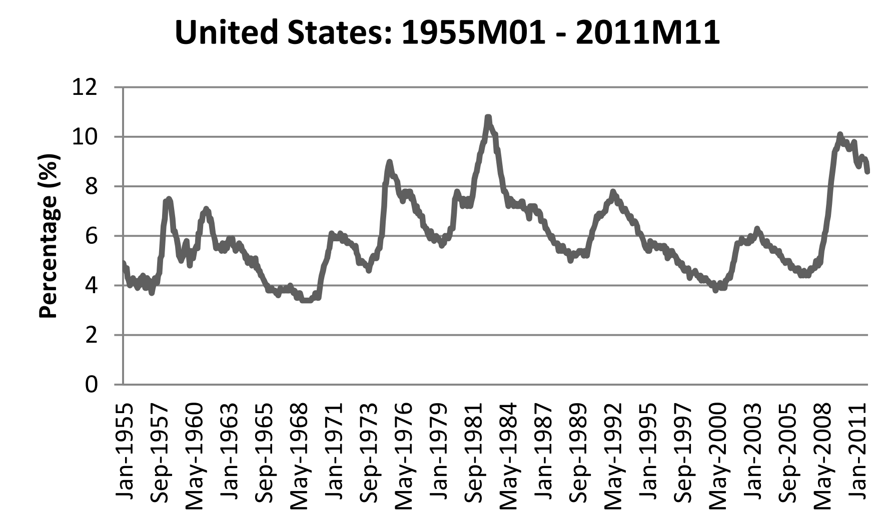

Figure 8: Time series graph, United States

4. Analysis

To get an overview of the time series properties of unemployment rates, some ADF and KPSS test are conducted as an initial step of the analysis. Hereafter, the ARFIMA estimates are considered and a more thorough analysis of the unemployment rates is performed.

4.1 Unit root tests

The unit root tests used in this paper are built upon asymptotic theory (i.e. using infinitely large samples) which implies that large samples are required to obtain reliable estimates. Size distortions are often a problem when ADF type unit root tests are used, especially when smaller sample sizes are used (see e,g, Blough, 1992; Cochrane, 1991; Diebold and Rudebusch, 1991). When using small samples low power is implied and the result of the unit root tests is unreliable. By this argument, the data material is not divided into smaller time periods to account for structural breaks when the ADF and KPSS tests are conducted. The loss in power imposed by the decreased sample size is considered as greater than the win obtained by detecting the influence of structural breaks. Several complementary KPSS tests are however conducted as a way to lend some credibility to the results found by the ADF tests.

In Table 3, the result of several ADF test are presented. In the ADF tests with only a constant, the result is ambiguous to some extent. The unit root null hypothesis (i.e. $d = 1$) is rejected at different levels of significance in three cases, while it is not rejected in the remaining five tests. The null hypothesis is rejected for Canada and France at the 10% significance level and at the 5% level for the United States. Therefore, at this point, there is some evidence that there is no hysteresis in unemployment in these countries. On the other hand, there are some indications of hysteresis in unemployment in Australia, Belgium, Netherlands, Sweden and the United Kingdom.

TABLE 3: ADF-tests

		ADF test (constant)		ADF test (constant+trend)	
Country	**Time period**	**Value**	**Prob.**	**Value**	**Prob.**
Australia	1978M01 - 2011M11	-1.75	0.41	-2.31	0.43
Belgium	1970M01 - 2011M11	-2.33	0.16	-1.64	0.77
Canada	1956M01 - 2011M11	-2.88	0.05*	-2.86	0.18
France	1978M01 - 2011M11	-2.90	0.05*	-2.62	0.27
Netherlands	1970M01 - 2011M11	-2.57	0.10	-2.67	0.25
Sweden	1970M01 - 2011M11	-1.74	0.41	-2.62	0.27
United Kingdom	1971M01 - 2011M11	-2.21	0.20	-2.24	0.47
United States	1955M01 - 2011M11	-3.22	0.02**	-3.30	0.07*

*** 1% significance ** 5% significance * 10% significance
Remark: The null hypothesis of the ADF test is that the time series has a unit root (see Dickey and Fuller, 1979).

When a trend component is added as a deterministic component in the tests, the result becomes explicit. In seven of the studied countries, the unit root null is not rejected; hence, the hysteresis hypothesis applies for these countries. Furthermore, the null hypothesis is rejected in the United States at the 10% significance level and hence indicates that there is no hysteresis in the US unemployment rate. The overall result this far indicates that there is some evidence of hysteresis in unemployment rates. Next, KPSS tests are conducted in order to strengthen the result from the ADF tests. The KPSS tests are found in Table 4.

TABLE 4: KPSS-tests

		KPSS test (constant)	KPSS test (constant+trend)
Country	**Time period**	**Statistic**	**Statistic**
Australia	1978M01 - 2011M11	0.87***	0.38***
Belgium	1970M01 - 2011M11	0.68**	0.41***
Canada	1956M01 - 2011M11	0.44*	0.44***
France	1978M01 - 2011M11	0.76***	0.42***
Netherlands	1970M01 - 2011M11	0.43*	0.40***
Sweden	1970M01 - 2011M11	1.97***	0.16**
United Kingdom	1971M01 - 2011M11	0.44*	0.45***
United States	1955M01 - 2011M11	0.36*	0.20**

*** 1% significance ** 5% significance * 10% significance
Remark: The null hypothesis of the KPSS test is that the time series is stationary and has no unit root (see Kwiatkowski et al., 1992).

The null hypothesis of stationarity in the KPSS test is rejected at varying levels of significance according to Table 4. The result is independent of which components that are included as regressors in the tests (i.e. it does not matter if only a constant or a constant + trend is considered). To some extent, these findings contradict the result of the ADF tests. This is at least true for Canada, France and the United States.

To summarize, the overall result from the univariate unit root tests is that unemployment rates likely are hysteresis time series. Most of the tests performed point in this direction even though there are some ambiguities when some countries are considered. In all, however, the result is the same as in several earlier papers (see e.g. León-Ledesma, 2002; Mitchell, 1993; Roed, 1996).

In order to assess the magnitude of the memory parameter and provide a more thorough analysis of the unemployment series, ARFIMA estimates are provided in the next section. The result found up to this point may be incorrect since it is possible that the unemployment series are fractionally integrated of a higher order but still is mean-reverting. For instance, an unemployment series could have the integration order I(0.99). Such series would likely, due to the weak power of traditional unit root tests, be considered as a process with hysteresis when the series in fact is a mean-reverting process.

4.2 ARFIMA estimates

Several papers detect that structural breaks affects the persistence in unemployment and hence the possibility to detect hysteresis (see e.g. Arestis and Mariscal, 1999; Lee and Chang, 2008; Yilanci, 2009). Since structural breaks have occurred during history, the data material is split into two parts when estimating the ARFIMA models. The break date chosen is the year 1990; the first sample consists of observations until 1989M12, and the second samples involve the observations 1990M01-2011M11. The rationale of choosing this year as a break point is manifold. First of all, the period after 1990 is characterized by global slowdowns (1990-1993, 1998, 2001-2002 and the financial crisis in 2007; Miller, 2008) as well as economic expansions (such as the IT boom around year 2000 as well as the flourishing economy in the middle of 2000). Secondly, the data material is approximately divided into relatively equal and large sample sizes. This is advantageous since the time periods are comparable and the estimates are consistent since the samples likely are large enough. It is difficult to assess when structural breaks have occurred and other structural breaks could have been considered in the paper; for instance, before and after the first and/or second oil crises. Unfortunately, such approach is problematic to achieve due to the lack of observations for some countries. Besides, dividing the data with respect to further structural breaks also results in small samples for other countries.

The memory parameter is estimated for the complete samples as well as for the sub-samples of the data material. The models in Table 5 are an excerpt of the most suitable models (according to the model selection algorithm described in section 2.3) when the complete data material is used. The complete specification of all models, from an ARFIMA (3, *d*, 3) down to an ARFIMA (0, *d*, 0) is found in the Appendix (see Table A.1-A.8). However, the corresponding tables for estimates concerning the period before and after 1990 are not reported in the paper. Since the intuition and model selection procedure is the same as when all observations are used, tables with estimates for the other time periods would not contribute to the analysis.

TABLE 5: ARFIMA-estimates (all observations)

Country	ARMA (p, q)	d	$t_{d=0}$	$t_{d=1}$	Normality test Prob.	BG-test Prob.
Australia	(2, 2)	1.01	13.85	0.10	0.95	0.01
Belgium	(0, 0)	1.22	29.23	5.31	0.00	0.93
Canada	(1, 2)	1.27	15.83	3.34	0.01	0.12
France	(0, 0)	1.35	16.48	4.27	0.26	0.18
Netherlands	(1, 0)	1.30	24.06	5.56	0.00	0.92
Sweden	(2, 0)	1.40	32.22	9.27	0.02	0.33
United Kingdom	(2, 2)	0.96	6.15	-0.28	0.13	0.82
United States	(2, 0)	0.49	8.37	-8.74	0.00	0.01

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

The ARFIMA estimates in Table 5 are estimated when the long data material is used (i.e. all observations available are used). In general, an AR part is included when the most appropriate model is selected (six out of eight models have an AR part) while a MA part seldom is included. Noticeable is that the white noise specification of the ARMA part is preferred when modeling unemployment in Belgium and France. Furthermore, higher order models, i.e. models that have AR and MA parts of order three, is avoided; such models seem to be less suitable to model unemployment series when all observations are used.

Table 5 reveals that most of the selected models fulfill the tests of normality and more importantly, pass the Breusch-Godfrey autocorrelation test. However, some observations are inaccurately modeled when the ARFIMA models are fitted and therefore cause non-normality in the residuals. Therefore the normality tests only are conducted after the data material has been cleansed from extreme outlying observations [i.e. observations located three interquartile ranges (IQR) from the 1st or 3rd quartile]. In some cases, the null hypothesis of normality in the residuals was strongly rejected before the data set was cleansed from outliers. After removing only a few observations (0-3% of the total observations for each country) and re-running the models, the null hypothesis of normality could not be rejected at any reasonable level of significance.

Long memory estimates are by convention often analyzed at the 5% significance level (see e.g. Gil-Alana, 2001); therefore, the 5% significance level is used in the upcoming analysis of the memory estimates. In the analysis, the hypotheses that $d = 0$ and $d = 1$ are tested using t-tests (the rejection area at the 5% significance level is when the absolute t-value is greater than or equal to 2). Hereafter, the individual point estimates of the memory parameters are evaluated. If both hypotheses are rejected and $\hat{d}$ is between zero and one, the natural rate of unemployment applies. If $\hat{d}$ however is equal to or greater than one, and both hypotheses are rejected, there is hysteresis in the series. If some of the hypotheses are not rejected, the result is inconclusive and nothing can be said about the estimated memory parameter.

According to Table 5, the null hypothesis that the long run parameter is equal to zero ($d = 0$) is rejected in all models; none of the unemployment series have short memory. This result is expected due to the unit root tests conducted earlier. In most cases, the estimates of the long memory parameter are greater than or equal to one. This provides some credibility to the hysteresis hypothesis. Table 5 shows that there is hysteresis in unemployment rates for Belgium, Canada, France, Netherlands and Sweden. The hysteresis result is strengthen since $d = 1$ is rejected and the parameter estimates are greater than to one. Therefore it is likely to believe that $d > 1$.

The maybe most interesting unemployment rates belong to Australia, the United Kingdom and the United States (see Table 5). When considering Australia, the estimated memory parameter is 1.01. Since the hypothesis $d = 1$ is not rejected it is not possible to conclude anything about the characteristics of the Australian unemployment rate. A similar argument as for Australia goes for the United Kingdom. The memory parameter is equal to 0.96 but the hypothesis that $d = 1$ cannot be rejected. The estimated memory parameter for the United States unemployment rate is by far the smallest memory estimate obtained (since $\hat{d} = 0.49$). Both the hypotheses that $d = 0$ and $d = 1$ are rejected (see the t-values 8.37 and -8.74) and since the estimate of d is between zero and one, there is evidence in favor of the natural rate of unemployment hypothesis for the US unemployment rate.

So far, the findings suggest that there likely is hysteresis in unemployment rates. In five of the countries (Belgium, Canada, France, Netherlands, Sweden), the hysteresis hypothesis applies, while no conclusion can be drawn for two countries (Australia and the United Kingdom). The only country in which the unemployment rate is mean-reverting and therefore follows the natural rate of unemployment hypothesis is the United States.

Table 6 provides the estimates of the memory parameters for the data sample using observations up to 1989M12. A similar pattern as before is seen concerning the AR and MA structure of the models. The models often have at least one AR component while existence of MA component is rare. Moreover, there are several models in which the white noise specification of the short memory components is preferred (see Belgium, France and Netherlands). Most of the models pass the normality test as well as the test for autocorrelation. The only model that not passes both tests is the one for the US unemployment rate (the model fails the JB normality test).

TABLE 6: ARFIMA-estimates (up to 1989M12)

Country	ARMA (p, q)	d	$t_{d=0}$	$t_{d=1}$	Normality test Prob.	BG-test Prob.
Australia	(1, 1)	1.09	7.45	0.60	0.70	0.90
Belgium	(0, 0)	1.15	19.62	2.60	0.16	0.93
Canada	(0, 1)	1.26	12.74	2.65	0.12	0.83
France	(0, 0)	1.24	7.77	1.49	0.82	0.71
Netherlands	(0, 0)	1.16	22.09	2.98	0.02	0.98
Sweden	(3, 2)	1.07	15.38	0.98	0.03	0.23
United Kingdom	(2, 0)	0.66	4.88	-2.57	0.16	0.23
United States	(2, 0)	0.53	6.33	-5.66	0.00	0.02

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

According to Table 6, the estimates often are greater than or equal to one. The hysteresis hypothesis is supported when considering Belgium, Canada and Netherlands since $d > 1$ and the hypothesis that $d = 1$ is rejected. Consequently, there is room for the governments to take measures against unemployment in these countries. Inconclusive result occurs for Australia, France and Sweden. It is reasonable to believe that the smaller sample sizes have caused larger standard errors hence yielding less significant results. Since the estimates are larger than one, there is weak support of hysteresis in these unemployment rates as well.

The persistence in unemployment is low before 1990s in both the United Kingdom and the United States. The parameter estimates of the integration orders are 0.66 and 0.53, respectively. Furthermore, for both countries, the hypotheses that d = 0 as well as d = 1 is rejected. Hence, it is likely to believe that the true value of the memory parameter lies within the interval $0 < d < 1$. In other words, the series follow a natural unemployment rate since the unemployment returns to its equilibrium after a disturbance or shock.

In Table 7 estimates based on the sample 1990M01-2011M11 is found. A change in comparison with the models examined so far, is that the models in Table 7 contain more ARMA parts and of higher orders than before. The AR part is often of order two or three while the same is true for the MA part. Evident as well, is that none of the unemployment series are modeled with the white noise specification of the ARMA part. Furthermore, the only model that fails any of the tests is the ARFIMA (3, 0.82, 0) for Netherlands (the residuals are not normally distributed).

TABLE 7: ARFIMA-estimates (1990M01-2011M11)

Country	**ARMA (p, q)**	d	$t_{d=0}$	$t_{d=1}$	**Normality test Prob.**	**BG-test Prob.**
Australia	(0, 1)	1.40	15.58	4.48	0.48	0.01
Belgium	(2, 3)	1.37	11.74	3.16	0.07	0.40
Canada	(3, 3)	0.98	11.27	-0.18	0.01	0.28
France	(3, 0)	0.93	5.41	-0.40	0.50	0.03
Netherlands	(3, 0)	0.82	7.87	-1.73	0.00	0.36
Sweden	(2, 0)	1.42	27.68	8.23	0.93	0.42
United Kingdom	(2, 2)	1.24	22.33	4.33	0.50	0.07
United States	(1, 0)	1.37	23.09	6.21	0.14	0.69

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

Just as before, the hypothesis that $d = 0$ is clearly rejected in all models according to Table 7. The hysteresis hypothesis is strengthened when considering the estimates in Table 7; the estimated values of d are mainly greater than or equal to one in the cases where both $d = 0$ and $d = 1$ are rejected (see Australia, Belgium, Sweden, the United Kingdom and the United

States). For Canada, France and Netherlands the result is inconclusive. Worth to notice is that the estimates of *d* are smaller than one for all of the inconclusive countries. The hypothesis that $d = 1$ cannot be rejected and it is therefore difficult to assess if there is hysteresis in the unemployment rates of Canada, France and Netherlands.

If the estimates in Table 6 and Table 7 are compared it is seen that the integration order often is higher when the estimates of the period 1990M01 - 2011M11 are considered. In five countries (Australia, Belgium, Sweden, the United Kingdom and the United States) the estimated memory parameters are larger in the 1990M01-2011M11 sample. The estimates of *d* are smaller in three of the studied countries (Canada, France and Netherlands). The reliability of the parameter estimates is however questioned for all countries in which the degree of persistence is lower in the 1990M01-2011M11 sample (according to the discussion above and Table 7). One plausible explanation to why the unemployment persistence is higher after 1990 is the turbulent economic conditions at this point in time. Since recessions tend to increase the integration order of unemployment (OECD, 2011b) the results is not surprising. Another possible explanation to the increased unemployment persistence is that unemployment during a recession may change people's attitudes toward seeking jobs. Some of the unemployed people lose motivation of finding a job (discouraged workers) and others suffer from deterioration of human capital after a long spell of unemployment (Song and Wu, 1998). Therefore it takes a long time for these individuals to enter the labor market. Increased integration order in unemployment sometimes is explained by changes in social institutions (i.e. welfare programs, unemployment insurances and so on; see e.g. Song and Wu, 1998 or Candelon et al., 2009), demographic changes or changes in the industrial decomposition (Clemente et al., 2004). So in other words, it is plausible that changes in institutions caused the higher integration orders after 1990.

To summarize the analysis of the ARFIMA estimates, there likely is hysteresis in unemployment rates. When all estimates of the memory parameters are considered (in Table 5, 6, 7) the finding of hysteresis in unemployment rates is strengthened. Thirteen out of 24 estimates of *d* are greater than or equal to one (as the hypotheses $d = 0$ and $d = 1$ are rejected as well). As a further confirmation of the hysteresis finding, only three out of 24 estimates of the memory parameter is smaller than one while the extreme hypotheses are rejected. In total, eight of the 24 estimated values of *d* are inconclusive since both hypotheses cannot be rejected at the same time.

5. Conclusions

The countries studied in this paper are Australia, Belgium, Canada, France, Netherlands, Sweden, the United Kingdom and the United States. Univariate unit root tests (ADF and KPSS tests) indicate that unemployment rates for these countries in general have at least one unit root. According to these results, there is hysteresis in unemployment rates and as a consequence policymakers and governments need to take actions in order to manage the unemployment levels in the long run. It is however well-known that univariate unit root tests perform poor when dealing with fractionally integrated time series (see e.g. Cochrane, 1991; Diebold and Rudebusch, 1991; Hassler and Wolters, 1994). Due to the univariate unit root tests' low power against fractionally integrated alternatives, such tests tend to favor the unit root hypothesis. In fact, series that not contain a unit root is incorrectly considered as a unit root process. In the context of this paper, a series that follows a natural rate of unemployment is incorrectly taken as a process with hysteresis. Hence, in order to establish if there is hysteresis in unemployment rates, Sowell's (1992) exact maximum likelihood (EML) ARFIMA estimator is used in the paper as well. The main advantage with ARFIMA estimators is that the fractional integration order can be estimated (see e.g. Sowell, 1992). A quantitative measure of the fractional integration order is given and it is possible to conduct a more refined analysis of the time series properties.

The analysis of the estimated memory parameters indicates that there is hysteresis in unemployment rates. Thirteen out of 24 estimated memory parameters are greater than or equal to one while only three estimates of the fractional integration parameters are between zero and one (the hypotheses $d = 0$ and $d = 1$ are rejected at the same time in both cases). The remaining estimates of the memory parameters return inconclusive results and further analysis is required to draw any reliable conclusions about these unemployment rates. The main result that there in general is hysteresis unemployment rates imply that politicians and governments carefully need design appropriate policy actions in order to control the unemployment rate in the long run.

Earlier literature suggests that structural breaks bias the estimated memory parameters (see e.g. Arestis and Mariscal, 1999; Lee and Chang, 2008). In order to account for structural breaks, the data material is therefore divided into smaller subsamples in the paper. The break date chosen is 1990 and the general result of hysteresis is unaffected no matter which sample that is used. Just as for the whole data material, the claim of hysteresis still remains.

Five out of eight estimated memory parameters are smaller before 1990 in comparison with the ones obtained from the data material containing observations for 1990M01-2011M11. This result could occur by chance but can also be explained by other factors. Earlier literature suggests that differences in integration order between countries (and also between different time periods concerning the same country) can be explained by discrepancies/changes in welfare benefits, unemployment benefits and the industrial decomposition; (see e.g. Song and Wu, 1998; Candelon et al., 2009).

Since the countries in the study are OECD countries and they are quite similar, there is a high internal validity when considering OECD countries. However, some evidence in the paper supports the idea that there is no hysteresis in the US unemployment rate. This is somewhat expected since numerous papers in the past have concluded this (see e.g. Romero-Ávila and Usabiaga, 2007, 2009). To be able to say anything about why integration orders differ between countries and why there is hysteresis in some unemployment rates, the analysis need to be taken down to a country specific level. Up to the author's knowledge very few studies investigate specific countries. This is therefore a suggestion for future research. Moreover, there exist numerous other countries, areas or continents that are unexplored and could be studied. Such studies would contribute to the existing literature on unemployment rates. However, one thing is sure, the debate whether there is hysteresis in unemployment rates is far from closed.

References

Arestis, P. and Mariscal, I. B-F. (1999), "Unit roots and structural breaks in OECD unemployment", *Economic Letters*, Vol.65; 149-156

Baillie, R.T. (1996), "Long memory processes and fractional integration in econometrics", *Journal of Econometrics*, Vol.73; 5-59

Blanchard, O.J. and Summers, L.H. (1986), "Hysteresis and the European unemployment problem", *NBER working paper series*, No.1950

Blough, S.R. (1992), "The relationship between power and level for generic unit root tests in finite samples", *Journal of Applied Econometrics,* Vol.7, No.3; 295-308

Box, G.E.P. and Jenkins, G.M. (1976), *Time series analysis: forecasting and control*, Rev.ED, Holden-Day, Oakland

Breusch, T.S. (1978), "Testing for autocorrelation in dynamic linear models", *Australian Economic Papers*, Vol.17; 334-355

Brockwell, P.J. and Davis, R.A. (2002), *Introduction to time series and forecasting*, 2ED, Springer-Verlag, New York

Camarero, M., Carrion-I-Silvestre, J.L. and 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*, Vol. 68, No. 2; 167-182

Candelon, B., Dupuy, A. and Gil-Alana, L. (2009), "The nature of occupational unemployment rates in the United States: hysteresis or structural?", *Applied Economics*, Vol.41, No.19; 2483-2493

Clark, A.E. and Oswald, J.A. (1994), "Unhappiness and unemployment", *The Economic Journal*, Vol.104, No.424; 648-659

Clemente, J., Lanaspa, L. and Montanés, A. (2004), "The unemployment structure of the US States", *The Quarterly Review of Economics and Finance*, Vol. 45; 848-868

Cochrane, J.H. (1991), "A critique of the application of unit root tests", *Journal of Economic Dynamics and Control*, Vol.15; 275-284

Cohen, B.J. (2001), "Bretton Woods system" in Jones, R.J.B (ed), *Routledge Encyclopedia of International Political Economy*, Routledge, London

Dalhaus, R. (1989), "Correction efficient parameter estimation for self-similar processes", The Annals of Statistics, Vol. 17; 1749-1766

Dickey, D.A. and Fuller, W.A. (1979), "Distribution of the estimators for autoregressive time series with a unit root", *Journal of the American Statistical Association*, Vol.74, No.366; 427-431

Diebold, F.X. and Rudebusch, G.D. (1991), “On the power of Dickey-Fuller tests against fractional alternatives”, *Economic Letters*, Vol.35; 155-160

Franke, F., Hardle, W.K. and Hafner, C.M. (2008), *Statistics of financial markets – an introduction*, 2ED, Springer-Verlag, Berlin

Friedman, M. (1968), “The role of monetary policy”, *The American Economic Review*, Vol.58, No.1; 1-17

Funke, M. (1998), “The long memory property of the US unemployment rate”, Hamburg University / London Business School, *Discussion paper*, No. 19-98

Gil-Alana, L.A. (1999), “Estimation of fractional ARIMA models for the UK unemployment”, University of Navarre, *Discussion paper*, No.05-99

Gil-Alana, L.A. (2001), “The persistence of unemployment in the USA and Europe in terms of fractionally ARIMA models”, *Applied Economics*, Vol.33, No.10; 1263-1269

Gil-Alana, L.A. (2002), “Modeling the persistence of unemployment in Canada”, *International Review of Applied Economics*, Vol.16, No.4; 465-477

Godfrey, L.G. (1978), “Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables”, *Econometrica*, Vol.46, No.6; 1293-1301

Granger, C.W.J. (1980), “Long memory relationships and the aggregation of dynamic models”, *Journal of Econometrics,* Vol.14; 227-238

Granger, C.W.J. and Joyeux, R. (1980), “An introduction to long-memory time series models and fractional differencing”, *Journal of Time Series analysis*, Vol.1, No.1; 15-29

Greene, W.H. (2003), *Econometric Analysis*, 5ED, Prentice Hall, Upper Saddle River, New Jersey

Gustavsson, M. and Österholm, P. (2006), ”Hysteresis and non-linearities in unemployment rates”, *Applied Economics Letters*, Vol.13, No.9; 545-548

Hassler, U. and Wolters, J. (1994), “On the power of unit root tests against fractional alternatives”, *Economic Letters*, Vol.45; 1-5

Hauser, M.A. and Kunst, R.M. (1998), “Fractionally integrated models with ARCH errors: With an application to the Swiss 1-month Euromarket interest rate”, Review of quantitative finance and accounting, Vol.10; 95-113

Holl, J. and Kunst, R.M. (2011), “Unit root in unemployment – new evidence from nonparametric tests, *Applied Economics Letters*, Vol.18, No.6; 509-512

Hosking, J.R.M. (1981), “Fractional differencing”, *Biometrika*, Vol.68, No.1; 165-176

Jarque, C.M. and Bera, A.K (1987), “A test for normality of observations and regressions residuals”, *International Statistical Review*, Vol.55, No.2; 163-172

Koustas, Z. and Veloce, W. (1996), “Unemployment hysteresis in Canada: an approach based on long memory time series models”, *Applied Economics*, Vol.28; 823-831

Kurita, T. (2010), “A forecasting model for Japan´s unemployment rate”, Eurasian Journal of Business and Economics, Vol.3, No.5; 127-134

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P and Shin, Y. (1992), “Testing the null hypothesis of stationarity against the alternative of a unit root”, *Journal of Econometrics*, Vol.54; 159-178

Lee, C-F. (2010), “Testing for unemployment hysteresis in nonlinear heterogeneous panels: international evidence”, *Economic Modelling*, Vol.27; 1097-1102

Lee, C-C. and Chang, C-P. (2008), “Unemployment hysteresis in OECD countries: centurial time series evidence with structural breaks”, *Economic Modelling*, Vol.25; 312-325

Lemieux, T. and MacLeod, B.W. (2000), “Supply side hysteresis: the case of the Canadian unemployment insurance system”, *Journal of Public Economics*, Vol. 78; 139-170

León-Ledesma, M.A. (2002), “Unemployment hysteresis in the US stats and the EU: a panel approach”, *Bulletin of Economic Research*, Vol.54, No.2; 95-103

Leybourne, S. and Newbold, P. (1999), “On the size properties of Phillips-Perron tests”, *Journal of Time Series analysis*, Vol.20, No. 1; 51-61

Lin, C-H., Kuo, N-F. and Y,C-D. (2008), “Nonlinear vs. nonstationary of hysteresis in unemployment: evidence from OECD countries”, *Applied Economics Letters*, Vol.15, No.11; 905-909

Meyer, B.D. (1990), “Unemployment insurance and unemployment spells”, *Econometrica*, Vol. 58, No. 4; 757-782

Miller, I. and Miller, M. (2003), *John E. Freund’s mathematical statistics with applications*, 7ED, Pearson Education, Prentice Hall, New Jersey

Miller, R. (2008), “Global recession risk grows as U.S ‘damage’ spreads”, Bloomberg, http://www.bloomberg.com. Retrieved 2011/12/24

Mitchell, F.M. (1993), “Testing for unit roots and persistence in OECD unemployment rates”, *Applied Economics*, Vol.25; 1489-1501

OECD (2011a), Organization for Economic Cooperation and Development, In *OECD Statistics Database*. Database name: Key short-term economic indicators. Dataset: Harmonized unemployment rates: all persons, s.a. Parameters: Ratios. Retrieved at http://stats.oecd.org. Data generated: 2011/12/24

OECD (2011b), “Persistence of high unemployment: what risks? what policies?”, *OECD Economics Department Policy Notes*, No.5

Phelps, E.S. (1967), “Phillips curves, expectations of inflation and optimal unemployment over time”, *Economica*, Vol.34, No.135; 254-281

Roed, K. (1996), “Unemployment hysteresis – macro evidence from 16 OECD countries”, *Empirical Economics*, Vol. 21; 589-600

Romero-Ávila, D. and Usabiaga, C. (2007), “Unit root tests and persistence of unemployment: Spain vs. the United States”, *Applied Economic Letters*, Vol.14, No.6; 457-461

Romero-Ávila, D. and Usabiaga, C. (2009), “The unemployment paradigms revisited: a comparative analysis of US State and European unemployment”, *Contemporary Economic Policy*, Vol. 27, No.3; 321-334

Song, F.M. and Wu, Y. (1998), “Hysteresis in unemployment: evidence from OECD countries”, *The Quarterly Review of Economics and Finance*, Vol. 38, No.2; 181-192

Sowell, F.B. (1992), “Maximum likelihood estimation of stationary univariate fractionally-integrated time-series models”, *Journal of Econometrics*, Vol. 53; 165-188

Yajima, Y. (1985), “On estimation of long-memory time series models”, *Australian Journal of Statistics*, Vol. 3; 303-320

Yilanci, V. (2008), “Are unemployment rates nonstationary or nonlinear? Evidence from 19 OECD countries”, *Economics Bulletin*, Vol.3, No.47; 1-5

Yilanci, V. (2009), “Analyzing the unemployment hysteresis for Turkey under structural breaks”, *Doğuş Üniversitesi Dergisi,* Vol.10, No.2; 324-335

APPENDIX - ARFIMA estimates (Complete dataset)

AUSTRALIA 1978M01-2011M11

	AR-coefficients			**MA-coefficients**					**Normality**		**Autocorrelation**	
ARMA (p, q)	**1**	**2**	**3**	**1**	**2**	**3**	**d**	**AIC**	**Statistic**	**Prob.**	**Statistic**	**Prob.**
(0, 0)	-	-	-	-	-	-	1.01	94.22	0.84	0.66	9.09	0.97
(1, 0)	-0.27^a	-	-	-	-	-	1.22	93.00	0.23	0.89	10.99	0.89
(2, 0)	-0.36^a	-0.13	-	-	-	-	1.31	93.25	0.29	0.86	8.57	0.95
(3, 0)	-0.41^a	-0.19	-0.06	-	-	-	1.35	93.69	0.43	0.81	6.39	0.98
(0, 1)	-	-	-	-0.38^b	-	-	1.33	92.82	0.16	0.92	6.45	0.99
(1, 1)	-0.05	-	-	-0.32	-	-	1.32	93.29	0.32	0.85	7.62	0.97
(2, 1)	-0.22	-0.09	-	-0.17	-	-	1.33	93.72	0.41	0.82	7.37	0.97
(3, 1)	-0.64	-0.27	-0.09	0.24	-	-	1.34	94.17	0.46	0.79	7.06	0.96
(0, 2)	-	-	-	-0.38^b	0.03	-	1.32	93.29	0.43	0.81	8.03	0.97
(1, 2)	0.93^a	-	-	-0.82^b	0.13	-	0.83	92.97	0.39	0.82	12.72	0.69
(2, 2)	1.68^a	-0.84^a	-	-1.76^a	1.00^a	-	**1.01**	91.60	0.11	0.95	29.79	0.01
(3, 2)	-0.05	-0.90^a	-0.31^a	-0.23^a	1.00^a	-	1.23	93.22	0.12	0.94	11.49	0.65
(0, 3)	-	-	-	-0.38^b	-0.02	0.09	1.32	93.57	0.55	0.76	10.94	0.81
(1, 3)	0.88^a	-	-	-0.90^a	0.08	0.14	0.95	93.05	0.25	0.88	18.23	0.25
(2, 3)	1.39^a	-0.46	-	-1.31^a	0.46	0.09	0.85	93.31	0.16	0.92	21.52	0.09
(3, 3)	0.68^a	0.82^a	-0.84^a	-0.78^c	-0.74^c	0.99^c	1.01	92.48	0.17	0.92	29.94	0.00

Table A.1: ARFIMA-estimates, Australia 1978M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

BELGIUM: 1970M01-2011M11

	AR-coefficients			MA-coefficients					Normality		Autocorrelation	
ARMA (p, q)	**1**	**2**	**3**	**1**	**2**	**3**	**d**	**AIC**	**Statistic**	**Prob.**	**Statistic**	**Prob.**
(0, 0)	-	-	-	-	-	-	**1.22**	-45.20	12.92	0.00	12.37	0.93
(1, 0)	0.01	-	-	-	-	-	1.22	-44.80	20.00	0.00	12.54	0.90
(2, 0)	0.05	0.06	-	-	-	-	1.17	-44.58	20.49	0.00	11.89	0.89
(3, 0)	0.01	0.04	-0.07^{a}	-	-	-	1.22	-44.46	17.07	0.00	10.56	0.91
(0, 1)	-	-	-	0.01	-	-	1.22	-44.80	18.83	0.00	12.50	0.90
(1, 1)	0.23	-	-	-0.20	-	-	1.20	-44.42	20.11	0.00	12.80	0.85
(2, 1)	-0.20	0.08	-	0.25	-	-	1.17	-44.26	19.38	0.00	11.59	0.87
(3, 1)	0.32	0.03	-0.10	-0.38	-	-	1.28	-44.22	16.46	0.00	8.96	0.94
(0, 2)	-	-	-	0.08	0.09	-	1.15	-44.69	18.68	0.00	11.83	0.89
(1, 2)	-0.22	-	-	0.29	0.10	-	1.16	-44.36	19.14	0.00	11.46	0.87
(2, 2)	-0.42	-0.42	-	0.46	0.51	-	1.19	-44.18	20.44	0.00	10.69	0.87
(3, 2)	0.88^{b}	-0.47	-0.12^{c}	-0.97^{b}	0.54^{c}	-	1.30	-44.12	17.12	0.00	6.88	0.98
(0, 3)	-	-	-	-0.02	0.04	-0.09	1.24	-44.54	17.82	0.00	9.86	0.94
(1, 3)	0.24	-	-	-0.28	0.04	-0.10	1.26	-44.22	17.82	0.00	9.25	0.93
(2, 3)	1.04^{a}	-0.68^{b}	-	-1.11^{a}	0.75^{a}	-0.11	1.29	-44.05	17.93	0.00	7.41	0.96
(3, 3)	0.80	-0.30	-0.24	-0.88^{c}	0.37	0.12	1.30	-43.76	9.10	0.01	6.79	0.96

Table A.2: ARFIMA-estimates, Belgium 1970M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

CANADA: 1956M01-2011M11

	AR-coefficients			MA-coefficients					Normality		Autocorrelation	
ARMA (p, q)	1	2	3	1	2	3	d	AIC	Statistic	Prob.	Statistic	Prob.
(0, 0)	-	-	-	-	-	-	1.10	20.98	10.72	0.00	40.02	0.02
(1, 0)	-0.17^{a}	-	-	-	-	-	1.19	20.39	11.12	0.00	35.63	0.05
(2, 0)	-0.27^{a}	-0.12^{c}	-	-	-	-	1.29	20.21	10.76	0.00	30.51	0.11
(3, 0)	-0.21	-0.07	0.05	-	-	-	1.23	20.44	13.20	0.00	30.72	0.08
(0, 1)	-	-	-	-0.24^{b}	-	-	1.26	20.18	10.88	0.00	33.54	0.07
(1, 1)	0.02	-	-	-0.27	-	-	1.26	20.48	10.87	0.00	33.34	0.06
(2, 1)	-0.64	-0.18	-	0.40	-	-	1.25	20.46	9.43	0.01	30.71	0.08
(3, 1)	0.75^{a}	0.04	0.12	-0.33	-	-	0.58	19.81	12.96	0.00	27.34	0.13
(0, 2)	-	-	-	-0.24^{b}	-0.01	-	1.26	20.48	10.86	0.00	33.18	0.06
(1, 2)	-0.95^{a}	-	-	0.73^{a}	-0.27^{a}	-	**1.27**	19.30	8.93	0.01	28.85	0.12
(2, 2)	1.60^{a}	-0.66^{a}	-	-1.50^{a}	0.64^{a}	-	0.89	19.95	12.80	0.00	29.25	0.08
(3, 2)	1.48^{a}	-0.79	0.29	-0.81^{b}	0.42	-	0.33	20.01	11.46	0.00	28.54	0.07
(0, 3)	-	-	-	-0.19^{b}	-0.02	0.10^{b}	1.20	20.17	11.19	0.00	29.87	0.09
(1, 3)	0.95^{a}	-	-	-0.42	-0.04	0.08	0.47	19.73	12.87	0.00	27.56	0.12
(2, 3)	-0.01	0.89^{a}	-	0.46^{c}	-0.51^{b}	0.02	0.57	18.94	11.47	0.00	26.05	0.13
(3, 3)	0.55	0.87^{a}	-0.53	-0.31	-0.81^{a}	0.48	0.77	19.10	10.64	0.00	25.41	0.11

Table A.3: ARFIMA-estimates, Canada 1956M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

FRANCE: 1978M01-2011M11

ARMA (p, q)	AR-coefficients 1	2	3	MA-coefficients 1	2	3	d	AIC	Normality Statistic	Prob.	Autocorrelation Statistic	Prob.
(0, 0)	-	-	-	-	-	-	**1.35**	-9.16	2.72	0.26	24.31	0.18
(1, 0)	-0.17	-	-	-	-	-	1.42	-8.90	3.76	0.15	8.32	0.97
(2, 0)	-0.11	0.10	-	-	-	-	1.38	-8.49	3.78	0.15	16.35	0.50
(3, 0)	-0.12	0.10	-0.01	-	-	-	1.38	-8.00	2.65	0.27	15.99	0.45
(0, 1)	-	-	-	-0.14	-	-	1.41	-8.85	3.44	0.18	9.66	0.94
(1, 1)	-0.67	-	-	0.53	-	-	1.40	-8.52	3.77	0.15	14.67	0.62
(2, 1)	0.91^{a}	0.05	-	-0.60^{c}	-	-	0.91	-8.73	4.00	0.14	6.55	0.98
(3, 1)	0.94^{a}	0.05	-0.03	-0.62^{c}	-	-	0.90	-8.25	3.49	0.17	7.02	0.96
(0, 2)	-	-	-	-0.12	0.09	-	1.38	-8.46	3.91	0.14	14.17	0.65
(1, 2)	0.96^{a}	-	-	-0.65	0.03	-	0.91	-8.73	4.40	0.11	6.52	0.98
(2, 2)	0.25	0.68^{c}	-	0.02	-0.41	-	0.95	-8.30	5.92	0.05	8.07	0.92
(3, 2)	0.37	0.65	-0.07	-0.02	-0.37	-	0.88	-7.83	2.73	0.25	9.04	0.83
(0, 3)	-	-	-	-0.12	0.09	-0.02	1.38	-7.97	2.71	0.26	14.45	0.57
(1, 3)	-0.66	-	-	0.53	0.01	0.00	1.39	-7.53	3.99	0.14	15.41	0.42
(2, 3)	0.30	0.66	-	0.10	-0.33	-0.06	0.84	-7.84	3.02	0.22	9.95	0.77
(3, 3)	-0.01	0.01	0.92^{a}	0.41	0.41	-0.59^{c}	0.84	-8.71	3.09	0.21	8.84	0.79

Table A.4: ARFIMA-estimates, France 1978M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

NETHERLANDS 1970M01-2011M11

	AR-coefficients			MA-coefficients					Normality		Autocorrelation	
ARMA (p, q)	**1**	**2**	**3**	**1**	**2**	**3**	**d**	**AIC**	**Statistic**	**Prob.**	**Statistic**	**Prob.**
(0, 0)	-	-	-	-	-	-	1.21	-95.11	14.99	0.00	17.40	0.69
(1, 0)	-0.16^b	-	-	-	-	-	**1.30**	-95.68	20.04	0.00	11.80	0.92
(2, 0)	-0.17^c	-0.01	-	-	-	-	1.31	-95.29	19.94	0.00	11.56	0.90
(3, 0)	-0.11	0.04	0.07	-	-	-	1.25	-95.13	19.97	0.00	10.69	0.91
(0, 1)	-	-	-	-0.18^b	-	-	1.32	-95.63	18.84	0.00	11.61	0.93
(1, 1)	-0.13	-	-	-0.04	-	-	1.31	-95.29	19.99	0.00	11.68	0.90
(2, 1)	0.96^c	0.01	-	-0.26	-	-	0.42	-95.56	24.45	0.00	11.76	0.86
(3, 1)	1.14^c	-0.10	-0.08	-0.68^a	-	-	0.67	-95.24	24.84	0.00	11.77	0.81
(0, 2)	-	-	-	-0.15^c	0.04	-	1.29	-95.32	20.90	0.00	11.76	0.90
(1, 2)	0.92^a	-	-	-0.64	-0.01	-	0.85	-95.46	25.01	0.00	12.54	0.82
(2, 2)	-0.64^a	-0.79^a	-	0.55^a	0.78^a	-	1.25	-95.36	25.02	0.00	8.73	0.95
(3, 2)	0.06	0.92^a	-0.03	0.63	-0.29	-	0.44	-94.80	24.55	0.00	11.74	0.76
(0, 3)	-	-	-	-0.18	0.01	0.00	1.32	-94.86	22.29	0.00	11.55	0.87
(1, 3)	0.93^a	-	-	-0.63^c	0.01	-0.04	0.82	-95.18	25.02	0.00	12.20	0.79
(2, 3)	-0.63^a	-0.83^a	-	0.49^a	0.79^a	-0.06	1.28	-95.04	24.69	0.00	7.50	0.96
(3, 3)	0.30	-0.30^b	0.76^a	-0.22	0.44^a	-0.65^a	1.04	-95.20	24.59	0.00	7.26	0.95

Table A.5: ARFIMA-estimates, Netherlands 1970M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

SWEDEN: 1970M01-2011M11

	AR-coefficients			MA-coefficients					Normality		Autocorrelation	
ARMA (p, q)	1	2	3	1	2	3	d	AIC	Statistic	Prob.	Statistic	Prob.
(0, 0)	-	-	-	-	-	-	0.97	13.01	4.19	0.12	150.76	0.00
(1, 0)	-0.38^{a}	-	-	-	-	-	1.14	4.68	7.82	0.02	106.94	0.00
(2, 0)	-0.72^{a}	-0.45^{a}	-	-	-	-	**1.40**	-7.40	8.09	0.02	21.19	0.33
(3, 0)	-0.76^{a}	-0.49^{a}	-0.04	-	-	-	1.43	-7.07	8.37	0.02	20.30	0.32
(0, 1)	-	-	-	-0.58^{a}	-	-	1.36	1.11	7.48	0.02	82.48	0.00
(1, 1)	-0.22^{a}	-	-	-0.53^{a}	-	-	1.43	-0.78	8.18	0.02	64.81	0.00
(2, 1)	-0.68^{a}	-0.43^{a}	-	-0.08	-	-	1.43	-7.06	8.58	0.01	20.33	0.31
(3, 1)	-1.40^{a}	-0.96^{a}	-0.33^{b}	0.65^{b}	-	-	1.42	-6.78	8.33	0.02	19.94	0.28
(0, 2)	-	-	-	-0.84^{a}	0.26^{a}	-	1.44	-3.13	8.52	0.01	44.50	0.00
(1, 2)	0.94^{a}	-	-	-1.09^{a}	0.36^{a}	-	0.75	-5.03	6.68	0.04	45.32	0.00
(2, 2)	-0.66^{a}	-0.43^{a}	-	-0.09	0.03	-	1.42	-6.68	8.48	0.01	20.46	0.25
(3, 2)	0.27^{b}	0.24^{b}	0.39^{a}	-0.34	-0.08	-	0.72	-8.23	8.29	0.02	19.94	0.22
(0, 3)	-	-	-	-0.56^{a}	0.04	0.22^{a}	1.25	-5.28	8.60	0.01	37.74	0.00
(1, 3)	-0.41^{b}	-	-	-0.25^{c}	-0.20^{c}	0.27^{a}	1.33	-5.78	8.64	0.01	28.91	0.04
(2, 3)	-0.66^{a}	-0.44^{a}	-	-0.08	0.03	-0.01	1.42	-6.28	6.61	0.04	20.39	0.20
(3, 3)	-1.51^{a}	-1.13^{a}	-0.34^{a}	0.76^{a}	0.07	-0.14	1.44	-6.17	7.62	0.02	18.32	0.25

Table A.6: ARFIMA-estimates, Sweden 1970M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

UNITED KINGDOM: 1971M01-2011M11

	AR-coefficients			MA-coefficients					Normality		Autocorrelation	
ARMA (p, q)	**1**	**2**	**3**	**1**	**2**	**3**	**d**	**AIC**	**Statistic**	**Prob.**	**Statistic**	**Prob.**
(0, 0)	-	-	-	-	-	-	1.39	-84.21	6.12	0.05	33.96	0.04
(1, 0)	-0.21^{a}	-	-	-	-	-	1.47	-84.94	4.11	0.13	43.93	0.00
(2, 0)	-0.22^{a}	-0.01	-	-	-	-	1.47	-84.54	4.13	0.13	46.60	0.00
(3, 0)	-0.21^{a}	-0.01	0.01	-	-	-	1.47	-84.13	4.00	0.14	43.64	0.00
(0, 1)	-	-	-	-0.20^{a}	-	-	1.47	-84.86	4.46	0.11	50.54	0.00
(1, 1)	-0.18	-	-	-0.04	-	-	1.47	-84.54	4.14	0.13	46.20	0.00
(2, 1)	0.84^{a}	0.11	-	-0.62^{a}	-	-	1.00	-86.53	4.11	0.13	11.37	0.88
(3, 1)	0.76^{a}	0.09	0.09	-0.58^{a}	-	-	1.04	-86.30	3.99	0.14	9.87	0.91
(0, 2)	-	-	-	-0.22^{b}	0.05	-	1.47	-84.57	4.27	0.12	44.63	0.00
(1, 2)	0.94^{a}	-	-	-0.84^{a}	0.14	-	1.09	-86.64	5.49	0.06	10.44	0.92
(2, 2)	1.92^{a}	-0.93^{a}	-	-1.68^{a}	0.73^{a}	-	**0.96**	-87.42	4.05	0.13	11.74	0.82
(3, 2)	1.88^{a}	-0.85^{a}	-0.04	-1.70^{a}	0.74^{a}	-	1.01	-87.03	4.09	0.13	11.75	0.76
(0, 3)	-	-	-	-0.26^{a}	0.02	0.11	1.47	-84.52	4.35	0.11	42.56	0.00
(1, 3)	0.93^{a}	-	-	-0.85^{a}	0.08	0.09	1.11	-86.61	4.06	0.13	9.71	0.92
(2, 3)	1.93	-0.94^{a}	-	-1.78^{a}	0.87	-0.05	1.04	-87.03	5.23	0.07	11.75	0.76
(3, 3)	0.92^{a}	0.99^{a}	-0.93^{a}	-0.68^{a}	-0.95^{a}	0.72^{a}	0.96	-86.63	9.16	0.01	11.80	0.69

Table A.7: ARFIMA-estimates, United Kingdom 1971M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.

UNITED STATES 1955M01-2011M11

	AR-coefficients			MA-coefficients					Normality		Autocorrelation	
ARMA (p, q)	**1**	**2**	**3**	**1**	**2**	**3**	**d**	**AIC**	**Statistic**	**Prob.**	**Statistic**	**Prob.**
(0, 0)	-	-	-	-	-	-	1.18	-53.07	21.43	0.00	96.88	0.00
(1, 0)	-0.32^{a}	-	-	-	-	-	1.39	-57.18	24.06	0.00	50.23	0.00
(2, 0)	0.57^{a}	0.35^{a}	-	-	-	-	**0.49**	-59.46	24.15	0.00	43.77	0.01
(3, 0)	0.47^{a}	0.34^{a}	0.07	-	-	-	0.57	-59.37	20.61	0.00	43.27	0.00
(0, 1)	-	-	-	-0.30^{a}	-	-	1.40	-55.92	23.80	0.00	64.82	0.00
(1, 1)	-0.37^{a}	-	-	0.07	-	-	1.37	-56.91	24.13	0.00	50.26	0.00
(2, 1)	0.66^{a}	0.23^{a}	-	-0.27	-	-	0.65	-59.46	20.63	0.00	42.75	0.01
(3, 1)	0.67^{a}	0.22^{c}	0.00	-0.28	-	-	0.65	-59.17	20.67	0.00	42.72	0.00
(0, 2)	-	-	-	-0.26^{a}	0.12^{a}	-	1.32	-57.05	24.46	0.00	50.31	0.00
(1, 2)	0.91^{a}	-	-	-0.51^{a}	0.12^{a}	-	0.64	-59.44	20.86	0.00	43.62	0.00
(2, 2)	0.57	0.32	-	-0.18	-0.05	-	0.66	-59.17	17.69	0.00	42.52	0.00
(3, 2)	0.21	0.55	0.08	0.18	-0.15	-	0.66	-58.89	15.91	0.00	42.34	0.00
(0, 3)	-	-	-	-0.24^{a}	0.12^{a}	0.02	1.30	-56.81	24.89	0.00	50.69	0.00
(1, 3)	0.91^{a}	-	-	-0.51^{a}	0.12^{a}	0.00	0.64	-59.15	20.98	0.00	43.64	0.00
(2, 3)	0.32^{c}	0.54	-	0.07	-0.18	0.04	0.66	-58.94	17.64	0.00	42.11	0.00
(3, 3)	1.21^{a}	0.27	-0.50^{b}	-0.41	-0.30	0.12	0.23	-59.60	19.48	0.00	39.81	0.00

Table A.8: ARFIMA-estimates, United Kingdom 1955M01 - 2011M11.
a = 1% significance, *b* = 5% significance, *c* = 10% significance.

Remark: The Jarque-Bera normality tests are conducted after extreme outliers have been removed from the data material. The null hypothesis in the JB test is that the residuals belong to a normal distribution. In the Breusch-Godfrey test for autocorrelation, the null hypothesis is that there is no autocorrelation in the residuals.