MONEY, OUTPUT AND THE UNITED STATES’ INTER-WAR FINANCIAL CRISIS: AN EMPIRICAL ANALYSIS

by

PATRICK JAMES COE

B.A., University of Essex, 1990
M.A., University of Essex, 1992

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Economics

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

March 1998

© Patrick James Coe, 1998
ABSTRACT

In the first essay of this thesis I test long-run monetary neutrality (LRMN) using the long-horizon approach of Fisher and Seater [18]. Using United States' data on M2 and Net National Product they reject LRMN for the sample 1869-1975. However, I show that this result is not robust to the use of the monetary base instead of M2. Nor is it robust to the use of United Kingdom data instead of United States data. These results are consistent with the interpretation that Fisher and Seater's result is a consequence of the financial crisis of the 1930s causing inside money and output to move together. Using a Monte Carlo study I show that Fisher and Seater's rejection of LRMN can also be accounted for by size distortion in their test statistic. This study also shows that at longer horizons, power is very low.

In the second essay I consider the financial crisis of the 1930s in the United States as change in regime. Using a bivariate version of Hamilton's [24] Markov switching model I estimate the probability that the underlying regime was one of financial crisis at each point in time. I argue that there was a shift to the financial crisis regime following the first banking crisis of 1930. The crucial reform in ending the financial crisis appears to have been the introduction of the Federal Deposit Insurance Corporation in January 1934. I also find that the time series of probabilities over the state of the financial sector contain marginal explanatory power for output fluctuations in the inter-war period.

A problem when testing the null hypothesis of a linear model against the alternative of the Markov switching model is the presence of nuisance parameters. Consequently, the likelihood ratio test statistic does not possess the standard chi-squared distribution. In my third essay I perform a Monte Carlo experiment to explore the small sample properties of the
pseudo likelihood ratio test statistic under the non-standard conditions. I find no evidence of size distortion. However, I do find that size adjusted power is very poor in small samples.
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Long-run Monetary Neutrality in the United States and United Kingdom, 1869-1995</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>2.1 Introduction</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>2.2 Methodology</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2.3 Time Series Properties of the Data</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.4 Long-run Neutrality in the United States</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>2.5 Long-run Neutrality in the United Kingdom</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2.6 Small Sample Considerations</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>2.7 Conclusions</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>Financial Crisis and Financial Reform During the Great Depression</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>3.1 Introduction</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>3.2 Financial Crisis and the Great Depression</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>3.3 Banking and Monetary Reform in the New Deal</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>3.3.1 Banking Reform</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>3.3.2 Monetary Reform</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>3.4 A Model of Time Series Subject to a Change in Regime</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>3.5 Estimation Results</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>3.6 Financial Crisis and Output Growth</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>3.7 Conclusions</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>The Markov Switching Model and Linear Time Series</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>4.1 Introduction</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>4.2 A Two State and a One State Model of GDP</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>4.3 Estimates Using Canadian GDP</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>4.4 Monte Carlo Experiment</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td>4.4.1 Distribution of Test Statistic</td>
<td>77</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENT

I would like to thank my supervisory committee Anji Redish, Jim Nason, Bob Allen and Mick Devereux for the support and guidance they have given me while writing this thesis. While at the University of British Columbia I have also benefited from the advice of John Cragg and Ron Shearer. Some of the work on this thesis has been undertaken while I have been at the University of Essex. While I have been at Essex, Tim Hatton and Gordon Kemp have provided helpful comments on parts of this thesis. I would also like to thank my wife and daughter, Pam and Katie, and my parents, John and Sheila, for the encouragement and support they have given me over the years. Financial support from the Albert Whitley Memorial Scholarship, a UBC Graduate Fellowship and the Anthony Scott Fellowship are gratefully acknowledged.
Chapter 1

Introduction

The severity of the Great Depression in the United States, relative to other periods of economic contraction, has sometimes been attributed to the financial crisis experienced by the United States in the early 1930s. It has long been argued that the financial crisis played a role through monetary channels, see for example, Friedman and Schwartz [24]. However, it has also been argued that the financial crisis played a role through non-monetary channels, see Bernanke [3]. Although many economists have some sympathy for the latter view, the empirical support for this argument is mixed at best. This thesis is an empirical study to help clarify the role of the financial crisis experienced by the United States economy in the 1930s.

This thesis consists of three related essays in empirical macroeconomics. In the first, I argue that the results of long-run monetary neutrality (LRMN) tests which include United States’ data from the 1930s can be influenced by the financial crisis. I also undertake a Monte Carlo study into the small sample properties of the test statistic I employ in this essay. In the second essay, I model the financial crisis as a change in regime. I argue that the state of the financial sector contains additional explanatory power for United States’ output fluctuations in the inter-war period. I also argue that the establishment of the Federal Deposit Insurance Corporation was the reform that ended the financial crisis. This is consistent with a claim made by Friedman and Schwartz. One problem in testing
between a model incorporating changes in regime and a linear model is the presence of nuisance parameters. Therefore, in my final essay, I explore these problems in a Monte Carlo experiment.

In a recent paper, Fisher and Seater [22] find that long-run monetary neutrality does not hold for the United States over the period 1869-1975. They test the null hypothesis that a change in the current money stock, as measured by M2, has no effect on future real income at horizons of one to thirty years. They find that they can reject this null hypothesis in favor of the alternative that the effect is non-zero for almost all horizons. This leads them to reject LRMN.

In my first essay, I examine the robustness of this result to a change in the monetary aggregate and a change in country. Their result is not robust to either change. Using United States data on the monetary base I am unable to reject the long-run neutrality of money. Using United Kingdom data on the monetary base and on a broad monetary aggregate I am again unable to reject LRMN.

The fact that the result of non-neutrality in the United States is critically sensitive to the monetary aggregate, and that the result is specific to the United States, suggest that it reflects the financial crisis there. Boschen and Otrok [7] show that LRMN holds for M2 in United States both before and after the 1930s. They suggest that Fisher and Seater's result is a function of either the Great Depression or the financial crisis of the 1930s. My results suggest that it is the latter of these. That is, the rejection of LRMN found by Fisher and Seater is a result of the 1930s financial crisis in the United States. As Friedman and Schwartz point out, the United States monetary base was largely unaffected by the financial crisis. They argue that the decline in the money stock came through the money multiplier as the public substituted away from deposits and banks increased reserve ratios. According to Bernanke, the financial crisis also had real effects as the breakdown in credit intermediation led to a decline in real income. As a result, during the 1930s, changes in the broad money stock and real income became highly correlated. The fact that I do not reject LRMN using the monetary base is consistent
with Boschen and Otrok’s second suggestion. My results for the UK are also consistent with their explanation. Although there was a recession in the United Kingdom, there was no financial crisis in the United Kingdom and I cannot reject LRMN for either the broad or the narrow monetary aggregate.

In the second half of this essay, I look at the small sample properties of the test statistic. The test developed by Fisher and Seater is similar to the long-horizon regressions which have proved relatively successful in detecting predictable components in stock prices, interest rates and inflation.\(^1\) Campbell [10] suggests that this relative success is either due to power advantages or to greater size distortion. Since size distortion is a potential explanation for Fisher and Seater’s rejection of LRMN, I perform a Monte Carlo experiment to investigate the size and power properties of Fisher and Seater’s test statistic. I find that there are serious size distortions at the longer horizons. In fact, this size distortion is sufficient to account for Fisher and Seater’s rejection of LRMN. Using empirical critical values, generated by Monte Carlo simulation, I am unable to reject LRMN using their dataset and their sample period. On the other hand, I also find that size adjusted power is poor at long horizons.\(^2\)

The theme of financial crisis is continued in the second essay. During the Great Depression the United States experienced several waves of bank failure. As mentioned above, financial crisis can have both real and monetary effects. After Roosevelt took office in March 1933 he undertook a number of reforms intended to stabilize the financial sector, and the economy in general. Friedman and Schwartz argue that the crucial reform that delivered stability to the financial sector was the introduction of the Federal Deposit Insurance Corporation in January 1934. On the other hand, Wigmore [56] argues that it was the decision to abandon the gold standard in March 1933 that ended the financial

---

\(^1\)For stock prices, see, for example, Fama and French [20] [21]; for interest rates, see Campbell and Shiller [12]; for inflation see Mishkin [41] [40]. This success is relative to short-horizon regressions. Consider the case of stock returns, a long horizon regression requires measuring the return on the stock over a period that is longer than the sampling frequency of the data.

\(^2\)Kilian [33] also finds a lack of high power at long-horizons when looking at exchange rate predictability for conventional long-horizon tests. He suggests that a bootstrap test has smaller size distortions than the conventional long-horizon regression tests.
In my second essay, I model the financial sector as being in one of two states, financial crisis or financial calm. This underlying state of the financial sector is unobservable. However, by using Hamilton's [28] Markov switching model and observable data, I am able to estimate the conditional probability of the financial crisis state at each point in time. The observable data I use is on the growth rate of real deposits and a yield spread. The latter proxies the cost of credit intermediation. My estimated conditional probabilities suggest that the financial crisis began in late 1930, following the First Banking Crisis, and ended in early 1934, following the introduction of federal deposit insurance. This is consistent with Friedman and Schwartz's view, but inconsistent with Wigmore's view.

Testing between Hamilton's Markov switching model and a linear model is problematic due to the presence of nuisance parameters. I tackle this problem in two ways. Firstly, I generate an empirical distribution for the pseudo likelihood ratio test statistic under the null hypothesis that the data were generated by the linear model. I then compare my pseudo likelihood ratio test statistic from my sample data to this distribution. Using the empirical distribution suggests that the null of a linear model can be rejected in favor of the Markov switching model at the 1% significance level. The second way I tackle this problem is to follow the suggestion of Hamilton [27]. I conduct a Lagrange multiplier test on the one state model to see if a two state model is required. Again the evidence supports the use of a two state model.

In this essay, I also ask whether treating the financial crisis as a change in regime provides additional explanatory power for output fluctuations during the interwar period. I add the conditional probability of financial crisis to a reduced form equation for the growth rate of real output. This equation also contains current and lagged values of the growth rate of the money supply, the linear measures of financial crisis mentioned above, and lagged values of real output. F-tests based on two-stage least squares estimation show the null hypothesis that these conditional probabilities contain no additional explanatory power can be rejected at the 5% significance level.
Recently non-linear models such as Hamilton's have been used to model the business cycle. However, as noted above, testing between the Markov switching model and a linear model is problematic. Under the null hypothesis that the data are generated by a single state, linear model, the parameters which define the second state are undefined. As a result, the information matrix becomes singular, and the likelihood ratio test statistic does not possess an asymptotic $\chi^2$ distribution. Therefore, in my final essay I look at the problem of testing between the non-linear model and a linear model. I also assess the extent of bias induced from fitting a one state model, when the true process has two states and vice versa. In both cases model mis-specification is a source of bias. In the latter case there is an additional source of bias due to Jensen's inequality.

I explore these issues using artificial data series for the growth rate of quarterly Canadian real per capita GDP. These artificial series are generated using parameter estimates from a linear AR(1) model and Hamilton's Markov switching model using data from 1960 to 1995. I generate artificial samples of sizes 100, 200, 300 and 400. For all the sample sizes I consider, I find no evidence of size distortion in the pseudo likelihood ratio test statistic. However, for all sample sizes with the exception of 400, I find that the power of this test statistic is poor. Currently, most quarterly macroeconomic time series consist of between 100 and 200 observations. For a sample size of 100, size adjusted power is approximately 0.41 and for a sample size of 200 it is approximately 0.79. It is only when the sample size reaches 400 that size adjusted power reaches 0.95. 400 is approximately double the number of quarterly observations currently available for most macroeconomic aggregates.

I evaluate the overall effect of bias induced from estimating the mis-specified model on the basis of out-of-sample predictive accuracy. My results suggest that the extent of bias induced from estimating the Markov switching model when the true process has only one state is minimal. The loss in forecast accuracy is typically less than 1%. This is also true, though to a slightly lesser extent, for the estimation of a single state model when the true process is the Markov switching model. Here the loss in forecast accuracy
is typically less than 3%.
Chapter 2

Long-run Monetary Neutrality in the United States and the United Kingdom, 1869-1995

2.1 Introduction

Classical macroeconomics asserts that permanent movements in nominal-side economic variables, such as the money stock and the rate of inflation, will have no long-run effect on real-side economic variables, such as real income and the real interest rate. Economists refer to these conjectures as the long-run neutrality propositions. These neutrality propositions have important implications for policy-makers. For example, according to LRMN, a permanent change in the level of the money stock will have no long-run effect on real income. However, if the world is non-classical in the long-run, then changes in the money stock do not produce proportionate changes in the price level. Instead these changes lead to changes in both the price level and real income. In this world, the policymaker can use monetary policy to alter the long-run trend of real income. In this chapter I investigate the long-run neutrality of money using a time span of data of over a hundred years for both the United States and the United Kingdom. I test the proposition that permanent
changes in the level of the money stock have no permanent effect on real income.

In their recent paper, Fisher and Seater [22] estimate long-run derivatives of real income with respect to the money stock using annual United States data on M2 and Net National Product from 1869 to 1975. They find that they can reject the hypothesis that the derivatives of real income with respect to the money stock are equal to zero. Hence, using their dataset, they reject LRMN. Boschen and Otrok [7] argue that this result is driven by observations from the Great Depression. Conducting similar tests, but omitting the observations of the 1930s, they find that LRMN cannot be rejected for the United States for most of the period since 1869.1 One interpretation of this result is that the Great Depression was a unique event in US monetary history. Consequently, one might conclude that any long-run monetary policy implications drawn from this episode would be irrelevant.

I offer evidence of the robustness of Fisher and Seater’s result in two ways. To begin, I re-estimate their long-run derivative using the monetary base, M0, as the monetary aggregate instead of M2. I find that their result is not robust to this change. Next, I re-estimate their long-run derivative using data from the United Kingdom over a similar sample period. Again their result is not robust to the change. This is true regardless of whether I use the monetary base or a broad measure of the money stock (M3).2 I argue that these two results, and that of Boschen and Otrok, are consistent with the explanation that Fisher and Seater’s rejection of long-run neutrality is due to the existence of an omitted variable, that omitted variable being the financial crisis experienced by the United States in the 1930s.

Campbell [10] suggests that long-horizon tests of predictability, such as those used by Fisher and Seater, Boschen and Otrok, and in this chapter, may have greater power than their short-horizon counterparts or greater size distortions. Therefore, I also perform a

---

1 They re-estimate the long-run derivative using the same dataset, but split into two sub-periods. They find that long-run neutrality holds for both the 1869-1929 and the 1940-92 sub-periods.

2 Haug and Lucas [32] show that, in the absence of a financial crisis, LRMN also holds for the broad monetary aggregate in Canada. They use Canadian data on M2 and GNP for the period 1914 to 1994.
small Monte Carlo experiment to investigate the small sample properties of Fisher and Seater’s long-run neutrality test. I find that there is considerable size distortion in the test statistic. This offers an alternative explanation for Fisher and Seater’s rejection of long-run neutrality. I use empirical critical values rather than asymptotic critical values to calculate the confidence bands around their long-run derivatives. Using these empirical confidence bands I am unable to reject long-run neutrality using their dataset. I also find that in addition to size distortion, Fisher and Seater’s test has low power at the longer horizons. This suggests the test contains very little economic information.

In section 2.2 I discuss the problem with reduced form neutrality tests identified by Sargent [49], and Lucas [36] [37], and, using the example of King and Watson [34], show that there is a straightforward test of long-run neutrality if the process generating money contains a unit root. In this section, I also set up the framework of Fisher and Seater [22]. Section 2.3 presents the results of unit root tests on the data. In section 2.4, I show that the result of Fisher and Seater is not robust to a change in the monetary aggregate. Section 2.5 presents the neutrality results for the United Kingdom. In section 2.6, I report the results from the Monte Carlo study. Finally, section 2.7 concludes.

2.2 Methodology

Consider a rational expectations model in which real income is the dependent variable and money is the independent variable. Sargent [49] and Lucas [36] [37] argue that in such a model, estimated reduced form coefficients are not solely the monetary policy parameters. Instead, these estimated coefficients are a combination of monetary policy and structural parameters. Consequently, in a rational expectations model, with short run non-neutrality of money, the use of reduced form equations to test long-run neutrality can lead to inappropriate conclusions. However, it has recently been shown that in the case of integrated variables these tests are valid under certain restrictions. These restrictions involve the order of integration of the money stock and real income processes,
and the endogeneity of money.\footnote{For example see Fisher and Seater [22]; King and Watson [34]; Boschen and Mills [8].}

This point, and the Lucas-Sargent critique, is illustrated by the following example from King and Watson [34]. Consider a simple macroeconomic model consisting of the following three equations:

\begin{align*}
    y_t &= \delta (p_t - E_{t-1}p_t), \\
    p_t &= m_t - \gamma y_t,
\end{align*}

and

\begin{equation}
    m_t = \rho m_{t-1} + u_t. \tag{2.3}
\end{equation}

Here $y_t$ is the natural logarithm of real income, $p_t$ is the natural logarithm of the price level, $E_{t-1}p_t$ is its rational expectation based on the information available at time $t-1$ and $m_t$ is the natural logarithm of the nominal money stock. $\delta$, $\gamma$ and $\rho$ are fixed parameters.

The mean zero random variable, $u_t$ represents a temporary, unanticipated component of monetary policy. Equation (2.1) is an aggregate supply function, equation (2.2) is a monetary equilibrium condition and equation (2.3) is a money supply rule. The model can be solved to give the reduced form for income:

\begin{equation}
    y_t = \lambda (1 - \rho L) m_t, \tag{2.4}
\end{equation}

where $\lambda = \delta / (1 + \gamma \delta)$ and $L$ is the lag operator.\footnote{The model is solved by substituting (2.2) into (2.1) to give $y_t = \delta [m_t - \gamma y_t - E_{t-1} (m_t - \gamma y_t)]$. Then, using (2.3), $E_{t-1}u_t = 0$ and from (2.1), $E_{t-1}y_t = 0$, gives $y_t = \delta [m_t - \gamma y_t - \rho m_{t-1}].$ This can be rearranged to give (2.4).}

King and Watson construct this model in such a way that only the unanticipated component of monetary policy (i.e. $u_t$) has any effect on real income. LRMN requires that a permanent change in the level of the money stock has no permanent effect on the level of real income. However, equation (2.4) shows that a one unit increase in the money stock will have a non-zero effect on income. Suppose one were to estimate a
distributed lag model based on (2.4). Typically monetary neutrality would be rejected, even though the model has been constructed such that only transitory components of the money stock can affect income. In this case, a test of long-run neutrality should be conducted by looking at the structural parameters of the model rather than the reduced form coefficients.

There are two exceptions to this, when $\delta = 0$ and when $\rho = 1$. The second of these cases is the more interesting. Equation (2.4) shows that when $\rho = 1$ an increase in the money stock will have a zero effect on real income. This implies that if the money stock process contains a unit root then there is a straightforward test of the LRMN proposition. This result is exploited by Fisher and Seater. The remainder of this section summarizes their methodology for the case where the money stock and real income are both I(1). For a fuller treatment see Fisher and Seater [22].

Consider the following stationary, unrestricted VAR where $m$ is the natural logarithm of the nominal money stock, $y$ is the natural logarithm of real income and $\Delta = (1 - L)$,

$$
\begin{align*}
    a(L) \Delta m_t &= b(L) \Delta y_t + u_t, \\
    d(L) \Delta y_t &= c(L) \Delta m_t + v_t.
\end{align*}
$$

(2.5)

The error terms are distributed as follows:

$$
\begin{bmatrix}
  u_t \\
  v_t
\end{bmatrix} \sim i.i.d. (0, \Omega), \quad \text{where } \Omega =
\begin{bmatrix}
  \sigma_{uu}^2 & \sigma_{uv}^2 \\
  \sigma_{vu}^2 & \sigma_{vv}^2
\end{bmatrix}.
$$

(2.6)

Fisher and Seater express the long-run derivative of $y$ with respect to a permanent change in $m$ as:

$$
LRD_{y,m} \equiv \lim_{k \to \infty} \frac{\partial \Delta y_{t+k}}{\partial \Delta m_{t+k}} / \partial u_t = \theta,
$$

(2.7)

which is defined when $\lim_{k \to \infty} \partial m_{t+k} / \partial u_t \neq 0$. That is, when permanent changes to the
money stock process exist. Here money is long-run neutral when $\theta = 0$.\(^6\)

As suggested earlier in this section the degree of integration of the two time series plays an important role in the interpretation of the estimates of these long-run derivatives. Firstly, consider the case where the nominal money stock is integrated of order zero. Here no permanent changes to the money stock process exist, $\lim_{k \to \infty} \partial m_{t+k} / \partial u_t = 0$. Therefore, the long-run derivative is undefined and so long-run neutrality cannot be addressed. Next, consider the case where the nominal money stock is I(1) but real income is I(0). Here a permanent change in the money stock cannot lead to a permanent change in real income as no permanent changes in real income exist.\(^7\) Finally, consider the case where both the nominal money stock and real income are I(1). Now tests of long-run neutrality are valid because permanent changes to both $m$ and $y$ exist. Here long-run neutrality of money implies the following restriction:

$$\frac{c(1)}{d(1)} = \theta = 0. \quad (2.8)$$

The restriction for long-run neutrality above contains only parameters from the second equation of (2.5) which can be consistently estimated under either of two recursive identification schemes by ordinary least squares. However, as Fisher and Seater [22] point out the individual parameters of $c(L)$ and $b(L)$ are not of interest. What is of interest is the long-run derivative, that is $c(1)/d(1)$. Therefore, I follow Fisher and Seater and allow $b_0$ and $c_0$ to be unrestricted. I use the long-run exogeneity of money as an identifying restriction. This implies

$$b(1) = 0, \quad (2.9)$$

which states that a permanent change in $y$ has no long-run effect on $m$.

This is also one of the identifying restrictions considered by King and Watson [34]. An alternative long-run identifying restrictions is the long-run neutrality of money, $c(1) = 0$.

\(^6\)More generally, long-run neutrality implies that $\theta = 0$ when $y$ is a real variable or the nominal interest rate. It implies that $\theta = 1$ when $y$ is a nominal variable.

\(^7\)If a series is stationary around a deterministic trend it is treated as being I(0).
Alternatively, short-run identifying restrictions can be used, \( c_0 = 0 \) or \( b_0 = 0 \). However, in this chapter I am interested in the robustness of Fisher and Seater’s result and the performance of their test statistic in small sample. For this reason, I restrict my attention to their identifying restriction.

The second identifying restriction is placed on the covariance matrix. That is \( \sigma_{uv} = 0 \). Under these restrictions I now have a structural VAR. Fisher and Seater show, that with \( b(1) = \sigma_{uv} = 0 \), equation (2.8) can be estimated by \( \lim_{k \to \infty} \beta_k \) where \( \beta_k \) comes from the estimation of the following equation,

\[
\sum_{j=0}^{k} \Delta y_{t-j} = \alpha_k + \beta_k \sum_{j=0}^{k} \Delta m_{t-j} + \epsilon_{kt}.
\] (2.10)

This can be simplified to

\[
(y_t - y_{t-k-1}) = \alpha_k + \beta_k (m_t - m_{t-k-1}) + \epsilon_{kt}.
\] (2.11)

Under the null hypothesis of LRMN, \( \beta_k = 0 \).

Fisher and Seater estimate equation (2.11) for values of \( k \) from one to thirty. They then plot the \( \hat{\beta}_k \) together with 95% asymptotic confidence intervals obtained using Newey and West’s [46] procedure. These confidence intervals are from a t-distribution with \( T/k \) degrees of freedom, where \( T \) is the sample size.\(^8\) They find that zero is below the lower 95% confidence band for all values of \( k \) except \( 25 < k < 27 \). This implies a rejection of LRMN. As Boschen and Otrok [7] point out this is especially striking as the degrees of freedom used to calculate these confidence bands decline dramatically as \( k \) becomes large.

In this chapter I examine whether this result extends to another monetary aggregate and another country. However, as pointed out earlier in this section, to be able to test LRMN using this framework, I require that the money stock process and the real income

\(^8\)The number of degrees of freedom is \( T/k \) rather than \( T - k \) as this is the number of non-overlapping observations. See, for example, Hansen and Hodrick [31].
process are integrated of at least order one. The next section looks at the order of integration of the money stock and real income series.

2.3 Time Series Properties of the Data

This section presents some time series properties of United States and United Kingdom, annual data on the nominal monetary base ($b_t$), the nominal money stock ($m_t$), real income ($y_t$), the price level ($p_t$) and the nominal interest rate ($i_t$). The time period for this data is 1869 to 1995 for the United States and 1871 to 1982 for the United Kingdom. All variables, except the nominal interest rate, are expressed in natural logarithms. Appendix one describes the data and lists sources.

Using a long span of data is particularly advantageous when the presence or otherwise of a unit root has important implications for the long-run neutrality tests. This is because the power of unit root tests increases not so much with the number of observations, as with the time span of the data. See, for example, Pierse and Snell [47]. Table 2.1 presents the results of unit root tests. These are Augmented Dickey-Fuller (ADF) tests statistics. I also report 95% confidence intervals for the largest autoregressive root. These confidence intervals are calculated using the procedure developed by Stock [51]. The number of lags used to correct for autocorrelation in the ADF regression is chosen according to the procedure outlined in Campbell and Perron [11].

These statistics show that I can reject the null hypotheses that real income, the monetary base and the money stock are I(2) for both countries in favor of the alternative that they are I(1) or less. For the United Kingdom, I am unable to reject that these three series are I(1) in favor of the alternative that they are stationary around a deterministic trend. Therefore, the long-run neutrality of money can be tested using equation (2.11)

---

9I begin with a maximum lag order of $k_{\text{max}} = 10$. If this 10th lag is significant, at the 10% level, I set $k = 10$. If not, I reduce the number of lags by one until the last lag is significant. If no lags are significant, I set $k = 0$.  

10This means that the issue of the superneutrality of money cannot be addressed using this dataset, because there are no permanent changes to the growth rate of the money stock.
for both the monetary base and the money stock. For the United States I am unable to reject the null hypothesis of an I(1) process against the trend stationary alternative for either the monetary base or the money stock. However, for the United States real income series, I am able to reject the null hypothesis of an I(1) series against this alternative. This suggests that there are no permanent changes in real income and therefore, that LRMN holds. However, the debate that began with Nelson and Plosser [45] over whether or not US real income contains a stochastic or deterministic trend remains unresolved. Therefore, I also report the results of neutrality tests under the assumption that there are permanent changes to income. That is, that the process for real income does contain a unit root. As I show in the next section, the results based on this assumption are not inconsistent with the long-run neutrality of money.\footnote{In addition to the test statistics and confidence bands reported in table (2.1) I also calculated the test statistic proposed by Elliot, Rothenberg and Stock [18]. This tests the null of a unit root against the alternative with the most power. The conclusions drawn from these tests are the same as those drawn from the results reported in table 2.1.}

2.4 Long-run Neutrality in the United States

Figure 2-1 shows the long-run derivatives of real income with respect to the money stock. It contains the estimates of $\beta_k$ from equation (2.11) for values of $k$ from 1 to 30, and 95\% confidence bands calculated using the Newey-West [46] procedure. Panel (a) shows the long-run derivative of real income with respect to the money stock for the United States for the period 1869 to 1975. This replicates the result of Fisher and Seater, that is, when using asymptotic confidence bands I reject the null hypothesis that the long-run derivative is equal to zero for most values of $k$. The only exception I find is for $21 \leq k \leq 27$. These positive and statistically significant estimates of the long-run derivatives are inconsistent with the hypothesis of LRMN. The dataset I use in this chapter is slightly different from that of Fisher and Seater, and this explains why the range of $k$ for which the null
hypothesis cannot be rejected differs slightly.\textsuperscript{12}

Next I extend the sample period to 1995. Panel (b) shows the result of the estimate of the same equation using data from 1869 to 1995. Again the result does not appear to be completely consistent with the notion of long-run neutrality. For all values of $k < 18$, zero does not lie between the 95\% asymptotic confidence bands. However, it does lie between these bands for all values of $k \geq 18$. This result is consistent with the observation made by Friedman and Kuttner [23] using quarterly, post-war data. Although the frequency and the time frame of the data are different, the general conclusion is the same. That is, including data from the 1980s weakens the evidence that there is a significant relationship between money and income. In addition to this, for the higher values of $k$, the point estimates for $\beta_k$ are lower than those for the 1869-1975 sample.

I also estimate the long-run derivatives of output with respect to money using this dataset for the two sub-samples that Boschen and Otrok [7] identify. They show that if the sample period is split into two sub-samples, one before the Great Depression and one after, then zero lies within the 95\% asymptotic confidence bands in both sub-samples. I find that this result also extends to this dataset. However, one should note that the degrees of freedom used to construct the confidence intervals around $\beta_k$ are calculated using $T/k$, and therefore decline rapidly as $k$ increases. For example, the confidence bands on $\beta_{30}$ for the pre-1930 sub-period, for example, are calculated using only two degrees of freedom.

The results of Boschen and Otrok suggest that Fisher and Seater's result is not robust to a change in the sample period. Here I show that Fisher and Seater's result is also not robust to a change in the monetary aggregate. Figure 2-2 shows the estimation of the long-run derivatives of real income with respect to the monetary base for the same two sample periods as in figure 2-1. Unlike Fisher and Seater's result for the money stock, I

\textsuperscript{12}Fisher and Seater used data on M2 and Net National Product from Friedman and Schwartz [25]. Here I use data the data from Balke and Gordon [2] as it provided a cleaner match with the data from the Economic Report of the President [17]. The Balke and Gordon dataset is also used in other studies. For a recent example, see Bordo, Goldin and White [6].
find that in general, zero is contained within the asymptotic 95% confidence bands.\(^{13}\)

Panel (a) shows that for the sample 1869 to 1975 I am unable to reject the null hypothesis that the long-run derivative is equal to zero for all values of \(k\) except \(k \leq 4\) and \(k = 13, 14\). This is Fisher and Seater’s sample period. When the sample is extended to 1995 the lower band is only above zero for \(k < 4\). This result is shown in panel (b).\(^{14}\) Therefore Fisher and Seater’s result is dependent not only on the sample period, as suggested by Boschen and Otrok, but also on the choice of monetary aggregate.

The point estimates of the long-run derivatives also differ across monetary aggregate. When the money stock is used the point estimate is approximately 0.4 for all values of \(k = 1, \ldots, 30\). When the monetary base is used the point estimate of the long-run derivative is 0.25 for horizons of up to 10 years. After this it falls, and is at zero by the 20 year horizon.

Boschen and Otrok\(^{7}\) point out that with the exception of the 1930s, M2 and income both exhibited a positive trend between 1869 and 1992, and that these trends were at different rates. During the 1930s however the two series both moved downwards in a parallel fashion. They offer two possible explanations for Fisher and Seater’s non-neutrality result. The first is that exogenous monetary shocks during the Great Depression were sufficient to cause a breakdown of long-run neutrality. The second is that the events in the financial markets during the 1930s were reflected in the broad monetary aggregate. This implies that the assumption that changes in this monetary aggregate reflect purely exogenous monetary shocks may be inappropriate. Boschen and Otrok suggest that if one believes the second explanation, the maintained hypothesis of the long-run exogeneity of money is inappropriate.

The results presented above using the monetary base help shed some light upon this issue. If shocks to outside money during the Great Depression were in fact sufficient to

\(^{13}\)I also use the data on high powered money from Friedman and Schwartz\(^ {25}\) to represent the monetary base, and their Net National Product series used by Fisher and Seater\(^ {22}\) to represent output and find that I am unable to reject long-run neutrality for the 1869-1975 sample.

\(^{14}\)I also find that zero lies between the two asymptotic confidence bands for the two sub-periods considered by Boschen and Otrok when using the monetary base.
break long-run neutrality, then LRMN should be rejected for the monetary base as well as M2. However, if the rejection is a result of shocks to the financial sector affecting inside money and real income in a similar fashion, then one would expect to see long-run neutrality hold for the monetary base.

Consider Friedman and Schwartz’s [24] proximate determinants approach to the money stock,

\[ M = \psi \left( \frac{D}{C}, \frac{R}{D} \right) \times B \]

(2.12)

where \( M \) is the broad money stock, \( B \) is the monetary base and \( \psi \) is the money multiplier. The latter depends positively on the deposit-currency ratio, \( \frac{D}{C} \), and negatively on the reserve-deposit ratio, \( \frac{R}{D} \). If, in a financial crisis, the public reduces its deposit-currency ratios and the banks increase reserve ratios, then, for a given level of the monetary base, the money stock must fall. Therefore, a financial crisis can have a negative effect on the money stock through the adjustment of agents’ portfolios.

Bernanke [3] argues that the financial crisis of the 1930s had non-monetary effects in addition to its monetary effects. He argues that the financial crisis led to a breakdown in credit intermediation. This in turn meant that less lending took place in equilibrium. The lower levels of lending translated into lower investment and lower real income. Consequently, the financial crisis also had real, non-monetary effects.

Therefore during a financial crisis it is possible that the broad money stock and real income could fall while the monetary base remains unchanged, or even rises. Figure 2-3 shows the growth rates of real income, the money stock and the monetary base for the period 1927 to 1941. It also shows the spread between the yield on bonds rated Baa by Moodys and the yield on long-term government bonds for the same period.\(^{15}\) This spread is used as a proxy for financial crisis.\(^{16}\) It is apparent that, beginning in 1931 the

---

\(^{15}\) The data on these yields comes from table 128 of The Board of Governors of the Federal Reserve System [4].

\(^{16}\) Banks perform an intermediary role between lenders and borrowers. They screen and monitor borrowers. The charge for doing this is the cost of credit intermediation. This is the difference between the interest rate a borrower pays and the interest rate the depositor receives. If, during a financial crisis, banks fear a run they will increase reserve-deposit ratios and substitute into safer and more liquid
growth rates of real income and the money stock become very similar, while this is not the case for the growth rates of real income and the monetary base. At this point the yield spread also begins to rise from about 2% in 1927-30 to a peak of almost 6% in 1932. This is consistent with a period of financial crisis.

If the financial crisis caused the money stock to fall through the money multiplier, and also had an effect on real income through the channel proposed by Bernanke, then, one would expect to see the growth rates of real income and the money stock follow a similar path. That fact the monetary base does not share this path is consistent with the financial crisis causing the declines in the money stock. However, it is not consistent with an explanation in which the money stock falls through an exogenous fall in high powered money. This suggests that the apparent non-neutrality Fisher and Seater find is produced by an omitted variable.\(^\text{17}\) That is, the co-movement of the money stock and real income is not because the money stock causes income, but because, during the 1930s, the financial crisis caused both variables.

### 2.5 Long-run Neutrality in the United Kingdom

Assuming that the rejection of long-run neutrality Fisher and Seater find is due to the financial crisis in the United States in the early 1930s, one would expect to find that long-run neutrality would hold in the absence of a financial crisis. The United Kingdom did not experience such a financial crisis during the sample period. This suggests that long-run neutrality should not be rejected using United Kingdom data regardless of the choice of monetary aggregate. Therefore, in this section, I repeat the LRMN tests using data from the United Kingdom.

---

\(^{17}\)See, for example, Lütkepohl [38].
Figure 2-4 contains estimates of the long-run derivatives of income with respect to the money stock and the money base using United Kingdom data from 1871 to 1982. These figures show that zero lies between the asymptotic 95% confidence bands for both the money stock or the monetary base. That is, in each case, this is true for all values of $k$. In fact the point estimates of these long-run derivatives are typically close to zero, and never above 0.2.

The United Kingdom did not experience a financial crisis as severe as that in the United States during the Great Depression throughout this sample period and long-run neutrality holds for the broad money stock as well as for the monetary base. For example, while the United States experienced a loss of depositor confidence in its financial sector in the early 1930s with the deposit-currency ratio falling from 11.2 in 1930 to 5.1 in 1933, depositor confidence in the UK appeared unchanged with the deposit-currency ratio remaining constant at about 7.5 during these years. In the previous section I used a yield spread as an indicator of financial crisis for the United States. Unfortunately, to the best of my knowledge, no comparable spread exists for the United Kingdom for the interwar period. In the next chapter, I suggest that the negative growth rate of real deposits in the United States in the early 1930s is consistent with a period of financial crisis. There I present data which show that real deposits grew steadily in the early 1930s in the United Kingdom. A result in which long-run neutrality holds for the money stock in the United Kingdom is consistent with an apparent rejection of long-run neutrality of M2 in the United States being the result of the financial crisis of the 1930s.

---

18 The figures for the United States are from Friedman and Schwartz [24] appendix B. The figures for the United Kingdom are from Capie and Webber [14] table I(3). While some of the fall in deposits in the United States was due to bank suspensions, this amount counted for only $0.8bn of the $12.3bn decline. Therefore the decline in deposits was mostly due to portfolio readjustment. These two figures are also from Friedman and Schwarz, table 16 and appendix A, respectively.

19 Capie and Collins [13] provide a summary of some of the data on interest rates available for the United Kingdom during the interwar period.
2.6 Small Sample Considerations

Campbell [10] points to the relative success of long-horizon regressions, such as those used in this chapter, in detecting predictable components in stock returns, interest rates and inflation. For stock prices, see for example, Fama and French [20] [21]; for interest rates, see Campbell and Shiller [12]; for inflation, see Mishkin [41] [40]. He suggests that these findings indicate that long-horizon tests of predictability have either greater power than their short-horizon counterparts or that they have greater size distortions. Size distortion in the test statistic offers an alternative explanation for Fisher and Seater's [22] rejection of LRMN. On the other hand, low power could account for my failure to reject LRMN. Therefore, in this section I perform a small Monte Carlo experiment to investigate the small sample properties of the test statistic.

I generate artificial data for the Monte Carlo experiment based on parameter estimates from the system described by equation (2.5). To estimate this system I follow the instrumental variables estimation procedure of King and Watson [34], which I outline in appendix two.20 I use the same data on real output and the monetary base from the United States as earlier in this chapter.21

First I estimate the system with only the identifying restrictions, \( b(1) = \sigma_{uv} = 0 \) imposed. These parameter estimates are then used to generate artificial data for the monetary base and real income under the alternative hypothesis. I then estimate the system with the long-run neutrality restriction, \( c(1) = 0 \), also imposed. This second set of parameter estimates are then used to generate artificial data for the two series under the null hypothesis. In each case I generate \( T + 200 \) observations, where \( T \) is the sample size. The first 200 observations are discarded to minimize the influence of starting values.

20Instrumental variables estimation is used to account for the potential correlation between the current value of the growth rate of output and the error term in the equation for the growth rate of money, and between the current value of the growth rate of money and the error term in the equation for the growth rate of output.

21I also estimated the system using the money supply series as opposed to the monetary base series, and then generated artificial data using these parameter estimates. This did not alter the results of the Monte Carlo experiment.
I generate samples of sizes, $T = 50, 75, 100, 125$ and $150$.

Using these artificial data series I then estimate equation (2.11) for all values of $k = 1, 2, \ldots, 30$ and calculate the $t$ statistic for test of the null hypothesis that $\beta_k = 0$, against the alternative that $\beta_k \neq 0$. For each sample size I repeat the experiment 2500 times. I then calculate the type I error rate and the empirical 95% critical value when the artificial data is generated under the null hypothesis, and the size adjusted type II error rate when it is generated under the alternative hypothesis.

For each value of $k$ Fisher and Seater construct a 95% confidence interval around $\beta_k$, where $\beta_k$ is estimated from equation (2.11). This confidence interval is constructed using the asymptotic $t$ distribution with $T/k$ degrees of freedom. The null hypothesis, that $c(1) = 0$, is then rejected if zero does not lie within the confidence interval. This is equivalent to a $t$ test of the null hypothesis of $\beta_k = 0$ against the alternative that $\beta_k \neq 0$. The critical value for this test comes from the asymptotic $t$ distribution with $T/k$ degrees of freedom.

Consider the case where the sample size is $T = 150$ and $k = 30$. For a two sided test, the 95% critical value from the asymptotic $t$ distribution with $T/k = 5$ degrees of freedom is 2.571. For each set of artificial data I calculate the $t$ statistic described above. If, when the artificial data are generated under the null hypothesis only 5% of the artificial $t$ statistics are above 2.571, then there is no size distortion in the test statistic. However, if more than 5% of the artificial test statistics are above 2.571 then the type I error rate is above 0.05. This would imply that there is size distortion in the test statistic. On the other hand when the data are generated under the alternative then only 5% of the artificial test statistics should be below 2.571 if the test is to have the correct power.

Table 2.2 contains the results of this Monte Carlo study. Panel (a) shows the type I error rates for values of $k = 1, 2, 5, 10, 20$ and 30. These results show that there are size distortions for all combinations of $k$ and $T$. That is, whatever the horizon, $k$, and the sample size, $T$, the type I error rate is greater than 0.05. I also find that these size
distortions are greater at the longer horizons.\textsuperscript{22} When \( k = 1 \) the true null hypothesis is rejected for between 15\% and 20\% of the simulated time series. When \( k = 30 \) this figure rises to between 25\% and 45\%. This is consistent with Campbell's suggestion that the observed relative success of long-horizon regressions in detecting predictable components maybe due to more serious size distortions than in short-horizon regressions. Panel (b) of table 2.2 contains the empirical 95\% critical values for the test statistics. Not surprisingly, given the size distortion, these are all above their asymptotic counterparts. Campbell also finds that empirical critical values are above their asymptotic counterparts.\textsuperscript{23}

Campbell suggests that an alternative reason for the relative success of long-horizon regression is that they have power advantages over their short-horizon counterparts. I find evidence to suggest that this is not the case. I use my empirical critical values from table 2.2 to calculate the size adjusted power of the test statistic for different values of \( k \) and \( T \). These figures are also reported in panel (c) of table 2.2. These results show that power is poorer at the longer horizons. For example, consider a sample size of \( T = 100 \). For the shorter horizons of \( k = 1 \) and \( k = 2 \), the size adjusted type II error rate is below 0.10. However, for the longer horizons of \( k = 20 \) and \( k = 30 \) the size adjusted type II error rate is above 0.50. When \( k = 30 \), the false null hypothesis is rejected in only 35\% of the simulations. Power at all horizons increases with sample size. However, even at a sample size of \( T = 150 \), larger than any considered in this chapter, power is only 0.51 at the longest horizon of \( k = 30 \).

This result, combined with the evidence of size distortion suggests that the relative success of long-horizon regressions in detecting predictable components is the result of size distortion rather than power advantages. It also suggests the possibility that Fisher and Seater's \textsuperscript{22} rejection of long-run neutrality could be the result of size distortion.

\textsuperscript{22}There is one exception to this, for \( T = 50 \) the type I error rate is lower for \( k = 30 \) than it is for \( k = 20 \). This can be explained by the fact that the degrees of freedom for these test statistics are calculated as \( T/k \). Therefore for \( T = 50 \) there is a large jump in the asymptotic critical value between \( k = 20 \) and \( k = 30 \). As a result a much higher value of the test statistic is required to reject the null hypothesis.

\textsuperscript{23}Campbell presents results for weighted and unweighted long-horizon regressions. The results presented here are analogous to his results for unweighted regressions.
For this reason I generate empirical critical values using Fisher and Seater's dataset. To do this, I estimate the system described by (2.5) under the null hypothesis, using their dataset. I then use these parameter estimates to generate artificial data for M2 and Net National Product. Finally, I estimate equation (2.11) using this artificial data and calculate the $t$ statistic for the test of $\beta_k = 0$ for $k = 1, 2, \ldots, 30$. I repeat this experiment 2500 times and generate an empirical 95% critical value for the test statistic for each value of $k$.

With these empirical critical values I then calculate 95% confidence intervals for Fisher and Seater's long-run derivative. Figure 2-5 shows their long-run derivative, first with asymptotic confidence intervals, and second with the empirical confidence intervals. By using empirical confidence intervals I am able to overturn their rejection of long-run neutrality. I am only able to reject the null hypothesis that the long-run derivative is equal to zero for $k \leq 3$ and $10 \leq k \leq 15$. For all other values of $k$ the null hypothesis cannot be rejected. This suggests another reason for Fisher and Seater's rejection of LRMN. That is, size distortion in the test statistic.

I also generate empirical confidence intervals for my long-run derivatives contained in figures 2-2 and 2-4. That is, the long-run derivatives for the United States data using the monetary base, and for the United Kingdom, using both the money stock and the monetary base. In each of these cases I am unable to reject the null hypothesis when using asymptotic confidence intervals. Unsurprisingly, given the results of table 2.2, these empirical critical values are wider than their asymptotic counterparts and so do not alter my conclusions. To summarize these results, panel (a) of table 2.3 reports the ratio of the empirical to the asymptotic critical value for values for these long-run derivatives.

While size distortion may account for Fisher and Seater's rejection of the null hypothesis, the power properties of the test statistic in table 2.2 suggest that my failure to reject long-run neutrality could be a result of the low power of the test statistic. As An-

---

24Fisher and Seater use a sample size of $T = 107$. As before, I generate an artificial data series of length $T + 200$ and discard the first 200 observations in order to minimize the effect of starting values.
Andrews [1] points out, failure to reject the null hypothesis does not mean that the whole of the alternative parameter space is inconsistent with the data. Andrews shows that by using the inverse power function it is possible to determine the region of the alternative parameter space that is inconsistent with the data. He also shows that it is possible to determine regions of high probability of type II error. Andrews refers to this as the region of low power. Even when an alternative parameter value in this region of low power is true, one still has a greater chance of failing to reject the false null hypothesis than one does of rejecting it. That is, the probability of type II error in this region is greater than 50%.

Panel (b) of table 2.3 shows the regions of low power for the cases in which I could not reject the null hypothesis of LRMN. Following Andrews, I calculate the region of low power as:

\[
\{ \beta : 0 < |\beta| < b \},
\]

where \( b \) is calculated as

\[
b = \lambda_{1,\alpha} \left( \frac{1}{2} \right) \hat{\sigma}_\beta,
\]

\( \hat{\sigma}_\beta \) is the estimate of the standard error of \( \hat{\beta} \) and \( \lambda_{1,\alpha} \left( \frac{1}{2} \right) \) is a positive constant whose value is taken from table 1 of Andrews [1]. \( \alpha \) refers to the significance level, in this case \( \alpha = 0.05 \). When the test fails to reject the null hypothesis it does not provide any evidence that the true parameter value lies outside the region described by equation (2.13).

For the cases where I am unable to reject the null hypothesis of \( \beta_k = 0 \), the upper bound on the regions of low power are typically between 0.1 and 0.2. For example, for the United Kingdom results using the monetary base, the inverse power function implies that the test provides no evidence that the long-run derivative is less than 0.14 for \( k = 30 \). Assuming that the velocity of circulation of the monetary base is constant, this figure of \( \beta_{30} = 0.14 \) implies that a 10% increase in the monetary base today would cause real

---

25 Inconsistent with the data in the following sense. The probability of type two error when the null cannot be rejected at the 5% significance level is also 5%.
output to be 1.4% higher, and the price level to be 8.6% higher in 30 years time. In other words, a change in the monetary base today has a non-zero long-run effect on real output.

The figures are calculated using an inverse power function derived using the asymptotic $\chi^2(1)$ distribution. In small samples the student's $t$ distribution approximates the $\chi^2$ distribution. If I use my empirical $t$ distribution to calculate the inverse power function the regions of low power are much larger. This is especially true as $k$ becomes larger. For example, the figure of 0.14 above becomes 0.55.

In summary, the results of this section suggest that Fisher and Seater's test does not perform well in small samples. I present evidence to show that there is serious size distortion in the test statistic at all horizons and that power is low at the longer horizons. This implies that in small samples there is very little economic information in their methodology.

2.7 Conclusions

In this chapter, I argue that the apparent violation of long-run neutrality of money found by Fisher and Seater [22] is dependent upon their choice of monetary aggregate. I find that their result is not robust to replacing the broad money stock with the monetary base. This is consistent with Boschen and Otrok's [7] explanation that long-run neutrality is violated when using M2 because the financial crisis of the 1930s had non-monetary effects as well as monetary effects. This led income and broad money changes to become highly correlated during the Great Depression and therefore caused an apparent violation of long-run neutrality. In addition, I show Fisher and Seater's result does not carry over to the United Kingdom, where there was no financial crisis. This provides further evidence that Fisher and Seater's rejection of long-run neutrality in the United States is a consequence of the financial crisis of the 1930s.

I also present a Monte Carlo experiment to examine the size and power of Fisher and
Seater's test statistic. The results of this experiment show that these tests have serious size distortions. In fact these size distortions are great enough that when using empirical critical values the confidence intervals on Fisher and Seater's long-run derivatives become sufficiently wide that I am unable to reject LRMN using their dataset. In general the test also has poor power properties. Power decreases with the length of the long-run derivative. For example, at the longer horizons of 20 and 30 years power is typically below 50%, while for the shorter horizons of one and two years it is typically above 90%. These results suggest that any observed success of long-horizon regressions, relative to short-horizon regressions, in detecting predictable components is due to size distortion rather than power advantages. This implies that in small samples there is very little economic information in the methodology proposed by Fisher and Seater.
Table 2.1: Unit Root Statistics

<table>
<thead>
<tr>
<th></th>
<th>Demeaned</th>
<th></th>
<th></th>
<th>Detrended</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\tau}^{\mu}$</td>
<td>95% C.I.</td>
<td>$k,T$</td>
<td>$\hat{\tau}$</td>
<td>95% C.I.</td>
<td>$k,T$</td>
</tr>
<tr>
<td>(a) United States 1869-1995</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_t$</td>
<td>-3.802</td>
<td>(-, 0.971)</td>
<td>1,125</td>
<td>$b_t$</td>
<td>-2.770</td>
<td>(0.819, 1.029)</td>
</tr>
<tr>
<td>$m_t$</td>
<td>-3.568</td>
<td>(0.715, 1.014)</td>
<td>1,125</td>
<td>$\Delta y_t$</td>
<td>-4.760</td>
<td>(-, 0.81)</td>
</tr>
<tr>
<td>$\Delta b_t$</td>
<td>-3.395</td>
<td>(0.707, 0.960)</td>
<td>9,116</td>
<td>$\Delta m_t$</td>
<td>-6.016</td>
<td>(-, -)</td>
</tr>
<tr>
<td>$p_t$</td>
<td>-1.645</td>
<td>(0.927, 1.040)</td>
<td>5,121</td>
<td>$i_t$</td>
<td>-2.237</td>
<td>(0.860, 1.023)</td>
</tr>
<tr>
<td></td>
<td>-2.189</td>
<td>(0.870, 1.035)</td>
<td>2,124</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) United Kingdom 1871-1982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_t$</td>
<td>-2.065</td>
<td>(0.878, 1.041)</td>
<td>1,110</td>
<td>$b_t$</td>
<td>-2.314</td>
<td>(0.851, 1.039)</td>
</tr>
<tr>
<td>$m_t$</td>
<td>-0.130</td>
<td>(1.015, 1.049)</td>
<td>2,109</td>
<td>$\Delta y_t$</td>
<td>-5.572</td>
<td>(-, 0.674)</td>
</tr>
<tr>
<td>$\Delta b_t$</td>
<td>-5.127</td>
<td>(-, 0.748)</td>
<td>0,110</td>
<td>$\Delta m_t$</td>
<td>-3.545</td>
<td>(0.665, 0.942)</td>
</tr>
<tr>
<td>$p_t$</td>
<td>-1.618</td>
<td>(1.021, 1.057)</td>
<td>7,104</td>
<td>$i_t$</td>
<td>1522</td>
<td>(1.012, 1.050)</td>
</tr>
<tr>
<td></td>
<td>0.403</td>
<td>(1.018, 1.053)</td>
<td>4,107</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All statistics are based on the Augmented Dickey-Fuller regression

$$\Delta x_t = \mu_0 + \mu_1 \tau + \rho x_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta x_{t-i} + \varepsilon_t$$

where $\tau$ is a linear time trend and $\varepsilon_t \sim i.i.d. N(0, \sigma^2)$. The term demeaned refers to the case where $\mu_0$ is unrestricted, but $\mu_1 = 0$. The term detrended refers to the case where
both $\mu_0$ and $\mu_1$ are unrestricted. $\hat{\theta}_\mu$ and $\hat{\theta}_T$ are the test statistics in the demeaned and detrended cases respectively. The figures in parentheses are 95% confidence intervals for the largest autoregressive root calculated using the procedure outlined by Stock [51]. The lag length $k$ was chosen using the procedure described in the text. $T$ refers to the number of observations used to estimate the statistics.
Table 2.2: Results from Monte Carlo Simulations

(a) Type I Error Rate

<table>
<thead>
<tr>
<th>$T$</th>
<th>$k = 1$</th>
<th>$k = 2$</th>
<th>$k = 5$</th>
<th>$k = 10$</th>
<th>$k = 20$</th>
<th>$k = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.15</td>
<td>0.16</td>
<td>0.19</td>
<td>0.28</td>
<td>0.37</td>
<td>0.26</td>
</tr>
<tr>
<td>75</td>
<td>0.15</td>
<td>0.14</td>
<td>0.17</td>
<td>0.27</td>
<td>0.36</td>
<td>0.37</td>
</tr>
<tr>
<td>100</td>
<td>0.15</td>
<td>0.13</td>
<td>0.16</td>
<td>0.24</td>
<td>0.34</td>
<td>0.36</td>
</tr>
<tr>
<td>125</td>
<td>0.15</td>
<td>0.13</td>
<td>0.15</td>
<td>0.23</td>
<td>0.36</td>
<td>0.41</td>
</tr>
<tr>
<td>150</td>
<td>0.17</td>
<td>0.14</td>
<td>0.15</td>
<td>0.23</td>
<td>0.37</td>
<td>0.43</td>
</tr>
</tbody>
</table>

(b) Empirical 95% Critical Values

<table>
<thead>
<tr>
<th>$T$</th>
<th>$k = 1$</th>
<th>$k = 2$</th>
<th>$k = 5$</th>
<th>$k = 10$</th>
<th>$k = 20$</th>
<th>$k = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.92</td>
<td>3.00</td>
<td>3.59</td>
<td>5.54</td>
<td>8.17</td>
<td>8.96</td>
</tr>
<tr>
<td>75</td>
<td>2.85</td>
<td>2.85</td>
<td>3.22</td>
<td>4.52</td>
<td>7.09</td>
<td>8.48</td>
</tr>
<tr>
<td>100</td>
<td>2.73</td>
<td>2.68</td>
<td>3.03</td>
<td>4.19</td>
<td>6.40</td>
<td>7.89</td>
</tr>
<tr>
<td>125</td>
<td>2.74</td>
<td>2.63</td>
<td>2.88</td>
<td>3.82</td>
<td>5.67</td>
<td>7.49</td>
</tr>
<tr>
<td>150</td>
<td>2.78</td>
<td>2.66</td>
<td>2.77</td>
<td>3.84</td>
<td>5.67</td>
<td>7.21</td>
</tr>
</tbody>
</table>

(c) Size Adjusted Power

<table>
<thead>
<tr>
<th>$T$</th>
<th>$k = 1$</th>
<th>$k = 2$</th>
<th>$k = 5$</th>
<th>$k = 10$</th>
<th>$k = 20$</th>
<th>$k = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.62</td>
<td>0.64</td>
<td>0.56</td>
<td>0.37</td>
<td>0.27</td>
<td>0.21</td>
</tr>
<tr>
<td>75</td>
<td>0.82</td>
<td>0.84</td>
<td>0.77</td>
<td>0.57</td>
<td>0.36</td>
<td>0.30</td>
</tr>
<tr>
<td>100</td>
<td>0.92</td>
<td>0.94</td>
<td>0.88</td>
<td>0.69</td>
<td>0.44</td>
<td>0.35</td>
</tr>
<tr>
<td>125</td>
<td>0.97</td>
<td>0.98</td>
<td>0.95</td>
<td>0.82</td>
<td>0.57</td>
<td>0.44</td>
</tr>
<tr>
<td>150</td>
<td>0.99</td>
<td>0.99</td>
<td>0.97</td>
<td>0.88</td>
<td>0.63</td>
<td>0.51</td>
</tr>
</tbody>
</table>

These results are based on 2500 repititions of the Monte Carlo experiment described in the text. The type I error rate is based on the 95% asymptotic critical values. The size adjusted power is based on a nominal size of 0.05.
The ratio of empirical to asymptotic critical values are based on a Monte Carlo experiment which is the similar to that which produced the results of table 2.3. For the United States results, artificial samples of size $T = 107$ and $T = 127$ are produced from parameter estimates of equation (2.5) using United States data on real income and the monetary base. For the United Kingdom results, two samples of size $T = 112$ are produced. One based on parameter estimates using United Kingdom data on real income and the monetary base and the other based on parameter estimates using real income and the money supply. In each case 2500 repetitions of the experiment are performed. The region of high probability of type II error is the region of the alternative parameter space for which the probability of type two error is greater than one half. In each case, the first two rows refer to the long run derivatives shown in figure 2-2 and the last two rows to the long-run derivatives shown in figure 2-4.
Figure 2.1: United States Long-run Derivatives: Money Stock

Fig 2-1a: Income on Money 1869-1975

Fig 2-1b: Income on Money 1869-1995
Figure 2.2: United States Long-run Derivatives: Monetary Base

Fig 2-2a: Income on Monetary Base 1869-1975

Fig 2-2b: Income on Monetary Base 1869-1995
Figure 2.3: United States Growth Rates and Yield Spread

Fig 2-3a: Growth Rates for y, m & b; U.S. 1919-1941

Fig 2-3b: Yield Spread; U.S. 1919-1941
Figure 2.4: United Kingdom Long-run Derivatives

Fig 2-4a Income on Money 1871-1982

Fig 2-4b Income on Monetary Base 1871-1982
Figure 2.5: Fisher and Seater's Long-run Derivatives

Fig 2-5a Income on Money; Asymptotic Critical Values

Fig 2-5a Income on Money; Empirical Critical Values
Chapter 3

Financial Crisis and Financial Reform during the Great Depression.

3.1 Introduction

In this chapter I argue that the financial crisis experienced in the United States during the Great Depression should be considered as a change in regime. I provide evidence to suggest that the disruptions to the financial sector during the early 1930s should be thought of as a shift to a regime of financial crisis from one in which the financial sector was relatively calm. I find that treating the financial crisis as a change in regime provides additional explanatory power for output during the interwar period. My results are consistent with Friedman and Schwartz's [24] view that the most important policy reform for the ending of the financial crisis was the introduction of the Federal Deposit Insurance Corporation (F.D.I.C.) in January 1934.

I model the financial sector as being in one of two possible regimes. The first is a regime of financial crisis, and the second a regime I call financial calm. Clearly, the current state of the financial sector in the 1930s is not directly observable by the mod-
ern day econometrician, or even by the econometrician of the 1930s. However, using a bivariate version of Hamilton's [28], Markov switching model I am able to estimate probabilities over the underlying state of the financial system using observable time series. The behavior of these two observable time series is dependent on the underlying state of the financial system. More specifically, the growth rate of real deposits is lower in the financial crisis regime and the cost of credit intermediation is higher in the financial crisis regime. The evidence suggests that the financial sector was in a state of crisis from late 1930 until early 1934. I find that these estimated probabilities over the state of the financial sector contain additional explanatory power for output fluctuations, above their linear counterparts, during the interwar period.

One of the more striking features of the Great Depression was the high incidence of bank failure. In 1930 a total of 1350 banks suspended payments; in 1931 this figure was 2293; in 1932 it was 1453 and by April 1933 over 3800 banks had suspended payments. However, in the remainder of 1933 less than 200 banks suspended payments and the following two years saw a total of only 91 suspensions.¹

Friedman and Schwartz [24] attribute this sudden decline in the number of bank suspensions to the introduction of the F.D.I.C. in January 1934. Wigmore [56], however, sees the crucial reforms as being the gold embargo and the exchange controls which began to be introduced by Roosevelt in the Spring of 1933. He argues that the introduction of these foreign exchange controls insulated the US banking system from further runs on its gold reserves. Although the results presented in this chapter are inconsistent with the view proposed by Wigmore, they are consistent with the view of Friedman and Schwartz. That is, the introduction of the F.D.I.C. was crucial to the ending of the financial crisis and the recovery from the Great Depression.

The remainder of this chapter is organized as follows. In the next section I discuss the implications of the financial crisis for the Great Depression and suggest two measures of financial crisis. In section 3.3, I give a summary of New Deal policy which respect

¹Board of Governors of the Federal Reserve System [5], p907.
to the financial sector. In section 3.4, I present the version of Hamilton’s model I use to estimate the conditional probabilities over the state of the financial system, and, in section 3.5, present the results of the estimation of this model. In section 3.6, I show that the estimated state of the financial system can help to explain output fluctuations during the interwar period. Section 3.7 concludes.

3.2 Financial Crisis and the Great Depression

Friedman and Schwartz [24] argue that the bank failures of the 1930s had an adverse effect on income during the Great Depression through two channels. A negative wealth effect on the value of bank shareholders and more importantly through a contraction in the money supply. According to Friedman and Schwartz the bank failures of the 1930s made deposits a less attractive proposition to the public. Therefore they substituted away from deposits and into other assets such as currency. This led to a large drop in deposits during the financial crisis and, as a result, a fall in the money stock through the money multiplier. Therefore, I assume that the growth rate of real deposits is lower in the financial crisis state than in the financial calm state.

One potential problem with using the growth rate of real deposits is that it not only reflects depositor confidence in the banking sector, but it also reflects bank failures. If a bank fails with losses to depositors, then deposits will fall. Therefore, any observed

---

2Temin [52] finds little evidence to support the hypothesis that there was a negative wealth effect following the stock market crash of 1929. He attributes this to the fact that stocks were only a small proportion of wealth. This also suggests a minor effect on income for a negative wealth effect following bank failure.

3This is the ‘contagion of fear’ argument. Suppose that following the failure of one insolvent bank the public begin to fear that other banks may be insolvent. When they have imperfect information about the solvency of a particular bank, their best strategy maybe to withdraw their deposits from that bank. Therefore, the overall fall in deposits maybe much greater than that due to the initial bank failure. A similar argument is made by Wicker [55].

4For a full explanation of this see Friedman and Schwartz [24], appendix B.

5One could also use the level of real deposits or the deposit currency ratio to infer when the financial sector was in a state of crisis. Both of these variables would also be lower in the financial crisis state. The problem with these two measures is that they are not stationary over the sample period.
reduction in the growth rate of real deposits will also reflect losses to depositors in failing banks. However, while deposits at commercial banks fell by $12.3bn between December 1929 and December 1932 only approximately $0.8bn of this was in losses due to bank suspensions.\textsuperscript{6} This suggests that most of the decline in the deposits was due to portfolio reallocation.

Figure 3.1a shows the quarterly growth rate of real deposits between 1919 and 1941.\textsuperscript{7} As one can see, there is a sharp fall in the growth rate of real deposits in the early 1930s. This is consistent with a period of financial crisis. There is a prolonged period in which, with the exception of one quarter, the growth rate of real deposits was negative. This period is from early 1931 until late 1933. In contrast, in the United Kingdom where there was no financial crisis, real deposits grew steadily between 1929 and 1933. The average, annual growth rate being 4.35\%.\textsuperscript{8}

Bernanke \cite{Bernanke2009} offers a third channel through which the financial crisis could have an adverse affect on income: through the collapse of credit intermediation. Given that loans markets are incomplete there is a role for intermediation between borrowers and lenders. Banks specialize in this activity by screening and monitoring borrowers. The charge for this service is the cost of credit intermediation. This is the differential between the interest rate that the lender receives from the bank and the interest rate that the borrower pays the bank. Bernanke argues that during the financial crisis of early 1930s the cost of credit intermediation began to rise. This meant that some borrowers found it difficult and expensive to obtain credit.

Banks typically hold a significant proportion of their assets in the form of illiquid

\textsuperscript{6}The first figure is from Friedman and Schwartz \cite{Friedman1963}, appendix B. The second figure and the one which follows are from Friedman and Schwartz, table 16. Total deposits in suspended banks between these dates totalled approximately $3.2bn.

\textsuperscript{7}The data series for deposits comes from Friedman and Schwartz \cite{Friedman1963}, appendix A. It is deflated using the GNP deflator from Balke and Gordon \cite{Balke1993}.

\textsuperscript{8}In 1929 the real net deposits in the United Kingdom were £2244m; in 1930 they were £2269; in 1931, £2279; in 1932, £2412 and in 1933, £2661. The data on nominal net deposits are from Capie and Webber \cite{Capie1999}, table II(2). I deflate this data using a GDP deflator with 1958=100. I construct this GDP deflator from nominal and real GDP data in Mitchell \cite{Mitchell1988}.
loans. On the other hand their liabilities are for the most part in the form of deposits and therefore payable on demand. This can make their position somewhat tenuous. As a result, during a bank run an illiquid, yet solvent bank could be forced to suspend payments.

Bernanke argued that as banks began to fail, non-failing banks began to fear a run. In response to this they increased their reserve-deposit ratios and substituted into safer and more liquid assets. Both of these factors led to a reduction in their role of credit intermediation. This led to a rise in the cost of credit intermediation. Consequently the borrower’s rate rose and the number of loans fell. Therefore, I assume that when the economy is in a state of financial crisis the cost of credit intermediation is higher than when it is in a state of financial calm.

Unfortunately, as Bernanke points out there is no direct, satisfactory measure of the cost of credit intermediation. Here I follow Bernanke [3] and Mishkin [42], and use, as a proxy, the spread between the yield on corporate bonds rated Baa by Moody’s and the yield on long-term government bonds. Admittedly this is not a perfect measure of the cost of credit intermediation. However, it does represent the different costs of funds to two different classes of borrowers. Bernanke argues that the only classes of borrower unaffected by the disruption of credit intermediation were the Federal government and some blue-chip corporations. Therefore, all else equal, we would not expect the yield on these government bonds to rise during the financial crisis. On the other hand, in their 1919 explanation of the different bond ratings Moody’s described Baa bonds as having “...an element of uncertainty regarding the permanently strong position of such issues.” ([44]; p13). As a result we would expect the yields on these bonds to rise as lenders began

9If banks were only making the safest and highest quality loans, then the reported rates for commercial loans do not reflect the shadow cost of bank funds to a representative borrower. In addition to this, in periods of high default on loans we would expect to see a rise in the cost of borrowing as lenders attempted to prevent any decline in the expected value of the loan.

10Both of these yields come from the Board of Governors of the Federal Reserve System [4], table 128.

11I also used the yield on Moody’s highest rated bonds (Aaa) in the place of the yield on long-term government bonds. The results were unaffected.
to prefer safer assets. Therefore, I use the spread between these two classes of bonds as a proxy for the cost of credit intermediation.\textsuperscript{12}

Figure 3.1b shows the spread between Baa rated bonds and long-term government bonds. This indicates that there is a sharp increase in this spread in the early 1930s, again this is consistent with a period of financial crisis. Also consistent with a financial crisis is the fact that this increase came from an increase in the yield on the risky asset rather than a fall in the yield on the risk-free asset.\textsuperscript{13}

In this chapter I propose that in the interwar period in the United States there were two states of the world. One in which the incidence of bank failure was high and all but the safest class of borrower found it expensive to obtain credit. This was reflected in a lower growth rate of real deposits and a higher cost of credit intermediation in this state. I refer to this state as the financial crisis state. A second state had a low incidence of bank failure and more equal costs of credit across classes of borrower.\textsuperscript{14} I refer to this as the financial calm state. In this state the growth rate of real deposits was higher and the cost of credit intermediation was lower.

### 3.3 Banking & Monetary Reform in the New Deal

At midnight on March 6th 1933, President Roosevelt declared that there would be a banking holiday from the 6th to the 9th of March. On the 9th of March, Congress passed the Emergency Banking Act (E.B.A.) which was the first of many banking and

\textsuperscript{12}As Bernanke points out, this measure is biased downwards as it does not allow for the reclassification of firms into higher risk categories. Assuming that the riskiest Baa rated bonds have the highest yields among Baa rated bonds, then, when they are downgraded the average yield on Baa rated bonds will fall. Therefore, downgrading would bias the yield spread downwards. However, despite this, the spread was considerably higher during the years of the Great Depression than during the rest of the interwar period.

\textsuperscript{13}In January 1930 the yield on Baa rated bonds was 5.29%. It reached a peak at 11.63% in May 1932. This is also when the spread peaked. In January 1930 the return on long-term government bonds was 3.43%. In May 1932 it was relatively unchanged at 3.76%. (Board of Governors of the Federal Reserve System; [4], table 128.)

\textsuperscript{14}That is, relative to the cost of funds to the safest borrowers, the cost of funds to other borrowers was lower than in the financial crisis state.
monetary reforms contained in the New Deal. It appears that these reforms were intended to stabilize the financial system, and coupled with Roosevelt's "fireside chats," restore confidence in the financial system. One of the questions asked in this chapter is whether or not these reforms were successful in restoring confidence. Therefore, this section summarizes the main aspects of these policies.\textsuperscript{15}

\subsection*{3.3.1 Banking Reform}

The E.B.A. extended the wartime measures that allowed the President to control the payment of deposits by banking institutions in periods of national emergency. Any institution that wished to do bank business during a period of emergency needed the authorization of the Secretary of the Treasury and Presidential Approval. The E.B.A. also facilitated the re-opening of national banks. On the 10th of March 1933, Roosevelt issued an executive order which gave the Secretary of the Treasury the power to grant licenses to member banks allowing them to reopen. To obtain these licenses member banks had to apply to their local Federal Reserve Bank which acted as the Secretary's agent. In a press statement on the 11th March and a radio address on the 12th, Roosevelt announced a program for the re-opening of licensed banks on the 13th, 14th and 15th of March. This order also allowed state banking authorities to open sound, non-member banks. Of the 17,796 commercial banks in operation at the beginning of 1933, 447 had disappeared by the end of the banking holiday. Of the remainder less than 12,000 were granted licenses to re-open. Although another 3398 reopened later, 2132 were suspended, liquidated or merged with other banks.\textsuperscript{16}

The E.B.A. gave the Reconstruction Finance Corporation (R.F.C.) the power to invest equity in banks without taking collateral. Mason [39] argues that the overcollateralization of R.F.C. loans up to this point in time had meant that R.F.C. loans were creating a

\textsuperscript{15}A more comprehensive treatment can be found in Chapter 8 of Friedman & Schwartz [24] or in Wicker [55].

\textsuperscript{16}Friedman & Schwartz [24], pp422-28. Those licensed banks accounted for $23.3bn of deposits, while the unlicensed banks accounted for $3.4bn of deposits, also Friedman & Schwartz [24], p423.
liquidity problem for the very banks they were supposed to help. After March 1933 the R.F.C. was able to invest in the equity of a bank through the preferred stock program. After examining panel data on banks in the Illinois portion of the Chicago Federal reserve district, Mason concludes that this program helped banks survive the depression.

The E.B.A. also removed the R.F.C.’s obligation to publish the names of the banks it was assisting. By publishing the names of banks it was assisting, the R.F.C. would have had two opposite effects on depositor confidence. On one hand it would have indicated to the depositor that his or her bank had Federal backing. However, on the other hand it would have suggested that the bank might not be able to meet all of its liabilities. The first was likely to increase confidence thus reducing the probability of the bank failing, while the latter would have the opposite effect. Butkiewicz [9] presented evidence to suggest that it was the latter effect that dominated and that the obligation to publish the names of banks receiving assistance increased bank failures. This view was shared by Friedman and Schwartz, who argued that the appearance of a bank’s name on the list was often interpreted as a sign of weakness and led to a run on the bank. They went on to argue that by restoring confidence in the monetary and economic system the E.B.A. contributed to recovery from the depression. They pointed to falling currency balances relative to income as evidence of increased confidence in the banking system. However they also argued that the introduction of the F.D.I.C. was the structural change that did the most to restore monetary stability.

The Banking Act of June 1933 announced the introduction of deposit insurance to the banking system. All member banks of the Federal Reserve System had to have their deposits insured by the F.D.I.C. Those banks that were not members could also apply for coverage. The introduction of deposit insurance occurred on the 1st January 1934. By July almost 14,000 of the 15,348 commercial banks were covered. These banks accounted

---

17 A provision of the Emergency Relief and Construction Act of July 1932 was interpreted by Speaker of the House, John N. Garner, as requiring the RFC to make public the names of the banks it had made loans to in the previous month. In fact the RFC was only required to inform the President and Congress to which banks it had extended loans.
for 97% of all deposits.\textsuperscript{18} To receive coverage banks had to pay a premium which was a percentage of their deposits. If they were not members of the Federal Reserve System they were also subject to examination by the F.D.I.C.\textsuperscript{19} Initially there was a limit on insured funds of $2500, however it was doubled on 1st July 1934.

Friedman and Schwartz argued that the number of bank failures was greatly reduced in 1934 for two reasons. Firstly, because small depositors knew that if their bank failed they would be reimbursed, and therefore bank runs did not spread from one bank to another. Secondly, bad banks were not allowed to fail, instead they were merged with good banks or re-organized under new management, with the F.D.I.C. assuming any losses.

Wigmore \textsuperscript{[56]} questions whether the banking reforms of the E.B.A. and the 1933 Banking Act were sufficient to explain the subsequent calm in the financial sector. He points out that the F.D.I.C. only covered accounts up to $2500, leaving $25bn in large accounts uncovered (two-thirds of total deposits in insured banks) and the fact that it did not become effective until 1934. The ability of the R.F.C. to provide capital is also seen as relatively unimportant by Wigmore; only $15m of authorizations had been made by the end of March 1933. Wigmore goes on to argue that much of the calm can be attributed to the gold embargo and foreign exchange controls that began in April 1933.

3.3.2 Monetary Reform

By Roosevelt's proclamation of March 6th 1933 banks were prohibited from paying out gold or dealing in foreign exchange during the Banking Holiday. The E.B.A. extended this and gave the President emergency powers over foreign exchange dealings, currency and gold movements, and banking transactions. On the 10th March, Roosevelt issued an executive order which forbade gold payments by banking and non-banking institutions

\textsuperscript{19}Until 1950 the FDIC required the permission of the Comptroller of the Currency or the Governor of the Federal Reserve System to examine member banks.
except under license and with the permission of the Secretary of the Treasury. However, despite this effective suspension of gold payments the exchange rate between the dollar and other currencies remained around its previous gold standard level for over a month. Friedman and Schwartz attributed this to the fact that the suspension was probably regarded as part of the emergency measures and therefore considered to be temporary. They point to the fact that there was no official announcement of devaluation, that after a few weeks several licenses had been granted and that the United States' technical gold position was strong enough to defend parity.

On the 5th April 1933 public hoarding of gold was prohibited and all gold, had to be delivered to the Federal Reserve Banks by the 1st May. Then, on the 20th April, Roosevelt made it clear that it was the intent of his administration to allow the dollar to devalue against other currencies in order to raise prices. This became official with the Thomas Amendment to the Agricultural Amendment Act on 12th May. This allowed Roosevelt to lower the gold content of the dollar to as little as 50% of its previous weight. By the end of July the dollar had devalued by almost 60%. On the 31st January 1934 Roosevelt's administration fixed the price of gold at $35 per fine ounce under the authority of the Gold Reserve Act.

Note that this did not represent a devaluation of the dollar. It had already fallen to this price by July of 1933. However, the Treasury had previously valued its gold holdings at $20.67, and so in January 1934 realized a paper profit from the devaluation of the dollar. It could print additional paper money in the form of gold certificates which, although they could not be held by the public, could be held by the Federal Reserve Banks. These certificates had a nominal value of nearly $3bn and did not require the acquisition of additional gold to maintain backing. Therefore they represented the potential for significant monetary expansion.

If these reforms did have a positive effect on the financial system, then this would be

---

20 There were some exceptions for the arts and industry.
21 Wigmore [56], p750.
reflected in the time series of estimated probabilities over the current state of the financial system. If these reforms were credible, then, given that the early 1930s was a period of financial crisis, one would expect that the crisis would end in 1933 or 1934. Wigmore's view that it was the gold embargoes and exchange controls that were the more important reform would suggest that there would be a regime change in the Spring of 1933. On the other hand, Friedman and Schwartz's view that the introduction of the F.D.I.C. was the more important reform would date the regime change as early 1934.

### 3.4 A Model of Time Series Subject to a Change in Regime

The model presented in this section is a straightforward variant of Hamilton’s [27], [28], [29], Markov switching model. Two time series follow a single regime. Assume the vector, \( x_t \), consisting of the growth rate of real deposits \( (x_{1t}) \) and the yield spread \( (x_{2t}) \) follows the stochastic process:

\[
\begin{bmatrix}
    x_{1t} - \mu_1 (S_t) \\
    x_{2t} - \mu_2 (S_t)
\end{bmatrix}
= 
\begin{bmatrix}
    \rho_{11} & 0 \\
    0 & \rho_{21}
\end{bmatrix}
\begin{bmatrix}
    x_{1,t-1} - \mu_1 (S_{t-1}) \\
    x_{2,t-1} - \mu_2 (S_{t-1})
\end{bmatrix}
+ 
\begin{bmatrix}
    \rho_{12} & 0 \\
    0 & \rho_{22}
\end{bmatrix}
\begin{bmatrix}
    x_{1,t-2} - \mu_1 (S_{t-2}) \\
    x_{2,t-2} - \mu_2 (S_{t-2})
\end{bmatrix}
+ 
\begin{bmatrix}
    \sigma_1 (S_t) v_{1,t} \\
    \sigma_2 (S_t) v_{2,t}
\end{bmatrix},
\]

(3.1)

where \( v_{it} \sim i.i.d. N (0, 1) \) and \( (1 - \rho_i z) \neq 0 \) for any \( z \) where \( |z| \leq 1 \), for \( i = 1, 2 \). Here \( \mu_1 \) refers to the mean of the first series, the growth rate of real deposits and \( \mu_2 \) refers to the mean of the yield spread. \( \rho_{11} \) and \( \rho_{21} \) are the first order, autoregressive parameters for these two series, \( \rho_{12} \) and \( \rho_{22} \) are the second order, autoregressive parameters for these two series.\(^{22}\) \( \sigma_1 \) and \( \sigma_2 \) are the standard deviations of the innovations, \( v_{1,t} \) and \( v_{2,t} \). This

---

\(^{22}\)This model described here contains two lagged dependent variables for each series. This matches the model chosen in the empirical section by sequentially testing downwards from a maximum lag order.
model differs from the usual AR(2) model as it allows the means of the two series and the standard deviations of their innovations to vary with the underlying regime, \( S_t \). I constrain this unobserved regime to be the same across the two series. This allows me to use observations on the two observed time series to draw inference about the common underlying regime.

This regime is modelled as the outcome of an unobserved, discrete time, discrete state, first order Markov process. There are two states, the financial crisis state and the financial calm state. Let state zero be the state of financial calm and state one be the state of financial crisis. In each period there is a fixed probability of switching between the states. These transition probabilities are

\[
P(\mathcal{S}_t = 1 | \mathcal{S}_{t-1} = 1) = p; \quad P(\mathcal{S}_t = 0 | \mathcal{S}_{t-1} = 1) = 1 - p; \\
P(\mathcal{S}_t = 1 | \mathcal{S}_{t-1} = 0) = 1 - q; \quad P(\mathcal{S}_t = 0 | \mathcal{S}_{t-1} = 0) = q.
\]  

(3.2)

The means of the two observed series and the standard deviations of their innovations are given by

\[
\mu_i(S_t) = \alpha_{i0} + \alpha_{i1} S_t; \quad \sigma_i(S_t) = \omega_{i0} + \omega_{i1} S_t; \quad i = 1, 2.
\]  

(3.3)

In the financial crisis state the growth rate of real deposits is lower and the yield spread is higher than in the financial calm state. This implies that \( \alpha_{11} < 0 \) and \( \alpha_{21} > 0 \).

As mentioned before the econometrician does not directly observe the current regime, or indeed any past regimes. Instead he or she draws inferences about these regimes based on what they have observed, in this case the growth rate of real deposits and the yield spread. This is done using the non-linear filter described below. Throughout the five steps of the filter, the parameter values \((p, q, \rho_{11}, \rho_{21}, \rho_{12}, \rho_{22}, \alpha_{10}, \alpha_{20}, \alpha_{11}, \alpha_{21}, \omega_{10}, \omega_{20}, \omega_{11}, \omega_{21})\) are assumed by the econometrician to be known constants. Using these, he or she then calculates a time series of probabilities over the current regime.
The five steps of the filter are given below. In order to keep notation simpler, \( P(z) \) denotes \( P(Z = z) \) when \( z \) is a discrete valued variable and the density function \( f(z) \) when \( z \) is a continuous variable. Iteration \( t \) of the filter takes as an input the conditional probability \( P(s_{t-1}|x_{t-1}, \ldots, x_0) \) and gives as output \( P(s_t|x_t, x_{t-1} \ldots, x_0) \), where \( x_t \) is a vector containing \( x_{1t} \) and \( x_{2t} \).

**Step One;** Using the Markov transition probabilities in equation (3.2) calculate

\[
P(s_t, s_{t-1}, s_{t-2}|x_{t-1}, \ldots, x_0) = P(s_t|s_{t-1}) \times P(s_{t-1}, s_{t-2}|x_{t-1}, \ldots, x_0), \tag{3.4}
\]

where the second term on the right hand side is the output from the previous iteration of the filter.

**Step Two;** The joint conditional density of \( x_t \) and \( (S_t, S_{t-1}, S_{t-2}) \) is given by

\[
P(x_t, s_t, s_{t-1}, s_{t-2}|x_{t-1}, \ldots, x_0) = P(x_t|s_t, s_{t-1}, s_{t-2}, x_{t-1}, \ldots, x_0) \times P(s_t, s_{t-1}, s_{t-2}|x_{t-1}, \ldots, x_0), \tag{3.5}
\]

where

\[
P(x_t|s_t, s_{t-1}, s_{t-2}, x_{t-1} \ldots x_0) = \frac{1}{2\pi \sigma_1(S_t) \sigma_2(S_t)} \times \exp \left\{ -\frac{1}{\sigma_1(S_t)} \right\} \times \left\{ \frac{1}{\sigma_2(S_t)} \right\} \times \left\{ \left( x_{1t} - \mu_1(S_t) \right)^2 + \left( x_{1t-1} - \mu_1(S_{t-1}) \right)^2 - \rho_{11} \left( x_{1t} - \mu_1(S_t) \right) \left( x_{1t-1} - \mu_1(S_{t-1}) \right) \right\} \times \left\{ \left( x_{2t} - \mu_2(S_t) \right)^2 + \left( x_{2t-1} - \mu_2(S_{t-1}) \right)^2 - \rho_{21} \left( x_{2t} - \mu_2(S_t) \right) \left( x_{2t-1} - \mu_2(S_{t-1}) \right) \right\}. \tag{3.6}
\]

**Step Three;** The branch of the likelihood function for observation \( t \) is

\[
P(x_t|x_{t-1}, \ldots, x_0) = \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} \sum_{s_{t-2}=0}^{1} P(x_t, s_t, s_{t-1}, s_{t-2}|x_{t-1}, \ldots, x_0). \tag{3.7}
\]

**Step Four;** Steps two and three combined give

\[
P(s_t, s_{t-1}, s_{t-2}|x_t, x_{t-1}, \ldots, x_0) = \frac{P(x_t, s_t, s_{t-1}, s_{t-2}|x_{t-1}, \ldots, x_0)}{P(x_t|x_{t-1}, \ldots, x_0)}. \tag{3.8}
\]
Step Five; The output of iteration $t$ of the filter is

$$P(s_t, s_{t-1} | x_t, x_{t-1}, \ldots, x_0) = \sum_{s_{t-2}=0}^1 P(s_t, s_{t-1}, s_{t-2} | x_t, x_{t-1}, \ldots, x_0).$$  \hfill (3.9)

As mentioned before each iteration of the filter assumes that the parameters of the model are known constants. However step three of the filter gives as a by-product the conditional likelihood for those constants for observation $t$. This can be summed over the $T$ observations to give

$$\ln P(x_T, x_{T-1}, \ldots, x_2 | x_1, x_0) = \sum_{t=2}^T \ln P(x_t | x_{t-1}, \ldots, x_0),$$  \hfill (3.10)

which allows the econometrician to estimate the parameters of the model by maximum likelihood.

Also note, that if in step five of the filter he or she were to sum over $s_{t-1}$ as well as $s_{t-2}$, the econometrician would calculate $P(s_t | x_t, x_{t-1}, \ldots, x_0)$. These are the conditional probabilities over the current state given what has been observed up to, and including, the current period. Alternatively, if he or she were to sum over $s_t$ and $s_{t-2}$ they would calculate $P(s_{t-1} | x_t, x_{t-1}, \ldots, x_0)$. These are the conditional probabilities over the state in the previous period given what has been observed up to, and including, the current period.

If the disruptions to the financial markets during the Great Depression can be thought of as a shift to a regime of financial crisis then this would be reflected in the estimated probabilities from steps four and five of the filter. More specifically, one would expect to see a probability of almost one assigned to the financial crisis regime during the early 1930s. Then, if the banking and monetary reforms did cause the financial crisis to be over, one would expect to see a probability of almost zero assigned to the financial crisis beginning in 1933 or early 1934, depending upon which were the crucial reforms. If Wigmore is correct, and leaving the gold standard was the more important reform, the regime change back to the state of financial crisis should occur in the Spring of 1933.
However, if Friedman and Schwartz are correct, and the more important reform was the introduction of the FDIC, then it should occur in early 1934.

3.5 Estimation Results

In this section I present the results of the estimation of the model described in the previous section. To estimate the model I use quarterly data on the growth rate of real deposits and the yield spread from 1919 to 1941. The results from maximum likelihood estimation of the Markov switching model and a single state model appear in table 3.1. This single state model is a special case of equation (3.1) in which there is only one state. In this model, the means of the two series and the standard deviations of their innovations are constant.

A likelihood ratio test of the null hypothesis that the data were generated by the single state model against the alternative that they were generated by the Markov switching model yields a test statistic of statistic of 127.8. This suggests a strong rejection of the null hypothesis. However, under the null hypothesis of a one state model the parameters which describe the second state are unidentified. This means that the likelihood ratio test statistic does not have a standard $\chi^2$ distribution. I tackle this problem in two ways. In the first instance I perform a small Monte Carlo experiment to generate an empirical critical value for this test statistic. In this experiment, I generate artificial data under the null hypothesis that the data were generated by the single state model. To do this I use the parameter estimates for the single state model in table one to create artificial time series for the growth rate of real deposits and

---

23To determine the lag order, I estimate a version of the model with one, two, three and four autoregressive parameters. Likelihood ratio tests do not allow the rejection of the model with three lags in favour of the one with four, or the model with two lags in favour of the one with three. However the one lag model can be rejected in favour of the two lag model.

24For the likelihood ratio test to have an asymptotic $\chi^2$ distribution we require that the information matrix is non-singular. If one tries to fit an $n$ state model when the true process has $n-1$ states, this condition will not hold. See, for example, chapter four of this thesis for a more detailed discussion of this issue.
the yield spread.\textsuperscript{25} I generate series of length $200 + T$, where $T$ is the sample size. I then discard the first 200 observations to minimize the influence of starting values. Next, I use the remaining artificial data to estimate both the single state model and the Markov switching model by maximum likelihood estimation. Finally, I calculate the likelihood ratio test statistic for the test of the null hypothesis of the single state model against the alternative of the Markov switching model. I repeat this procedure 500 times.

The percentiles of my simulated distribution are slightly above those of the asymptotic $\chi^2 (6)$ distribution. For example, the empirical 95% and 99% critical values are 17.21 and 27.34 respectively. However, these are still much lower than the test statistic calculated using the sample data. In fact, of the 500 artificial test statistics I generate by Monte Carlo simulation, not one is above my sample statistic of 127.8.

The second way I tackle the problem is to follow the suggestion of Hamilton \textsuperscript{[27], [30]}. That is, I take the one state model and conduct a test on that model to see whether the two state model is required. I use a Lagrange Multiplier test in which the null hypothesis is the single state model with homoskedastic errors from table 3.1. This is tested against the alternative in which the variance of the residuals from that model depend on the lagged filter output. For comparison I also test this null hypothesis against two other alternatives. One in which the variance of the residuals depends on the lagged level of the series and one in which it depends on the lagged squared residuals. The latter of these being a test for ARCH effects.\textsuperscript{26} In each case the test statistic is

$$TR^2 \sim \chi^2 (1),$$

where $T$ is the sample size and $R^2$ is the coefficient of determination from the estimation of

$$\hat{u}_i^2 = \alpha + \beta z_i, \quad i = 1, 2,$$

\textsuperscript{25}The standard deviations of the innovations to these series are $\omega_{10} = 3.107$ and $\omega_{20} = 0.641$ respectively.

\textsuperscript{26}See Engle [19].
where \( \hat{u}_t^2 \) are the residuals from the equation for variable \( i \) in the single state model and 
\( z_t \) is either the lagged level of the series, the lagged squared residuals or the lagged filter 
output.

The test statistics in table 3.2 provide further evidence to support of the use of a two 
state model. The null hypothesis of homoskedastic errors can be rejected in favor of the 
alternative where their variance depends on the lagged output from the filter described 
in the previous section. This applies to the residuals from the estimation of both series 
under the single state specification. For the other two alternatives, the lagged level and 
the ARCH effects, the null hypothesis can only be rejected for one of the series. This is 
the yield spread when the alternative is lagged level effects, and the growth rate of real 
deposits when the alternative is ARCH effects.

The estimated Markov probabilities show the unconditional probabilities for a change 
in regime from financial calm to financial crisis and vice-versa. If the state at time \( t \) was 
financial calm then there is a 97% chance that state at time \( t + 1 \) was also financial calm. 
This implies a 3% chance that it was financial crisis. Therefore, the expected length of 
a period of financial calm is \( 33\frac{1}{3} \) quarters. On the other hand, if the state at time \( t \) was 
financial crisis then there is an 86% chance that the state at time \( t + 1 \) was also financial 
crisis. This implies a 14% chance that it was financial calm. Therefore, the expected 
length of a financial crisis is just over 7 quarters.

The means of the growth rate of real deposits and the yield spread when the economy 
was in the state of financial calm were 1.491 and 2.446 respectively. When the economy 
was in the state of financial crisis the mean growth rate of real deposits fell to \(-1.357\) and 
the mean of the yield spread rose to 3.279. A test that the mean of a series is the same in 
the two regimes is just a simple test of the null hypothesis that the incremental parameter, 
\( \alpha_{11} \) or \( \alpha_{21} \), is equal to zero. This test produces a rejection of the null hypothesis at the 
1% level for both series. The lower growth rate of real deposits and higher yield spread 
suggest a state of financial crisis.\(^{27}\)

\(^{27}\)For the growth rate of real deposits to be negative in the financial crisis state it must be the case
One can pass the parameter estimates from table one through the filter of the previous section to give a time series of conditional probabilities that the economy was in each of the two states. Figure 3.2a shows the conditional probability that the current state was financial crisis at each point in time from the 3rd quarter of 1919 to the end of 1941. This probability is conditional on information up to and including the current period. As mentioned earlier, at time $t$ the econometrician is able to update his or her conditional probability over previous states. In this model, as there are two autoregressive parameters the econometrician can revise his or her conditional probability of the state at time $t - 1$ and time $t - 2$. The revised conditional probabilities, $P(s_{t-2} | x_t, x_{t-1}, \ldots, x_0)$ are shown in figure 3.2b. These are probabilities that the state at time $t - 2$ was financial crisis conditional on information from up to and including period $t$. Appendix three gives the full time series of all the estimated probabilities.

These diagrams show that a shift away from the financial calm state and into the financial crisis state occurred in late 1930. This coincides with the first banking crisis of October 1930. It is at this point that Friedman and Schwartz argued that a "...contagion of fear spread among depositors." It is interesting to note that there is no change in regime with the stock market crash in late 1929. Immediately after the stock market crash the economy does not switch to the state of financial crisis, but remains in the state of financial calm. This is consistent with the point emphasized by Mishkin [42] that a stock market crash alone does not necessarily imply a financial crisis.

Moving on to the end of the financial crisis, if Wigmore is correct and the exchange controls of the Spring of 1933 were more important than the introduction of deposit

---

28 According to figure 3.2a the conditional probability assigned to the financial crisis state reasonably high in 1920. Examination of figure 3.1a shows that this coincides with a decline in real deposits. However looking at the updated conditional probabilities in figure 3.2b shows that it is unlikely that the economy was in the financial crisis state at this time.

29 Friedman and Schwartz, [24], p308.
insurance then one would expect to see a sharp fall in the probability of the financial crisis state in either the second or third quarter of 1933. Figure 3.2a shows the conditional probability of financial crisis only falls to 0.776 in the third quarter of 1933. In the fourth quarter this probability is back to one. This suggests that there was no change in regime immediately following the introduction of exchange controls in the Spring of 1933. This point is emphasized by looking at the updated probabilities in figure 3.2b. The updated probability for the financial crisis state for the third quarter of 1933 only falls to 0.936. For the rest of 1933 this probability is one. This result is inconsistent with Wigmore's view that the exchange controls were the more important reform for bringing stability to the financial sector.

The conditional probabilities suggest that the financial crisis ended in the first quarter of 1934. This is shown much more clearly using the updated conditional probabilities of figure 3.2b than the conditional probabilities of figure 3.2a. The probabilities of figure 3.2a suggest that the probability of financial crisis was 0.812 in the first quarter of 1934, 0.260 in the second quarter and 0.631 in the third quarter. After this they quickly drop to almost zero. However using the updated probabilities of figure 3.2b shows a probability of 0.459 that the economy was in a state of financial crisis in the first quarter of 1934. This probability falls to almost zero in the second quarter of 1934 and then remains at zero for the remainder of 1934, and until 1937. This result is consistent with Friedman and Schwartz's view that the introduction of the F.D.I.C. was the crucial reform to the financial sector.

During the first quarter of 1934 the introduction of the F.D.I.C. was not the only major policy change. On the 31st January 1934, Roosevelt announced that the dollar price of gold was to be fixed at $35 per fine ounce of gold. This returned the United States to a de facto gold standard. Although this did not represent any devaluation of the dollar from its floating level, the Treasury was able to revalue its gold holdings.30 Therefore, fixing the gold price of the dollar represented the potential for a significant expansion of

---

30 As Wigmore [56] points out, the dollar had already devalued 60% by the end of July 1933.
high powered money. Friedman and Schwartz argue that, during the Great Depression, the Federal Reserve should have provided a significant expansion of high powered money to protect the banking sector. This might lead one to argue that monetary expansion in early 1934 was the most important factor in ending the financial crisis. However looking at data on high powered money shows that the increase in high powered money in the first quarter of 1934 was only sufficient to restore the stock of high powered money to approximately its February 1933 level.\textsuperscript{31} Therefore, this was not the monetary expansion that the Federal Reserve failed to provide during the early 1930s. This suggests that it was not monetary expansion in early 1934 that ended the financial crisis.

There is also a brief period of financial crisis in late 1937 and early 1938. Between 1936 and 1937 the Federal Reserve more than doubled the required reserve ratio for member banks from 6.2\% of total assets to 12.6\% of total assets.\textsuperscript{32} If banks are not holding excess reserves, or do not wish their level of excess reserves to fall, then an increase in required reserves can lead to a fall in the amount of loans that they will make. As the banking system reduces its role as a credit intermediary the cost of credit intermediation will rise. As a result, raising required reserves can lead to a fall in inside money.

Therefore, it is possible that by raising reserve requirements the Federal Reserve would have caused the growth rate of real deposits to fall and the cost of credit intermediation to rise. That is, these series moved in a fashion similar to that in which they would move during a financial crisis. In order for this to happen it must have been the case that banks were either holding no excess reserves or that they wished to keep their level of excess reserves constant after the change in required reserves. In 1937 and 1938 the first of these scenarios did not hold. However, Friedman and Schwartz suggest that the second of these scenarios did hold.

\textsuperscript{31}In February 1933 the stock of high powered money was $8.807bn, in March 1934 it was $8.998bn, Friedman and Schwartz [24], appendix B. These are nominal figures. The real stock of high powered money does not regain its level of the first quarter of 1933 until 1935. See the figures for M0 and the GNP deflator in Balke and Gordon [2].

\textsuperscript{32}Friedman and Schwartz [24], table 19.
"When the rise in reserve requirements immobilized the accumulated cash, they (member banks) proceeded rather promptly to accumulate additional cash for liquidity purposes." Friedman and Schwartz [24], p458.

### 3.6 Financial Crisis and Output Growth

In this section, I use the estimated conditional probability of financial crisis as an explanatory variable for output growth in the interwar period. I show that including the estimated conditional probability of financial crisis in a reduced form equation provides additional explanatory power for output. This is explanatory power in addition to that provided by lagged values of output, current and lagged values of the money stock, current and lagged values of the growth rate of real deposits and current and lagged values of the yield spread.\(^{33}\)

I estimate the following reduced form equation using quarterly data from 1920 to 1941 using two stage least squares:

\[
\Delta y_t = \alpha + \sum_{i=1}^{n} \beta_i \Delta y_{t-i} + \sum_{j=0}^{n} \gamma_j \Delta m_{t-j} + \sum_{j=0}^{n} \delta_j x_{1,t-i} + \sum_{j=0}^{n} \theta_j x_{2,t-j} + \lambda_0 P(S_t = 1 | x_t, x_{t-1}, x_0) + \lambda_1 P(S_{t-1} = 1 | x_t, x_{t-1}, x_0) + \lambda_2 P(S_{t-2} = 1 | x_t, x_{t-1}, x_0) + e_t. \tag{3.13}
\]

Here \(\Delta y_t\) is the change in the natural logarithm of real GNP, \(\Delta m_t\) is the change in the natural logarithm of the money stock, \(x_{1,t}\), \(x_{2,t}\) and \(x_t\) are as described earlier and \(e_t \sim i.i.d. N(0, \sigma_e^2)\). I use two stage least squares estimation because the conditional probabilities on the right hand side of equation (3.13) are potentially correlated with the error term, \(e_t\). This is because the conditional probabilities are not predetermined at time

:\(^{33}\)Both the real GNP series and the money stock series are taken from Balke and Gordon [2]. As deposits form a significant proportion of M2 there is the possibility of collinearity between the growth rate of real deposits and the the growth rate of M2. Therefore when using M2 as the monetary aggregate I estimate equation (3.13) with and without the growth rate of real deposits. The results are unchanged.
t, they are dependent on $x_{1,t}$ and $x_{2,t}$.

The results reported below are for $i = j = 1$. Under the null hypothesis that the state of the financial sector has no additional in sample explanatory power for output growth, $\lambda_0 = \lambda_1 = \lambda_2 = 0$. This test statistic is distributed $F(3, 77)$. I estimate this equation twice. Once using M2 to represent the money stock and once using M0. The result is similar for both measures of the money stock. When M2 represents money the test statistic is 3.61, and when M0 represents money the test statistic is 2.79. These statistics have p-values of 0.0170 and 0.0459 respectively and so both imply a rejection of the null hypothesis at the 5% level. In both cases the sum of the point estimates, $\hat{\lambda}_0 + \hat{\lambda}_1 + \hat{\lambda}_2$ is approximately -0.20, suggesting the growth rate of output is lower during a financial crisis. Therefore, treating the financial crisis as a change in regime provides additional, in-sample, explanatory power for output fluctuations during the interwar period.

Staiger and Stock [50] note that the two-stage least squares estimator maybe biased when the instruments are weakly correlated with the regressors. Therefore, I perform a small Monte Carlo experiment to investigate whether my rejection of the null hypothesis, $\lambda_0 = \lambda_1 = \lambda_2 = 0$, is a result of a bias in the two-stage least squares estimator. First I estimate equation (3.13) by ordinary least squares with the null hypothesis of $\lambda_0 = \lambda_1 = \lambda_2 = 0$ imposed. Then I generate artificial data for output under the null hypothesis using these parameter estimates, the sample data on $\Delta m_t$, $x_{1,t}$, $x_{2,t}$, an initial value for $\Delta y_t$ and random innovations, $u_t \sim N(0, s^2)$. Here $s^2$ is the estimate of the variance of the residuals from the OLS estimation of the restricted version of equation (3.13). Next, I use the artificial time series for output together with the data on $\Delta m_t$, $x_{1,t}$, $x_{2,t}$ and the conditional probabilities to estimate equation (3.13) by two-stage least squares. Finally

---

34 In addition to the predetermined variables in equation (3.13), $\Delta y_t$, $\Delta m_t$, $\Delta m_{t-1}$, $x_{1,t}$, $x_{1,t-1}$, $x_{2,t}$, $x_{2,t-1}$ and a constant, I use lagged values of each of the conditional probabilities as instruments. The F-statistic from each of the the 1st stage regressions is above 25 and the $R^2$ is always above 0.70. I experiment with different instruments and the results remain unchanged.

35 I also consider values of $i$ and $j$ up to and including $i = j = 4$. The results are unaffected by the choice of $i$ and $j$. I also estimate equation (3.13) with only the current values of $\Delta m_t$, $x_{1,t}$ and $x_{2,t}$ and no lagged values of $\Delta y_t$, again the results are unaffected.
I calculate the F-statistic for $\lambda_0 = \lambda_1 = \lambda_2 = 0$. This procedure is repeated 10,000 times.

Table 3.3 contains the type I error rates, the proportion of times the artificial test statistic is greater than the sample test statistic and the 90th, 95th and 99th percentiles of the simulated distribution for the test statistic. The results from the Monte Carlo experiment show that my rejection of the null hypothesis is not the result of bias in the two-stage least square estimator. Based on an asymptotic 5% critical value the type I error rate is 0.0513 when money is represented by M2. When money is represented by M0 this error rate is 0.0488. These figures suggest that there is no size distortion in the test statistic. When money is M2 the proportion of times that the artificial test statistic is above the statistic from the sample data is 0.0174. When money is M0 the proportion of times that the artificial test statistic is above the statistic from the sample data is 0.0428. In both cases these figures are close to the p-values of the sample test statistics.

3.7 Conclusions

In this chapter, I present evidence to support the hypothesis that the financial crisis of the early 1930s in the United States should be modelled as a change in regime. I argue that the financial reforms contained in the New Deal played a role in ending the financial crisis. My results are inconsistent with the view of Wigmore. He argues that the financial crisis ended with the decision to abandon the gold standard in the Spring of 1933. My results suggest that the financial crisis did not end until early 1934. This is consistent with the view of Friedman and Schwartz. They argue that the introduction of the F.D.I.C. in January 1934 was crucial to ending the financial crisis. Temin and Wigmore [53] argue that the New Deal represented a change in policy regime that played a crucial role in ending the Great Depression. I present evidence to show that the state of the financial sector contains significant explanatory power for output fluctuations during the interwar period. This is consistent with the view that the financial crisis played a role in the Great Depression and that the end of the financial crisis was important for the
recovery from the Great Depression. It is also consistent with Bernanke’s view that the financial crisis had an adverse effect on output above and beyond the monetary effects.
<table>
<thead>
<tr>
<th></th>
<th>Markov Switching</th>
<th>Linear Alternative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>( p )</td>
<td>0.861</td>
<td>(0.097)</td>
</tr>
<tr>
<td>( q )</td>
<td>0.971</td>
<td>(0.021)</td>
</tr>
<tr>
<td>( \alpha_{10} )</td>
<td>1.491</td>
<td>(0.292)</td>
</tr>
<tr>
<td>( \alpha_{11} )</td>
<td>-2.848</td>
<td>(1.034)</td>
</tr>
<tr>
<td>( \alpha_{20} )</td>
<td>2.446</td>
<td>(0.152)</td>
</tr>
<tr>
<td>( \alpha_{21} )</td>
<td>0.833</td>
<td>(0.123)</td>
</tr>
<tr>
<td>( \rho_{11} )</td>
<td>0.565</td>
<td>(0.104)</td>
</tr>
<tr>
<td>( \rho_{12} )</td>
<td>-0.300</td>
<td>(0.101)</td>
</tr>
<tr>
<td>( \rho_{21} )</td>
<td>0.953</td>
<td>(0.112)</td>
</tr>
<tr>
<td>( \rho_{22} )</td>
<td>-0.078</td>
<td>(0.109)</td>
</tr>
<tr>
<td>( \omega_{10} )</td>
<td>2.578</td>
<td>(0.213)</td>
</tr>
<tr>
<td>( \omega_{11} )</td>
<td>1.482</td>
<td>(0.792)</td>
</tr>
<tr>
<td>( \omega_{20} )</td>
<td>0.211</td>
<td>(0.018)</td>
</tr>
<tr>
<td>( \omega_{21} )</td>
<td>1.216</td>
<td>(0.263)</td>
</tr>
<tr>
<td>log-likelihood</td>
<td>-253.44</td>
<td>log-likelihood</td>
</tr>
</tbody>
</table>
Table 3.2 LM Tests for Conditional Heteroscedasticity of Residuals

<table>
<thead>
<tr>
<th>Yield Spread</th>
<th>Growth of Real Deposits</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_t$ is the lagged level of the series.</td>
<td></td>
</tr>
<tr>
<td>15.16</td>
<td>0.29</td>
</tr>
<tr>
<td>$z_t$ is lagged squared residuals (ARCH)</td>
<td></td>
</tr>
<tr>
<td>1.069</td>
<td>4.04</td>
</tr>
<tr>
<td>$z_t$ is lagged filter output (regime changes)</td>
<td></td>
</tr>
<tr>
<td>13.43</td>
<td>5.97</td>
</tr>
</tbody>
</table>

5% Critical Value is 3.84 sample size: $T = 87$

The test statistic is $T \times R^2$ from the estimation of $\hat{u}_t^2 = \alpha + \beta z_t$, where $T$ is the sample size and $z_t$ is as described above. The test statistic is distributed $\chi^2(1)$. 
Table 3.3: Results from Monte Carlo Experiment on Bias in the 2SLS Estimator

(a) Money is M2

<table>
<thead>
<tr>
<th>Type I Error Rate:</th>
<th>0.0513</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of Artificial Statistics Above Sample Statistic:</td>
<td>0.0179</td>
</tr>
<tr>
<td>90% Percentile:</td>
<td>2.117</td>
</tr>
<tr>
<td>95% Percentile:</td>
<td>2.736</td>
</tr>
<tr>
<td>99% Percentile:</td>
<td>4.064</td>
</tr>
</tbody>
</table>

(b) Money is M0

<table>
<thead>
<tr>
<th>Type I Error Rate:</th>
<th>0.0488</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of Artificial Statistics Above Sample Statistic:</td>
<td>0.0428</td>
</tr>
<tr>
<td>90% Quantile:</td>
<td>2.152</td>
</tr>
<tr>
<td>95% Quantile:</td>
<td>2.693</td>
</tr>
<tr>
<td>99% Quantile:</td>
<td>3.919</td>
</tr>
</tbody>
</table>

These results are based on 10,000 replications of the experiment described in the text.
Figure 3.1: Real Deposits and Bond Yield Spread 1919-1941

Fig 3.1a: Quarterly Growth Rate of Real Deposits

Fig 3.1b: Baa Yield - Yield on L/T Government Bonds
Figure 3.2: Conditional Probabilities 1919-1941

Fig 3.2a: Conditional Probability of Financial Crisis

Fig 3.2b: Updated Conditional Probability of Financial Crisis
Chapter 4

The Markov Switching Model and Linear Time Series

4.1 Introduction

The most common approach to studying the business cycle has been to consider fluctuations around a single long-run rate of growth. Recently, however, it has been suggested that the trend rate of growth itself is subject to stochastic shifts; see for example Hamilton [28] or Potter [48]. One interpretation of this latter approach is that the economy moves between states of high and low growth. However, testing between a model with a single long-run growth rate and one with two states is problematic. Since the parameters which define the second state are undefined under the null hypothesis that the economy has only one state, the likelihood ratio test statistic does not possess the standard $\chi^2$ distribution.\(^1\)

Also, in addition to the standard problem of model mis-specification bias, there is an additional source of bias that arises if one estimates a two state model when the true data generating process possesses only one state. The source of this bias is Jensen's

\(^1\)This problem also applies more generally. When testing a model with $k$ states against an alternative with $k+j$ states ($j \geq 1$) the parameters which define the last $j$ states will not be defined under the null.
inequality. That is, the expectation of a function of a parameter vector, $\theta$, is not equal to the function of the expectation of that same parameter vector, when that function is non-linear. Under certain parameter restrictions a non-linear model, such as the Markov switching model, maybe a reparameterization of a linear model. However, the maximum likelihood estimator will not be invariant to a reparameterization that involves a non-linear function.

In this chapter, I investigate these issues using Hamilton’s Markov switching model for Canadian real per-capita GDP. I calculate the size and power properties of the pseudo likelihood ratio test statistic for the null hypothesis of a one state model against the alternative of a two state model, under the non-standard distribution. I find that there is no size distortion in the test statistic. At the 95% significance level, depending on sample size, the probability of type I error lies between 0.036 and 0.027. However, I find that power is poor in smaller samples. It is only when the sample size reaches 400 that the size adjusted type II error rate falls below 0.05.

I also evaluate the degree of bias in parameter estimates due to model mis-specification and Jensen’s inequality. This is done using out-of-sample predictive accuracy. I find that when the true process is a single state AR(1) process, there is no loss in predictive accuracy from using the two state Markov switching model. That is the predictions from using the two state model are as accurate as those generated by using a single state AR(1) model. I also compare the predictive accuracy of the single state and the two state models when the true process has two states. I find that, for horizons of one year and under, there is a slight loss in predictive accuracy when using the single state model.

These results are consistent with the assertion that there is little or no bias induced by using the two state model when the true process has only one state, or vice-versa. This is supported by examining the estimated long-run growth rates from the artificial samples. In each case the mean of the estimated long-run growth rate from the mis-specified model is as close to the true value as that from the correctly specified model. Estimating the mis-specified model does, however, induce bias in the estimated autoregressive parameter.
This is more severe when the true model is the two state Markov switching model and the mis-specified model is the single state model.

4.2 A Two State and a One State Model of GDP

In this section I present a simple, two state version of Hamilton's [28] Markov switching model, a linear, single state counterpart and discuss the problem of testing between the two models. I also consider under what circumstances do the estimates from the Markov switching model imply a linear model and compare the processes for producing out-of-sample predictions from both models.

Consider first the two state model. A single time series, $y_t$, has the following stochastic process,

$$y_t - \mu(S_t) = \rho [y_{t-1} - \mu(S_{t-1})] + u_t,$$

(4.1)

where $u_t \sim i.i.d. N(0, \sigma_u^2)$ and $|\rho| < 1$. Note that this model varies from the usual, linear AR(1) model in that it allows for discrete shifts in the mean of the growth rate of GDP. More specifically, the mean, $\mu(S_t)$, can vary with the underlying regime, $S_t$. This regime is not observed by the econometrician. I consider a model with two regimes, identified by $S_t = 0$ and $S_t = 1$. The mean of the series in these two regimes is given by

$$\mu(S_t) = \alpha + \gamma S_t.$$  

(4.2)

Therefore, in regime zero the mean of $y$ is $\alpha$ and in regime one the mean of $y$ is $\alpha + \gamma$.

The regime is modelled as the outcome of an unobserved, discrete time, discrete state, first order Markov process. Each period there is a fixed probability of switching between

---

2More complicated versions of the model can allow for shifts in the residual variance and the autoregressive parameters.
the states. These transition probabilities are

\[
\begin{align*}
P(S_t = 1|S_{t-1} = 1) &= p; \quad P(S_t = 0|S_{t-1} = 1) = 1 - p \\
P(S_t = 1|S_{t-1} = 0) &= 1 - q; \quad P(S_t = 0|S_{t-1} = 0) = q
\end{align*}
\] (4.3)

As mentioned before the econometrician does not directly observe the current regime, or indeed any past regimes. Instead he or she draws inferences about these regimes based on what they have observed, in this case \( \{y_t\} \). This inference is based on a non-linear filter. At time \( t \), the filter takes as its input the conditional probability,

\[
P(S_{t-1} = 1|y_{t-1}, y_{t-2}, \ldots, y_1)
\] (4.4)

and gives as output the conditional probability,

\[
P(S_t = 1|y_t, y_{t-1}, \ldots, y_1)
\] (4.5)

This filter is similar to that in the previous chapter and is described in appendix four. Step three of the filter gives as a by-product the conditional likelihood which allows the econometrician to estimate the parameters of the model by maximum likelihood. To begin the filter, I assume the probability that the economy is in state one at time zero to be the unconditional probability

\[
P(S_0 = 1) = \pi = \frac{(1 - q)}{(1 - p) + (1 - q)}
\] (4.6)

Certain parameter restrictions from this two state model would imply a one state model. If either of the switching probabilities is equal to one then a single state model is implied. Recall \( p = P(S_t = 1|S_{t-1} = 1) \). Therefore, \( p = 1 \) implies that the economy begins in state one, see equation (4.6), and never switches to state zero. In this case the single long-run rate of growth for the economy is given by \( \alpha + \gamma \). Alternatively, \( q = 1 \) implies that the economy begins in state zero, again see equation (4.6), and never switches
to state one. In this case the single long-run growth rate is given by \( \alpha \). Finally if \( \gamma = 0 \) then there is no difference the growth rate of GDP between the two states.

A single state counterpart to the model described by equations (4.1) to (4.3) is given by

\[
(y_t - \alpha) = \rho (y_{t-1} - \alpha) + v_t, \tag{4.7}
\]

again \( v_t \sim i.i.d. N(0, \sigma_v^2) \) and \(|\rho| < 1\). Here instead of the trend rate of growth shifting stochastically with the underlying state, it is constant. In other words, there is only a single state for GDP growth. Note that the parameters \( \gamma, p \) and \( q \) are not defined for this model.

Both of these models can be estimated using maximum likelihood. Let \( l(y, \theta) \) be the log-likelihood for time series \( y \) and parameter vector \( \theta \). The vector \( y \) contains \( n \) elements, where \( n \) is the sample size, and the vector \( \theta, k \) elements. For the Markov switching model \( k = 6 \) and for the linear model \( k = 3 \). Now define the matrix of contributions to the gradient as \( G(y, \theta) \) where\(^3\)

\[
G_{ii}(y, \theta) = \frac{\partial l_t(y, \theta)}{\partial \theta_i}. \tag{4.8}
\]

As mentioned above, under the null hypothesis of a single state model, the parameters which describe the second state, \( \gamma, p \) and \( q \) are not defined. Therefore, the elements of \( G(y, \theta) \) which describe the contribution of those three parameters to the gradient will be zeros under the null hypothesis. Define the information matrix as

\[
\Gamma^n = \sum_{i=1}^{n} I_t(\theta), \tag{4.9}
\]

where element \( i, j \) of the \( k \times k \) matrix \( I_t(\theta) \) is given by

\[
(I_t(\theta))_{ij} = E_p (G_{ti}(y, \theta) G_{tj}(y, \theta)). \tag{4.10}
\]

Here \( E_p \) refers to the expectation calculated using the data generating process charac-
characterized by the particular parameter vector $\theta$. Hamilton [30] points out that one of the regularity conditions for the likelihood ratio test statistic to have a standard, asymptotic $\chi^2$ distribution is that the information matrix be non-singular. Clearly this condition will fail to hold under the null hypothesis of a one state model.

Putting aside the problem of testing for the number of states, there exists the potential problem of model mis-specification. When estimating a model that incorporates changes in regime there is the possibility that a change in regime may be used to explain 'outliers' from a single regime. For example, suppose the true data generating process for the economy has a single long-run rate of growth. However, an econometrician estimates a two state Markov switching model using this data. By explaining an observed growth rate which is far from the single, long-run growth rate as a change in regime, this non-linear model maybe able to produce a significant improvement in fit over the linear model. As a result, the econometrician may reject the true, linear model in favor of the mis-specified non-linear alternative. If the econometrician then uses this non-linear model for out-of-sample prediction, he or she may generate less accurate predictions than they would have done had they used the true linear model.

Therefore, one desirable feature of models that incorporate the possibility of a shift in the trend rate of growth, is that when estimated using observations generated from a linear time series, they return parameter estimates which are consistent with a single long-run rate of growth. As mentioned earlier, in Hamilton's Markov switching model this can happen in one of two ways. One way is that the estimates of the parameters which describe the movement between the different regimes can imply that no such change in regime takes place.\textsuperscript{4} The other, is that there is no difference in the growth rate of the economy in the two regimes.

However, mis-specification bias is not the only potential source of bias that comes from using a two-state model when the true data generating process has only one state. One of the desirable properties of maximum likelihood estimators is their invariance to

\textsuperscript{4}The probability of switching to state $i$ conditional on being in state $j \neq i$ is estimated as being zero.
reparameterizations of the model. Let \( \hat{\theta} \) be the unbiased maximum likelihood estimator of the parameter vector \( \theta \), that is \( E\hat{\theta} = \theta \). Now consider a reparameterization of the model such that \( \lambda = g(\theta) \). Invariance to reparameterizations implies that \( \hat{\lambda} = g(\hat{\theta}) \). However, if \( g \) is a non-linear function, this will not hold. This is a consequence of Jensen’s inequality.\(^5\) Therefore the choice of a non-linear reparameterization can lead to biased parameter estimates. As a result, even if the estimation of the two state model returns parameter estimates which imply a one state model, the parameters which describe the single state model may still be biased.

In what follows, I examine the degree of bias on the basis of out-of-sample predictive accuracy. Suppose that the extent of bias induced by fitting the two-state model when the true data generating process has only one state is small. In this case the out-of-sample predictive performance of the non-linear model should be similar to that of a linear model. The greater the extent of the bias, the poorer the predictive performance of the non-linear model. Similarly, if the degree of bias induced by fitting a one state model when the true data generating process has two states is small, the predictive accuracy of the linear model should be similar to that of the Markov switching model.

Prediction using the one state model is straightforward. Let \( E[y_{t+h} | y_t] \) be the \( h \) step ahead prediction of \( y \) conditional on the observation of \( y \) until time \( t \). Clearly from equation (4.7)

\[
E[y_{t+1} | y_t] = \alpha + \rho (y_t - \alpha),
\]

and it is easily shown that in general

\[
E[y_{t+h} | y_t] = \alpha + \rho^h (y_t - \alpha).
\]

Therefore, given \( |\rho| < 1 \), \( \lim_{h \to \infty} E[y_{t+h} | y_t] = \alpha \).

Prediction using the Markov Switching model is less straightforward as it also re-

\(^5\)More specifically \( Eg(\hat{\beta}) > g(E\hat{\beta}) \) if \( g \) is a strictly convex function and \( Eg(\hat{\beta}) < g(E\hat{\beta}) \) if \( g \) is strictly concave.
quires the prediction of future states. Using equations (4.1) and (4.2) we can see that the conditional one-step ahead prediction of \( y \) will depend now on the conditional probabilities assigned to the two states in that period. This probability is conditional on the observation of \( y \) until the current period. Therefore the one-step ahead prediction can be written as

\[
E[y_{t+1} | y_t] = \alpha + P (S_{t+1} = 1 | y_t, \ldots, y_0) \times \gamma + \rho (y_t - \alpha - P (S_t = 1 | y_t, \ldots, y_0) \times \gamma).
\] (4.13)

As with the linear model, using forward iteration and the law of iterated projections, it can be shown that

\[
E[y_{t+h} | y_t] = \alpha + P (S_{t+h} = 1 | y_t, \ldots, y_0) \times \gamma + \rho^h (y_t - \alpha - P (S_t = 1 | y_t, \ldots, y_0) \times \gamma).
\] (4.14)

Hamilton [28] shows that the prediction of future states conditional on current information is governed by the following equation:

\[
P (S_{t+h} = 1 | y_t, \ldots, y_0) = \pi + \lambda^h (P (S_t = 1 | y_t, \ldots, y_0) - \pi),
\] (4.15)

where \( \lambda = p + q - 1 \). From equation (4.15) it can be seen that as the prediction horizon tends to infinity the conditional prediction of the future state tends to the unconditional probability, \( \pi \).\(^6\) Therefore, from equation (4.14), given that \( |\rho| < 1 \), \( \lim_{h \to \infty} E [y_{t+h} | y_t] = \alpha + \pi \gamma \). That is, as the prediction horizon tends to infinity the prediction of \( y \) tends to a constant which is a weighted sum of the means in the two states.

### 4.3 Estimates Using Canadian GDP

In this section I estimate the Markov switching model and the linear AR(1) model of the previous section using quarterly Canadian real per capita GNP \( (Y_t) \) from 1960 to

\(^6\)Assuming that either \( p \) or \( q \) is strictly less than unity implies that \( \lambda < 1 \).
1995. I estimate the model using annualized growth rates of this series \((y_t)\), which are shown in figure 4.1a. The mean growth rate over this period is 2.24% and the standard deviation is 4.05%. Panel (a) of table 4.1 presents the maximum likelihood estimates of the parameters of the Markov switching model and panel (b) of the same table presents maximum likelihood estimates of the linear model.

The estimates from the Markov switching model suggest two different states. State zero is a low growth state with an average, annual growth rate of \(-2.26\%\) and state one is a high growth state with an average, annual growth rate of 3.07%. In comparison the linear model implies a single long-run growth rate of 2.20%. The Markov switching probabilities imply that the expected duration of recessions is about \(7\frac{1}{2}\) quarters and the expected length of booms is about 26 quarters. Figure 4.1b shows the time series of estimated probabilities that the economy was in the high growth state conditional on what has been observed to date, that is \(P(S_t = 1 \mid y_t, \ldots, y_0)\). This diagram shows two main periods of low growth, the recessions of the early 1980s and the early 1990s.

Note that the Markov switching model contains three more parameters than the linear model. It also gives an improvement in the log-likelihood of 5.15 relative to the linear model. Under standard conditions this would imply a likelihood ratio test statistic of 10.30. Using the standard \(\chi^2(3)\) distribution, this would lead to a rejection of the linear model in favor of the Markov switching model at the 5% significance level. However, as

---

7 An Augmented Dickey-Fuller (ADF) test of the null hypothesis that this series is \(I(2)\) against the alternative that it is \(I(1)\) with a trend gives a test statistic of -3.9879. Using Stock’s [51] procedure for calculating the confidence intervals of the largest autoregressive root gives an upper bound on this root of 0.952. The ADF test of the null that the series is \(I(1)\) against the alternative that it is \(I(0)\) with a trend gives a test statistic of -0.6521. This translates into a confidence interval for the largest autoregressive root of (1.002, 1.036). These figures imply that the series is \(I(1)\).

8 More specifically \(y_t = 400 \times \ln \left( \frac{Y_t}{Y_{t-1}} \right)\). The data are from CANSIM. Series D10373 (Real GDP) divided by series D1 (population).

9 These results suggest less frequent periods of low growth for Canada than Hamilton’s [28] results imply for the United States. The periods of low growth in the early 1960s and early 1980s that Hamilton identifies for the United States also appear in Canadian per capita real GDP. In between these dates Hamilton identifies two other periods of low growth for the United States. These occur in the early and mid 1970s. However, for Canada, only a probability of about 0.5 is attached to the low growth state in the early 1970s, and even less in the mid 1970s.
mentioned earlier this test statistic does not possess a standard $\chi^2$ distribution and so the rejection of the linear model on this basis may be erroneous.

4.4 Monte Carlo Experiment

In this section, I generate artificial data under both the null hypothesis that the data generating process has only one state and the alternative hypothesis that it follows a two state Markov switching process. In each case I then estimate both the one state and the two state model and compare their predictive accuracy. I also calculate the size and power properties of the pseudo likelihood ratio test under the non-standard conditions discussed earlier.

Generating data under the null hypothesis that there is only one state is straightforward. Using equation (4.7) and the parameter estimates in panel (b) of table 4.1, I generate $T + H + 200$ observations for $y$, where $T$ is the sample size of interest and $H$ is the maximum length of the prediction horizon. The first 200 observations are then discarded to minimize the influence of initial values. Here the artificial data for $y_t$ is generated according to:

$$y_t = \bar{\alpha} + \hat{\rho} (y_{t-1} - \bar{\alpha}) + \varepsilon_t$$

(4.16)

where $\varepsilon_t \sim i.i.d. N(0, \sigma^2_t)$.

Generating data under the alternative hypothesis is more complex. In Hamilton's model the unobserved state of the economy is one of many factors that determine the dynamic process for output. As a result, it is possible that one could observe a high rate of growth for output even when the economy is in the low growth state.\textsuperscript{10} Here I use an approach based on Gallant et al [26]. Using $s_{t-1}$ and the estimates of the Markov transition probabilities, $\hat{\rho}$ and $\hat{q}$, from panel (a) of table 4.1, I generate the conditional

\textsuperscript{10}This is in contrast to the threshold models proposed by Tong [54] in which the current state is directly observable. Often, as is the case in self-exciting threshold autoregressive models, the state is a function of the level of output. In these models, if output was below the fixed threshold at time $t$ then the economy was in the low growth state at time $t$. 

75
distribution for $s_t$. If $s_{t-1} = 1$ then proportion $\hat{p}$ of the elements of the conditional distribution for $s_t$ will be ones and proportion $1 - \hat{p}$ will be zeros. Similarly, if $s_{t-1} = 0$ then proportion $1 - \hat{q}$ of the elements of the conditional distribution will be ones and proportion $\hat{q}$ will be zeros. Having generated this conditional distribution for $s_t$, I then randomly draw a value of $s_t$ from that distribution. For the initial value of the state at time one I draw $s_1$ from an unconditional distribution with proportion $\pi$ of its elements being ones and proportion $1 - \pi$ being zeros.

A similar method is used to generate the distribution for $y_t$ conditional on $s_{t-1}$. If $s_{t-1} = 1$ then proportion $\hat{p}$ of the elements of the conditional distribution for $y_t$ are equal to

$$\hat{\alpha} + \hat{\gamma} + \hat{p} (y_{t-1} - \hat{\alpha} - \hat{\gamma}) + w_{it}$$

and proportion $1 - \hat{p}$ are equal to

$$\hat{\alpha} + \hat{\rho} (y_{t-1} - \hat{\alpha} - \hat{\gamma}) + w_{it}$$

where $\hat{\alpha}$, $\hat{\gamma}$ and $\hat{\rho}$ are the parameter estimates in table 4.1, and $w_{it}$ are drawn from a mean-zero normal distribution with standard deviation $\hat{\sigma}$ which is also reported in table 4.1. The subscript $i$ refers to the individual element of the distribution. Similarly, if $s_{t-1} = 0$ then proportion $\hat{q}$ are

$$\hat{\alpha} + \hat{\gamma} + \hat{q} (y_{t-1} - \hat{\alpha} - \hat{\gamma}) + w_{it}$$

and proportion $1 - \hat{q}$ are

$$\hat{\alpha} + \hat{\gamma} + \hat{p} (y_{t-1} - \hat{\alpha}) + w_{it}.$$  

As before, having generated the conditional distribution for $y_t$, I then randomly draw the artificial observation $y_t$ from that distribution. To generate $y_1$, I draw from an unconditional distribution with proportion $\pi$ of its elements given by $\hat{\alpha} + \hat{\gamma} + w_{i1}$ and proportion $1 - \pi$ given by $\hat{\alpha} + w_{i1}$. As with the generation of the linear time series I
generate $T + H + 200$ observations for both the state and $y_t$ and discard the first 200.

I use the first $T$ observations of the remaining artificial time series for $y_t$ to estimate the one state and the two state models described in section 4.2. I then calculate the pseudo likelihood ratio test statistic under the null and the alternative hypotheses. Finally, I then use the two estimated models to predict $y_{t+1}, y_{t+2}, \ldots, y_{t+H}$. I consider sample sizes of 100, 200, 300 and 400. $H$ is set at 32, thus representing predictions of up to eight years ahead. The experiment is repeated 1000 times for each sample size.

### 4.4.1 Distribution of Test Statistic

Table 4.2 gives the small sample properties of the pseudo likelihood ratio test statistic. The type I error rate, based on a nominal size of 0.05, is below 0.05 for all four sample sizes. This suggests that there is no size distortion. The next three rows of table 4.2 show the 90%, 95% and 99% critical values from the empirical distribution. Typically these figures lie between the critical values from the asymptotic $\chi^2(2)$ and $\chi^2(3)$ distributions. This is not surprising given the type I error rates.

Using linear interpolation between the critical values for the sample sizes of 100 and 200 gives a 95% critical value of 6.848 and a 99% critical value of 10.16 for the sample size of 144. This is the length of the sample used in the previous section to estimate the two models using Canadian data on per capita real GDP. The pseudo likelihood ratio test statistic calculated in the previous section is 10.30. Therefore, using these empirical distributions suggests that the linear AR(1) model for Canadian real, per-capita GDP growth can be rejected in favor of the Markov switching alternative at the 1% level.

These results are similar to the results of Lam [35] and Cecchetti, Lam and Mark [15]. Lam generates the empirical distribution under the null hypothesis for the pseudo likelihood ratio test statistic for the null hypothesis of an ARIMA model against the alternative of Hamilton's [28] original model. Based on 100 replications and using a sample size of 131 observations, he calculates a 90% critical value of 6.65 and a 95% critical value of 9.14. Cecchetti, Lam and Mark generate data under the null of a normal
distribution and compare this with the alternative of a Markov switching model. They perform this experiment for United States annual data. They use sample sizes of 96 for consumption and 116 observations for both real dividends and GNP. They find that from 1000 replications 99.2% are below their pseudo likelihood ratio test statistic for consumption of 11.39. The 99% critical values for asymptotic $\chi^2(3)$ distribution is 11.34. The p-values reported for their test statistics for real dividends and GNP also suggest that there is no size distortion.\textsuperscript{11}

As discussed earlier, in addition to generating data under the null hypothesis of a linear AR(1), I also generate data under the alternative hypothesis of a two state Markov switching model. This allows me to calculate the power of the pseudo likelihood ratio test statistic. Size adjusted power is given in the final row of table 4.2. These figures show that the power of the pseudo likelihood ratio test statistic is poor in small samples. In this experiment, using a 95% significance level, when the sample size is 100 the false null hypothesis is rejected in favor of the true alternative in only 40.6% of the simulations. When the sample size is 200 this figure rises to 78.6%. However, it is only when the sample size reaches 400 that size adjusted power reaches 95%.

Note that these size adjusted power figures are calculated using the empirical 95% critical values from table 4.2. These are all below their asymptotic $\chi^2(3)$ counterparts. When asymptotic critical values are used, power is even lower for each sample size. For a sample size of 100 power calculated using the asymptotic 95% critical value is 0.353. When the sample size rises to 200 this figure rises to 0.69; at $T = 300$ it is 0.841; and at $T = 400$ it is 0.939.

Most quarterly macroeconomic time series currently consist of somewhere between 100 and 200 observations. Based on these figures, with a sample size of 100, there is a much higher chance of failing to reject the single state model in favor of the two state model than rejecting it, even when the two state model is the true model. When the

\textsuperscript{11}These test statistics are 17.27 and 27.87 and have p-values of 0.000 and 0.001 respectively. Therefore while this is consistent with no size distortion, these tests statistics are too far above the asymptotic critical values to allow us to completely rule it out.

78
sample size is 200 there is still a 30% chance of failing to reject the false null. These results show that, at the 95% significance level in order for the type II error rate for the pseudo likelihood ratio test statistic to fall to 5% we require a sample size of 400. In other words, in order for this pseudo likelihood ratio test statistic to have the correct power against the single state null hypothesis, we require a quarterly time series of more that twice the length than that currently available for most macroeconomic aggregates.

4.4.2 Predictive Accuracy

To compare the predictive accuracy of the single state and the two state models I use two statistics. These are mean squared prediction error and mean absolute prediction error. For each repetition of the experiment I obtain a series of predictions, $E[y_{t+h} | y_t]$ for $h = 1, 2, \ldots, 32$, using the single state model. I also obtain a similar series of predictions for the two state model. The mean of the predictions across the 1000 repititions are shown in figure 4.2. Panel (a) contains the means of the predictions when the one state model is the true model and panel (b) the means when the two state model is the true model. In each case $m_1$ refers to the mean prediction from the one state model and $m_2$ the mean prediction from the two state model. $s_1$ and $s_2$ are the respective standard deviations. This figure shows that in each case the means of the predictions from the true model and the mis-specified model are almost identical.

Using these predictions and the artificial data, I calculate prediction errors at each horizon for both models. I then average across repetitions to calculate the mean squared prediction errors and mean absolute prediction errors. The mean squared prediction error at forecast horizon $h$ for the model with $k$ states is given by $MSPE_{k,h}$. Similarly the mean absolute prediction error at forecast horizon $h$ for the model with $k$ states is given by $MAPE_{k,h}$. I report these statistics as the ratio of the statistic for the two state model to the statistic of the one state model. For a given forecast horizon, if the average of the prediction errors from the single state model are less than those from the Markov switching model, this ratio will be greater than one. If, on average they are greater, this
figure will be less than one.

Table 4.3 contains these ratios for the case when the true model is the single state AR(1) model. Here I report the results for forecast horizons of \( h = 1, 2, 4, 8, 16 \) and 32 and for each sample size. Panel (a) contains the ratios of mean squared prediction errors and panel (b) the ratios of mean absolute prediction errors. These results show that when the true model has only one state there is almost no loss in predictive accuracy from using the over parameterized, two state Markov switching model. The ratio of \( MSPE_{2,h} \) to \( MSPE_{1,h} \) is only above 1.01 on one occasion, for \( T = 300 \) and \( h = 2 \). On every other occasion there is less than a 1\% increase in mean squared prediction error when using the two state model. The results are similar when looking at mean absolute prediction error. As discussed earlier there are two sources of bias that can arise from fitting a two state model when the true process has only a single state. These are due to model mis-specification and Jensen's inequality. The results in table 4.3 suggest that the degree of bias arising from these two sources is small.\(^{12}\)

This is supported by looking at the long-run average growth rate and the autoregressive parameter. As the forecast horizon increases the forecasts from both the one and the two state model converge to the long-run growth rate implied by their parameter estimates. For the one state model this is just equal to \( \alpha \). For the two state model this is equal to \( \alpha + \pi \gamma \). The true long-run growth rate when the data are generated under the null hypothesis of a single state for GDP growth is 2.204\%. For the smallest sample size of 100, the average estimated long-run growth rate for the single state model is 2.198\%. For the two state model this figure is 2.201\%. The results are similar for the larger sample sizes.\(^{13}\) This suggests that there is very little bias resulting from fitting the over parameterized two state model. When the process has only one state, the true value of

\(^{12}\)However, these results are also consistent with the interpretation that the degree of bias arising from these two sources is large, and the two sources are offsetting. While there is no reason to suppose that this is the case, it cannot be determined whether or not this is so using the MSPE and MAPE.

\(^{13}\)For \( T = 200 \) the average estimated long-run growth rate for the single state model is 2.207 and for the two state model it is 2.208. For \( T = 300 \), these figures are 2.211 and 2.210; for \( T = 400 \) they are 2.210 and 2.210.
the autoregressive parameter is 0.238. There is some downward bias in the estimate of this parameter when using the two state model. When the sample size is 100 the average estimate is 0.185. This estimate rises with the sample size, to reach 0.212 when \( T = 400 \).

Table 4.4 contains the ratios of the mean squared prediction errors and the mean absolute prediction errors when the true model is the two state Markov switching model. This table shows that there is a slight loss in predictive accuracy if one uses a single state model when the true process has two states. However, this is confined to horizons of up to one year. Even then, this loss in predictive accuracy is never more than 3%. The average estimates of the long-run growth rates for the two models are both very close to the true value of 2.046. For the sample size of \( T = 100 \), the average estimates are 2.011 and 2.012 for the one state and the two state model respectively.\(^{14}\) Both of these results suggest that there is little bias induced through fitting a one state model when the true process has two states. As is the case when the true process has one state, estimating the mis-specified model does induce bias in the autoregressive parameter. In this case, there is a large degree of upward bias in the parameter estimate. The true parameter value is 0.019. The average estimates from the single state model are 0.166, 0.180, 0.180 and 0.186 for sample sizes of \( T = 100, 200, 300 \) and 400 respectively.

### 4.5 Conclusions

The likelihood ratio test statistic for the test of a linear AR(1) model against the alternative of a Markov switching model does not have the standard \( \chi^2 \) distribution. This is because some of the parameters of the Markov switching model are not defined under the null hypothesis of a linear model. In this chapter I examine the size and power properties of the pseudo likelihood ratio test under the non-standard conditions. I find that there are no size distortions that arise from using this test statistic under non-standard conditions. However, I find that the power properties of the test statistic are poor in

\(^{14}\)When the sample sizes is \( T = 200 \) these figures are 1.975 and 1.975; for \( T = 300 \) they are 1.980 and 1.980, and for \( T = 400 \), 1.980 and 1.979.
small samples. Size adjusted power does not reach 95% until the sample size reaches 400. Clearly 400 is larger than the number of quarterly observations than is typically available on most macroeconomic aggregates. Therefore, the econometrician who fails to reject the null of a linear model in favor of the Markov switching model may do so not because the null model is true, but because he or she has insufficient observations for the test statistic to have the correct power.

I also investigate the bias resulting from estimating a two state model when the true model has only one state, and from estimating a single state model when the true model has two states. In both cases I find that there is little loss of predictive accuracy from using the mis-specified model. In the former case, this suggests that the degree of bias resulting from using an over parameterized model and the non-linear reparameterization is small. The latter case suggests, though to a slightly lesser extent, that the degree of bias from using the under parameterized model is also small.
Table 4.1: Markov Switching and AR(1) Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Markov Switching Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>-2.255</td>
<td>0.952</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>5.521</td>
<td>0.907</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.019</td>
<td>0.092</td>
</tr>
<tr>
<td>$p$</td>
<td>0.962</td>
<td>0.025</td>
</tr>
<tr>
<td>$q$</td>
<td>0.866</td>
<td>0.084</td>
</tr>
<tr>
<td>$\sigma_u$</td>
<td>3.412</td>
<td>0.227</td>
</tr>
<tr>
<td>log-likelihood</td>
<td>-393.40</td>
<td></td>
</tr>
<tr>
<td>(b) Linear AR(1) Model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>2.204</td>
<td>0.431</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.238</td>
<td>0.081</td>
</tr>
<tr>
<td>$\sigma_v$</td>
<td>3.940</td>
<td>0.232</td>
</tr>
<tr>
<td>log-likelihood</td>
<td>-398.55</td>
<td></td>
</tr>
</tbody>
</table>

Both of these models are estimated using quarterly data on the annualized growth rate of Canadian per-capita, real GDP. The beginning of the sample period is 1960Q1 and the end is 1995Q4. This gives 144 observations.
The error rates are based on the 95% critical value from the $\chi^2(3)$ distribution. This asymptotic critical value is 7.81. The 90% and 99% asymptotic critical values are 6.25 and 11.34 respectively. Under the null hypothesis the data are generated by a linear AR(1) model. Under the alternative the data are generated by the Markov Switching Model with one autoregressive parameter. These error rates and empirical critical values are based on 1000 replications of the Monte Carlo study described in the text.
Table 4.3: Relative Forecasting Performance: True Model has 1 State

(a) $MSPE_{2,h}/MSPE_{1,h}$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$T = 100$</th>
<th>$T = 200$</th>
<th>$T = 300$</th>
<th>$T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.009</td>
<td>1.001</td>
<td>0.998</td>
<td>1.001</td>
</tr>
<tr>
<td>2</td>
<td>1.005</td>
<td>1.003</td>
<td>1.012</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>1.004</td>
<td>1.002</td>
<td>1.004</td>
<td>1.002</td>
</tr>
<tr>
<td>8</td>
<td>1.006</td>
<td>1.001</td>
<td>0.998</td>
<td>0.999</td>
</tr>
<tr>
<td>16</td>
<td>0.998</td>
<td>1.002</td>
<td>1.002</td>
<td>1.000</td>
</tr>
<tr>
<td>32</td>
<td>1.000</td>
<td>0.999</td>
<td>1.001</td>
<td>1.001</td>
</tr>
</tbody>
</table>

(b) $MAPE_{2,h}/MAPE_{1,h}$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$T = 100$</th>
<th>$T = 200$</th>
<th>$T = 300$</th>
<th>$T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>1.006</td>
<td>1.000</td>
<td>1.003</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>1.003</td>
<td>1.000</td>
<td>1.003</td>
<td>1.003</td>
</tr>
<tr>
<td>8</td>
<td>1.003</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>16</td>
<td>1.000</td>
<td>1.000</td>
<td>1.003</td>
<td>1.000</td>
</tr>
<tr>
<td>32</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Relative forecast performance is measured as the ratio of the mean squared prediction error from the two state Markov Switching model at forecast horizon $h$, denoted $MSPE_{2,h}$, to the mean squared prediction error from the single state, linear AR(1) model, denoted $MSPE_{1,h}$. This ratio is also calculated using the mean absolute prediction error.
Table 4.4 Relative Forecasting Performance: True Model has 2 States

(a) $MSPE_{2,h}/MSPE_{1,h}$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$T = 100$</th>
<th>$T = 200$</th>
<th>$T = 300$</th>
<th>$T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.970</td>
<td>0.977</td>
<td>0.973</td>
<td>0.979</td>
</tr>
<tr>
<td>2</td>
<td>0.973</td>
<td>0.975</td>
<td>0.975</td>
<td>0.983</td>
</tr>
<tr>
<td>4</td>
<td>1.011</td>
<td>0.977</td>
<td>0.988</td>
<td>0.968</td>
</tr>
<tr>
<td>8</td>
<td>1.006</td>
<td>1.014</td>
<td>0.997</td>
<td>0.992</td>
</tr>
<tr>
<td>16</td>
<td>1.005</td>
<td>1.002</td>
<td>1.002</td>
<td>1.002</td>
</tr>
<tr>
<td>32</td>
<td>1.000</td>
<td>0.998</td>
<td>1.001</td>
<td>1.001</td>
</tr>
</tbody>
</table>

(b) $MAPE_{2,h}/MAPE_{1,h}$

<table>
<thead>
<tr>
<th>$h$</th>
<th>$T = 100$</th>
<th>$T = 200$</th>
<th>$T = 300$</th>
<th>$T = 400$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.976</td>
<td>0.991</td>
<td>0.981</td>
<td>0.981</td>
</tr>
<tr>
<td>2</td>
<td>0.984</td>
<td>0.986</td>
<td>0.988</td>
<td>0.988</td>
</tr>
<tr>
<td>4</td>
<td>1.009</td>
<td>0.991</td>
<td>0.994</td>
<td>0.981</td>
</tr>
<tr>
<td>8</td>
<td>1.009</td>
<td>1.003</td>
<td>1.000</td>
<td>0.994</td>
</tr>
<tr>
<td>16</td>
<td>1.003</td>
<td>1.003</td>
<td>1.000</td>
<td>1.003</td>
</tr>
<tr>
<td>32</td>
<td>0.997</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

See notes at the foot of table 4.3.
Figure 4.1: Canadian Real Per-Capita GDP 1960-1995

Fig 4.1a: Annualized Growth Rate

Fig 4.1b: Conditional Probability of High Growth State
Figure 4.2: Mean & Standard Deviations of Forecasts ($T=100$)

**Fig 4.2a: One State Model is the True Model**

**Fig 4.2b: Two State Model is the True Model**
Bibliography


Appendix One - Data for Chapter Two

This appendix describes the data used in chapter two and the sources from which it was compiled.

(1) United States

The monetary base is currency held by the public plus bank vault cash, plus, after October 1914, member bank deposits and non-member clearing accounts at the Federal Reserve. The money stock (M2) is currency held by the public plus adjusted deposits at all commercial banks. Both are measured in billions of current dollars. Real income is measured by real GNP in billions of 1972 dollars. The source for the monetary base, the money stock and real income from 1869 to 1983 is Balke and Gordon [2]. For all series, the source from 1984 to 1995 is the Economic Report of the President [17].

(2) United Kingdom

The monetary base is currency held by the public plus bank reserves. The money supply (M3) is non-bank holdings of notes and coins plus the deposits of all residents. Both are measured in millions of current pounds sterling. Output is measured by real GDP in millions of 1958 pounds sterling. For the UK the source of both monetary series is Capie and Webber [14]. The GDP data is from Mitchell [43].

Appendix 2: Estimation of VAR Under Null and Alternative

Assuming that both \( y_t \) and \( m_t \) are stationary in first differences the VAR of equation (2.5) can be written as:

\[
\begin{align*}
    a(L) \Delta m_t &= b(L) \Delta y_t + u_t \\
    d(L) \Delta y_t &= c(L) \Delta m_t + v_t.
\end{align*}
\]  
(A2.1)

For simplicity assume a first order VAR, and so equation (A2.1) becomes:

\[
\begin{align*}
    \Delta m_t &= b_0 \Delta y_t + b_1 \Delta y_{t-1} + a_1 \Delta m_{t-1} + u_t \\
    \Delta y_t &= c_0 \Delta m_t + c_1 \Delta m_{t-1} + d_1 \Delta y_{t-1} + v_t.
\end{align*}
\]  
(A2.2)

King and Watson show that the top line of this equation can be rewritten as

\[
\Delta m_t - \gamma \Delta y_t = a_1 (\Delta m_{t-1} - \gamma \Delta y_t) + b \Delta^2 y_t + u_t
\]  
(A2.3)

where \( \gamma = b(1)/(1-a_1) = b(1)/a(1) \) is the long-run multiplier. Under Fisher and Seater's identification scheme \( b(1) = 0 \) and therefore \( \gamma = 0 \). Thus equation (A2.3) becomes

\[
\Delta m_t = a_1 \Delta m_{t-1} + b \Delta^2 y_t + u_t.
\]  
(A2.4)

Potentially there is correlation between \( \Delta^2 y_t \) and \( u_t \). Therefore I estimate this equation by instrumental variables, with the instruments being \{\( \Delta m_{t-1}, \Delta y_{t-1} \)\}. Under the identification scheme that \( \text{cov}(u_t, v_t) = 0 \) the estimated residuals for this equation are a valid instrument in the estimation of the second equation in (A2.1). The second equation of (A2.1) is then estimated by instrumental variables with the instruments being \{\( \Delta m_{t-1}, \Delta y_{t-1}, \hat{u}_t \)\}. Using these parameter estimates I generate the data under the alternative hypothesis.

To generate data under the null hypothesis, I estimate the money equation in exactly the same way. King and Watson show that the second line of equation (A2.1) can be written as

\[
\Delta y_t - \delta \Delta m_t = a_1 (\Delta y_{t-1} - \delta \Delta m_t) + b \Delta^2 m_t + v_t
\]  
(A2.5)

Where \( \delta \) is the long-run multiplier of income with respect to money. The long-run neutrality restriction is that \( \delta = 0 \). Thus, this equation becomes

\[
\Delta y_t = a_1 \Delta y_{t-1} + b \Delta^2 m_t + v_t.
\]  
(A2.6)

I estimate this equation by instrumental variables, with instruments \{\( \Delta y_{t-1}, \Delta m_{t-1}, \hat{u}_t \)\}. Using these parameter estimates I generate data under the null hypothesis.
### Appendix 3: Estimated Conditional Probabilities 1919Q3 - 1941Q4

#### Conditional Probabilities of Financial Crisis

<p>| Year (Q)   | $P(s_t = 1 | x_t, x_{t-1}, x_o)$ | $P(s_{t-1} = 1 | x_t, x_{t-1}, x_o)$ | $P(s_{t-2} = 1 | x_t, x_{t-1}, x_o)$ |
|------------|---------------------------------|-----------------------------------|-----------------------------------|
| 1919Q3     | 0.011                           | 0.003                             | 0.000                             |
| 1919Q4     | 0.067                           | 0.008                             | 0.002                             |
| 1920Q1     | 0.786                           | 0.164                             | 0.029                             |
| 1920Q2     | 0.380                           | 0.368                             | 0.080                             |
| 1920Q3     | 0.041                           | 0.036                             | 0.035                             |
| 1920Q4     | 0.252                           | 0.113                             | 0.085                             |
| 1921Q1     | 0.109                           | 0.088                             | 0.033                             |
| 1921Q2     | 0.036                           | 0.018                             | 0.015                             |
| 1921Q3     | 0.008                           | 0.003                             | 0.001                             |
| 1921Q4     | 0.009                           | 0.001                             | 0.000                             |
| 1922Q1     | 0.005                           | 0.001                             | 0.000                             |
| 1922Q2     | 0.019                           | 0.002                             | 0.000                             |
| 1922Q3     | 0.008                           | 0.002                             | 0.000                             |
| 1922Q4     | 0.005                           | 0.001                             | 0.000                             |
| 1923Q1     | 0.028                           | 0.001                             | 0.000                             |
| 1923Q2     | 0.014                           | 0.004                             | 0.000                             |
| 1923Q3     | 0.033                           | 0.004                             | 0.001                             |
| 1923Q4     | 0.011                           | 0.003                             | 0.000                             |
| 1924Q1     | 0.011                           | 0.001                             | 0.000                             |
| 1924Q2     | 0.006                           | 0.001                             | 0.000                             |
| 1924Q3     | 0.005                           | 0.000                             | 0.000                             |
| 1924Q4     | 0.009                           | 0.001                             | 0.000                             |
| 1925Q1     | 0.013                           | 0.001                             | 0.000                             |
| 1925Q2     | 0.005                           | 0.001                             | 0.000                             |
| 1925Q3     | 0.008                           | 0.000                             | 0.000                             |
| 1925Q4     | 0.006                           | 0.001                             | 0.000                             |
| 1926Q1     | 0.006                           | 0.000                             | 0.000                             |
| 1926Q2     | 0.010                           | 0.001                             | 0.000                             |
| 1926Q3     | 0.010                           | 0.001                             | 0.000                             |
| 1926Q4     | 0.011                           | 0.001                             | 0.000                             |
| 1927Q1     | 0.005                           | 0.001                             | 0.000                             |
| 1927Q2     | 0.007                           | 0.000                             | 0.000                             |
| 1927Q3     | 0.014                           | 0.001                             | 0.000                             |
| 1927Q4     | 0.007                           | 0.001                             | 0.000                             |
| 1928Q1     | 0.005                           | 0.000                             | 0.000                             |
| 1928Q2     | 0.018                           | 0.001                             | 0.000                             |
| 1928Q3     | 0.021                           | 0.003                             | 0.000                             |
| 1928Q4     | 0.006                           | 0.002                             | 0.000                             |</p>
<table>
<thead>
<tr>
<th>Year</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1929Q1</td>
<td>0.052</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>1929Q2</td>
<td>0.016</td>
<td>0.006</td>
<td>0.000</td>
</tr>
<tr>
<td>1929Q3</td>
<td>0.015</td>
<td>0.002</td>
<td>0.001</td>
</tr>
<tr>
<td>1929Q4</td>
<td>0.024</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>1930Q1</td>
<td>0.032</td>
<td>0.006</td>
<td>0.001</td>
</tr>
<tr>
<td>1930Q2</td>
<td>0.009</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>1930Q3</td>
<td>0.006</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1930Q4</td>
<td>0.993</td>
<td>0.032</td>
<td>0.003</td>
</tr>
<tr>
<td>1931Q1</td>
<td>0.926</td>
<td>0.926</td>
<td>0.026</td>
</tr>
<tr>
<td>1931Q2</td>
<td>1.000</td>
<td>0.986</td>
<td>0.985</td>
</tr>
<tr>
<td>1931Q3</td>
<td>1.000</td>
<td>1.000</td>
<td>0.988</td>
</tr>
<tr>
<td>1931Q4</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1932Q1</td>
<td>0.852</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1932Q2</td>
<td>1.000</td>
<td>0.857</td>
<td>1.000</td>
</tr>
<tr>
<td>1932Q3</td>
<td>1.000</td>
<td>1.000</td>
<td>0.874</td>
</tr>
<tr>
<td>1932Q4</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1933Q1</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1933Q2</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1933Q3</td>
<td>0.776</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1933Q4</td>
<td>1.000</td>
<td>0.952</td>
<td>1.000</td>
</tr>
<tr>
<td>1934Q1</td>
<td>0.812</td>
<td>1.000</td>
<td>0.936</td>
</tr>
<tr>
<td>1934Q2</td>
<td>0.260</td>
<td>0.256</td>
<td>1.000</td>
</tr>
<tr>
<td>1934Q3</td>
<td>0.631</td>
<td>0.466</td>
<td>0.459</td>
</tr>
<tr>
<td>1934Q4</td>
<td>0.101</td>
<td>0.098</td>
<td>0.064</td>
</tr>
<tr>
<td>1935Q1</td>
<td>0.012</td>
<td>0.007</td>
<td>0.001</td>
</tr>
<tr>
<td>1935Q2</td>
<td>0.007</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1935Q3</td>
<td>0.026</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>1935Q4</td>
<td>0.013</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>1936Q1</td>
<td>0.188</td>
<td>0.037</td>
<td>0.009</td>
</tr>
<tr>
<td>1936Q2</td>
<td>0.057</td>
<td>0.048</td>
<td>0.006</td>
</tr>
<tr>
<td>1936Q3</td>
<td>0.049</td>
<td>0.019</td>
<td>0.016</td>
</tr>
<tr>
<td>1936Q4</td>
<td>0.015</td>
<td>0.007</td>
<td>0.002</td>
</tr>
<tr>
<td>1937Q1</td>
<td>0.022</td>
<td>0.003</td>
<td>0.001</td>
</tr>
<tr>
<td>1937Q2</td>
<td>0.012</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>1937Q3</td>
<td>0.026</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>1937Q4</td>
<td>0.998</td>
<td>0.144</td>
<td>0.012</td>
</tr>
<tr>
<td>1938Q1</td>
<td>1.000</td>
<td>0.999</td>
<td>0.125</td>
</tr>
<tr>
<td>1938Q2</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1938Q3</td>
<td>0.764</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>1938Q4</td>
<td>0.383</td>
<td>0.378</td>
<td>1.000</td>
</tr>
<tr>
<td>Year</td>
<td>Value1</td>
<td>Value2</td>
<td>Value3</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>1939Q1</td>
<td>0.084</td>
<td>0.073</td>
<td>0.072</td>
</tr>
<tr>
<td>1939Q2</td>
<td>0.053</td>
<td>0.026</td>
<td>0.022</td>
</tr>
<tr>
<td>1939Q3</td>
<td>0.029</td>
<td>0.015</td>
<td>0.006</td>
</tr>
<tr>
<td>1939Q4</td>
<td>0.013</td>
<td>0.003</td>
<td>0.002</td>
</tr>
<tr>
<td>1940Q1</td>
<td>0.008</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1940Q2</td>
<td>0.007</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1940Q3</td>
<td>0.005</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1940Q4</td>
<td>0.007</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>1941Q1</td>
<td>0.007</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>1941Q2</td>
<td>0.038</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>1941Q3</td>
<td>0.016</td>
<td>0.005</td>
<td>0.000</td>
</tr>
<tr>
<td>1941Q4</td>
<td>0.021</td>
<td>0.003</td>
<td>0.001</td>
</tr>
</tbody>
</table>
Appendix 4: Filter for Two State Model of Chapter Four

Throughout the five steps of this filter, the parameter values \(p, q, \alpha_0, \alpha_1, \gamma, \sigma_u^2\) are assumed by the econometrician to be known constants. Using these, she then calculates a time series of probabilities over the current regime and past regimes. The five steps of the filter are given below. In order to keep notation simpler, \(P(z)\) denotes \(P(Z = z)\) when \(z\) is a discrete valued variable and the density function \(f(z)\) when \(z\) is a continuous variable. Iteration \(t\) of the filter takes as an input the conditional probability \(P(s_{t-1} | y_{t-1}, \ldots, y_0)\) and gives as output \(P(s_t | y_t, y_{t-1}, \ldots, y_0)\). That is, each iteration of the filter updates the conditional probabilities over the current state one period.

**Step One;** Using the Markov transition probabilities in equation (4.3) calculate

\[
P(s_t, s_{t-1} | y_{t-1}, \ldots, y_0) = P(s_t | s_{t-1}) \times P(s_{t-1} | y_{t-1}, \ldots, y_0),
\]

where the second term on the right hand side is the output from the previous iteration of the filter.

**Step Two;** The joint conditional density of \(x_t\) and \((S_t, S_{t-1})\) is given by

\[
P(x_t, s_t, s_{t-1} | y_{t-1}, \ldots, y_0) = P(x_t | s_t, s_{t-1}, y_{t-1}, \ldots, y_0) \times P(s_t, s_{t-1} | y_{t-1}, \ldots, y_0),
\]

where

\[
P(y_t | s_t, s_{t-1}, y_{t-1}, \ldots, y_0) = \frac{1}{\sqrt{2\pi\sigma_u}} \exp\left\{-\frac{1}{2\sigma_u} \left[ (y_t - \alpha - \gamma s_t) - \rho_1 (y_{t-1} - \alpha - \gamma s_{t-1}) \right]^2 \right\}
\]

**Step Three;** The branch of the likelihood function for observation \(t\) is

\[
P(y_t | y_{t-1}, \ldots, y_0) = \sum_{s_t=0}^{1} \sum_{s_{t-1}=0}^{1} P(y_t, s_t, s_{t-1} | y_{t-1}, \ldots, y_0).
\]

**Step Four;** Steps two and three combined give

\[
P(s_t, s_{t-1} | y_t, y_{t-1}, \ldots, y_0) = \frac{P(y_t, s_t, s_{t-1} | y_{t-1}, \ldots, y_0)}{P(x_t | y_{t-1}, \ldots, y_0)}.
\]

**Step Five;** The output of iteration \(t\) of the filter is

\[
P(s_t | y_t, y_{t-1}, \ldots, y_0) = \sum_{s_{t-1}=0}^{1} P(s_t, s_{t-1} | y_t, y_{t-1}, \ldots, y_0).
\]