Erasmus School of Economics

The Stability of the Phillips Curve
Evidence from the United States and the United Kingdom

Raihan Pradhana – 479856

Master Thesis Policy Economics (2017–2018)

Dr Agnieszka Markiewicz

24 August 2018

Abstract

An analysis of data from the United States and the United Kingdom supports the view that the traditional Phillips curve is unstable over time. The first stability test suggests that there is weak evidence of stability in the inflation-unemployment trade-off. Three out of four stability tests suggest that the Phillips curve is stable over the examined time period. However, the out-sample dynamic simulation test fails to predict an inflation pattern during the Global Financial Crisis. The second stability test strongly rejects the evidence of symmetry in the inflation-output relationship. The L-shaped Phillips equation outperforms the other non-linear and linear models, despite a very small difference in the R-squared and the Log-Likelihood function. The failure in the Phillips curve throughout the examined time period illustrates that relying on the statistical relationships between inflation and unemployment/output is not enough to formulate a proper monetary policy, especially during large economic shocks.

Keywords: Traditional Phillips curve, Inflation, Output Gap, Unemployment, Stability, Non-linearity, Asymmetry, United States, United Kingdom

Acknowledgments

I would like to sincerely acknowledge and thank the following people who have supported me, not only during the course of this thesis, but throughout my Master's programme.

Firstly, I would like to thank Ibu, Ayah, Raisa and Jenny, for their continuous support during my four years of studying abroad. This entire journey would have been impossible to complete without your prayers and encouragement. I would also like to thank my supervisor, Dr. Agnieszka Markiewicz, for her unwavering support and input. I am extremely grateful for her invaluable guidance and her ideas for improving the output of this research.

I would also like to acknowledge the academic and technical support that I received from Erasmus University Rotterdam. The computer facilities and the study space where I spent my days working on this thesis were indispensable.

Finally, I would like to thank all of my close friends and classmates, as well as the entire Indonesian community in Rotterdam, for your immense support and encouragement. You have all helped me to focus on what has been a hugely rewarding and enriching process.

Content

Chapter 1: Introduction

The dynamics of inflation have been the subject of many economic debates over the past 60 years. The economic literature in this area is often associated with an analysis of the Phillips curve, which explains the empirical relationship between inflation and unemployment. This empirical relationship was first discovered by William Phillips in 1958, where he found evidence that an increased level of employment within an economy will correspond to a higher rate in the wage increase in the short run.

Although William Phillips himself did not postulate the relationship between the unemployment and inflation rate, the notion was later made popular by Milton Friedman in 1967. The essence of the Friedman's Phillips curve is that the expected rate of change in overall price level depends upon the rate on change in nominal wages. Labour wages need to keep pace with inflation and unemployment rate that is close to the natural rate. This leads to an interdependent relationship between the rate of change in output and inflation. This is the basis of the modern Phillips curve that has been widely used in many macroeconomic literatures in the past 60 years.

Until the mid-1970s, it has been widely believed that the relationship between inflation and unemployment was permanently stable. As a result, the Central Bank, in many developed countries, practiced its dual mandate through direct intervention to the economy. More specifically, in a case where the unemployment rate is high, the monetary policy could be used as an instrument to stimulate the economy, thus lowering the rate of unemployment. If the Phillips equation were stable, this would correspond to a higher rate of inflation with an equally significant reduction in the unemployment rate.[1] However, this is not the case when the trade-off between inflation and unemployment is unstable.

According to Robert Lucas (1978), using a purely empirical relationship with highly aggregated historical data is often problematic and inappropriate for a policy evaluation. Such relationships are often not structural, as they are assumed to be constant, with respect to changes in government policies. A policy evaluation based on an empirical relationship

[1] Implying a stable and linear relationship between inflation and unemployment.

between two corresponding variables will often lead to a misleading policy implication, if it is not backed by proper economic theory.

In the context of the Phillips curve, Lucas argued that the agents' expectations of future events may depend on the behaviour of other agents in the economy. The estimated Phillips curve implies a particular trend for inflation, which ultimately depends on the historical trend of the variable. However, if economic agents understand the implications of an inflation altering policy behaviour, then presumably their cost of living will adjust according to their expectations. This will disrupt the structural relationship between the two underlying variables, and, as a result, will become an inappropriate predictor of inflation. According to the Lucas critique, the Phillips curve is unstable.

The primary aim of this thesis is to test the hypothesis of a stable Phillips curve in the United States and the United Kingdom. This thesis extends the time period tested in much of the seminal literature from the 1990s (Fuhrer 1995 & Fitzgerald et al. 2013), which failed to consider the effect of the global financial crisis in 2008. Prior to the 1990s, several investigations on the stability of the curve discovered a stable relationship between inflation and unemployment. The relationship was found to be alive and well in Jeffrey Fuhrer's (1995) study on US data from 1960 to 1995. The relationship was found to be stable during the 1977's oil shock and did not convey instability throughout the examined time period.

On the contrary, most studies that tested the stability through the linearity approach have found evidence of instability in the trade-off. Most papers that evaluate the shape of the Phillips curve tend towards the conclusion that there is non-linearity and asymmetry in the Phillips equation. These investigations compare an extensive form of the Phillips equation with the conventional Phillips curve model to evaluate the shape of the inflation-unemployment relationships within the examined country. Studies comparing the flexible approximation of the equation with the conventional Phillips model (Laxton et al., 1994; Razzak, 1997; De Veirmann, 2007; Chadha, Masson and Meredith (1992)[2], 1992; Wimanda et al., 2011) found evidence of non-linearity and asymmetry in the Phillips equation[3].

---

[2] Referred to as CMM (1992) from this point onwards.

[3] Note that instability implies asymmetry, but not necessarily non-linearity. Asymmetry means that positive demand shocks tend to increase inflation by more than the negative demand shocks of a similar magnitude that would reduce it.

The tested Phillips curves are based on the traditional form, where the equation only employs the backward-looking expectation of inflation, as well as the unemployment variable. Testing the stability requires two separate approaches. In the first test of stability, the Phillips curve's construction is based on the inflation-unemployment relationship. The relationship is tested according to the in-sample and out-sample test, which employs both the sample fit and dynamic simulation approach. The second test of stability involves determining the correct shape of the Phillips curve. There are four functional forms that will be used to describe the Phillips curve: linear, quadratic, cubic and L-shaped. The non-linear shape of the Phillips curve is compared with the benchmark linear function to determine the shape that best describes the inflation-output relationships. As shown in later chapters, several results stand out and appear to be quite robust.

This thesis will contribute to the estimation of the traditional Phillips curve for the United States (US) and the United Kingdom (UK). The popularity of testing the stability and the shape of the Phillips equation has been diminishing since the 1990s. This thesis is an extension of the results found in the 1990s Phillips curve literature. During the mid-1980s recessions, the failure of the monetary policy to contain the so-called stagflation raised some scepticism regarding the use of the Phillips equation in designing monetary policy. Designing monetary policy during that time required careful attention to prevent undesirable long-run consequences. As a result, plenty of studies have emphasised the topic of the stability of the Phillips curve.

The number of studies regarding the linearity and the shape of the Phillips curve has been limited over the past few years. Moreover, there has been less emphasis on this research subject during the global financial crisis. This is perhaps due to the scale of the monetary failure in the current financial crisis that was comparatively smaller than the failure during the oil crisis of 1977. Consequently, this lack of research during the past decade is the main motivation behind this thesis.

The first stability test is mainly inspired from the results found in Fuhrer (1995). Fuhrer discovered a stable relationship between inflation and unemployment during the oil crisis of 1977. However, this stability test did not extend beyond 1995. This thesis extends the time period examined by Fuhrer and replicates the method used to evaluate the stability in the Phillips equation. In particular, it is interesting to determine whether the Phillips relationship was stable during the most recent financial crisis, since there has been plenty of evidence of the instability in the relationship during that time.

The second stability test was inspired by Laxton et al. (1994), Razzak (1997) and CMM (1992). Few studies determined the shape of the Phillips curve after the 1990s. Tests for linearity, especially after the global financial crisis, will assist policymakers in designing an appropriate policy by considering the current shape of the Phillips curve. Moreover, the shape of the Phillips curve, after the recent crisis, might have changed due to shock. As a result, the older literature might have little relevance in explaining the shape of the Phillips curve today.

The rest of this thesis is organized as follows. Chapter 2 discusses the theoretical foundation of this thesis. Chapter 3 sheds light on the type of data used to describe inflation, unemployment and the output gap in the US and the UK. Chapter 3 also provides brief discussion regarding the trend and anomalies in the data series. Chapter 4 discusses the methodologies and the model selected in evaluating the stability and shape of the Phillips curve. Finally, in Chapter 5, the obtained results are presented. The conclusions are accordingly derived in Chapter 6.

Chapter 2: Theoretical Framework

In the following section, conflicting arguments regarding the Phillips curve's stability are discussed. Several articles have found evidence of stability in the Phillips equation over time. Fuhrer (1995), for instance, found clear evidence of stability in the inflation-unemployment relationships in the US data. Fuhrer (1995) employs the in-sample and out-sample fit test, as well as the dynamic simulation test, to evaluate the stability. Both tests found evidence of stability, where the Phillips curve estimated on and off the data did not illustrate any sign of getting off track from the actual inflation pattern.

The literature also employs the breakpoint test to test the hypothesis that the estimated coefficients on lagged inflation shifted across time when there was a structural shock. The coefficient stability test indicates no sign of instability for a variety of breakpoints centred around the 1980s. Fitzgerald et al. (2013) found similar evidence. They argue that using US national data to evaluate the stability of the inflation-unemployment relationship is insufficient. National data on inflation and unemployment provides minimal evidence of stability, as the Central Bank is actively pursuing its dual mandate and, at the same time, dampening the structural links between the trade-offs. That being said, the regional Phillips curve is found to be stable, as it is immune to the national scale policy alterations. An analysis of the regional data shows that the estimates of the Phillips curve trade-offs have remained relatively stable over the examined periods of time. The results also indicate a consistent negative relationship between inflation and unemployment within the regional data.

The opposing view regarding the stability of the Phillips curve trade-off has also been discussed and evaluated using a completely different technique. The most widely used technique to argue against stability is testing the linearity and symmetry of the Phillips equation. Several studies have used a flexible approximation of the non-linear function, where the functions have been compared with respect to the linear approximation.

CMM (1992) modelled inflation for seven economies from 1966 to 1988. They constructed a Phillips curve such that the inflation-output trade-off became increasingly steep as the output gap approached the upper limit of the achievable output number in the short-run. This is known as the L-shaped function. The inclusion of such a variable in the Phillips equation would imply asymmetry and instability. By using the estimated hybrid form of the Phillips curve, they found evidence of the L-shaped function explaining the inflation-output gap trade-off in the Group of Seven economies.

As an extension of CMM (1992), Laxton et al. (1994) employed a similar method in evaluating the shape of the Phillips curve. They compared the linear approximation with the L-shaped approximation using the Log-likelihood function (LLF) as a mode of statistical comparison. They found that the value of the LLF is maximised when the L-shaped function is employed, implying instability in the inflation-output trade-off. In addition to the L-shaped function, they also tested the quadratic and cubic function to further test the evidence that the instability in the Phillips curve is not due to a misspecification in the L-shaped function. The test confirmed this hypothesis.

Wimanda et al. (2012) evaluated the determinants of inflation and the shape of the Phillips curve in Indonesia by conducting a statistical comparison between the linear and the non-linear models. The non-linear model includes the Quadratic, Kink-1, Kink-2, L-shaped, Trend and Variance functions. After comparing the LLF and the R-squared results, they found evidence of a non-linear and asymmetric relationship, where the L-shaped function outperforms the linear model. This implies instability in the Phillips curve trade-off in the Indonesian data over the examined time period.

Razzak (1997) found the relationship between output and inflation in New Zealand to be asymmetric. He estimated the Phillips curve using the generalized method of moments (GMM) with a two linear segment kinked function. He then tested the hypothesis that the coefficient of the kinked function is zero, where the rejection of the hypothesis would imply asymmetry in the Phillips curve. The resulting estimate rejected the hypothesis, implying instability in the inflation-output relationship.

A series of other literatures like De Veirmann (2007), Ball-Mankiw-Romer (1998) and Dotsey-King-Wolman (1999) tested the hypothesis of the flat Phillips curve. More specifically, they evaluated the declining trend in inflation created by a situation where the price level was sticky, which, in turn, caused the Phillips curve to flatten. A flatter Phillips curve implies that a decline in inflation would result in a proportionally larger reduction in output than the linear model. Evidence from the Japanese economy in the late 1990s tends to support this hypothesis, where the inflation-output relationship is flatter than the standard linear model. The model implies that, in such an economy, the output effects of disinflation are much larger, in comparison to the linear model.

Chapter 3: Description of the Data Series

This chapter discusses the variables included in the constructed Phillips equation and the stationarity test. The data series consists of the quarterly figures of both the US and the UK's percentage change in the GDP deflator, the unemployment rate and the output gap. All data were seasonally adjusted and limited to the 1975–2016 period.

Table 3 presents some of the basic descriptive statistics of the data. One interesting pattern is that the standard deviation of the UK series is larger than that of the US. This implies that, over time, inflation, unemployment and the output gap in the UK are more volatile than in the US. The following section will describe the data series used in more detail. The Augmented Dicky-Fuller test will be used to test for stationarity.

Table 3

	The United States		The United Kingdom	
Data	Mean	Standard Deviation	Mean	Standard Deviation
GDP Deflator (Percentage Change %)	0.790	0.566	1.258	1.448
Unemployment Rate (%)	6.452	1.575	5.476	2.653
Output Gap (%)	-0.026	1.363	-0.030	1.387

There are two patterns that emerge from the data. Firstly, both countries suffered from the impact of the oil crisis of the mid-1970s and the global financial crisis of 2008. This similarity makes it easier to conduct a comparative analysis of the Phillips curve stability across countries, since they both faced a similar exogenous shock. Secondly, the UK data starts from 1960, while the US data starts from 1947. To obtain an appropriate comparison, the data was trimmed to start from 1975 for both countries. Reducing the observed data period could also simplify the stability test without having to reduce the accuracy of the estimation. This is a slight extension of the Fuhrer (1995) estimation, who accounts for the Phillips curve from 1960 to 1995.

3.1 GDP Deflator

The implicit price deflator, or better known as the GDP deflator, is a price metric that converts the measured output at current prices into a constant dollar GDP. The GDP deflator measures the extent to which the price level fluctuates within the economy. This includes the prices of consumer, business and government goods that are purchased. The choice of the base year is arbitrary, and the price level will be normalised to one during that specific year. The GDP deflator is calculated as the ratio of seasonally adjusted nominal GDP to real GDP.

In this thesis, the percentage change in the GDP deflator is used as a measure for the inflation level[4]. The data for the US GDP deflator is collected from the US Bureau of Economic Analysis, while the UK GDP deflator is collected from the Organization for Economic Co-operation and Development (OECD). The percentage change in the GDP deflator is calculated in terms of quarterly annualised rates. The percentage change during the current time period is calculated with respect to its previous quarter, where the growth rate is then annualised. This is consistent with the price metric used in Fuhrer (1995), CMM (1992) and Laxton (1994), where the GDP deflator, as an inflation measure, is also considered.

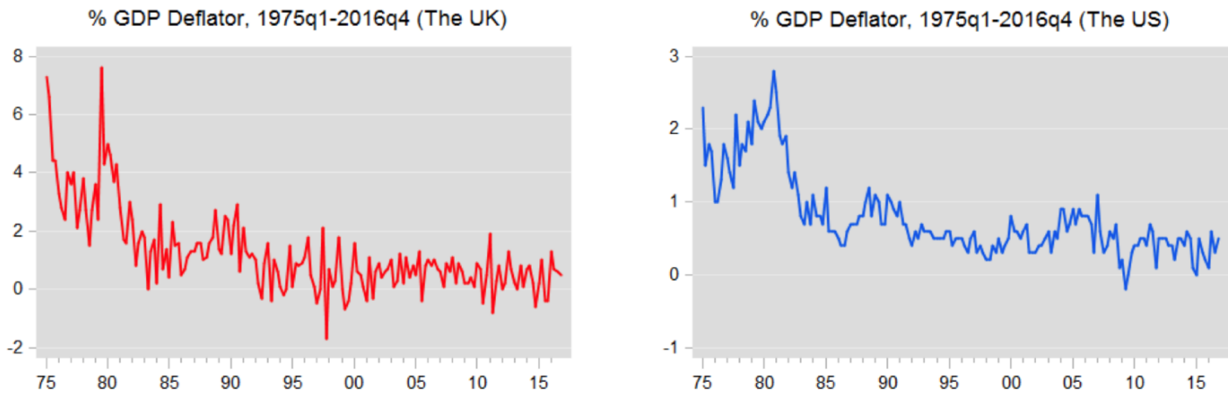
For the sake of simplicity, the consumer price index (CPI) and the core CPI will not be considered in this thesis, as opposed to Fuhrer (1995). The two versions of the Phillips curve provide a qualitatively similar result. While the GDP deflator is the weighted average price of final goods produced in the economy, the CPI employs the weighted average of the price of the consumer goods. The similarity between the two measures is that both consider the basket of goods to be fixed over time. However, the CPI provides a higher measure of inflation relative to the GDP deflator. The CPI is also more volatile. The standard deviation of the CPI inflation is usually much higher in any time series, in comparison to the GDP deflator inflation.

The inflation figure for the US and the UK is particularly interesting during the examined periods of 1975–2016. Regardless of the similarity in their economic system, the

[4] The percentage change in the GDP deflator will simply be referred to as the GDP deflator from now on.

inflation dynamics are quite different in the two countries. Figure 3.1 represents the GDP deflator of the UK and the US respectively.

Figure 3.1



The inflation pattern in the US data can be divided into three sub-periods. In the first time period (i.e., mid-1970s to mid-1980s), the US economy was characterised by high inflation and low economic growth. The GDP deflator peaked in 1980, when the economy fell into a deep recession. The second time period (i.e., mid-1980s to late-1990s) was characterised by another episode of high inflation. Inflation started out relatively low in 1986; however, the early crisis of the 1990s triggered yet another series of high inflation rates. The economy turned in an increasingly healthy performance as the 1990s progressed, while inflation remained at a much more stable level. The final time period, that covers the 2000s, was characterised by uncertainty and deflationary pressure. Inflation peaked in 2007 and declined steadily over the next seven years. Due to the impact of the financial crisis, along with the failure of a conventional monetary policy in stimulating growth, inflation remained below 0% in 2008.

On the other hand, the inflation pattern in the UK can be divided into four sub-periods. In the first period, the economy was generally considered to be facing severe structural problems. Inflation in the late 1970s almost reached double figure due to the failure in expansionary fiscal policy in stimulating growth. These rates of inflation, were at the time, costly to the economy and led to an inevitable recession in the early 1980s. In the second period, inflation is brought under control with tighter monetary policy and stricter policy on money supply growth. The economy enjoyed a period of low inflation during the Lawson boom of the late 1980s, but it fell into another recession in the early 1990s due to an unsustainable economic growth. Inflation is once again, increasing, but not at a significant rate. The third period that covers mid-1990s to 2007 was a period known as the great moderation. Inflation was relatively low within the period. The final period that covers 2008 to 2016 was associated with a cost push inflation. This is due to the devaluation of the British pound and rising commodity prices in the international market.

3.1.1 Unit Root Test of the GDP Deflator

One of the assumptions of the distributed lag model (e.g., Phillips curve) is that the outcome and explanatory variables are stationary random variables. It is possible to estimate a regression for two non-stationary variables and establish a statistically significant relationship. However, such relationships are invalid and said to be spurious. Hence, a non-stationary time series random variable needs to be converted into a stationary random variable to minimise such a problem.

To test for stationarity, it is necessary to conduct a unit root test for the underlying variable. In this case, the Augmented Dickey-Fuller test (ADF) is used under the null hypothesis that the GDP deflator has a unit root, where the lag length is determined by the Schwarz Info Criterion (SIC). The intercept is included in the test equation for each country. The results have been reported in Table 3.1.

Table 3.1

Country	Augmented Dicky-Fuller (ADF) test		Integrated of Order
	t-statistic	p-value	
The United States	-3.76431	0.004	I(0)
The United Kingdom	-3.391637	0.0126	I(0)

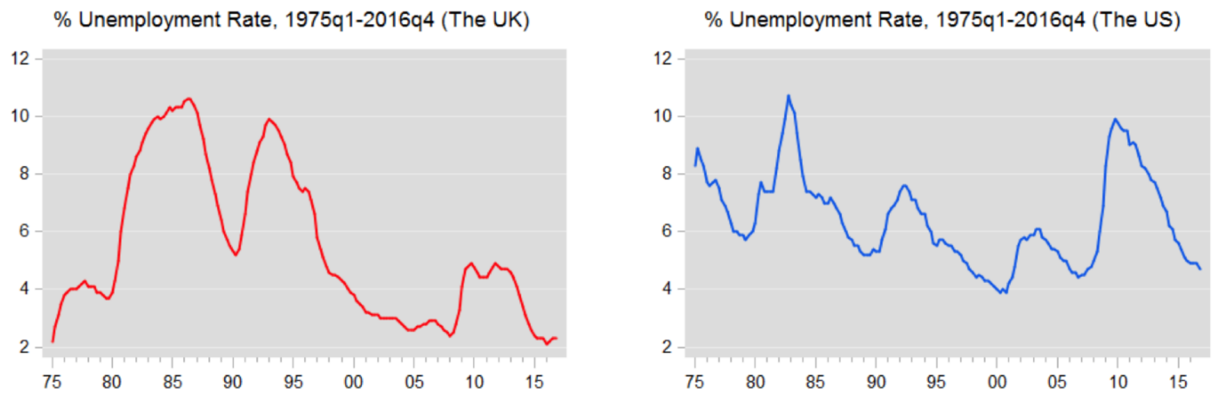
As can be seen from the table above, the ADF test for both the UK and the US rejects the hypothesis that there are unit roots in the series. Thus, it can be concluded that the GDP deflator is integrated of order 0.

3.2 Unemployment Rate

In general, the unemployment rate refers to the number of unemployed individuals, as a percentage of the labour force. In this thesis, two types of unemployment measures have been used to construct the Phillips equation. For the US data, the seasonally adjusted civilian unemployment rate is used. The civilian unemployment rate essentially follows the technical definition of the unemployment rate, which is calculated as the ratio of the number of unemployed people in the labour force and the total size of the labour force. The US unemployment series has been collected from the Bureau of Labor Statistics (BLS), which sub-categorised the data based on the age of the individuals, as well as the race. The BLS collects data on the civilian unemployment rate of all individuals, categorizing them based on ethnic origins, as well as gender and age-specific unemployment in the US.

For the UK data, the seasonally adjusted registered unemployment rate is used. Similar to the civilian unemployment rate, the registered rate of unemployment is calculated as the ratio of the number of unemployed people in the labour force to the total size of the labour force. The UK series has been obtained from the data gathered by the OECD.

Figure 3.2



The unemployment figures for the US and the UK are particularly interesting during the 1975–2016 time period. The unemployment dynamics are quite similar in the two countries, with the exception of the global financial crisis period, where the unemployment rate soared higher in the US. Figure 3.2 demonstrates the GDP deflator of the UK and the US respectively.

The unemployment pattern in the US data is divided into three sub-periods. In the first time period (i.e., the mid-1970s to late-1980s), unemployment was declining, due to the massive economic expansion regarded as the long boom. However, the increasing federal funds rate in the early 1980s caused setbacks in the manufacturing sector, which, in turn, triggered massive job losses. Despite a slight recovery, the unemployment rate continued to soar. It reached 10.8% in 1982, its highest level since the Great Depression. The unemployment rate declined steadily between 1982 and the early 1990s.

The second time period (i.e., the 1990s) was characterised by a decade of economic prosperity. However, job creation remained weak throughout late 1992. A turn came in 1994, when the unemployment rate declined steadily throughout the rest of the decade, due to the dot-com bubble. During the final time period (i.e., the 2000s), the US economy faced a prolonged recessionary period. The unemployment rate soared in the early part of the decade, but fell quickly in 2005. The financial crisis triggered another dramatic rise in unemployment in 2008, reaching a rate of 9.9%, at the highest.

The unemployment pattern in the UK is also divided into three sub-periods. In the first time period (i.e., the mid-1970s to the mid-1980s), unemployment started out at a relatively high level, with many jobs being lost during the recession of 1975. As the economy was hit by another recession in the early 1980s, unemployment continued to soar, until the economic boom during the second half of the 1980s. The majority of these job losses were due to the decline in the demand for products from the heavy industrial sector.

The unemployment rate remained high until the second time period (i.e., the mid-1980s to the late 1990s). The economic boom during the second half of 1980 pulled down the unemployment rate to 7%. However, the economy entered another recessionary period and the unemployment rate soared to almost 10% in 1994. The economic situation improved after 1994. Unemployment was kept under control in 1997. The final time period (i.e., the 2000s) was characterised by a much lower rate of unemployment. The unemployment rate was kept below 5% throughout the decade, with the exception of the financial crisis period.

3.2.1 Unit Root Test of the Unemployment Rate

To test for stationarity, a unit root test of the series is required. The Augmented Dickey-Fuller test (ADF) can be used under the null hypothesis that the unemployment rate has a unit root, where the lag length is determined by the Schwarz Info Criterion (SIC). The intercept is included in the test equation of each country. The results have been reported in Table 3.2.

Table 3.2

Country	Augmented Dicky-Fuller (ADF) test		Integrated of Order
	t-statistic	p-value	
The United States	-3.086	0.030	I(0)
The United Kingdom	-2.759	0.066	I(0)

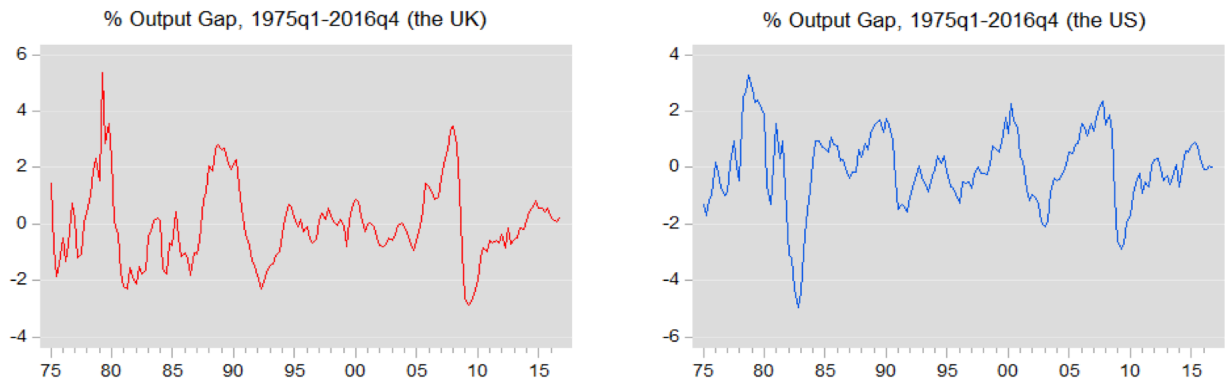
Table 3.2 illustrates that the ADF test for both the UK and the US rejects the hypothesis that there are unit roots in the series. Thus, it can be concluded that the unemployment rate is integrated of order 0.

3.3 Output Gap

The output gap is defined as the difference between the actual and potential output of an economy. The level of output gap that reaches the potential output level is achieved when the maximum amount of goods and services produced is at its most efficient level. That is, the state in which the potential output reaches its full capacity. In most cases, and similarly in this thesis, the output is measured by the real GDP. The seasonally adjusted real GDP data for the US is collected from the US Bureau of Economic Analysis. The real GDP is measured in billion dollars, chained with inflation in 2009. The seasonally adjusted real GDP data for the UK is collected from Eurostat, where real GDP is measured in million dollars, chained with inflation in 2010.

Similar to real GDP, the output gap can move in two different directions: positive and negative. A positive output gap occurs when the actual output exceeds its full capacity, indicating a high demand for goods and services in the economy. This might be considered beneficial in the aggregate. However, workers and businesses must operate far beyond their most efficient capacity to satisfy this huge demand. This would consequently spur inflation due to the increasing costs of labour. On the other hand, a negative output gap occurs when the actual output falls short of its full capacity. A negative gap indicates a low demand for goods and services in the economy, which implies that workers and businesses are operating below their maximum efficiency level. This, in turn, creates slack in the labour market (IMF, 2013). These two premises, which will be tested further in Chapter 5, are the building blocks of this thesis.

Figure 3.3



The Hodrick-Prescott filter (HP filter) is widely used to compute the output gap. The HP filter is similar to the approach that CMM (1992), Wimanda et al. (2012) and Razzak (1997) used in computing the output gap. The series of the computed output gap has been displayed in Figure 3.3. This figure illustrates that the US output gap mirrors the UK output gap closely throughout the examined time period, indicating that the real GDP in both countries is moving in a similar direction. However, during the early 1980s recessions, the US output gap fell much lower, in comparison to the UK's, indicating that the recession, triggered by the oil crisis, influenced real GDP much more significantly in the US. On the contrary, during the global financial crisis of 2008, the output gap for both countries fell to -3%. This trend implies that the current financial crisis resulted in a similar reduction in the real GDP in both countries, unlike the 1980s oil shock.

The output gap in the constructed Phillips curve is computed by applying the HP filter. The output gap is assumed to be stationary in the HP filter approach. Therefore, a unit root test for this variable is not required.

Chapter 4: Methodology and Model Selection

This chapter explores the methodology in more detail. To evaluate the stability of the Phillips curve, this thesis considered two distinct approaches. In the first approach, the Phillips curve was modelled based on the inflation-unemployment relationship. The stability of the Phillips curve was then evaluated based on the in- and out-of-sample test inspired by the methodology used in Fuhrer (1995). In the second approach, the Phillips curve was modelled based on the inflation-output gap relationship. The stability test was then based on the linearity test, where the benchmark linear model was compared with various non-linear and asymmetric approximations.

As opposed to some of the literature that has been considered for this thesis, the model used throughout this thesis did not include a forward-looking expectation variable of inflation[5]. As the primary aim of this thesis is to test the stability of the traditional Phillips curve, including a forward-looking variable would defeat the main purpose of the research. Moreover, the exclusion of the forward-looking variable made the study easier, as the dimension of the problem was lower.

4.1 Model Estimation

The Phillips curve employed the Autoregressive Distributed Lag (ARDL) model to express the Phillips curve trade-off. The model is autoregressive in the sense that inflation is explained, in part, by several lagged values of itself and the distributed lag component that takes the form of the successive lags of the unemployment variable. Despite the simple appearance, the ARDL model is categorised as an infinite lag model, that is, the ARDL(p, q) model is flexible enough to approximate any shape of an infinite lag distribution with sufficiently large values of p and q. The ARDL model of the Phillips curve is estimated by a least squares estimation, assuming conditional mean independence on the error term holds.

The number of lags chosen in the model were different from the number of lags used in the literature mentioned in Chapter 2. When the true model is not known, there is a risk of over or under-specifying the model by adding too many or too few lags. In this case, the number

[5] See Calvo (1983), Rudd and Whelan (2005), Roeger and Herz (2012)

of lags of inflation and unemployment was determined using the Akaike Information Criterion (AIC). Therefore, the combination of lags of inflation and unemployment that gives the lowest AIC value is the best model for the Phillips curve. The criterion attempts to balance the information gained when adding additional lags to the increase in the complexity of the model that the lags cause. This information criterion is designed specifically for model selection process. In most cases, a more parsimonious model is preferred, provided that the candidate models are correctly specified.

The AIC apparently does not always result in the best possible model. Davidson and McKinnon (2004) indicate that whenever two or more models are nested, the AIC may fail to choose the most parsimonious one. In another case, if two or more models are non-nested and only one is well specified, the AIC chooses the well specified model asymptotically, given that this model has the largest value of the log-likelihood. The Bayesian Information Criterion (BIC) is referred to avoid the problem discussed previously. However, Stock and Watson (2007) suggest that the AIC is much better, in terms of model selection, given that including parameters are, in many ways, better than omitting important and significant parameters.

4.2 Stability Test 1 – The Fit and Dynamic Simulation Test

The Lucas critique argues that the Phillips curve will not be stable in the long run, particularly as the behaviour of the monetary authority changes. To test this assumption, consider the first test of stability. The tests were structured accordingly, beginning with the least stringent test and going on to the more complex tests. In Chapter 2, it was shown that several papers have found an instability in the relationship between inflation and unemployment in the short run; however, this is not the case in Fuhrer (1995). The main reason why the method used by Fuhrer is employed is because it is the simplest and least stringent method to use and because they found evidence of stability in the inflation-unemployment relationships. The test results will be compared with that of Fuhrer (1995) and will be used to examine whether or not the stability found in Fuhrer extends to the 2000s' data. The same test will also be conducted within the UK data series.

4.2.1 Model Specification 1

The basic specification in this test follows the ‘price-price’ model, as proposed by Fuhrer (1995). This model explains the relationship between inflation and the unemployment rate. The specifications for the US and the UK Phillips curves are presented in Equations (1) and (2), respectively.

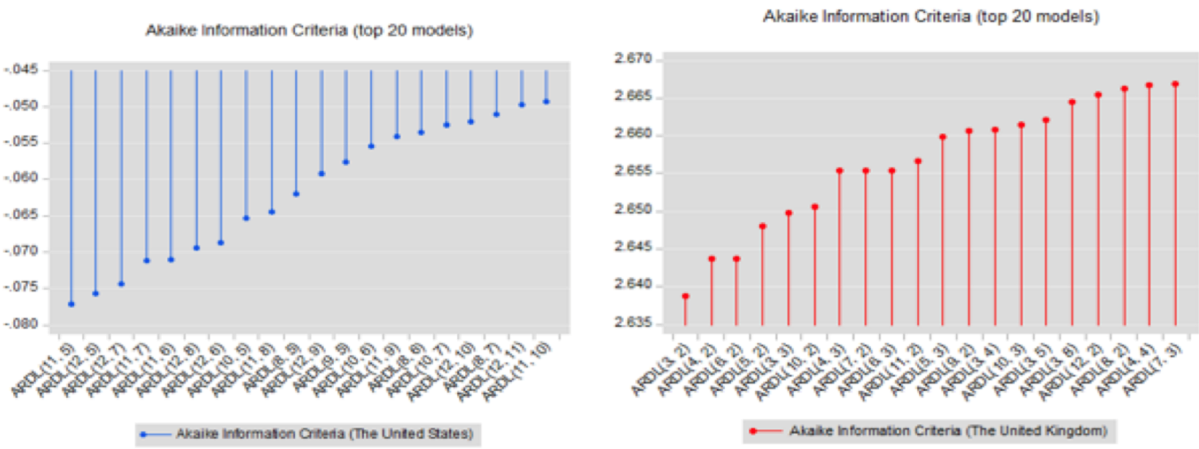
$$\pi_t = a_t + \sum_{i=1}^{11} \beta_i (\pi_{t-i}) + \sum_{j=1}^{5} \rho_j (U_{t-j}) + \epsilon_t \quad (1)$$

$$\pi_t = c_t + \sum_{i=1}^{3} \delta_i (\pi_{t-i}) + \sum_{j=1}^{2} \gamma_j (U_{t-j}) + \varepsilon_t \quad (2)$$

Here, π_t indicates the percentage change in inflation, as measured by the GDP deflator, while U_t represents the unemployment rate. ϵ_t and ε_t are the error terms for the country-specific model. As opposed to the model used in some of the referenced papers, the specification in this thesis employs the traditional Phillips curve model. Hence, the examined Phillips curve constitutes the backward expectations of inflation. Moreover, the model does not include any additional control variables. In this case, adding control variables does not improve the forecasting performance of the model. While it is sufficient in some of the literature, including such variables does not change the coefficient estimate of the inflation-unemployment model by that much.

Figure 4.1 illustrates the twenty models with the lowest criterion value for both countries. As can be seen, the AIC suggests that the ARDL(11,6) and ARDL(3,2) are the most appropriate specifications for the US and the UK Phillips curves, respectively. Thus, the Phillips curve will follow this specified model. In this particular case, the current value of unemployment is excluded from the distributed lag part of the model. The inclusion of the current value of unemployment messes up the significance, as well as the coefficient sign of the unemployment rate. Excluding the current value improves the estimation and forecasting performance of the model.

Figure 4.1



4.2.2 Stability Test

To evaluate the stability of the Phillips curve, several in-sample and out-of-sample tests were conducted. The first test considers the in-sample fit test of the Phillips curve. The test involves a graphical comparison between the actual GDP deflator data and the estimated fitted value of the Phillips curve using the entire sample set. The primary aim is to observe whether the Phillips curves estimated within the sample time period illustrate any sign of instability, which can be examined by how well the estimated fitted value follows the actual inflation data. The in-sample fit test employs the one-quarter-ahead forecast, where the model is solved based on the endogenous variables for the current time period. The in-sample fit test allows the Phillips equation to get back on track by feeding in actual lagged observations of inflation that the equation did not predict. This is a simple yet important first step in looking for signs of instability. It applies the information that was used to estimate the Phillips curve to measure the success within the estimation sample.

The other in-sample test involves an evaluation of the multi-period dynamic forecasting performance of the equation. Again, this involved a graphical comparison between the actual GDP deflator data and the estimated dynamic simulation of the Phillips curve, which was based on how well the trends fit with each other. Unlike the in-sample fit test, the dynamic simulation does not allow the equation to get back on track, as one large error in predicting inflation will be fed into all the subsequent time periods. The estimated Phillips curve will differ significantly from the trend if it is not well specified. The dynamic simulation also employed the one-step-ahead forecast, where the model is solved using the values of the endogenous variables prior to the sample solutions being used.

The second test considers the out-of-sample test. The test procedure is similar to that of the in-sample test, involving both the in-sample fit and dynamic performance test. However, the tests were modified to measure the performance of the equation outside the sample. To do this, the Phillips curve was first estimated within a limited sample size. In this case, the out-of-sample test employed the Phillips curve estimated from 1975–2000. Then, the test proceeded to graphically compare the fitted value, as well as the dynamic simulation, from the estimated Phillips curve of 1975-2000, with the actual inflation pattern in 2001–2016. The stability of the one-step-ahead forecast was then evaluated by whether or not the fitted value and the dynamic simulation from the estimated Phillips curve of 1975 - 2000 was running off track from the actual inflation fluctuation in 2001-2016. This was necessary to test whether or not the Phillips curve that was constructed prior to the 21st century was still able to predict the inflation pattern in the 2000s.

4.3 Stability Test 2 – Linearity and Asymmetry Test

With regard to the inflation-output Phillips curve, the series of tests conducted were slightly different from the previous section. The stability tests of the Phillips curve were inspired by the method used in Laxton et al. (1994), CMM (1992), Razzak (1997) and Wimanda et al. (2012). The results of the following stability test will be used to support the conclusion from the first test. In addition, the results of the second stability test will be used to determine the most appropriate shape of the Phillips curve in the examined countries.

4.3.1 Model Specification 2

Similar to the inflation-unemployment model, the basic specification in this test follows the 'price-price' model, as proposed by Fuhrer (1995). However, in this case, the unemployment variable was replaced by various forms of the output gap. The specifications for the US and the UK Phillips curves are presented in Equations (3) and (4), respectively.

$$\pi_t = a_t + \sum_{i=1}^{11} \beta_i(\pi_{t-i}) + \rho_j f(gap)_t + \epsilon_t \qquad (3)$$

$$\pi_t = c_t + \sum_{i=1}^{3} \delta_i(\pi_{t-i}) + \gamma_j f(gap)_t + \varepsilon_t \quad (4)$$

Here, π_t indicates the percentage change in inflation, as measured by the GDP deflator, while $f(gap)$ indicates the various forms of the output gap. ϵ_t and ε_t represent the error terms for the country-specific model. Similar to the model in the previous chapter, this model does not include any additional control variables, due to the previously explained reasons.

As a benchmark model, consider the linear approximation of the output gap function. As will be discussed in the next chapter, the linear approximation will be compared with the non-linear approximation to determine the underlying asymmetry and stability of the inflation-output relationship. A linear specification output model is presented in Equation (5) and illustrated in Figure 4.2.

$$f(gap)_t = \rho_1 gap_t \quad (5)$$

In addition to the linear approximation, three forms of non-linearity were considered: quadratic, cubic and L-shaped. The simplest forms of non-linearity were the quadratic and the cubic models. These two series are classified as the power series expansion, where the function for the output gap is interpreted as the deviation between the potential and actual output level. The quadratic specification consists of only the linear and quadratic terms. The output function of the quadratic specification is presented in Equation (6) and illustrated in Figure 4.3.

$$f(gap)_t = \rho_1 gap_t + \rho_2 gap_t^2 \quad (6)$$

The quadratic approximation implies that as the output gap becomes increasingly negative, the effect of inflation tapers off. Stated differently, as the output gap approaches a very small number, the relationship between the output gap and inflation diminishes, implying that the Phillips curve trade-off follows a convex shape. Another alternative is to raise the function of the output gap to an odd integer[6]. The simplest form, as considered by Wimanda and Laxton, is the cubic function. This approximation is presented in Equation (7) and illustrated in Figure 4.4.

[6] The power must be odd to preserve the inflation-output relationship. If the gap is raised to an even power, the effect on inflation does not have the same sign as the output gap.

$$f(gap)_t = \rho_3 gap_t^3 \quad (7)$$

There are three implications for employing the cubic approximation of the inflation-output relationship. Firstly, the effect of the rising output gap on inflation would increase rapidly, provided that there is a positive output gap. Secondly, the downward pressure on inflation will expand as the negative output gap increases in size. Finally, the slope of the function will approach zero when the output gap (regardless of whether it is positive or negative) is small. These remarks imply that as the output gap becomes increasingly positive/negative, the upward/downward impact on inflation significantly increases.

The power series expansion of the output gap, while non-linear, is not asymmetric. Hence, as strongly emphasised by Laxton, the quadratic and the cubic model cannot explain the apparent asymmetries (hence, instability) in the inflation-output relationship. It is empirically possible that the Phillips curve is non-linear, but symmetric and stable. Hence, another model that describes the underlying asymmetry is needed. Another approximation of the output gap has been provided by CMM (1992), where it takes on the form of the modified parabolic function. Unlike the quadratic and the cubic form, the so-called L-shaped function can capture the asymmetries in the inflation-output relationship. This approximation is presented in Equation (8) and illustrated in Figure 4.5.

$$f(gap)_t = \rho_4 \left[\frac{\omega^2}{(\omega - gap_t)} - \omega\right] \quad (8)$$

The L-shaped Phillips curve implies that the size of the inflation-output trade-off becomes steeper as the output gap approaches the capacity constraint. This means that if there is an upper limit on the achievable level of output (the capacity constraint) in the short run, the relationship between inflation and output immediately becomes linear when the output gap approaches this limit.

In the L-shaped output gap function, the parameter ω governs the degree of non-linearity in the Phillips curve[7]. A smaller coefficient estimate on ω implies a smaller distance

[7] ω is interpreted as the wall parameter, as it determines the values in which the relationship between inflation and the output gap effectively becomes linear.

between the zero output gap and the capacity constraint. Moreover, the slope of the inflation-output relationship following the L-shaped function can be interpreted by the underlying properties of the function[8] which, in this case, is the first derivative of Equation (8):

$$f'(gap)_t = \left[\frac{\rho_4 \omega^2}{(\omega - gap_t)^2}\right]$$

Referring to Appendix 4.2.1a, the L-shaped function can be interpreted in several ways. Firstly, Property 8a indicates that as the parameter ω becomes large, the L-shaped function approaches linearity[9]. Secondly, Property 8b indicates that the effect of the output gap on inflation rises towards infinity as the output gap approaches the wall parameter. Thus, the wall

Figure 4.2

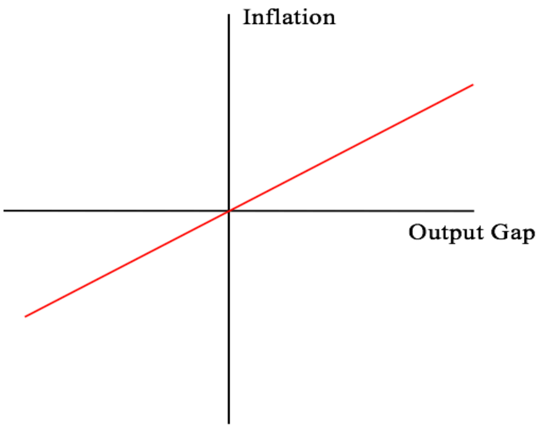


Figure 4.3

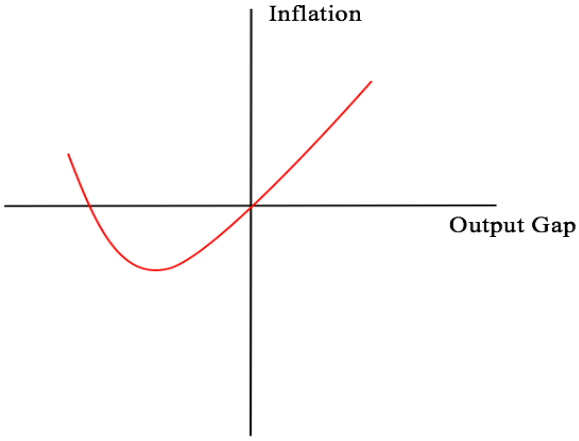


Figure 4.4

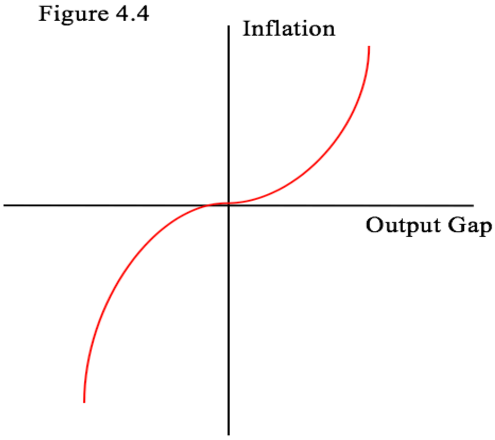
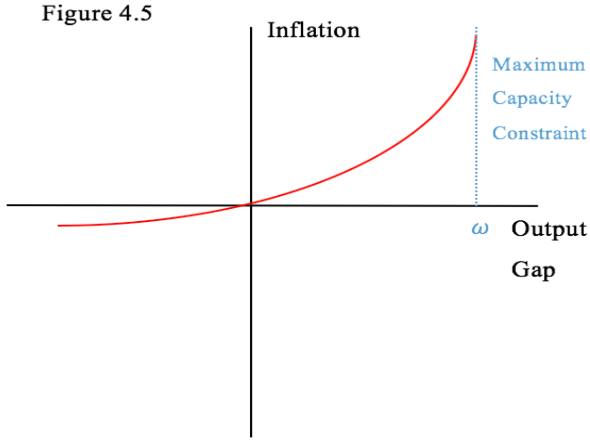


Figure 4.5



Phillips Curve

8 See Appendix 4.2.1a.
9 In contrast, as the wall parameter becomes small, the function approaches a reverse L-shaped relationship.

parameter acts as a limit beyond which the output cannot rise further in the short run. Thirdly, Property 8c indicates that there is a lower bound in which the output gap influences inflation. As the output gap becomes negative, the slope of the function decreases towards the lower bound of $-\beta\omega$. Finally, Property 8d indicates that when the effect of the output gap on inflation is 0 (that is, in the extreme case where inflation is not correlated with the output gap), the slope of the coefficient is just β.

The selected non-linear model also has several drawbacks. As previously mentioned, the theoretical literature on the non-linearities of the Phillips curve is very limited; therefore, there is little guidance on the preferred functional form. However, a preference towards one specific form can be conceptually and empirically determined. More specifically, the Phillips curve, consisting only of the cubic term, has conceptual drawbacks that are still considered symmetric. Similarly, the quadratic function also fails to impose an upper bound on output.

The upward sloping region, when the excess supply level is very high, is conceptually implausible and often difficult to interpret. The L-shaped function seems to be most plausible for conducting a linearity and symmetry test, seeing that it is asymmetric, non-linear, and, at the same time, imposes an upper bound on output. As argued by Laxton, the L-shaped model is the perfect model to test asymmetry (hence, instability) in the Phillips curve. However, in the absence of clear theoretical predictions, the Phillips curve employing the L-shaped function is often difficult to be empirically interpreted. Regardless of these drawbacks, all functional forms previously considered are tested accordingly.

The Phillips curve employed the Autoregressive Distributed Lag (ARDL) model to model the inflation-output relationship. This is similar to the approach used to model the inflation-unemployment relationship. Furthermore, the number of lags chosen in the model is determined by the AIC. Unlike the inflation-unemployment relationship, the number of lags for the output gap will be specified as 0. This is because the linearity test is interested in the coefficient of the specific functional form of the output gap. Thus, it is preferable to only include the current variable of the output gap to make a good comparison between each model.

In this case, only the most efficient number of inflation lags will be searched. The ARDL model that provides the lowest AIC value is, therefore, the best model for the Phillips curve. Figure 4.6 demonstrates the models with the lowest criterion value for both countries.

As can be seen, the AIC suggests that the ARDL(11,0) and ARDL(3,0) are the most appropriate specifications for the US and the UK Phillips curves, respectively. Thus, the Phillips curve followed these specified models.

4.3.3 Stability Test 2

Several tests were conducted to test the symmetry and linearity of the Phillips curve. Each model, along with the linear specification, was tested accordingly, starting with the hypothesis tests for the output gap coefficient and gradually going onto the model comparison.

The first test followed the simple methodology in Razzak (1997). To test the null hypothesis that the Phillips curve is symmetric and/or linear, consider the null hypothesis that the coefficient of the non-linear curve is 0. If the coefficient in the quadratic and the cubic functions are not zero, then there is evidence of a non-linear Phillips curve. Furthermore, if the coefficient on the L-shaped function is also not zero, then there is evidence of a non-linear and asymmetric Phillips curve. Therefore, the rejection of the null hypothesis implies the validity of the constructed non-linear model. This is an important first step in testing the linearity and symmetry of the inflation-output relationship, seeing that the failure to reject the null hypothesis would imply that the relationship between inflation and the non-linear output gap is zero.

The second test followed the methodology in Wimanda et al. (2011) and Laxton et al. (1994). This test estimated various forms of the non-linear model and compared them with the baseline linear specification. To select the most appropriate model, the R-squared value and LLF were compared. The specifications with the greatest R-squared and LLF values were preferred to the other alternatives.

As noted previously, the quadratic, cubic, and L-shaped Phillips curves together imply non-linearity, while the L-shaped Phillips curve by itself implies asymmetry. If the L-shaped curve has the highest R-squared and LLF values, in comparison to the others, then there is evidence of a non-linear and asymmetric (hence, instability) relationship between inflation and output. On the other hand, if the quadratic, or the cubic, functions have the highest R-squared and LLF values, in comparison with the others, then there is evidence of non-linearity in the curve, but not asymmetry. If the hypothesis is rejected, the R-squared and LLF values are the highest possible outcome for the linear function. Consequently, there is no evidence of symmetry and linearity in the Phillips curve.

The test ultimately suggests that a non-linear Phillips curve does not necessarily imply asymmetry. On the other hand, an asymmetric Phillips curve will always imply non-linearity. The prior expectation was that the coefficients on the output gaps should be positive. The tests were conducted with the following structure in mind. First, the hypothesis test for the coefficients of all the models were evaluated. Second, the benchmark linear model was compared with the non-linear specifications. Third, in the case that the linear model was outperformed by the three non-linear models, symmetry would have been tested by comparing the L-shaped model with the quadratic and the cubic models. The Phillips curve would have been concluded to be unstable, if and only if, the L-shaped Phillips curve performed better than any other specification.

To confirm the resulting asymmetry of the Phillips curve, the Diebold-Mariano test was taken into consideration. The Diebold-Mariano test was used to strengthen the conclusion of the previous test by comparing the forecasting performance of the linear and the non-linear model. The conventional approach to compare the predictive accuracy of the model is to select the forecast that has the smallest measurement error. The Diebold-Mariano test goes one step further than the conventional approach by determining whether the compared forecasts of the model are significantly different from each other. Consequently, the next hypothesis is:

$$H_0 : E(d_t) = 0 \ \forall t$$
$$H_1 : E(d_t) \neq 0$$

where d_t represents the loss differential between the two forecasts. The two forecasts have equal accuracy if, and only if, the loss differential has zero expectations[10]. The null hypothesis indicates that the two forecasts of the time series have the same accuracy, while the alternative indicates otherwise. In the context of this thesis, each form of the Phillips curve equation was tested for the forecasting accuracy. The forecasting accuracy of inflation, using the linear Phillips curve, was tested with the forecasting accuracies of the quadratic, cubic and L-shaped models for non-linearity and the asymmetry hypothesis.

[10] Dt is defined as the difference between the squared/absolute residuals of the forecast. A complete proof of the Diebold-Marino statistic is presented in Appendix 4.3a.

Chapter 5: Results

In this chapter, the Phillips curve estimation, along with the stability tests, are presented. The specification of the Phillips curve follows the traditional model, where it includes only the backward-looking expectation of inflation. The Phillips curve is based on the quarterly data for the US and the UK from 1975 to 2016. It is evaluated with the modified version of the ARDL model, whose lag values are determined by the AIC.

The first Phillips curve model employs the inflation-unemployment relationship. The stability of the estimated model is then tested, using various techniques inspired by Fuhrer (1995). The second Phillips curve model employs the inflation-output relationship. The stability of the estimated model is tested through a linearity approach.

5.1 Stability Test 1 – The Fit and Dynamic Simulation Test:

As explained in the previous chapter, the inflation-unemployment relationship is evaluated with the ARDL (11,6) and the ARDL (3,2) models for the US and the UK data, respectively. This means that the US model employs 11 lags of inflation and 6 lags of unemployment, while the UK model employs 3 lags of inflation and 2 lags of unemployment. The analysis will primarily focus on the unemployment coefficients, which are expected to have a negative summation. The estimated inflation-unemployment Phillips curves are presented in Tables 5.1 and 5.2 for the US and the UK, respectively.

Table 5.1
Inflation-Unemployment Phillips Curve (The United States)
Data: Quarterly, 1975 (Q1)–2016 (Q4)

Variable	Estimated Coefficient	Std. Error	T-statistic
a_t	0.093	0.085	1.096
β_1	0.594	0.077	7.724
β_2	0.137	0.088	1.562
β_3	0.025	0.088	0.288
β_4	0.299	0.085	3.503
β_5	-0.287	0.087	-3.307
β_6	0.182	0.089	2.058
β_7	-0.200	0.087	-2.306
β_8	0.128	0.087	1.474
β_9	0.149	0.087	1.716
β_{10}	-0.220	0.088	-2.513
β_{11}	0.137	0.071	1.913
ρ_1	-0.243	0.068	-3.546
ρ_2	0.326	0.129	2.521
ρ_3	0.063	0.138	0.461
ρ_4	-0.383	0.131	-2.932
ρ_5	0.225	0.069	3.247

R-squared	0.863
LLF	24.422
DW-Stat	1.999

Table 5.2
Inflation-Unemployment Phillips Curve (The United Kingdom)
Data: Quarterly, 1975 (Q1)–2016 (Q4)

Variable	Estimated Coefficient	Std. Error	T-statistic
c_t	0.172	0.178	0.967
δ_1	0.289	0.073	3.941
δ_2	0.261	0.074	3.543
δ_3	0.322	0.073	4.378
γ_1	-0.579	0.263	-2.199
γ_2	0.567	0.263	2.153

R-squared	0.63
LLF	-216.77
DW-Stat	2.11

In the US model, all unemployment coefficients are statistically significant, except for the third-period lag. The overall response to the unemployment rate, as calculated by the sum of the unemployment coefficients, is -0.011, implying that statistically significant evidence of a trade-off between inflation and unemployment in the US indeed exists. However, the value is relatively small, in comparison to the coefficients that were found in Fuhrer (1995), perhaps due to the different time period that was examined. While Fuhrer (1995) examined data prior to 1995, the Phillips curve, in this thesis, extends the analysed period to 2016. This indicates that as the examined series is extended, the magnitude of the inflation-unemployment trade-off diminishes. Furthermore, the Durbin-Watson test statistic of 1.99 indicates that there is no serial correlation in the residuals.

For the UK model, all unemployment coefficients are found to be statistically significant. Even in this model, the overall response to the unemployment rate, as calculated by the sum of the unemployment coefficients, is -0.011. Thus, the UK Phillips curve has the same implication as the US Phillips curve. The sum of the coefficient estimates of inflation demonstrate that, regardless of the economic condition in each country, rising inflation has responded similarly towards unemployment over the past 40 years. Furthermore, the Durbin-Watson test results in the test statistic of 2.11 indicate that there is no serial correlation in the residuals.

The stability analysis begins with the in-sample test. Here, the ability of the constructed Phillips curve to predict an inflationary movement is further examined. The stability of the Phillips curve under large shocks (e.g., oil crisis, global financial crisis) has been particularly highlighted. Figures 5.1 and 5.2 display the in-sample fit test of the UK and the US Phillips curve, respectively.

Figure 5.1

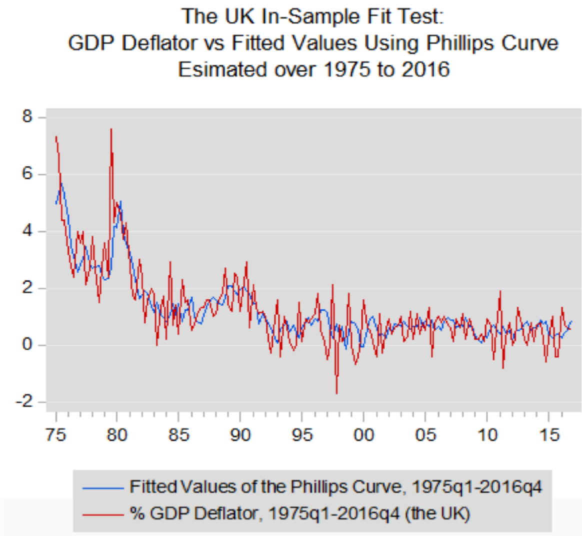
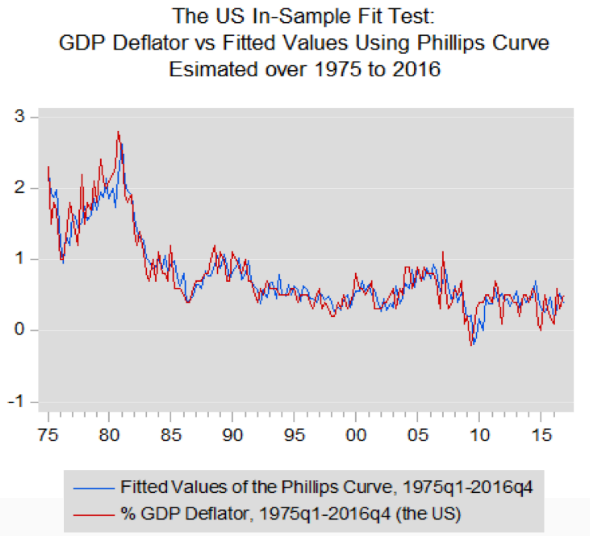


Figure 5.2



The figures compare the actual data for the inflation rate and the fitted values from the estimated inflation-unemployment Phillips curve. For the US data, the in-sample fit test provides evidence of the inflation-unemployment stability over the 40-year time period. As the figure indicates, there is no sign that the Phillips curve estimated within the sample is running off track. In particular, the Phillips curve predicts inflation during the oil crisis, the late 1990s recessions and the global financial crisis relatively well. This indicates that there is no over or under prediction of inflation by the Phillips curve. Similarly, the UK data also illustrates evidence of stability in the inflation-unemployment relationship over the past 40 years.

The estimated Phillips curve predicted the inflation pattern relatively well, especially during the oil crisis. However, the estimated Phillips curve failed to predict the inflationary pressure in the late 1970s; that being said, it did manage to capture the 1980s recessions much better. Though the volatility of inflation during the great moderation and the global financial crisis was not captured well, the estimated fitted value differs significantly from the trend. Overall, the in-sample fit test indicates that stability indeed exists during the 1990–2016 time period.

Another in-sample test involves the dynamic performance test. Unlike the in-sample fit test, the dynamic test simulates the Phillips curve over a long time period, feeding in the simulated values for the current time period, as the lagged values in the past time periods. In the dynamic simulation, one large error in predicting inflation can feed into all subsequent predictions, causing the constructed Phillips curve to differ significantly from the trend. Figures 5.3 and 5.4 display the in-sample dynamic simulation test for the UK and the US Phillips curves, respectively.

Figure 5.3

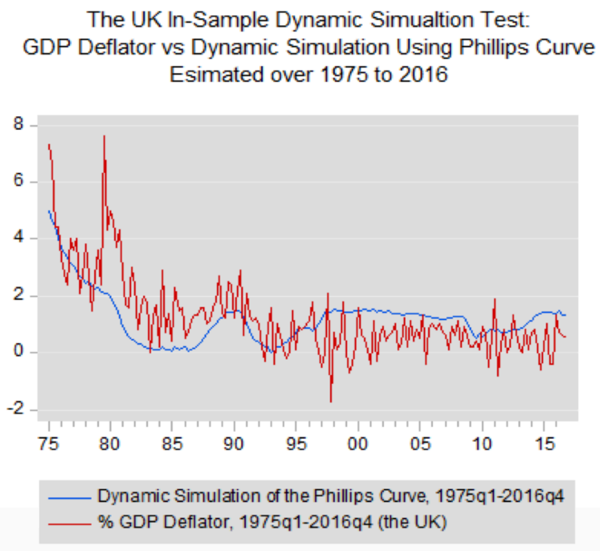
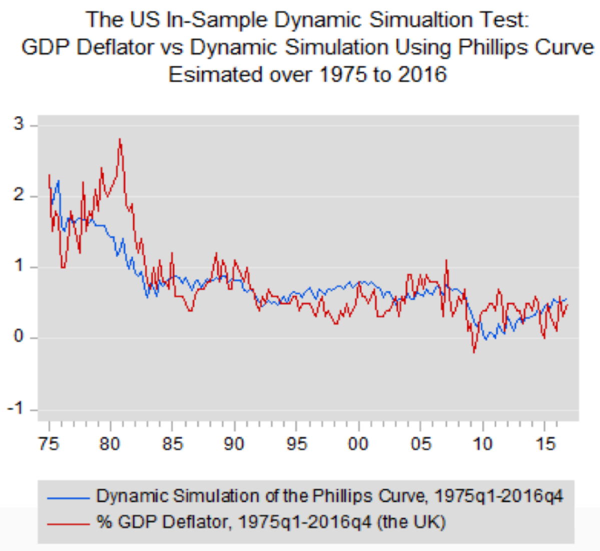


Figure 5.4



The figures compare the actual data for the inflation rate and the dynamic simulation from the estimated inflation-unemployment Phillips curve. Using the Phillips curve estimated over the entire sample, the dynamic simulation, based on this observation, is compared with the actual inflation pattern. As indicated by the US data, the Phillips curve fits the actual data well. Though the estimated Phillips curve under-predicted the inflationary shock during the oil crisis, it was able to correctly predict the inflation pattern of the 1990s recessions and the global financial crisis.

The UK Phillips curve also fits the data well. There are under-predictions of the inflation pattern in the 1980s recessions; however, the overall shape conveys stability in the Phillips curve trade-off. The in-sample tests conclude that there is evidence of implied stability in the inflation-unemployment relationship in both the US and the UK Phillips curves.

The second stability test involves an out-sample test. In this test, the Phillips curve is estimated from the 1975 to 2000 data and is used to construct the fitted value and the dynamic tracking performance. These values are then compared with the inflation pattern from the 2001–2016 period. The examined time period is particularly interesting, seeing that it can be used to explicitly analyse the stability during the global financial crisis. The estimated coefficients of the Phillips curve are presented in Tables 5.5(a) and 5.5(b) in the Appendix. The sum of the unemployment coefficients is -0.07 and -0.05 for the US and the UK, respectively. This indicates that when unemployment changes, the response of inflation is much weaker for the data series prior to the year 2000. Figures 5.5 and 5.6 display the out-of-sample fit test for the UK and the US Phillips curve, respectively.

Figure 5.5

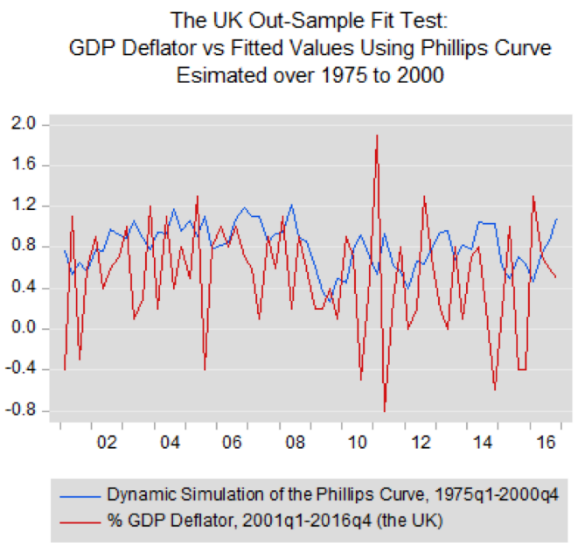
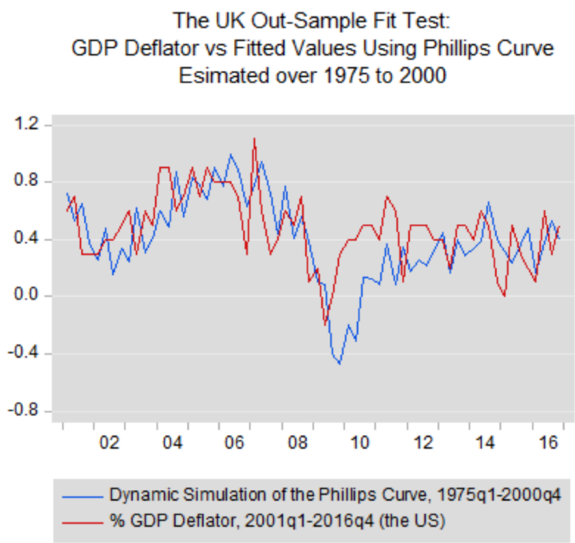


Figure 5.6



The figures compare the actual data for the inflation rate in the years 2001–2016 and the fitted values from the estimated inflation-unemployment Phillips curve in 1975-2000. As the figures indicate, there is no indication of the Phillips curve running amok in either country during the 2001–2016 sample time period. With regard to the US data, the Phillips curve predicted the fall in the inflation level during the global financial crisis. Though the estimated model under-predicts the disinflation, the overall fit conveys stability in the inflation-unemployment relationship. With regard to the UK data, the Phillips curve failed to predict the cost-push inflation in 2010. Regardless, the Phillips curve did not illustrate any sign of going off track from the inflation pattern. Hence, it still conveys stability in the inflation-unemployment relationships.

Another out-sample test involves a dynamic performance test. Figures 5.7 and 5.8 display the out-sample dynamic simulation test for the UK and the US Phillips curves, respectively. The figures compare the actual data for the inflation rate and the dynamic simulation from the estimated inflation-unemployment Phillips curve of the 1975 to 2000 period. The dynamic simulation for both countries fails to predict the inflation pattern for the 2001–2016 period. As indicated by the US data, the dynamic simulation fits relatively well with the inflation data until the global financial crisis occurs. Though the disinflation during the global financial crisis is captured relatively well, the Phillips curve fails to capture the recovery during the 2010–2011 periods. The dynamic simulation predicts that there would be a continuous disinflation in the economy, indicating that the Phillips curve deviates from the actual inflation data as a result of the under-prediction. Similarly, as indicated by the UK data, the dynamic simulation over-predicts the inflation pattern for the examined time periods of 2001 to 2016. As opposed to the US figure, the dynamic simulation begins to deviate off track at an early stage of the examined sample time period and continues to over-predict throughout the series.

Figure 5.7

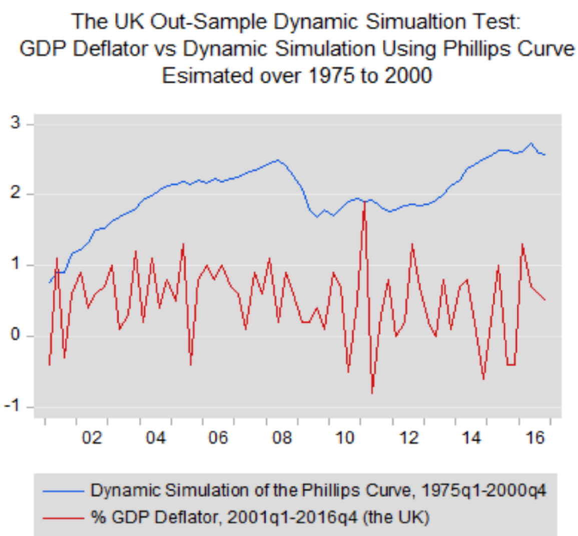
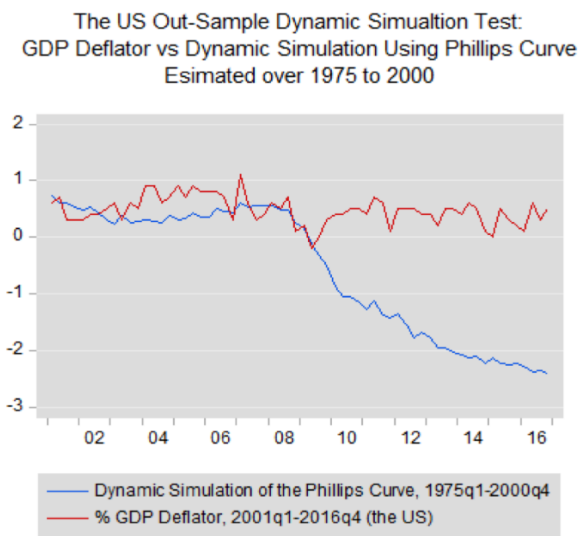


Figure 5.8



Overall, the in-sample and out-sample tests suggest that there is weak evidence of stability in the US and the UK data, especially in the light of the current financial crisis. The stability is considered weak, as the overall in-sample and the out-sample fit tests tend towards stability. However, the out-sample dynamic simulation contradicts this conclusion. The test shows that there is no evidence of stability over the period of 2001–2016, as opposed to the evidence found in the previous tests. Furthermore, based on the out-sample dynamic simulation, the UK data is implied to provide weaker evidence of stability, in comparison to the US, seeing that instability in the UK data persists for a longer time period.

5.2 Stability Test 2 – Linearity and Asymmetry Tests

In the following chapter, the Phillips curve will employ the inflation-output relationship, which will be tested for linearity and asymmetry simultaneously. The specification of the Phillips curve follows the traditional model, which only includes the backward-looking expectation of inflation and a different functional form of the output gap. The Phillips curve is based on the quarterly data for the US and the UK over the period of 1975–2016. The curve was evaluated using the modified version of the ARDL model, whose lag values are determined by the AIC. The Phillips equations are re-written and presented in Equations (12a) to (15a) for the US and Equations (16a) to (19a) for the UK in the Appendix. The estimated Phillips curve equation for the different functional forms of the output gap has been presented in Table 5.3 of this chapter.

For the first test, consider the test for the null hypothesis, where the coefficient on the output gap of the non-linear function is 0. The failure to reject this null, that the coefficients of the non-linear model are zero, implies that there is no evidence of a relationship between inflation and unemployment in the short run. As displayed in Table 5.3, the hypothesis is rejected in almost every model, except the quadratic model. Most of the coefficients on the output gap have a non-zero and statistically significant value, except for the quadratic term. The null hypothesis is that the coefficient on the quadratic output cannot be rejected, seeing that the coefficient is indeed close to 0 and, most importantly, not statistically significant. This result is consistent in both the US and the UK data.

The rejection in the quadratic model implies that there is minimal evidence of convexity in the trade-off, that is, there is no evidence to illustrate that the relationship between the output gap and inflation diminishes as the output gap approaches a very small number. This conclusion, however, does not imply a rejection of the non-linearity hypothesis of the Phillips curve, seeing that the cubic and the L-shaped functions are still valid. The hypothesis that the coefficients are zero for these functions is rejected. The result is consistent in both the examined countries, with the UK's cubic and L-shaped coefficient being much larger than the US's. With this conclusion, the next test will disregard the quadratic function and use the L-shaped and the cubic functions as a comparative model for symmetry and non-linearity.

For the second test, the linear, non-linear and the non-symmetric Phillips curves are compared to evaluate the most suitable functional form of the output gap. The linear model is used as a benchmark for the comparison, where it was evaluated with regard to the cubic and L-shaped model. The focus of the analysis is the coefficient of the various functional forms of the output gap, which is expected to have a positive and significant value. In selecting the best model, the model's R-squared and LLF values were compared. The specification with the greatest R-squared and LLF values was preferred to the compared alternatives.

As an opening note, the benchmark linear model has a positive and significant coefficient in both the examined countries. The coefficient, however, is greater in the UK than in the US, implying a much stronger inflation-unemployment relationship in the UK's historical data. This model is then compared with the cubic and L-shaped function for the linearity and symmetry test.

With regard to the cubic form of the Phillips curve, the coefficient of the cubic function is found to be positive and significant in both the examined countries. The coefficient is very small, much smaller in comparison to the linear model. This result is consistent with the results in Wimanda et al. (2011), which also found a significant but small coefficient estimate. The LLF values and the adjusted R-squared value is greater in comparison to the linear model. This implies that the cubic model outperforms the linear model in both the US and the UK. Based on the results, the inflation-output relationship is found to be non-linear, but not asymmetric. The symmetry test will involve a comparison between the L-shaped and the linear models.

Before comparing the L-shaped function with the cubic and linear functions, the value of ω, or the wall, needs to be parameterised. Similar to the approach adopted by Wimanda et al. (2011), CMM (1994) and Laxton et al. (1994), the wall parameter is evaluated on the grid search, where the value of ω varies from 0.02 to 0.10, with increments of 0.001. The value of the wall parameter that maximises the LLF is then used to estimate the coefficient of the L-shaped function. For both the US and the UK, the wall parameter that maximises the LLF is 0.04. This value implies that the relationship between inflation and output becomes vertical as the output gap reaches 0.04. In other words, as the output gap cannot increase above its capacity constraint of ω, the Phillips curve becomes increasingly vertical as the output gap approaches 0.04.

Table 5.3
Inflation-Output Phillips Curve (The United States)
Data: Quarterly, 1975 (Q1)–2016 (Q4)

Variable	Linear		Quadratic		Cubic		L-Shape	
	Coefficient	Std Error	Coefficient	Std Error	Coefficient	Std Error	Coefficient	Std Error
α_t	0.04058	0.03269	0.04074	0.03276	0.03498***	0.03288	0.05853*	0.03287
β_1	0.54323***	0.07760	0.54239***	0.07776	0.55673	0.07696	0.56253***	0.07670
β_2	0.06988	0.08731	0.07202	0.08754	0.07780	0.08728	0.11606	0.08744
β_3	0.07860	0.08766	0.07559	0.08796	0.06843***	0.08786	0.06501	0.08794
β_4	0.29972***	0.08755	0.29938***	0.08772	0.29041***	0.08764	0.31860***	0.08804
β_5	-0.27242**	0.08864	-0.27145**	0.08882	-0.28394**	0.08852	-0.29204**	0.08845
β_6	0.19005**	0.09007	0.18904**	0.09025	0.19430**	0.09018	0.14215	0.09154
β_7	-0.22769**	0.08828	-0.22095*	0.08908	-0.21461**	0.08866	-0.19294**	0.08949
β_8	0.09565	0.08823	0.10168	0.08890	0.11593	0.08881	0.03531	0.08948
β_9	0.14581*	0.08836	0.14681*	0.08855	0.15559*	0.08865	0.16522*	0.08894
β_{10}	-0.14253	0.08949	-0.14688*	0.08992	-0.16706*	0.08907	-0.18491**	0.08906
β_{11}	0.13437*	0.07155	0.13504*	0.07170	0.13524**	0.07164	0.15665**	0.07158
ρ_1	0.049425***	0.01454	0.046835**	0.01512				
ρ_2			-0.00391	0.00612				
ρ_3					0.004552***	0.00137		
ρ_4							0.015227**	0.00459
LLF	17.56		17.78		17.62		17.71	
R-squared	0.850871		0.851265		0.850464		0.85177	
DW-Stat	2.027109		2.026595		2.050085		2.033374	

Table 5.4
Inflation-Output Phillips Curve (The United Kingdom)
Data: Quarterly, 1975 (Q1)–2016 (Q4)

Variable	Linear		Quadratic		Cubic		L-Shape	
	Coefficient	Std Error	Coefficient	Std Error	Coefficient	Std Error	Coefficient	Std Error
c_t	0.17499*	0.09402	0.17645*	0.09923	0.16463*	0.09600	0.18239*	0.10029
δ_1	0.25515***	0.07281	0.25563***	0.07373	0.26382***	0.07534	0.29482***	0.07394
δ_2	0.25042***	0.07214	0.25064***	0.07251	0.24769***	0.07368	0.23779**	0.07669
δ_3	0.31797***	0.07101	0.31784***	0.07128	0.30507***	0.07271	0.29369***	0.0001
γ_1	0.158725**	0.04993	0.159743**	0.05456				
γ_2			-0.00112	0.02384				
γ_3					0.008626*	0.00501		
γ_4							0.398905**	0.08245
LLF	-214.2506		-214.2494		-213.7916		-213.1178	
R-squared	0.640076		0.640081		0.624579		0.658605	
DW-Stat	2.180175		2.17808		2.212079		2.080245	

* - significant at the 10% level
** - significant at the 5% level
*** - significant at the 1% level

The coefficient estimates in the L-shaped Phillips curves are positive and significant for the examined countries. The coefficient is significantly larger in the UK, in comparison to the US, implying a much stronger inflation-unemployment relationship in the UK's historical data. On the other hand, the LLF and the R-squared values are greater, in comparison to the linear model, implying that the L-shaped model outperforms the linear model in both the US and the UK. This result suggests that the L-shaped function is better at describing the US and the UK Phillips curves than the linear function. Conclusively, the results show that there is indeed evidence of asymmetry and non-linearity in the Phillips curve.

Table 5.4 – Diebold-Mariano Test

United States	Diebold-Mariano Test Statistic	P-value	Accuracy
Linear vs Cubic	-0.02329	0.981419	Same
Linear vs L-shape	-2.424559	0.0015327	L-shape is more accurate
Cubic vs L-shape	-2.261204	0.023747	L-shape is more accurate

United Kingdom	Diebold-Mariano Test Statistic	P-value	Accuracy
Linear vs Cubic	0.335661	0.737126	Same
Linear vs L-shape	-2.896513	0.003773	L-shape is more accurate
Cubic vs L-shape	-1.430061	0.152699	Same

For the final test, consider the output in Table 5.4, which demonstrates the results of the Diebold-Mariano test for the comparative model. The first table compares the model specification using the US data. The hypothesis, that the forecasts have the same accuracy, is rejected when comparing the linear and cubic function with the L-shaped function. However, the hypothesis cannot be rejected when comparing the linear and cubic functions. The Diebold-Mariano test for the US data indicates that the forecasting accuracy of the L-shaped function is better in comparison to the linear and the cubic functions, which strengthen the findings of the model comparison test.

The second table compares the model specification for the UK data. The hypothesis, that the forecasts have the same accuracy, is rejected when comparing the linear with the L-shaped function. However, the hypothesis failed to be rejected when comparing the L-shaped function with the cubic function, as well as when comparing the linear function with the cubic function. The Diebold-Mariano test for the US data indicates that the forecasting accuracy of the L-shaped function is better, in comparison to the linear model. However, there is no significant difference in the forecasting accuracy between the cubic and the L-shaped functions. Regardless, evidence from the UK data indicates asymmetry in the Phillips curve. Hence, the L-shaped function is still better than the linear function.

The study results, and Wimanda and Laxton's findings reveal similar evidence of asymmetry and non-linearity in the Phillips curve, which also tends towards the L-shaped function.

Chapter 6: Conclusions and Discussion

The primary objective of this thesis is to evaluate the stability of the Phillips curve. The Phillips curve employs the traditional form, where it was designed to only include the backward-looking expectation of inflation. The estimation results of the inflation-unemployment Phillips curve suggest that there is weak evidence of stability. Three out of four tests confirm stability during the examined time periods. These tests indicate that the inflation-unemployment relationship is relatively stable over the long run and through several recessionary periods characterised by inflationary pressure and disinflation. However, when conducting a more stringent out-sample test, the Phillips curve trade-off conveys instability. The dynamic simulation failed to predict the inflation pattern during the global financial crisis and, thus, rejects the hypothesis that there is a stable Phillips curve during the time period.

The linearity tests, on the other hand, provide evidence of an asymmetric and non-linear relationship between inflation and the output gap. There is evidence to prove that the L-shaped function outperforms the other non-linear, as well as linear, model, despite a very small difference in the R-squared and the LLF values. This specification fits the US and the UK data relatively well, indicating that the inflation-output relationship is best described by using an L-shaped function. The second stability test concludes that there is strong evidence of non-linearity in the Phillips curve trade-off. This is supported by evidence found in the Diebold-Mariano test, which conveys an overall instability in the Phillips curve.

The conclusions highlight several important points. Firstly, the Phillips curve is found to be weakly unstable during the examined time period. More importantly, there is evidence that, during the global financial crisis, the inflation-unemployment relationship became highly unstable for both countries. This contradicts the results in Fuhrer (1995), which revealed a stable relationship during the oil shock. The recent financial crisis had a much larger impact and was able to disrupt the relationship between inflation and unemployment during this particular time period. Moreover, by employing the same technique, extending the examined time period within the same economy would quickly result in an instability in the Phillips curve trade-off.

Secondly, the shape of the Phillips curve is found to be L-shaped for both countries. This is consistent with the results found in CMM (1992), Laxton et al. (1994) and Wimanda et al. (2012), where similar evidence was found. This implies that a positive demand shock tends to increase inflation more than the reduction caused by negative demand shocks of a similar magnitude. The result of both tests confirms that there is indeed an unstable Phillips relationship in both the US and the UK. This evidence is consistent with the characteristics of the business cycle of the past few decades in many developed economies. The unexpected and long-lasting economic downturn tends to be followed by a short-lived, but rapid increase in inflation.

The instability in the Phillips relationship has clear implications for the design of monetary policy. In a linear and stable world, the Phillips curve is a feasible tool for designing monetary policy. There is no direct link between monetary policy and the level of output, as negative and positive demand shocks will have a symmetric effect on inflation. In light of the Lucas critique, the net effect on output approaches zero, regardless of the policy stance. This is not the case in a non-linear and unstable world. Policymakers cannot exploit the statistical relationships to formulate policy, seeing that a positive output gap will raise inflation by more than the negative shocks reduce it. If immediate actions are taken to offset the negative output gap (hence, high unemployment rate), it will result in a significant and sudden rise in the inflation rate. When the output gap is moving in another direction, a large volume of negative demand shocks to slow down the economy will do very little to lower inflation. The failure in the Phillips curve illustrates that relying on statistical relationships between inflation and output itself is not enough to formulate a proper monetary policy, especially during large shocks. Thus, relying on the Phillips curve to formulate monetary policy can lead to costly policy errors, which would have a negative impact on the economy.

As a closing remark, it must be said that several improvements can be made to this thesis. This thesis conducted two tests for stability; however, it did not examine the cause of such instability. The first stability test concluded that the Phillips curve was unstable during the global financial crisis, which led to the belief that the crisis caused the relationship to break down. However, the reasoning behind this is weak, and further tests regarding this issue could be conducted. Furthermore, the non-linear and asymmetric model goes beyond the three non-linear models that were considered. These models are perhaps best summarised in Wimanda et al. (2011), where six non-linear models are considered. Regardless, the three models considered in this thesis were enough to represent non-linearity in the Phillips equation. Finally, as the breakpoint test conducted by Fuhrer is not presented in this thesis, the first stability test does not fully replicate the methodology found in Fuhrer[11]. The inclusion of the breakpoint test would perhaps be able to further support the instability of the relationships. However, the two tests conducted in this thesis are enough to point out the instability in the Phillips equation.

[11] Contrary to the statement made in the first chapter.

References

Abbas, S. K. et al. (2016). The New Keynesian Phillips curve in a small open economy: Empirical evidence from Australia. *Economic Record*, *92*(298), 409–34.

Ball, L. et al. (1988). The new Keynesian economics and the output-inflation trade-off. *Brookings Papers on Economic Activity*, *1988*(1), 1–82.

Ball, L. M., & Mazumder, S. (2011). *Inflation Dynamics and the Great Recession* (No. w17044). National Bureau of Economic Research.

Brayton, F., Levin, A., Lyon, R., & Williams, J. C. (1997, December). The evolution of macro models at the Federal Reserve Board. In *Carnegie-Rochester Conference Series on Public Policy* (Vol. 47, pp. 43–81). North-Holland.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, *12*(3), 383–98.

Chadha, B., Masson, P. R., & Meredith, G. (1992). Models of inflation and the costs of disinflation. *Staff Papers*, *39*(2), 395–431.

Clark, P., Laxton, L., & Rose. D. (1996). Asymmetry in the US output-inflation nexus. *Staff Papers*, *43*(1), 216–51.

Davidson, R., & MacKinnon, J. G. (2004). *Econometric Theory and Methods* (Vol. 5). New York: Oxford University Press.

De Veirman, E. (2007). *Which nonlinearity in the Phillips curve? The absence of accelerating deflation in Japan* (No. 536). The Johns Hopkins University, Department of Economics.

Debelle, G., & Laxton, D. (1997). Is the Phillips curve really a curve? Some evidence for Canada, the United Kingdom, and the United States. *Staff Papers*, *44*(2), 249–82.

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *20*(1) 134–44.

Dolado, J. J., Marıa-Dolores, R., & Naveira, M. (2005). Are monetary-policy reaction functions asymmetric? The role of nonlinearity in the Phillips curve. *European Economic Review*, *49*(2), 485–503.

Dotsey, M., King, R. G. & Wolman, A. L. (1999). State-dependent pricing and the general equilibrium dynamics of money and output. *The Quarterly Journal of Economics*, *114*(2), 655–90.

Eliasson, A. C. (2001). *Is the short-run Phillips curve nonlinear? Empirical evidence for Australia, Sweden and the United States* (No. 124). Sveriges Riksbank.

Fischer, S. (1977). Long-term contracting, sticky prices, and monetary policy: A comment. *Journal of Monetary Economics*, *3*(3), 317–23.

Fitzgerald, T. J., Holtemeyer, B., & Nicolini. J. P. (2013). Is there a stable Phillips Curve after all? *Economic Policy Paper*, *13*.

Fuhrer, J. C. (1995). The Phillips curve is alive and well. *New England Economic Review*, 41–57.

Fuhrer, J. C. (1997). The (un) importance of forward-looking behavior in price specifications. *Journal of Money, Credit, and Banking*, 338–50.

Fuhrer, J., & Moore, G. (1995). Inflation persistence. *The Quarterly Journal of Economics*, *110*(1), 127-59.

Galı, J., Gertler, M., & Lopez-Salido, J. D. (2001). European inflation dynamics. *European Economic Review*, *45*(7), 1237–70.

Huh, H. S., & Jang, I. (2007). Nonlinear Phillips curve, sacrifice ratio, and the natural rate of unemployment. *Economic Modelling*, *24*(5), 797–813.

Kobbi, Imen, and Foued-Badr Gabsi. "The Nonlinearity of the New Keynesian Phillips Curve: The Case of Tunisia." Economies 5.3 (2017): 24.

Koske, I., & Pain, N. (2008). The usefulness of output gaps for policy analysis.

Laxton, D., Rose, D., & Tambakis, D. (1999). The US Phillips curve: The case for asymmetry. *Journal of Economic Dynamics and Control*, *23*(9–10), 1459–85.

Fisher, J. D., Liu, C. T., & Zhou, R. (2002). When can we forecast inflation? *Economic Perspectives-Federal Reserve Bank of Chicago*, *26*(1), 32–44.

Moccero, D., Watanabe, S., & Cournède, B. (2011). What drives inflation in the major OECD economies?

Mladenović, Z., & Nojković, A. (2012). Inflation persistence in Central and South-Eastern Europe: Evidence from univariate and structural time series approaches. *Panoeconomicus*, *59*(2), 235–66.

Nell, K. S. (2006). Structural change and nonlinearities in a Phillips curve model for South Africa. *Contemporary Economic Policy*, *24*(4), 600–17.

Phillips, A. W. (1958). The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957 1. *Economica*, *25*(100), 283–99.

Roeger, W., & Herz, B. (2012). *Traditional versus New Keynesian Phillips Curves: Evidence from output effects*. BiblioGov.

Rudd, J., & Whelan, K. (2005). New tests of the new-Keynesian Phillips curve. *Journal of Monetary Economics*, *52*(6), 1167–81.

Rudd, J., & Whelan, K. (2007). Modeling inflation dynamics: A critical review of recent research. *Journal of Money, Credit and Banking*, *39*, 155–70.

Rudebusch, G., & Svensson, L. E. (1999). Policy rules for inflation targeting. In *Monetary Policy Rules* (pp. 203–262). University of Chicago Press.

Sergo, Z., Saftic, D., & Tezak, A. (2012). Stability of Phillips curve: The case of Croatia. *Economic Research-Ekonomska Istraživanja*, *25*(sup1), 65–85.

Turner, P. (1997). The Phillips curve, parameter instability and the Lucas critique. *Applied Economics*, *29*(1), 7–10.

Wimanda, R. E., Turner, P. M., & Hall, M. J. (2011). Expectations and the inertia of inflation: The case of Indonesia. *Journal of Policy Modeling*, *33*(3), 426–38.

Wimanda, R. E., Turner, P. M., & Hall, M. J. (2013). The shape of the Phillips curve: The case of Indonesia. *Applied Economics*, *45*(29), 4114–21.

Appendix

Appendix 4.2.1a: Properties of the L-shaped Function

$$\lim_{\omega \to \infty} f'(gap) = \beta \qquad (8a)$$

$$\lim_{\omega \to \infty} f'(gap) = \infty \,, f(\omega) = \infty \qquad (8b)$$

$$\lim_{gap \to -\infty} f'(gap) = 0 \,, f(-\infty) = -\beta\omega \qquad (8c)$$

$$f'(0) = \beta, f(0) = 0 \qquad (8d)$$

Appendix 4.3a: Diebold-Mariano Statistical Proof

Let $\{y_t; t = 1, \ldots., T\}$ be the vector space of the actual series of the variable of interest, yt.

It is often the case that two (or more) models are available for forecasting a specific variable of interest. The two forecasts can be defined as follows:

$\{\widetilde{y_{it}}; t = 1, \ldots., T\}$, $i = 1,2$

The Diebold Mariano test is used to determine whether the two forecasts $\widetilde{y_{1t}}$ and $\widetilde{y_{2t}}$ are equally good. Let e_{it} be the residuals of the two forecasts.

As such: $e_{it} = \widetilde{y_{it}} - y_t$, $i = 1,2$

The loss associated with forecast i is assumed to be the function of the forecast error e_{it}. It is denoted by $g(e_{it})$, such that it takes on the value of zero when no forecasting error is made, never becomes negative and increases in size as the errors approach ∞. $g(e_{it})$ is referred to as the loss function, where it takes the form of either the absolute error loss $g(e_{it}) = |e_{it}^2|$ or the squared error loss $g(e_{it}) = e_{it}^2$.

From the loss function, the loss differential between the two forecasts can be defined as d_t, such that $d_t = g(e_{1t}) - g(e_{2t})$.

The two forecasts, $\widetilde{y_{1t}}$ and $\widetilde{y_{2t}}$, have equal accuracy if and only if the loss differential has zero expectation for all t. It follows the hypothesis test as follows:

$$H_0 : E(d_t) = 0 \; \forall t$$

$$H_1 : E(d_t) \neq 0$$

where the null hypothesis is rejected when the two forecasts have different levels of accuracy and fail to be rejected when the two forecasts have the same accuracy. The sample and population mean of the loss differential are respectively defined as:

$$\mu = E(d_t)$$

$$\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$$

For $n > k \geq 1$:

$$\gamma_k = \frac{1}{n}\sum_{t=1}^{n}(d_t - \bar{d})(d_{t-k} - \bar{d})$$

γ_k is the autocovariance at lag k. For $h \geq 1$, the Diebold-Mariano (DM) statistic is defined as:

$$DM = \frac{\bar{d}}{\sqrt{(\gamma_0 + 2\sum_{k=1}^{h-1}\gamma_k)/n}}$$

Under the null hypothesis, the test statistic DM follows a standard normal distribution $N(0,1)$. The null hypothesis is rejected if the DM statistic fails outside of the range of $-z_{a/2}$ to $z_{a/2}$, such that $|DM| > z_{a/2}$.

Here, $z_{a/2}$ is the upper z-value from the standard normal table corresponding to half of the desired a level of the test.

The key assumption of the DM test is that the loss differential time series is stationary[12].

Appendix 5.2a: Non-Linear Model of the Phillips Curve

$$\pi_t = a_t + \sum_{i=1}^{11} \beta_i(\pi_{t-i}) + \rho_1 gap_t + \epsilon_t \quad (12a)$$

$$\pi_t = a_t + \sum_{i=1}^{11} \beta_i(\pi_{t-i}) + \rho_1 gap_t + \rho_2 gap_t^2 + \epsilon_t \quad (13a)$$

$$\pi_t = a_t + \sum_{i=1}^{11} \beta_i(\pi_{t-i}) + \rho_3 gap_t^3 + \epsilon_t \quad (14a)$$

$$\pi_t = a_t + \sum_{i=1}^{11} \beta_i(\pi_{t-i}) + \rho_4 \left[\frac{\omega^2}{(\omega - gap_t)} - \omega\right] + \epsilon_t \quad (15a)$$

$$\pi_t = c_t + \sum_{i=1}^{3} \delta_i(\pi_{t-i}) + \gamma_1 gap_t + \epsilon_t \quad (16a)$$

$$\pi_t = c_t + \sum_{i=1}^{3} \delta_i(\pi_{t-i}) + \gamma_1 gap_t + \gamma_2 gap_t^2 + \epsilon_t \quad (17a)$$

$$\pi_t = c_t + \sum_{i=1}^{3} \delta_i(\pi_{t-i}) + \gamma_3 gap_t^3 + \epsilon_t \quad (18a)$$

$$\pi_t = c_t + \sum_{i=1}^{3} \delta_i(\pi_{t-i}) + \gamma_4 \left[\frac{\omega^2}{(\omega - gap_t)} - \omega\right] + \epsilon_t \quad (19a)$$

[12] The proof is taken from the work of Diebold & Marianno in 2002.

Appendix 6a: Out-of-Sample Phillips Curve Estimation

Table 5.5(a)
Inflation-Unemployment Phillips Curve (The United States)
Data: Quarterly, 1975 (Q1)–2000 (Q4)

Variable	Estimated Coefficient	Std. Error	T-Statistic
a_t	0.320112	0.152804	2.094919
β_1	0.593332	0.098419	6.028662
β_2	0.122106	0.114271	1.068565
β_3	-0.024714	0.109441	-0.225818
β_4	0.386955	0.107264	3.607491
β_5	-0.339352	0.112698	-3.011152
β_6	0.217253	0.116349	1.867253
β_7	-0.176076	0.109671	-1.605501
β_8	0.128493	0.10797	1.190081
β_9	0.245686	0.108328	2.267981
β_{10}	-0.25432	0.110957	-2.29205
β_{11}	0.160536	0.090187	1.780034
ρ_1	-0.291914	0.084857	-3.440065
ρ_2	0.415242	0.156749	2.649089
ρ_3	-0.172991	0.170982	-1.011751
ρ_4	-0.267364	0.16114	-1.659204
ρ_5	0.250471	0.085081	2.943907

R-squared	0.896376
LLF	20.97757
DW-Stat	2.093864

Table 5.5(b)
Inflation-Unemployment Phillips Curve (The United States)
Data: Quarterly, 1975 (Q1)–2000 (Q4)

Variable	Estimated Coefficient	Std. Error	T-Statistic
c_t	0.555	0.459	1.210
δ_1	0.308	0.097	3.176
δ_2	0.272	0.098	2.777
δ_3	0.249	0.098	2.540

γ_1	-0.593	0.352	-1.683
γ_2	0.543	0.363	1.497

R-squared	0.61
LLF	-148.67
DW-Stat	2.07