CENTRAL BANK OF ICELAND

Working Papers No. 7

BUSINESS CYCLE FORECASTING AND REGIME SWITCHING

Thórarinn G. Pétursson*

April 2000

Abstract

This paper applies Hamilton's (1989) Markov-switching model of business cycle dynamics to real GDP in Iceland for the period 1945 to 1998. The resulting model gives a reasonable description of the data generating process for real GDP and produces business cycles that correspond quite well to conventional wisdom concerning the Icelandic business cycle. Although the model cannot be distinguished from a simple, linear time series model, it offers some improvements in terms of mean absolute forecast errors and in forecasting business cycle turning points to the official forecasts made by the National Economic Institute.

Keywords: Business cycles, Regime shifts, Forecast comparison
JEL: C22, E32, E37

*Division Chief, Economics Department, Central Bank of Iceland, Kalkofnsvegur 1, 150 Reykjavík, Iceland and Reykjavík University, Ofanleiti 2, 103 Reykjavík, Iceland. Email address: `thorar@centbk.is`. I am grateful to Anders Warne for generously providing me with his Gauss code for the EM algorithm and to seminar participants at the Central Bank of Iceland, Guðmundur Guðmundsson and Már Guðmundsson for valuable comments and suggestions. The usual disclaimer applies. The views expressed in this paper should not be interpreted as reflecting those of the Central Bank of Iceland.

1. Introduction

Modern economies undergo significant short-run variations in aggregate activity, commonly referred to as business cycles. Understanding the causes and nature of these fluctuations and the timing of turning points between periods of expansion and contraction is of central importance for government and private decision makers. In this context, certain "stylized facts" have been observed, which any model of business cycles would need to account for. These include the observation that a business cycle usually has a two to six year periodicity, rarely exceeding eight years in length (Canova, 1998). Another stylized fact is that aggregate fluctuations do not seem to exhibit any simple regular or cyclical pattern, with business cycles varying in size and spacing. Thus, business cycles cannot be thought of as combinations of deterministic cycles of different lengths. Rather the economy seems to be perturbed by disturbances of various types and sizes at more or less random intervals. A third important fact is that business cycle movements seem to be asymmetric. Periods of expansion usually have a longer duration than periods of contractions.

Standard, linear time series models will have a hard time explaining some of these stylized facts, such as the observed asymmetry between expansions and contractions, generating suboptimal forecasts of aggregate activity. An alternative framework would be a non-linear time series model. One such model is Hamilton's (1989) Markov-switching autoregressive model, which has fostered a great deal of interest as an empirical approach to characterizing observed business cycle dynamics. Hamilton's model is attractive since it can account for complicated dynamics such as asymmetry and conditional heteroscedasticity, yet is very flexible and simple to estimate and interpret.

In this model of business cycle fluctuations, turning points between expansions and contractions are treated as structural events that are inherent in the data generating process and are estimated jointly with other parameters of the process. Hamilton (1989) successfully applies this model to quarterly GDP in the US and finds that the model generates expansionary and contractionary periods which roughly correspond to the NBER business cycle phases of expansions and contractions. Furthermore, since the probability law governing shifts between expansions and contractions is explicitly modelled, meaningful forecasts of business cycle turning points can be conducted within this framework.

The Markov-switching model of business cycle dynamics has been extended to allow for transitory as well as permanent shocks to output (Lam, 1990); regime dependence of other parameters (Hansen, 1992b); increased number of regimes (Clements and Krolzig, 1998); and duration-dependent transition probabilities (Filardo and Gordon, 1998). It has also proven to be a promising approach for studying other economic phenomena. These include analyses of stock market volatility (Hamilton and Susmel, 1994) and its relationship to the business cycle (Hamilton and Lin, 1996); intrinsic bubbles and switching regimes in stock markets (Driffill and Sola, 1998); leading indicators and economic forecasting (Hamilton and Perez-Quiros, 1996); the relationship between output and prices over the business cycle (Ravn and Sola, 1995); the relationship between inflation rates and inflation uncertainty (Kim, 1993); regime switching in cointegrating relations (Hall *et al.*, 1997); the expectations hypothesis of the term structure (Sola and Driffill, 1994); and unemployment and real interest rate persistence (Bianchi and Zoega, 1998 and Garzia and Perron, 1996, respectively).

The Markov-switching model is used in this paper to analyse aggregate output in Iceland for the period 1945 to 1998. Section 2 briefly discusses Markov chains. Section 3 introduces a Markov-switching model of real GDP and Section 4 contains the estimation results. In Section 5 the forecasting ability of the Markov-switching specification is compared to other models. The final section concludes.

2. Markov Chains

Consider a latent, random variable $s_t \in \{1, 2, ..., r\}$ (with $r \geq 2$ and finite), where the probability that s_t takes the value j depends on the past only through the most recent value s_{t-1}

$$\Pr(s_t = j \,|s_{t-1} = i, s_{t-2} = k, ...) = \Pr(s_t = j \,|s_{t-1} = i\,) = p_{ij}; \quad i, j = 1, 2, ..., r \tag{2.1}$$

A process like (2.1) is called a homogenous first-order, r-state Markov chain with transition probabilities $\{p_{ij}\}_{i,j=1,2,...,r}$. The transition probability p_{ij} gives the probability that state i will be followed by state j (see Hamilton, 1994). The Markov transition probabilities satisfy

$$\sum_{j=1}^{r} p_{ij} = 1; \qquad \forall\ i \tag{2.2}$$

and by collecting all the transition probabilities together a $r \times r$ matrix $\mathbf{P}$, called the transition matrix, is obtained

$$\mathbf{P} = \begin{pmatrix} p_{11} & \cdots & p_{r1} \\ \vdots & \ddots & \vdots \\ p_{1r} & \cdots & p_{rr} \end{pmatrix}$$

It is assumed that the Markov process is irreducible, so that $p_{ii} < 1$ for all i (the states are therefore said to be non-absorbing). Furthermore, since from (2.2)

$\mathbf{P}'\boldsymbol{\iota} = \boldsymbol{\iota}$, with $\boldsymbol{\iota}$ being a $r \times 1$ vector of ones, it follows by construction that $\mathbf{P}$ has one eigenvalue equal to unity. Therefore, the remaining eigenvalues are assumed to lie inside the unit circle and the Markov process is said to be ergodic. The ergodic probabilities, $\Pr(s_t = j) = \pi_j$, can be collected into a $r \times 1$ vector $\boldsymbol{\pi}$ satisfying $\mathbf{P}\boldsymbol{\pi} = \boldsymbol{\pi}$ and $\boldsymbol{\iota}'\boldsymbol{\pi} = 1$. These ergodic probabilities can be interpreted as the unconditional probabilities for each of the r different states, that is π_j is the probability of being in state j in the next period, independent of the current state.

For example, for a two-state Markov chain

$$\boldsymbol{\pi} = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix} = \begin{pmatrix} (1 - p_{22})/(2 - p_{11} - p_{22}) \\ (1 - p_{11})/(2 - p_{11} - p_{22}) \end{pmatrix}$$

with $\pi_1 + \pi_2 = 1$.

3. A Markov-Switching Model for GDP

Consider the time series of the log of real output, $\{y_t\}_{t=1}^{T}$. Without loss of generality, y_t can be decomposed into a trend component τ_t and a residual x_t, commonly referred to as the "cyclical" component (see, for example, Clark, 1987 and Watson, 1986)[1]

$$y_t = \tau_t + x_t \tag{3.1}$$

As Canova (1998) notes there is a fundamental disagreement on the properties of the trend component and on its relationship with the cyclical component. The traditional approach was to specify the trend component as a deterministic polynomial function of time, $\tau_t = \beta(t)$, assumed to be independent of the cyclical component, cf. Blanchard (1981). This implies that shocks to output are only transitory. Thus, the cyclical component could easily be extracted from the original series using standard regression techniques. However, since the seminal paper by Nelson and Plosser (1982), who found evidence of a stochastic trend rather than a deterministic one in many key macroeconomic series suggesting a significant permanent component in these series, this approach has come under strong criticism.[2] Alternative definitions of the trend, different assumptions concerning the relationship between the trend and the cycle and methods for estimating the two components have therefore been proposed by Beveridge and Nelson (1981) and

[1] Typically, business cycle fluctuations are identified as the deviations of the process from trend. See, however, Burnside (1998) for a critique.

[2] The existence of a unit root in real aggregate output is an ongoing debate in the literature. See, for example, Murray and Nelson (1998) and Diebold and Senhadji (1996).

Campbell and Mankiw (1987), who use ARIMA models where the autoregressive process contains a unit root, Watson (1986) and Clark (1987), who use a linear unobserved component approach where the time series is a sum of a random walk and a stationary ARMA process, and King *et al.* (1991), who use a cointegration approach to identify common stochastic trends in an equilibrium business cycle model.[3]

In all the above studies, it is assumed that the growth rate of output follows a linear stationary process, so optimal forecasts of output growth are linear functions of their lagged values. An alternative description of the trend component in real output has been proposed by Hamilton (1989). As in the above papers the trend component is considered stochastic rather than deterministic and allows for permanent effects of shock to output. However, instead of following a linear stationary process, output growth is assumed to follow a non-linear stationary process. In particular, Hamilton (1989) assumes that the trend component can be written as

$$\tau_t = \tau_{t-1} + \mu_{s_t} \tag{3.2}$$

where μ_{s_t} is the growth rate of the trend component, which is assumed to follow a first-order, two-state ergodic Markov process

$$\mu_{s_t} = (1 - s_t)\mu_0 + s_t\mu_1 \tag{3.3}$$

with the unobserved state variable s_t taking on the values zero and one. The transitory probabilities are given as[4]

$$\Pr(s_t = 0\,|s_{t-1} = 0) = p; \qquad \Pr(s_t = 1\,|s_{t-1} = 1) = q \tag{3.4}$$

The stochastic trend will have an expected slope of $\pi\mu_0 + (1 - \pi)\mu_1$, where $\pi = \Pr(s_t = 0)$, and may either be viewed as a time trend with a Markov-switching growth rate or, alternatively, as a random walk component with a discrete shock

[3] See Guðmundsson *et al.* (2000) for an application of this approach to Icelandic data.

[4] The two states can be interpreted (but do not have to be) as representing "recessions" and "expansions". p is therefore the probability of continuing in a recession, while q is the probability of remaining in an expansion. $1 - p$ is the probability of switching from a recession to an expansion, with $1-q$ representing the switch in the opposite direction. In the specification in (3.4), the transition probabilities are constant and do not depend on the duration or strength of the business cycle. This is done for the sake of parsimonity. Diebold and Rudebusch (1990) argue that this is not a bad representation of historical US data. The results in Filardo and Gordon (1998), however, suggest that the probability of transition out of recession is increasing in the duration of the recession.

given by μ_{s_t}.[5] This specification of the trend in (3.2) is fundamentally different from the trend specification used in the papers referred to above in that the trend need only change in response to occasional, discrete events rather than changing every period. Furthermore, when added to a linear cyclical component, x_t, the process for y_t becomes non-linear for which an ARIMA representation exists, but does not generate optimal forecasts of the future value of the series, see Hamilton (1989).

In what follows, the cyclical component is assumed to evolve independently of the trend component and have the following autoregressive representation

$$\varphi(L)x_t = \varepsilon_t \tag{3.5}$$

where $L^n z_t \equiv z_{t-n}$ is the lag operator and $\varepsilon_t \sim \mathsf{NID}(0, \sigma_\varepsilon^2)$. It is further assumed that $\varphi(L) = 0$ has one root on the unit circle, but all the others outside. Therefore $\varphi(L)$ can be written as $(1-L)\phi(L)$, where $\phi(L) = 0$ has all roots outside the unit circle.[6]

Therefore, by taken first difference of (3.1), multiplying with $\phi(L) = 1 - \sum_{i=1}^{k} \phi^{(i)} L^i$ and rearranging, the following model is obtained

$$(g_t - \mu_{s_t}) = \sum_{i=1}^{k} \phi^{(i)}(g_{t-i} - \mu_{s_{t-i}}) + \varepsilon_t \tag{3.6}$$

where $g_t \equiv (1-L)y_t$ is the growth rate of output. Equation (3.6) is the Hamilton (1989) autoregressive model with a Markov-switching mean, sometimes called a centered Markov-switching model.

In the Markov-switching model in (3.6), the difference between the two states is completely captured by differences in the mean of the process. This implies that the conditional distribution of a realization depends upon the previous k values of the Markov process. This is equivalent to a r^{k+1}-state Markov process and makes the model quite difficult to estimate. More recently, Hamilton (1990) suggests an alternative Markov-switching model, where the intercept of the regression is

[5] The Markov trend can be solved backwards to give $\tau_t = \tau_0 + \mu_0 t + (\mu_1 - \mu_0)\sum_{j=0}^{t} s_{t-j}$, where τ_0 is the initial value of the Markov trend. Hence, the Markov trend consists of a deterministic trend with slope μ_0 and a stochastic trend with an impact value of $\mu_1 - \mu_0$.

[6] Since both τ_t and x_t have unit roots they are not individually identified. Note also that this specification of the trend and cyclical components implies that all the shocks to output are permanent and the business cycle asymmetry shows up in the growth rate of output. In Lam (1990) the cyclical component is stationary, thus allowing the economy to be subject to both permanent and transitory shocks. In this case the business cycle asymmetry shows up in the permanent component of output. The Kalman filter estimates of an unobserved component model from Elíasson (1998) suggest that the cyclical component indeed contains a unit root supporting the specification used here.

allowed to switch between regimes instead of the mean of the process. This model can be written as

$$g_t = \alpha_{s_t} + \sum_{i=1}^{k} \phi^{(i)} g_{t-i} + \varepsilon_t \tag{3.7}$$

Although (3.7) might seem as a minor modification of (3.6), the two models actually have quite different dynamic behaviour. The centered Markov-switching model in (3.6) is non-linear in some parameters even after conditioning on current and past states. In contrast the model discussed in Hamilton (1990) is linear after conditioning on the current state and is therefore much easier to estimate. Furthermore, the regime dependence in the centered model is asymmetric in that some parameters depend on the current state, while others depend on both current and past states. In the model in Hamilton (1990), regime dependence is treated symmetrically since all coefficients depend only on the current regime. As discussed in Warne (1996) these models are non-nested and it is not obvious how one should choose between them.

In the following the main attention will be on (3.7), denoted as MS(2)-AR(k), although both models will be discussed. The Markov-switching model is also compared to a conventional linear AR model of the form

$$g_t = \alpha + \sum_{i=1}^{k} \phi^{(i)} g_{t-i} + \eta_t \tag{3.8}$$

where $\eta_t \sim \mathsf{IID}(0, \sigma_\eta^2)$. This seems a natural benchmark to compare with the Markov model as (practically) all covariance stationary processes have an autoregressive representation, which can be written as (3.8) where the residual η_t is white noise.

Note that testing the single regime model in (3.8) against the switching regime model in (3.7), is problematic and cannot be done using the standard likelihood ratio (LR) test framework. The reason is that under the null hypothesis of (3.8), the transition probabilities in $\mathbf{P}$ are unidentified. This makes the asymptotic information matrix singular under the null, failing the standard regularity conditions for constructing an asymptotic valid test of the null hypothesis. However, Hansen (1992b, 1996) has suggested a testing approach that delivers valid inference when unidentified nuisance parameters are present under the null.[7] An alternative procedure recommended by Hamilton (1996) is to test for possible misspecification of each of the models of interest. This might involve testing for serial correlation, autoregressive conditional heteroscedasticity (ARCH) and parameter instability.

[7] Alternative solutions to this identification problem have been suggested by Davies (1977), Gallant (1977) and Bianchi and Zoega (1998).

Yet another approach for comparing the Markov-switching model with a linear AR model is to compare their forecasting ability. All these approaches are adopted here.

4. Estimation Results

4.1. The data

The data comprises of annual Icelandic GDP data from 1945 to 1998 at 1990 prices.[8] The variable used is 100 times the log difference of real GDP. The original data is obtained from *Sögulegt yfirlit hagtalna* 1995, published by the National Economic Institute, except for the period 1995 to 1998, where the data is obtained from *Hagtölur mánaðarins*, July 1998, published by the Central Bank of Iceland. Figure 1 plots the data, including shaded areas for periods commonly associated with recessions.

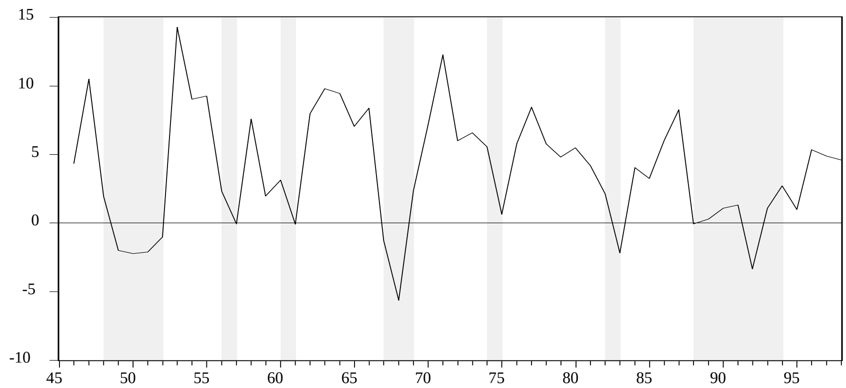

Figure 1. Real GDP growth 1946-1998

Table 1 gives an overview of these periods and the average growth in each period. According to these, there have been seven recessions and eight expansions over the period analysed here, implying a business cycle with a 7 year periodicity. The average duration of recessions according to Table 1 is about 3.3 years and

[8] Since quarterly national accounts are not available for Iceland, annual data was chosen rather than trying to construct quarterly data from the annual series. The advantage of using annual data is that using some constructed quarterly data may produce spurious business cycles. The disadvantage is, however, that annual data may obscure a substantial amount of the cyclical variation in economic activity. Thus, business cycles will tend to have longer duration on average in annual data than in quarterly data. Elíasson (1998) uses constructed quarterly data in an unobserved component model and his results suggest that using annual data leads to a relatively small loss of information, although the duration of expansions and contractions might be overstated.

Table 1. Chronology of business cycle downturns

Period	Duration	Average growth	Description of main sources
1948-52	5	-1.08%	Deteriorating terms of trade; the Korean war
1956-57	2	1.14%	Rising oil prices; the Suez conflict
1960-61	2	1.53%	Falling fish prices
1967-69	3	-1.52%	Falling fish catches and fish prices
1974-75	2	3.10%	Deteriorating terms of trade; first OPEC crisis
1982-83	2	-0.02%	Falling fish catches
1988-94	7	0.43%	Falling fish catches and fish prices

Sources: Magnússon and Einarsson (1985) and Snævarr (1993).

the average duration of an expansion is 3.8 years, which may suggest asymmetric business cycles, with expansions lasting longer than contractions. Looking at the average growth rate in these recessionary periods it is clear that the most severe recessions occurred in 1967-69, 1948-52, 1982-83 and 1988-94, although the last one is probably more severe in terms of how long it lasted rather than in terms of decline in output (although output declined by more than 3% in 1992). This recession can in fact be viewed as two joint recessions, the former due to falling fish prices and catches from 1988 to 1991 and the second from 1992 to 1994 due to falling fish catches. The least severe recession is the OPEC crisis in 1974-75, with output growth only slightly below the periods average (see Table 2).[9]

According to Table 1, the dominant source of shocks to the Icelandic economy are foreign shocks that affect the demand for Icelandic export products and the terms of trade. This is confirmed in a structural vector autoregressive analysis by Guðmundsson *et al.* (2000). Only three shocks can be described as domestic real shocks: 1967-1968, 1982-1983 and 1988. These shocks have, however, been the most severe ones. See Guðmundsson and Einarsson (1987) for a more detailed discussion.

Table 2 reports some descriptive statistics for real GDP growth for the period 1946-1998 and for two subperiods. Over the whole period the unconditional mean growth is 3.9% with a standard deviation of over 4%. It is therefore evident that there has been a substantial fluctuation in GDP growth over the whole period. In fact, Guðmundsson *et al.* (2000) show that output has fluctuated more in Iceland than in other industrial countries. They also find that the Icelandic business cycle is more or less independent of the business cycle in other industrial countries.

Table 2 also indicates that the average growth rate has fallen over the period and fluctuates less than in the earlier period, as in Guðmundsson *et al.* (2000).

[9] The second OPEC crisis in 1979-80, where output rose by more than 5%, does not register as a business cycle downturn in Iceland as it coincided with a substantial increase in fish catches.

Table 2. Descriptive statistics

	1946-1998	1946-1972	1973-1998
Mean	3.93%	4.46%	3.37%
Standard deviation	4.26%	5.20%	3.01%
Minimum	-5.63%	-5.63%	-3.35%
Maximum	14.29%	14.29%	8.46%
Observations	53	27	26
Unit root test	-2.184	-3.599	-1.906

Note: The unit root test is the Dickey and Fuller (1979) test for a single unit root in output. The regressions include a constant, a trend and lagged growth rates (1 lag for the period 1946-1998 and 1946-1972 but none for the period 1973-1998). The alternative hypothesis is therefore that output is trend stationary. The critical values are -3.497 (5%) and -4.142 (1%) and are obtained from MacKinnon (1991).

Finally, a unit root in the autoregressive representation of the log of real output cannot be rejected. This implies that output has a stochastic trend, thus supporting the findings in Nelson and Plosser (1982). Although the statistical power of the unit root test is notoriously low, this at least implies that shocks to output are very persistent.

4.2. A simple autoregressive model

Estimating a simple AR(k) model for the period 1946 to 1998 with $k = 2$ gives the following (standard errors in parenthesis)[10]

$$\widehat{g}_t = \underset{(0.811)}{2.758} + \underset{(0.138)}{0.453}\, g_{t-1} - \underset{(0.138)}{0.189}\, g_{t-2} \tag{4.1}$$

$\log L = -140.4$, $\sigma_\eta = 3.91$, $F_{ar1} = 0.01$, $F_{arch1} = 0.36$, $VS = 0.53^*$, $JS = 0.80$

There is no evidence of serial correlation or ARCH effects in the residual. There is, however, evidence of parameter instability using Hansen's (1992a) tests for in-sample parameter stability. This suggests that the linear AR(2) model is misspecified and that a non-linear model, such as the Markov-switching model, may be a more appropriate representation of the data.

The estimated AR(2) model has complex eigenvalues given by the conjugate pair $0.23 \pm 0.37i$ with modulus equal to 0.43. This implies a dynamic multiplier

[10] The AIC information criterion suggests a two lag model rather than a one lag model. The AR(2) model was also chosen to correspond to the lag length of the Markov-switching model below.

that follows a pattern of damped oscillation with frequency equal to 1.02. The cycles associated with the dynamic multipliers have a period of 6.1. Thus, the AR(2) model in (4.1) implies (for a given shock) a cycle with a 6 year periodicity.

The estimates indicate that innovations to output are very persistent, as suggested by the unit root tests above. A 1% unanticipated increase in output will increase the optimal forecast of y_{t+h} for $h \to \infty$ by 1.36%. If output was trend stationary the implied revision of the output forecast should be zero. These results are consistent with the results obtained by Campbell and Mankiw (1987), and imply a dominant permanent component in output.[11]

4.3. The Markov-switching AR model

The estimation of the Markov-switching AR model entails a numerical maximization of the conditional likelihood function to obtain estimates of the autoregression and transition probabilities. The iterative procedure used here is based on the EM algorithm of Dempster *et al.* (1977), as first suggested by Hamilton (1990). This algorithm is designed for a general class of models where the observed time series depends on some unobservable stochastic variable − the regime variables s_t in the Markov-switching model.

Table 3 reports the maximum likelihood estimates of the parameters of MS(2)-AR(k) with $k = 2$. The model was also estimated for $k = 1$ and on the basis of the AIC information criteria a two lag model was chosen. As expected, the estimated growth rate for the two states can be associated with slow and fast growth states for the Icelandic economy. The "recessionary" regime has a mean growth rate of 0.6% ($\mu_0 = \alpha_0/\phi(1)$), but this growth rate is not significantly different from zero. The recessionary state can therefore be said to correspond to a state of stagnation. The "expansionary" state, on the other hand, has a mean growth rate of 6.8% ($\mu_1 = \alpha_1/\phi(1)$).

The autoregressive parameters turn out to be quite small and in fact are not significantly different from zero. This suggests that the dynamics of Icelandic output growth are better captured by recurrent shifts between the recessionary and expansionary regimes than by the dynamics implied by the autoregressive representation. The quantitative importance of the Markov process can also be seen by calculating the fraction of the total variation of output growth explained by the Markov process. According to the estimates in Table 3 the variance of the Markov process is 9.7 which amounts to about 54% of the total variation of growth in real GDP.[12]

---

[11] This result is found robust to models allowing richer dynamics. For example, an ARMA(1,2) suggests a 1.63% increase in the optimal forecast.

[12] The variance of the Markov process is calculated as $\mathsf{var}(\mu_{s_t}) = (\mu_1 - \mu_0)^2\pi(1-\pi)$.

Table 3. ML estimates of MS(2)-AR(2)

Parameter	Estimate	Standard error
α_0	0.557	0.844
α_1	6.637	0.913
$\phi^{(1)}$	0.215	0.160
$\phi^{(2)}$	-0.187	0.133
σ_ε	2.495	0.287
p	0.704	0.136
q	0.744	0.123
π	0.464	0.137
d_0	3.4	1.56
d_1	3.9	1.87
$\log L$	-138.4	

Note: d_i is the expected duration of regime i. The standard errors of d_i are calculated using a Taylor approximation. Thus, for a recession, $\mathsf{var}(d_0) = f'(p)^2\mathsf{var}(p)$, where $f(p) = 1/(1-p)$.

The estimates of the transition probabilities indicate that the probability of remaining in a recession next year is about 70%, whereas the probability of remaining in a expansion is about 74%. Thus, both regimes are found quite persistent. Another way to see this is to calculate the expected duration of a typical recession or expansion, which is easily obtained from the maximum likelihood estimates. The expected duration of a recession, conditional on being in that state, is (see Hamilton, 1989)

$$d_0 = \sum_{m=1}^{\infty} mp^{m-1}(1-p) = \frac{1}{(1-p)}$$

Thus, the expected duration of a recession is 3.4 years, whereas the expected duration of an expansion is 3.9 years. This can be compared to the average 3.3 and 3.8 years duration from Table 1. This might imply that the business cycle is asymmetric, with expansions having a longer expected duration than recessions, although the estimated standard errors indicate that the implied durations are not significantly different from each other. The maximum likelihood estimates therefore suggest a business cycle with a 7 year periodicity, just as in Table 1.

The table also reports the ergodic probabilities. The estimates indicate that the probability of being in the recessionary regime, $\pi = \Pr(s_t = 0)$, is about 46%, independent of what state the economy was in the previous period.

An important issue considered by Hamilton (1989) is the change in the infinite horizon forecast of real output with respect to changes in the Markov state and the autoregressive process. The permanent effect attributed to the autoregressive component can be calculated as

$$\lim_{h\to\infty} \frac{\partial y_{t+h}}{\partial \varepsilon_t} = \phi(1)^{-1} = 1.03$$

Thus, a 1% unanticipated shock to output will lead to a 1% revision of output forecasts over a long horizon. This can be compared to the estimate of 1.36% from the linear AR(2) model.

The change in the infinite horizon forecast of real output attributed to the Markov process can be found by comparing forecasts at time $t+h$ as $h \to \infty$ when it is known that the economy is in a recession ($s_t = 0$) and when it is known that the economy is expanding ($s_t = 1$). This is given by (see Hamilton, 1989)

$$\lim_{h\to\infty} \left[\mathsf{E}(y_{t+h} \,|s_t = 1, \mathcal{I}_t) - \mathsf{E}(y_{t+h} \,|s_t = 0, \mathcal{I}_t) \right] = \frac{(\mu_1 - \mu_0)\lambda}{1-\lambda} = 5.07$$

where $\lambda \equiv -(1-p-q)$ and $\mathcal{I}_t = \{g_t, g_{t-1}, ...\}$ is the information set at time t. Thus, perfect knowledge that the economy has gone into a recession is associated with a 5% permanent drop in the infinite horizon forecast of the level of real output. Information about the state of the economy at time t therefore has a permanent effect on the level of real GDP growth.

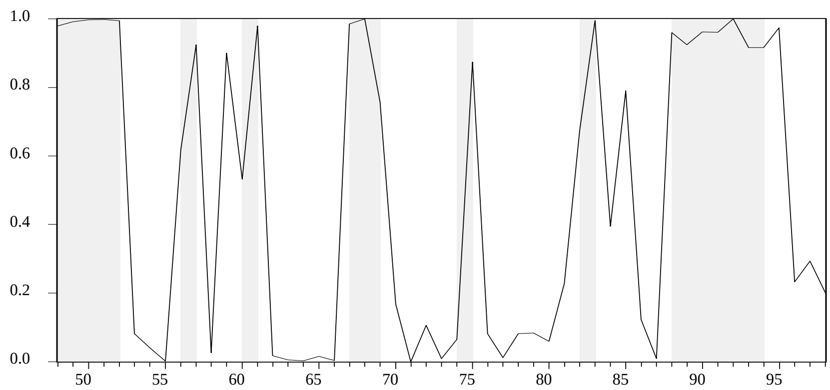

Figure 2. Smoothed probability of a recession, $\Pr(s_t = 0 \,|\mathcal{I}_T)$

Figure 2 plots the smoothed probabilities of being in the recessionary state, $\Pr(s_t = 0 \,|\mathcal{I}_T)$, where $\mathcal{I}_T$ is the full information set. Again, the shaded areas denote the recessionary periods from Table 1.[13] The figure shows that the algorithm

[13] It should be emphasized that the Markov-switching model does not try to "fit" the business cycle chronology from Table 1. Rather, the datings from Table 1 are used as a diagnostic tool to verify that the model's phase chronology is roughly consistent with what is generally considered as the business cycle.

Table 4. Business cycle datings from the MS(2)-AR(2) model

Period	Duration	Average growth
1948-52	5	-1.08%
1956-57	2	1.14%
1959-61	3	1.68%
1967-69	3	-1.52%
1975	1	0.64%
1982-83	2	-0.02%
1985	1	3.24%
1988-95	8	0.50%

matches the datings from Table 1 very well. The smoothed probabilities suggest that the recession in 1960-61 actually started a year earlier, the recession starting in 1974 only occurring in 1975, a small recession in 1985 and that the recession ending in 1994 ends a year later. The filtered probability, $\Pr(s_t = 0 \,|\mathcal{I}_t)$, gave practically the same results.[14]

It should be noted that the algorithm usually gives very strong signals for identifying when the economy is in a recession. Only on four occasions was the smoothed probability between 30 and 70%: 1956, 1960, 1982 and 1984.[15] This suggests that the filter captures the main part of the underlying pattern in the data of dichotomous shifts between the slow and fast growth regimes.

Table 4 summarizes the business cycle datings generated by the maximum likelihood estimates. The datings are based on the metric that given the data set $\mathcal{I}_T$, the economy is more likely to be in a recession, that is periods where $\Pr(s_t = 0 \,|\mathcal{I}_T) > 0.5$ (see Hamilton, 1989). As discussed above, the datings largely correspond to the datings in Table 1.

Finally, Figure 3 plots output growth along with the estimated mean growth from the Markov-switching model, μ_{s_t}. The figure also plots the Markov recessionary dates from Table 4 in shaded areas. It shows how well the shifts between the recessionary and expansionary regimes capture the business cycle dynamics of real

[14]The maximum likelihood estimates gives three different types of probability measures of a recession in period t, which differ in the information they condition on. The smooth probability, $\Pr(s_t = 0 \,|\mathcal{I}_T)$, conditions on the full sample information and is useful for business cycle dating. The filter probability, $\Pr(s_t = 0 \,|\mathcal{I}_t)$, conditions on the contemporaneous information set and is useful for evaluating the strength of the contemporaneous signal of a recession. Finally, the *ex ante* probability, $\Pr(s_t = 0 \,|\mathcal{I}_{t-1})$, conditions on previous information and is useful for evaluating the ability of the model to forecast business cycle turning points in real time exercises. The correlation between the filter and smooth probabilities is 0.99 while the correlation between these and the *ex ante* probability are 0.46 and 0.48, respectively.

[15]Thus, the algorithm might be suggesting that the recession in 1982-83 actually spans the period 1982-85.

output growth, supporting the claim that real output growth is well characterized by recurrent shifts between the slow and fast growth regimes.

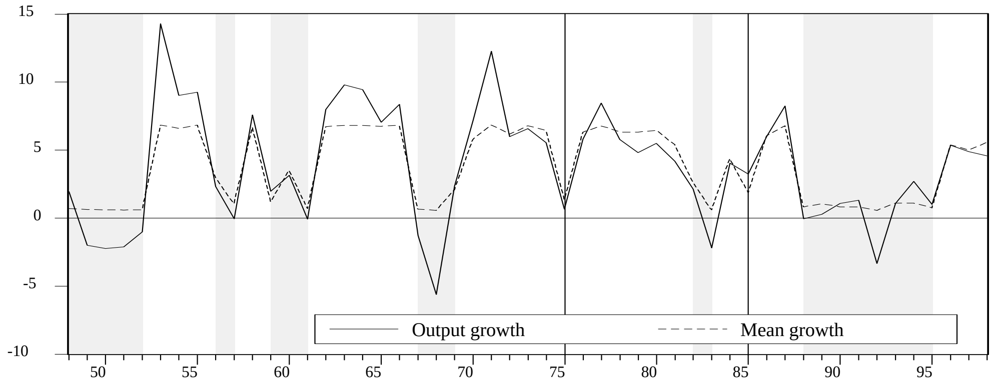

Figure 3. Output and mean growth 1948-1998

4.3.1. Diagnostic tests

Table 5 reports diagnostic tests for the MS(2)-AR(2) model. First, the table reports stability tests suggested by Nyblom (1989). These tests use the conditional scores of the likelihood function where the alternative hypothesis is that the parameters follow a martingale sequence. The first test reports a constancy test for the transition probabilities, i.e. the free parameters in $\mathbf{P}$. The number of restrictions for this test is given by $r(r-1) = 2$. The second test tests whether the intercept vector $(\alpha_0, \alpha_1)'$ is constant over time. The number of restrictions tested is given by $r = 2$. The third test tests whether the estimated variance is constant over time. The number of restrictions tested is 1. Finally, a constancy test for the model in a given state is reported. The number of restrictions tested in this case is given by $k + 2 = 4$.

These tests have non-standard limiting distributions, but critical values are tabulated in Nyblom (1989). As it is not clear whether the asymptotic theory used by Nyblom (1989) applies to Markov-switching models, these tests should only be interpreted as being indicative. The results in Table 5 show no evidence of instability of the estimated parameters, although there is some indication of instability of the intercept vector at the 95% critical level but the null of constant intercept vector is not rejected at the 99% level. This could suggest some instability in the intercept vector, not captured by the Markov-switching intercepts. One potential explanation, looking at Figure 3, is a decline in the high growth rate μ_2 in the period since the early 1970s, probably due to a productivity slowdown in the latter half of the sample.[16]

[16] A three-regime Markov-switching model with $\mu_{s_t} = (0.3, 5.5, 8.8)$ supports this claim. In

Table 5. Diagnostic tests for MS(2)-AR(2)

	Stability tests		
Transition probabilities	0.148	1.074[a]	0.748[b]
Intercept vector $(\alpha_0, \alpha_1)'$	0.836	1.074[a]	0.748[b]
Variance	0.167	0.743[a]	0.461[b]
Equation in state 0	0.878	1.623[a]	1.237[b]
Equation in state 1	0.845	1.623[a]	1.237[b]
	Uncorrelated Markov chain		
Wald test	4.548		0.033[c]
F-test	4.013		0.051[c]
	Misspecification tests		
Autocorrelation	0.620		0.651[c]
ARCH	1.237		0.272[c]
Higher-order Markov chain	0.160		0.958[c]

Note: The stability tests are suggested by Nyblom (1989), using the conditional scores of the likelihood function, testing for stability of the transition probabilities, the intercept, the residual variance and the equation in a given state. The test for uncorrelated Markov chain reports a Wald test and a F-version of the test. The misspecification tests are F-approximations of the conditional scores tests discussed in Hamilton (1996). Under the null hypothesis they are distributed as $F(m, 46)$, where m equals 4, 1 and 4 respectively. a, b 99%, 95% critical values from Nyblom (1989). c p-values.

The second part of Table 5 tests whether the Markov chain is uncorrelated, i.e. the hypothesis that $\Pr(s_t = j \,|s_{t-1} = i\,) = \Pr(s_t = j)$. This tests whether the transition probabilities equal the (long-run) ergodic probabilities, i.e. whether the probability of staying in a particular state is the same as the probability of returning to it from all other states. An alternative way to describe these restriction is to note that this implies that $p + q = 1$, i.e. the regime probabilities follow a Bernoulli process instead of a Markov process. The table reports a Wald test for this hypothesis and a F-approximation of the Wald test. The number of restrictions is given by $(r-1)^2 = 1$. The Wald test rejects at the 95% critical level and the F-test is close to rejecting.

Finally, Table 5 reports three misspecification tests for the Markov-switching model suggested by Hamilton (1996). These tests are based on the conditional

this model, the contractionary phase remains largelly unchanged from the two-regime model, whereas the expansionary phase is split into two phases: one high growth and one extreme growth regime, with the extreme growth regime occurring in 1953-55, 1958 and 1962-66. The high growth regime coincides with the expansionary regime in all the latter expansionary periods from the MS(2)-AR(2) model.

scores of the likelihood function and are F-approximations proposed by Newey (1985), Tauchen (1985) and White (1987) as suggested by Hamilton's (1996) Monte Carlo results. The first test tests for first-order autocorrelation in the residual and imposes $r^2 = 4$ restrictions on the model. The second tests for first-order ARCH effects in the residual and imposes 1 restriction on the model. Finally, the Markov test tests the specification of the first-order Markov chain, that is whether $\Pr(s_t = j \,|s_{t-1} = j) = \Pr(s_t = j \,|s_{t-1} = j, s_{t-2} = j)$. This test imposes $2r = 4$ restrictions on the model. In no cases do the test statistics reject the null hypothesis of a correctly specified model.

4.3.2. Alternative Markov-switching models

Table 6 reports estimation of alternative specifications of the Markov-switching model. Model 1 is the centered Markov-switching model in (3.6). As seen in the table this alternative specification of the Markov-switching model gives almost identical results to the previous model. The reason is that the autoregressive parameters are quite small and almost sum to zero, implying that $\alpha_{s_t} \approx \mu_{s_t}$. The centered model, however, suggests that the expected duration of an expansionary phase is over 5 years, which might seem implausibly long.[17] It would be possible to use formal non-nested LR tests, such as in Vuong (1989), to try to distinguish between the two models, but as evident from the log-likelihood such tests would almost surely be non-significant.

The Markov-switching model in Table 3 implies that the regime process s_t is positively serially correlated, i.e. an expansionary state is likely to be followed by another expansionary state, and vice versa. Model 2 in Table 6, however, restricts the Markov-process to be serially uncorrelated so that the probability of being in a given regime is independent of the previous state. From Table 4 these restrictions seem to be rejected but only marginally. This model is also of interest as this is the simple switching model suggested by Hansen (1992b). The parameter estimates are very similar to the estimates in Table 3, except that the regimes are less persistent. Comparing the log-likelihood with the one obtained in Table 3 suggests that the Markov-switching model in Table 3 is somewhat superior in describing the data to the simple switching model. A LR test for the restriction $p + q = 1$ gives $\chi^2(1) = 3.6$ with a p-value of 0.058, rejecting the simple switching model at the 6% critical level.

Further generalizations of the Markov-switching model are also possible. Table 6 reports a Markov-switching model where the intercept and the autoregressive

[17] The main reason for this result is that the centered Markov model assigns relatively low probability to the recession in 1975, about 40%. Thus, according to this model there is quite a high probability of an expansion prevailing for the period 1970 to 1982, generating this high expected duration of the expansionary state.

Table 6. Alternative Markov-switching models

	Model 1		*Model 2*		*Model 3*		*Model 4*	
Parameter	Est.	Std. err.	Est.	Std. err.	Est.	Std. err.	Est.	Std. err.
μ_0 or α_0	0.171	1.084	0.524	1.222	0.595	0.920	0.824	1.294
μ_1 or α_1	6.532	0.781	5.882	1.601	6.934	0.971	5.787	2.055
$\phi_0^{(1)}$	0.377	0.233	0.443	0.160	0.270	0.207	0.623	0.149
$\phi_0^{(2)}$	-0.250	0.185	-0.247	0.172	-0.215	0.175	-0.337	0.150
$\phi_1^{(1)}$	–	–	–	–	0.160	0.204	–	–
$\phi_1^{(2)}$	–	–	–	–	-0.165	0.177	–	–
$\sigma_{\varepsilon,0}$	2.565	0.283	2.711	0.342	2.473	0.291	2.854	0.358
$\sigma_{\varepsilon,1}$	–	–	–	–	–	–	3.338	0.379
p	0.711	0.145	0.525	0.179	0.707	0.130	0.403	0.452
q	0.807	0.101	0.475	0.179	0.730	0.136	0.000	0.390
$\log L$	-138.5		-140.2		-138.2		-139.0	

parameters are allowed to switch between states (Model 3). This could capture the notion that shocks to output are more persistent in recessions than in expansions. There is, however, very little improvement in terms of the log-likelihood (a LR test gives $\chi^2(2) = 0.4$ ($\mathsf{p} = 0.82$)) and the estimated autoregressive parameters are very similar across regimes. Other parameters are unchanged.

Finally, Table 6 reports a Markov switching model that allows the intercept and variance of the innovation process ε_t to switch between states (Model 4), capturing the notion that recessions coincide with greater uncertainty and thus greater fluctuations in output. This model actually leads to a deterioration of the log-likelihood and the standard errors are larger. There were also some problems with the algorithm encountered, with the final estimates somewhat sensitive to the initial values chosen, with numerical underflow occurring for many starting values.[18] Comparing the estimated variances between states also suggests that the difference in the variances between states is hardly significant, but this should be interpreted with care given the numerical problems encountered with this model.

Overall, these results, and the misspecification tests above, suggest that the MS(2)-AR(2) model is a reasonably good approximation of the underlying data generating process for real output in the period analysed here.

[18] This was the only model specification which was found sensitive to the choice of initial values.

Table 7. Standardized LR statistics for MS(2)-AR(2) model

	Bandwidth parameter				
LR test	$M = 0$	$M = 1$	$M = 2$	$M = 3$	$M = 4$
1.648	$\mathsf{p} = 0.755$	$\mathsf{p} = 0.745$	$\mathsf{p} = 0.681$	$\mathsf{p} = 0.672$	$\mathsf{p} = 0.656$

Note: See Hansen (1996) for details on the bandwidth parameter. The results were obtained using Bruce Hansen's Gauss code `markovp.prg` with the "Grid 3" option of Hansen (1996) and using 1,000 Monte Carlo replications.

4.3.3. Testing the Markov-switching model against a single-regime model

This sections compares the two-regime Markov-switching model against the single-regime, autoregressive model in (4.1). The testing procedure suggested by Hansen (1992b, 1996) is used. Hansen suggested a non-standard procedure for testing the null of a single-regime model against a two-regime Markov-switching model that avoids the identification problem which arises using the standard LR test framework, as discussed above. The procedure suggested by Hansen delivers a bound on the asymptotic distribution of the standardized LR test, and may therefore be quite conservative, i.e. tending to be under-sized in practice and of low power.[19] The tests should therefore only be considered as suggestive for the sample size used here.

Table 7 reports the p-values of the standardized LR test of a linear AR(2) model against the MS(2)-AR(2). The results suggest that the null hypothesis of a single regime AR(2) model cannot be rejected and are not found sensitive to the choice of bandwidth parameter. Thus, despite the evidence of non-stability of the AR(2) model, the data is not able to discriminate between the AR(2) and MS(2)-AR(2) models on the basis of the standardized LR test statistic. This is in line with results found on other data sets, cf. Hansen (1992b, 1996) and Clements and Krolzig (1998) for US data.

However, the inability of the linear AR(2) model to characterize some important aspects of the business cycle, such as the observed asymmetry between expansions and contractions in real GDP, suggests that the Markov-switching model may still be important. For forecasting purposes, the Markov-switching model may be better equipped for discerning business cycle turning points, since the linear AR model does not take into account the forecastable asymmetries in the business cycle. It is therefore important to evaluate the models on the basis

[19] See, for example Clements and Krolzig (1998). Hansen, in fact, argues that this is not the case, based on Monte Carlo calculations of the finite sample size and power of the standardized LR test.

of their forecasting performance.

5. Forecasting

This final section compares the forecasting ability of the Markov-switching model with the simple, linear AR alternative. The forecasting comparison is based on mean-square error (MSE) and mean-absolute error (MAE) loss functions, using both in-sample and out-of-sample forecasts. It has been argued that forecast performance comparison based on measures such as MSE of first moment forecasts may fail to discriminate between linear and non-linear models even if the non-linear model is the true data generating process, see e.g. Clements and Smith (1998). The results in Lam (1990), Hamilton and Perez-Quiros (1996), Clements and Krolzig (1998) and Clements and Smith (1998) show that, in general, no clear distinction can be made between linear autoregressive models and non-linear models, such as the Markov-switching model and the self-exciting threshold autoregressive (SETAR) model of Tong (1983), when the forecast comparison is based on MSE loss functions.

An alternative procedure for comparing these models could therefore be based on qualitative, direction-of-change statistics.[20] Two such measures of the models' ability to forecast the state of the business cycle are used. The first is to compare the Markov-switching model to a naive forecast that predicts a constant probability of a recession every year, as suggested by Hamilton and Perez-Quiros (1996) and Hamilton and Lin (1996). The second is to use the confusion rate suggested by Swanson and White (1995), which counts the number years the models predict the direction of output changes in the wrong direction.

5.1. Forecasts using Markov-switching models

In MS(2)-AR(k) models the conditional expectations, $\mathsf{E}(g_{t+h}\,|\mathcal{I}_t)$, is given by

$$\widehat{g}_{t+h|t} = \widehat{\alpha}_{t+h|t} + \sum_{i=1}^{k} \phi^{(i)} \widehat{g}_{t+h-i|t} \tag{5.1}$$

with initial values $\widehat{g}_{t+h|t} = g_{t+h}$ for $h \leq 0$. The predicted mean is given as the weighted average of the means from the recessionary and expansionary regimes, where the weights are the probabilities of being in a given regime conditional on the sample information set $\mathcal{I}_t$

[20] Clements and Smith (1998) use complete forecast densities to compare linear and non-linear models.

$$\widehat{\alpha}_{t+h|t} = \sum_{j=0}^{1} \alpha_j \Pr(s_{t+h} = j \,|\mathcal{I}_t) \tag{5.2}$$

The predicted regime probabilities are given by

$$\Pr(s_{t+h} = j \,|\mathcal{I}_t) = \sum_{i=0}^{1} \Pr(s_{t+h} = j \,|s_t = i) \Pr(s_t = i \,|\mathcal{I}_t)$$

with the transition probabilities (see Hamilton, 1994)

$$\begin{aligned}
\Pr(s_{t+h} &= 0 \,|s_t = 0) = \frac{(1-q) + \lambda^h(1-p)}{1-\lambda} \\
\Pr(s_{t+h} &= 1 \,|s_t = 1) = \frac{(1-p) + \lambda^h(1-q)}{1-\lambda} \\
\Pr(s_{t+h} &= 0 \,|s_t = 1) = \frac{(1-q) - \lambda^h(1-q)}{1-\lambda} \\
\Pr(s_{t+h} &= 1 \,|s_t = 0) = \frac{(1-p) - \lambda^h(1-p)}{1-\lambda}
\end{aligned}$$

where $\lambda \equiv -(1-p-q) < 1$, as before.

Thus, the optimal predictor of the Markov-switching model is linear in the last k observations and the last regime inference, but there exists no purely linear representation of the optimal predictor in the information set, see Clements and Krolzig (1998). Note, however, that the optimal forecasting rule becomes linear in the limit as the regimes become completely unpredictable. Thus, the transition probabilities converge to the unconditional ergodic probabilities

$$\lim_{h\to\infty} \Pr(s_{t+h} = 0 \,|s_t = 0) = \Pr(s_{t+h} = 0) = \pi$$

Hence, the long-run forecast for the Markov chain will be independent of the current state.

5.2. In-sample forecast comparison

In this case the models are estimated for the full sample and in-sample forecasts for the period 1950 to 1998 conducted. Three forecasting horizons are used: 1 year, 2 years and 5 years. Table 8 reports the root-mean-square error ($RMSE$) and mean-absolute error MAE for these three forecasting horizons.

From the table it is evident that the Markov-switching model has smaller $RMSE$s for short and long forecasting horizons but the linear model for the

Table 8. In-sample forecast comparison 1950-1998

	RMSE			*MAE*		
Models	$h = 1$	$h = 2$	$h = 5$	$h = 1$	$h = 2$	$h = 5$
Markov-switching model	3.68	3.94	4.11	2.89	3.24	3.34
Linear model	3.78	3.91	4.19	3.07	3.28	3.41

	Turning point predictions		
	Naive forecast	Markov model	Linear model
TP test	0.25	0.21	–
CR test	–	0.20	0.20

Note: The TP test is the turning point test suggested by Hamilton and Perez-Quiros (1996) and Hamilton and Lin (1996). The CR test is the confusion rate test suggested by Swanson and White (1995).

medium term. However, the Markov-switching model generates smaller MAEs for all horizons. However, given the relatively small difference between the forecast loss functions and the fact that the Markov-switching model uses three more parameters than the linear AR model, it becomes clear that the differences between the forecast loss functions are not statistically significant, consistent with the in-sample comparison using the likelihood principle above. These findings are consistent with the findings in papers such as Clements and Krolzig (1998), who find that linear AR models are a relatively robust forecasting device, even when the data are generated from non-linear models.

An alternative evaluation criteria for forecasting performance is to compare the models' ability to predict business cycle turning points, which is conceptually separate concern from minimizing the MSE or the MAE of a forecast, see Hamilton and Perez-Quiros (1996). Table 8 reports two measures of turning point prediction accuracy. The first test is the turning point test suggested by Hamilton and Perez-Quiros (1996) and Hamilton and Lin (1996), TP. They suggest using *ex post* dating of business cycle downturns, such as those from Table 1, to determine the turning points. This test compares turning point predictions from the Markov-switching model to a naive forecast which always predicts a constant probability of business cycle downturns. To calculate this naive forecast, a dummy variable D_t is constructed. D_t equals unity when the economy was in a recession according to Table 1 and zero otherwise. The average value of D_t is 0.45, implying that the economy is in a recession 45% of the time, according to the business cycle datings in Table 1. The naive forecast would therefore be that this reflects the constant probability that the economy will be in a recession next year, regardless of current economic conditions. This forecast can be evaluated on the basis of its average squared deviations from the *ex post* values from Table 1

$$TP = T^{-1}\sum_{t=1}^{T}(D_t - \widehat{D})^2$$

where $\widehat{D} = 0.45$. In contrast, the Markov-switching model provides an *ex ante* forecast of this magnitude in the form of a conditional probability that the unobserved variable s_t will take on the value zero conditional on $\mathcal{I}_{t-1}$, $\Pr(s_t = 0\,|\mathcal{I}_{t-1})$.

As seen from the table the value for TP for this forecast turns out to be 0.21, whereas the value for the naive forecast is 0.25. This amounts to $(= 1 - 0.21/0.25)$ 16% improvement over the naive forecast. It is also interesting to note that the ability to predict turning points increases substantially when the additional information from $\mathcal{I}_{t-1}$ to $\mathcal{I}_t$ is included. In that case the Markov-switching model gives a 56% improvement over the naive forecast. Although this is not a forecast strictly speaking, this gives an idea of how valuable the additional information from $t-1$ to t is for evaluating business cycle turning points.

An alternative measure of turning points prediction ability is the confusion rate (CR) measure suggested by Swanson and White (1985). This simply amounts to taking the ratio of periods in which the Markov-switching and linear models forecasted the direction of output growth in the wrong direction to the total number of forecast periods. As seen from the table, both models predict output changes in the wrong direction 20% of the time. Hence, the models are equally "confused".[21]

5.3. Out-of-sample forecast comparison

In-sample forecasting comparison can be misleading as an indication of out-of-sample forecasting ability, since information is utilized that would not be available for actual out-of-sample forecasts. It is therefore also important to compare the models out-of-sample forecasting ability.

Due to relatively few observations the out-of-sample forecasting comparison is only conducted for the period 1991 to 1998. Although this leaves only eigth observations for comparing the forecasts, and therefore makes the results less reliable, this should give some indications for the out-of-sample forecasting ability of the Markov-switching model.

Table 9 compares one-year ahead forecasts from the Markov-switching model to forecasts from the simple AR(2) model for the periods 1991 to 1998 and 1993 to 1998, based on a series of rolling regression windows. The table also includes

[21] Swanson and White (1985) calculate the probability that the confusion rates of different forecasts is statistically significant using the hypergeometric distribution. This is not done here since the confusion rates of the two forecasts evaluated here are practically the same.

Table 9. One-year ahead out-of-sample forecasts

	1991 − 1998			*1993 − 1998*		
Models	*Bias*	*RMSE*	*MAE*	*Bias*	*RMSE*	*MAE*
MS(2)-AR(2) model	0.97	2.67	1.88	-0.04	1.59	1.17
AR(2) model	1.23	2.94	2.10	0.12	1.71	1.28
NEI model	-1.41	2.56	2.02	-2.22	2.86	2.35
	Confusion rate			*Confusion rate*		
MS(2)-AR(2) model	0.13			0.00		
AR(2) model	0.13			0.00		
NEI model	0.25			0.33		

one-year ahead forecasts from the National Economic Institute (NEI) macroeconometric model, published annually in their macroeconomic outlook *Þjóðhagsáætlun*.

The Markov-switching model has the smallest bias of all the three models and smaller $RMSE$ and MAE than the linear AR model in all cases. Compared to the NEI model, the Markov-switching model does reasonably well. Its bias is smaller and it has a smaller loss based on a MAE loss function. Based on a MSE loss function, however, the NEI model does better. This is due to the failure of both autoregressive models to predict the large downturn in 1992. This is apparent when comparing the forecasts from the period 1993 to 1998, where the Markov-switching model is found to perform best on all accounts.

Analysing the ability of the Markov-switching model to predict out-of-sample turning points implies that the Markov-switching model gives a 20% improvement over the naive forecast, with $\widehat{D} = 0.50$ and $TP = 0.25$ for the naive forecast and $TP = 0.20$ for the Markov-switching model. Finally, the NEI forecasts have a larger confusion rate in both periods.

Overall, the in- and out-of-sample forecasting comparison suggests that although the Markov-switching model has marginally smaller forecast errors than the linear autoregressive model, there is little to choose between these two models on statistical grounds. Compared to the out-of-sample forecasts made by the National Economic Institute, however, the Markov-swithching model seems to offer some improvements, both in terms of mean absolute error and in terms of forecasting business cycle turning points. The Markov-switching model is also able to improve on turning point predictions made by a naive forecast that always predicts a constant probability of business cycle downturns.

6. Conclusions

This paper models the business cycle dynamics of real GDP in Iceland as stochastic, discrete shifts in the long-run trend of output, added to conventional autoregressive dynamics, using the Markov-switching model of Hamilton (1989). The estimation results suggest that the Icelandic business cycle is better captured by recurrent shifts between recessionary and expansionary regimes than by the dynamics implied by an autoregressive representation. The model generates business cycles that correspond quite well to conventional wisdom concerning the Icelandic business cycle, with expected duration of typical expansions and recessions equal to 3.9 and 3.4 years, respectively. The shifts between the expansionary and contractionary regimes account for more than half of the observed variation in output growth and the results indicate that perfect knowledge of a move from an expansion to a recession is associated with a 5% drop in the long-run forecast level of real GDP.

Comparing the Markov-switching model to a conventional linear, autoregressive model reveals that the Markov-switching model cannot be distinguished from the linear counterpart, based on the likelihood principle or in-sample forecasting ability. Furthermore, based on out-of-sample forecasting comparison, there seems little to choose between them, although the Markov-switching model has marginally smaller forecast errors. Compared to the forecasts made by the National Economic Institute, however, the Markov-switching model offers some improvements in terms of mean absolute error and in terms of predicting business cycle turning points. The Markov-switching model therefore seems to be a useful framework for describing Icelandic business cycle dynamics, and may be a valuable tool for forecasting aggregate output.

References

[1] Beveridge, S. and C. R. Nelson (1981), "A New Approach to Decomposition of Time Series in Permanent and Transitory Components with Particular Attention To Measurement of the 'Business Cycle"', *Journal of Monetary Economics*, **7**, 151-74.

[2] Bianchi, M. and G. Zoega (1998), "Unemployment Persistence: Does the Size of the Shock Matter?", *Journal of Applied Econometrics*, **13**, 283-304.

[3] Blanchard, O. J. (1981), "What is Left of the Multiplier-Accelerator?", *American Economic Review, Papers and Proceedings*, **71**, 150-4.

[4] Burnside, C. (1998), "Detrending and Business Cycle Facts: A Comment", *Journal of Monetary Economics*, **41**, 513-32.

[5] Campbell, J. Y. and N. G. Mankiw (1987), "Are Output Fluctuations Transitory?", *Quarterly Journal of Economics*, **102**, 857-80.

[6] Canova, F. (1998), "Detrending and Business Cycle Facts", *Journal of Monetary Economics*, **41**, 475-512.

[7] Clark, P. K. (1987), "The Cyclical Component of US Economic Activity", *Quarterly Journal of Economics*, **102**, 797-814.

[8] Clements, M. P. and H.-M. Krolzig (1998), "A Comparison of the Forecast Performance of Markov-Switching and Threshold Autoregressive Models of US GNP", *Econometrics Journal*, **1**, C47-75.

[9] Clements, M. P. and J. Smith (1998), "Evaluating the Forecast Densities of Linear and Non-Linear Models: Applications to Output Growth and Unemployment", University of Warwick, unpublished manuscript.

[10] Davies, R. B. (1977), "Hypothesis Testing When a Nuisance Parameter is Present Only under the Alternative", *Biometrika*, **64**, 247-54.

[11] Dempster, A. P., N. M. Laird and D. B. Rubin (1977), "Maximum Likelihood Estimation from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society B*, **39**, 1-38.

[12] Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of the American Statistical Association*, **74**, 427-31.

[13] Diebold, F. X. and G. Rudebusch (1990), "A Non-Parametric Investigation of Duration Dependence in the American Business Cycle", *Journal of Political Economy*, **98**, 596-616.

[14] Diebold, F. X. and A. S. Senhadji (1996), "The Uncertain Unit Root in Real GNP: A Comment", *American Economic Review*, **86**, 1291-8.

[15] Driffill, J. and M. Sola (1998), "Intrinsic Bubbles and Regime-Switching", *Journal of Monetary Economics*, **42**, 357-73.

[16] Elíasson, L. (1998), "Mæling á íslenskri hagsveiflu", Central Bank of Iceland, unpublished manuscript.

[17] Filardo, A. J. and S. F. Gordon (1998), "Business Cycle Durations", *Journal of Econometrics*, **85**, 99-123.

[18] Gallant, A. R. (1977), "Testing for a Nonlinear Regression Specification: A Nonregular Case", *Journal of the American Statistical Association*, **72**, 523-30.

[19] Garzia, R. and P. Perron (1996), "An Analysis of Real Interest Under Regime Shift", *Review of Economics and Statistics*, **78**, 111-25.

[20] Guðmundsson, M., T. G. Pétursson and A. Sighvatsson (2000), "Optimal Exchange Rate Policy: The Case of Iceland", in *Macroeconomic Policy: Small Open Economies in an Era of Global Integration,* eds. M. Guðmundsson, T. T. Herbertsson and G. Zoëga. Reykjavík: University of Iceland Press (forthcoming).

[21] Hall, S. G., Z. Psaradakis and M. Sola (1997), "Cointegration and Changes in Regime: The Japanese Consumption Function", *Journal of Applied Econometrics*, **12**, 151-68.

[22] Hamilton, J. D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, **57**, 357-84.

[23] Hamilton, J. D. (1990), "Analysis of Time Series Subject to Changes in Regime", *Journal of Econometrics*, **46**, 39-70.

[24] Hamilton, J. D. (1994), *Time Series Analysis*, Princeton: Princeton University Press.

[25] Hamilton, J. D. (1996), "Specification Testing in Markov-Switching Time-Series Models", *Journal of Econometrics*, **70**, 127-57.

[26] Hamilton, J. D. and C. Lin (1996), "Stock Market Volatility and the Business Cycle", *Journal of Applied Econometrics*, **11**, 573-93.

[27] Hamilton, J. D. and G. Perez-Quiros (1996), "What Do the Leading Indicators Lead?", *Journal of Business*, **69**, 27-49.

[28] Hamilton, J. D. and R. Susmel (1994), "Autoregressive Conditional Heteroscedasticity and Changes in Regime", *Journal of Econometrics*, **64**, 307-33.

[29] Hansen, B. E. (1992a), "Testing for Parameter Instability in Linear Models", *Journal of Policy Modeling*, **14**, 517-33.

[30] Hansen, B. E. (1992b), "The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP", *Journal of Applied Econometrics*, **7**, S61-82.

[31] Hansen, B. E. (1996), "Erratum: The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP", *Journal of Applied Econometrics*, **11**, 195-8.

[32] Kim, C.-J. (1993), "Unobserved-Component Time Series Models with Markov-Switching Heteroscedasticity: Changes in Regime and the Link Between Inflation Rates and Inflation Uncertainty", *Journal of Business and Economic Statistics*, **11**, 341-349.

[33] King, R. G., C. I. Plosser, J. H. Stock and M. W. Watson (1991), "Stochastic Trends and Economic Fluctuations", *American Economic Review*, **81**, 819-40.

[34] Lam, P.-S. (1990), "The Hamilton Model with a General Autoregressive Component: Estimation and Comparison with Other Models of Economic Time Series", *Journal of Monetary Economics*, **26**, 409-32.

[35] MacKinnon, J. G. (1991), "Critical Values for Cointegration Tests", in *Long-Run Economic Relationships: Readings in Cointegration*, eds. R. F. Engle and C. W. J. Granger. Oxford: Oxford University Press.

[36] Magnússon, G. and T. Einarsson (1985), "Hagsveiflur, gengismál og jöfnunarsjóðir", in *Klemensarbók*, ed. S. Snævarr. Reykjavík: Félag viðskipta- og hagfræðinga.

[37] Magnússon, G. and T. Einarsson (1987), "A Further Study in the Icelandic Business Cycle", *Icelandic Economic Papers*, W87:3, University of Iceland,

[38] Murray, C. J. and C. R. Nelson (1998), "The Uncertain Trend in US GDP", University of Washington, Seattle, unpublished manuscript.

[39] National Economic Institute, *Þjóðhagsáætlun*, 1990-1998. Reykjavík.

[40] Nelson, C. R. and C. I. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series", *Journal of Monetary Economics*, **10**, 139-62.

[41] Newey, W. K. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests", *Econometrica*, **53**, 1047-70.

[42] Nyblom, J. (1989), "Testing for the Constancy of Parameters over Time", *Journal of the American Statistical Association*, **84**, 223-30.

[43] Ravn, M. O. and M. Sola (1995), "Stylized Facts and Regime Changes: Are Prices Procyclical?", *Journal of Monetary Economics*, **36**, 497-526.

[44] Snævarr, S. (1993), *Haglýsing Íslands*, Reykjavík: Mál og Menning.

[45] Sola, M. and J. Driffill (1994), "Testing the Term Structure of Interest Rates from a Stationary Markov-Switching VAR", *Journal of Economic Dynamics and Control*, **18**, 601-28.

[46] Swanson, N. R. and H. White (1995), "A Model-Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks", *Journal of Business and Economic Statistics*, **13**, 265-75.

[47] Tauchen, G. (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood Models", *Journal of Econometrics*, **30**, 415-43.

[48] Tong, H. (1983), *Threshold Models in Non-Linear Time Series Analysis*, New York: Springer-Verlag.

[49] Vuong, Q. H. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypothesis", *Econometrica*, **57**, 307-33.

[50] Warne, A. (1996), "Autocovariance Functions and Maximum Likelihood in a VAR Model under Markov Switching", Institute for International Economic Studies, Stockholm University, unpublished manuscript.

[51] Watson, M. W. (1986), "Univariate Detrending Methods with Stochastic Trends", *Journal of Monetary Economics*, **18**, 49-75.

[52] White, H. (1987), "Specification Testing in Dynamic Models", in *Advances in Econometrics: Fifth World Congress*, Vol. II, ed. T. F. Bewley. Cambridge: Cambridge University Press.