Business Cycle Asymmetries: Characterisation and Testing based on Markov-Switching Autoregressions

Michael P. Clements and Hans-Martin Krolzig*
Department of Economics, University of Warwick, Coventry, CV4 7AL.
M.P.Clements@Warwick.ac.uk
Institute of Economics and Statistics, St Cross Building, Manor Road, Oxford, OX1 3UL.
Hans-Martin.Krolzig@nuffield.oxford.ac.uk

December 8, 1998

Abstract

We propose testing for business cycle asymmetries in Markov-switching autoregressive (MS-AR) models. We derive the parametric restrictions on MS-AR models that rule out types of asymmetries such as deepness, steepness, and sharpness, and set out a testing procedure based on Wald statistics which have standard asymptotics. For a two-regime model, such as that popularised by Hamilton (1989), we show that deepness implies sharpness (and vice versa) while the process is always non-steep. We illustrate with two and three-state MS models of US GNP growth, and with models of US output and employment. Our findings are compared with those obtained from standard non-parametric tests.

Journal of Economic Literature Classification: C32

Key words: Business cycle asymmetries, Markov-switching models, non-parametric tests.

1 Introduction

> '*the most violent declines exceed the most considerable advances [. . .] Business contractions appear to be a briefer and more violent process than business expansions*'
> Mitchell (1927, p. 290)

There has been much interest in whether macroeconomic variables behave differently over the phases of the business cycle. Sichel (1993, p. 224) defines an asymmetric cycle as 'one in which some phase of the cycle is different from the mirror image of the opposite phase'. A number of types of asymmetry, such as steepness, deepness, and sharpness (or turning point asymmetry) have been proposed, and tested for empirically using separate non-parametric tests. Other types of asymmetries, such as asymmetric persistence to shocks, have been explored in parametric models (see Beaudry and Koop, 1993 and Hess and Iwata, 1997a). Moreover, while business cycles were originally viewed as consisting of co-movements in many economic variables (see, e.g., Burns and Mitchell, 1946), the research to date on business cycle asymmetries has largely been univariate, based on testing individual series such as GNP, industrial production and unemployment, while ignoring possible co-dependencies or co-movements in the variables over the cycle.

*Financial support from the UK Economic and Social Research Council under grant L116251015 is gratefully acknowledged by both authors. All the computations reported in this paper were carried out in Ox: see Doornik (1996).

Since Hamilton (1989) the Markov-switching autoregressive model (MS-AR) has become popular in empirical business cycle research as a way of characterising the business-cycle phases of expansion and contraction. In this paper we analyse the conditions under which this class of model is capable of generating various types of business cycle asymmetry. The conditions are expressed as restrictions on the parameters of the MS-AR model which, if they hold, would rule out a particular type of asymmetry. We are then able to derive tests of these restrictions based on estimated MS-AR models, providing parametric tests as alternatives of the non-parametric tests typically used in the literature.

The rest of this introduction provides a motivation for focusing on MS-AR models as a way of testing business-cycle asymmetries. Firstly, a number of other papers test for certain types of asymmetries in parametric models. For example, Sichel (1991) looks for business cycle duration dependence using a parametric hazard model, while, e.g., Diebold, Rudebusch and Sichel (1993), Filardo (1994) and Filardo and Gordon (1998) explore the same issue in extensions of the Hamilton (1989) model that allow for time-varying transition probabilities. Generally there appears to be positive duration dependence in contractions in the US post-War period, so that the probability of moving out of recession increases with the duration of recession. However, there appears to have been little attempt to formally test for other types of asymmetries using MS-AR models, and it is that lacuna that this paper seeks to fill.

Our parametric approach would only be expected to yield more powerful tests of asymmetry if the model is a good representation of the process generating the data. This is the usual trade-off between parametric and non-parametric approaches — the latter are robust to model mis-specification, which may lead to misleading inferences in the former, but are likely to have lower power. To that end, we check that the model is reasonably well specified. We choose the MS models to test for asymmetries because they appear to offer a good fit to the data and to characterise typical business cycle features. An added attraction is that the MS-AR framework can be readily extended to multivariate settings (see Krolzig, 1997 for an overview), and a number of papers have sought to do so, e.g., Ravn and Sola (1995), Diebold and Rudebusch (1996), Hamilton and Lin (1996), Krolzig and Sensier (1998), and Krolzig and Toro (1998), but again, without directly addressing the issue of testing for asymmetries, or of analysing what possible types of asymmetry the models are capable of generating. For example, Ravn and Sola (1995) look at whether prices move counter-cyclically (i.e., whether the covariance between output growth and inflation is negative), and use MS models to control for changes in the (unconditional) means of the variables, which might otherwise lead to misleading inferences concerning the co-movements between the series. Hamilton and Lin (1996) consider a bivariate model for stock returns and output growth, where the dynamic linkage between the two variables centres on possible dependence between the latent processes. Diebold and Rudebusch (1996) consider dynamic-factor models with regime switching — the factor structure captures the idea that variables move together over the business cycle, and the regime switching allows for asymmetric behaviour over the cycle.

The empirical relevance of our results rests on the MS model being a good representation of the data. However, given the widespread popularity of the MS model in applied research, it is of interest to establish precisely which types of asymmetries these models are in principle capable of generating, particularly since some of the literature appears confused on this point [1]. As we shall show, the number of states turns out to be crucial. Deducing which types of asymmetry the MS models are capable of generating differs from Hess and Iwata (1997b), who investigate by simulation whether empirically-estimated models (ARIMA and popular scalar non-linear models) are able to replicate the ‘fundamental business cycle features’ of observed durations and amplitudes of contractions and expansions. Their features of interest differ from ours, and empirical durations and magnitudes are calculated for a two-state representation of

[1] For example, Sichel (1993, p. 232, footnote 19) states that the Hamilton (1989) two-state model implies steepness in US GNP. In fact, steepness (as defined formally below) can not arise in such a model.

the business cycle – contractions and expansions. This is somewhat crude, ignoring as it does the fast recovery phase that has characterised all US recessions except the last (the 1990 recession). They find their three-state MS model fails to generate contractions of sufficient duration or depth (given their definition of a contraction), but we note below that their model is estimated with a restricted transition matrix and assumes a homogeneous variance of shocks across regimes. Our preferred model is not restricted in this way. We are also interested in whether it is *feasible* for the MS class of models to generate certain asymmetric features, as well as whether empirically estimated models possess such features, which is the focus of the simulations in Hess and Iwata (1997b).

Further, because MS-AR models can be readily applied to multivariate settings, they can be used to model co-movements of variables over the business cycle, and thus more closely reflect what the originators of business cycle research had in mind. We show empirically that utilising the co-movements between series may provide useful tests for asymmetries in highly restricted models, but that for two related series x and y, we might in general expect asymmetries in x to help to explain those in y (and vice versa), so that multivariate models exhibit fewer asymmetries. This is obvious if one takes the view that asymmetries (and non-linearities more generally) in univariate time-series representations arise because of omitted variables. The exception is when the dependence between the series is only allowed via the Markov process (as in, e.g., Ravn and Sola, 1995 and Hamilton and Lin, 1996).

As a final motivation for our work, we note that McQueen and Thorley (1993, pp. 342 – 343) and Sichel (1993, pp. 225 – 226) discuss the importance, from both theoretical and empirical viewpoints, of establishing whether there are asymmetries in the business cycle.

The types of asymmetry of interest are generally taken to relate to the de-trended log of output. For example, Speight and McMillan (1998) consider the de-trended component (x_t) of the variable y_t, where $x_t = y_t - \tau_t$. τ_t is a non-stationary trend component, and x_t is stationary, possibly consisting of cycle and noise components. We assume the non-stationarity can be removed by differencing, i.e., $x_t = \Delta y_t$. Trend elimination by differencing is natural in our setup, because the MS model is typically estimated on the first difference of the log of output. However, none of the propositions on asymmetries in MS-AR models that follow, nor the testing procedures, require this method of de-trending, and remain valid whichever method is used. All we require is that a MS-AR model can be estimated for the de-trended series, howsoever obtained. The sensitivity of the findings of asymmetries to the method of trend elimination requires further research, and Gordon (1997) shows that in general the model of the short-run fluctuations in output may depend on the treatment of the trend component.

The plan of the paper is as follows. In section 2 we briefly review the literature on business cycle asymmetries. Section 3 formally defines the concepts of deepness, steepness and sharpness gleaned from the literature, and derives the corresponding parameter restrictions on the MS model, paying particular attention to the empirically relevant two and three regime models. Then, section 4 sketches out our proposed testing procedures based on Wald tests, which obviates the necessity of estimating the restricted (null) MS model. Section 5 notes the straightforward extension to multiple time series, and section 6 sets out the empirical illustrations. Section 7 concludes.

2 A brief review of the literature on business cycle asymmetries

2.1 Steepness and deepness

Sichel (1993) distinguishes two types of business cycle asymmetry: ‘steepness’ and ‘deepness’. The former relates to whether contractions are steeper (or less steep) than expansions, the latter to whether the amplitude of troughs exceeds that of peaks. Neftci (1984) found evidence of steepness in post-War US unemployment during contractions using a non-parametric test of whether there are longer runs of increases than decreases in the series. However, Falk (1986) failed to find evidence of steepness in other US quarterly macroeconomic series using Neftci’s procedure, and Sichel (1989) suggested an error in Neftci’s work and indicated that the procedure might fail to find steepness when in fact it is present. Neftci defines an indicator variable $I_t = 1$ if $\Delta y_t > 0$ and $I_t = -1$ if $\Delta y_t \leq 0$. Suppose I_t can be represented by a second-order Markov process, then steepness would imply that $p_{11} > p_{00}$, where $p_{11} = prob[I_t = 1 \mid I_{t-1} = 1, I_{t-2} = 1]$ and $p_{00} = prob[I_t = -1 \mid I_{t-1} = -1, I_{t-2} = -1]$. The problem with this procedure is its sensitivity to noise. If increases (decreases) are inadvertently measured as decreases (increases), then the counts of transitions from which the estimates of the transition probabilities are derived will be affected. Sichel (1989) finds strong evidence of asymmetry in annual unemployment, for which measurement error is presumably less important.

Sichel (1993) suggests a test of deepness based on the coefficient of skewness calculated for the detrended series. Deepness of contractions will show up as negative skewness, since it implies that the average deviation of observations below the mean will exceed that of observations above the mean. Steepness implies negative skewness in the first difference of the detrended series: decreases should be larger, though less frequent, than increases. On the basis of these tests, deepness is found to characterise quarterly post-War US unemployment and industrial production, with weaker evidence for GNP, while only unemployment (of the three) appears to exhibit steepness.

2.2 Sharpness

Sharpness or turning point asymmetry, as introduced by McQueen and Thorley (1993), would result if, e.g., troughs were sharp and peaks more rounded. They present two tests. The first is based on the magnitude of growth rate changes around NBER-dated peaks and troughs. The mean absolute changes are calculated for peaks and troughs separately, and the test for asymmetry is based on rejecting the null of the population mean changes in the variable at peaks and troughs being equal. McQueen and Thorley (1993) find the null of equal turning point sharpness can be rejected for both the unemployment rate and industrial production. Their second testing procedure is based on a second-order three state Markov chain. They distinguish between contraction (1), moderate (2) and high (3)(recovery) states. The hypothesis in Hicks (1950), that troughs are sharper than peaks, translates into $p_{113} > p_{331}$, where p_{113} is the probability of jumping from the contraction to high growth state ($p_{113} = prob[I_t = 3 \mid I_{t-1} = 1, I_{t-2} = 1]$) and p_{331} is the probability of jumping directly from high growth to contraction. ‘Complete’ TP symmetry requires $p_{113} = p_{331}$ as well as $p_{112} = p_{332}$ and $p_{223} = p_{221}$. They again find evidence of sharpness asymmetry for post-War unemployment and industrial production, but the susceptibility of the test to noise is evident when they consider pre-War industrial production and post-War agricultural unemployment: in both cases quarterly volatility in the series interrupts runs of ones and threes, reducing the number of sharp TPs and the power of the test.

3 Concepts of asymmetry and MS models

The (vector) time series is assumed to have been generated by a (V)AR(p) with M Markov-switching regimes in the mean of the process, which we label an MSM(M)-AR(p) process:

$$x_t - \mu(s_t) = \sum_{k=1}^{p} \alpha_k \left(x_{t-k} - \mu(s_{t-k})\right) + u_t, \quad u_t|s_t \sim \mathsf{NID}\left(0, \sigma^2\right). \tag{1}$$

We can order the regimes by the magnitude of μ such that $\mu_1 < \ldots < \mu_M$. The Markov chain is ergodic, irreducible, and there does not exist an absorbing state, i.e., $\bar{\xi}_m \in (0,1)$ for all $m = 1, \ldots, M$, where $\bar{\xi}_m$ is the ergodic or unconditional probability of regime m. The discussion below is for a univariate time series, but we also note the generalization to a vector process. The transition probabilities are time-invariant:

$$p_{ij} = prob(s_{t+1} = j | s_t = i), \quad \sum_{j=1}^{M} p_{ij} = 1 \quad \forall i, j \in \{1, \ldots, M\}. \tag{2}$$

3.1 Types of asymmetry

For clarity, we formally define the concepts of deepness, steepness and sharpness discussed and motivated in section 2. Then we derive the corresponding parameter restrictions on the MS-VAR and apply them to the case of two and three regimes processes.

Definition 1. *Deepness. Sichel (1993). The process $\{x_t\}$ is said to be* **non-deep (non-tall)** *iff x_t is not skewed:*

$$\mathsf{E}\left[(x_t - \bar{\mu})^3\right] = 0.$$

Analogously we can define steepness as skewness of the differences:

Definition 2. *Steepness. Sichel (1993). The process $\{x_t\}$ is said to be* **non-steep** *iff Δx_t is not skewed:*

$$\mathsf{E}\left[\Delta x_t^3\right] = 0.$$

The business cycle literature indicates the possibility of negative skewness of x_t and Δx_t — thus steep and deep contractions. The opposite case is of **tall** ($\mathsf{E}[(x_t - \mu)^3] > 0$) and **steep (**$\Delta x_t$ positively skewed) expansions.

Definition 3. *Sharpness. McQueen and Thorley (1993). The process $\{x_t\}$ is said to be* **non-sharp** *iff the transition probabilities to and from the two outer regimes are identical:*

$$p_{m1} = p_{mM} \text{ and } p_{1m} = p_{Mm}, \text{ for all } m \neq 1, M; \text{ and } p_{1M} = p_{M1}.$$

In a two-regime model, for example, non-sharpness implies that $p_{12} = p_{21}$. In a three-regime model, it requires $p_{13} = p_{31}$ and in addition $p_{12} = p_{32}$ and $p_{21} = p_{23}$. When $M = 4$ the following restrictions on the matrix of transition probabilities are required to hold for non-sharpness:

$$P = \begin{bmatrix} 1-a-b-c & a & b & c \\ d & * & * & d \\ e & * & * & e \\ c & a & b & 1-a-b-c \end{bmatrix} \tag{3}$$

3.2 Asymmetries in MS-AR processes

We now present the restrictions on the parameter space of the MSM-AR model that correspond to the concepts of asymmetry. Proofs of these propositions are confined to an appendix.

While the restrictions implied by sharpness follow immediately, testing for deepness and steepness is less obvious.

Proposition 1. *An MSM(M)-AR(p) process is* **non-deep** *iff*

$$\sum_{m=1}^{M} \bar{\xi}_m \mu_m^{*3} = \sum_{m=1}^{M-1} \bar{\xi}_m \mu_m^{*3} + \left(1 - \sum_{m=1}^{M-1} \bar{\xi}_m\right) \mu_M^{*3} = 0 \tag{4}$$

with $\mu_m^* = \mu_m - \mu_y = \sum_{i \neq m} (\mu_m - \mu_i)\, \bar{\xi}_i$ *and where* $\bar{\xi}_m$ *is the unconditional probability of regime* m*.*

Unfortunately the expression (4) is a highly complicated third-order polynomial in the regime-dependent parameters of the process $\mu_1, \ldots, \mu_M$ and the unconditional regime probabilities $\bar{\xi}_1, \ldots, \bar{\xi}_{M-1}$, which are non-linear functions of the transition parameters p_{ij}:

$$\begin{aligned} \mathsf{E}\left[\mu_t^k\right] &= \sum_{m=1}^{M} \bar{\xi}_m \left(\mu_m - \sum_{i=1}^{M} \bar{\xi}_i \mu_i\right)^k = \sum_{m=1}^{M} \bar{\xi}_m \left[\mu_m - \mu_M - \sum_{i=1}^{M-1} (\mu_i - \mu_M)\, \bar{\xi}_i\right]^k \\ &= \sum_{m=1}^{M-1} \bar{\xi}_m \left[(\mu_m - \mu_M) - \sum_{i=1}^{M-1} (\mu_i - \mu_M)\, \bar{\xi}_i\right]^k + \left(1 - \sum_{m=1}^{M-1} \bar{\xi}_m\right) \left[- \sum_{i=1}^{M-1} (\mu_i - \mu_M)\, \bar{\xi}_i\right]^k \end{aligned}$$

However, for $M = 2$ the problem becomes analytically tractable.

Example 1. *Consider the case of two regimes. The MSM(2)-AR(p) process can be written as the sum of two independent processes:* $y_t - \mu_y = \mu_t + z_t$*, where* μ_y *is the unconditional mean of* y_t*, such that* $\mathsf{E}[\mu_t] = \mathsf{E}[z_t] = 0$*. While the process* $z_t = \sum_{j=1}^{p} \alpha_j z_{t-j} + u_t$ *is gaussian,* μ_t *represents the contribution of the Markov chain,* $\mu_t = (\mu_1 - \mu_2)\zeta_t$*, with* $\zeta_t = \xi_{1t} - \bar{\xi}_1$*, which equals* $1 - \bar{\xi}_1$ *if the regime is* 1 *and* $-\bar{\xi}_1$ *otherwise. Invoking proposition 1, the skewness of the Markov chain is given by:*

$$\mathsf{E}\left[\mu_t^3\right] = \sum_{m=1}^{2} \bar{\xi}_m \mu_m^{*3} = \bar{\xi}_1 \mu_1^{*3} + (1 - \bar{\xi}_1) \mu_2^{*3}$$

where $\bar{\xi}_1 = p_{21}/(p_{12} + p_{21})$ *is the unconditional probability of regime one,* $\mu_1^* = \mu_1 - \mu_y = (1 - \bar{\xi}_1)(\mu_1 - \mu_2)$ *and* $\mu_2^* = \mu_2 - \mu_y = (-\bar{\xi}_1)(\mu_1 - \mu_2)$*. Thus:*

$$\begin{aligned} \mathsf{E}\left[\mu_t^3\right] &= \bar{\xi}_1 (1 - \bar{\xi}_1)^3 (\mu_1 - \mu_2)^3 + (1 - \bar{\xi}_1)(-\bar{\xi}_1)^3 (\mu_1 - \mu_2)^3 \\ &= \left[\bar{\xi}_1 (1 - \bar{\xi}_1)^3 - (1 - \bar{\xi}_1)(\bar{\xi}_1)^3\right] (\mu_1 - \mu_2)^3 \\ &= \bar{\xi}_1 (1 - \bar{\xi}_1) \left[(1 - \bar{\xi}_1)^2 - \bar{\xi}_1{}^2\right] (\mu_1 - \mu_2)^3 \\ &= \bar{\xi}_1 (1 - \bar{\xi}_1) \left[1 - 2\bar{\xi}_1\right] (\mu_1 - \mu_2)^3 . \end{aligned}$$

As the Markov-switching model implies that $\mu_1 \neq \mu_2$ *and* $\bar{\xi}_1 \in (0, 1)$*, non-deepness,* $\mathsf{E}[\mu_t^3] = 0$*, requires that* $\bar{\xi}_1 = 0.5$*. Hence the matrix of transition probabilities must be symmetric,* $p_{12} = p_{21}$*. This also implies that the regime-conditional means* μ_1 *and* μ_2 *are equidistant to the unconditional mean* μ_y*.*

Hence, in the case of two regimes we can test for non-deepness by testing the hypothesis $p_{12} = p_{21}$. This is equivalent to the test of non-sharpness. For processes with $M > 2$ we propose to test for non-deepness based on the μ_m^* conditional on μ_y and the $\bar{\xi}_m$.

Example 2. *Consider now an MSM(3)-AR(p) process. Again, by invoking proposition 1, the skewness of the Markov chain μ_t is given by:*

$$\mathsf{E}\left[\mu_t^3\right] = \sum_{m=1}^{3} \bar{\xi}_m \mu_m^{*3} = \bar{\xi}_1 \mu_1^{*3} + \bar{\xi}_2 \mu_2^{*3} + (1 - \bar{\xi}_1 - \bar{\xi}_2)\mu_3^{*3}$$

where $\mu_m^* = \mu_m - \mu_y = \mu_m - \bar{\xi}_1\mu_1 - \bar{\xi}_2\mu_2 - (1 - \bar{\xi}_1 - \bar{\xi}_2)\mu_3 = \sum_{i \neq m} \bar{\xi}_i(\mu_m - \mu_i)$. *Thus:*

$$\mathsf{E}\left[\mu_t^3\right] = \sum_{m=1}^{3} \bar{\xi}_m \left[\sum_{i \neq m} \bar{\xi}_i(\mu_m - \mu_i)\right]^3 .$$

Non-deepness, $\mathsf{E}\left[\mu_t^3\right] = 0$*, requires that:*

$$\mu_3^{*3} = \frac{\bar{\xi}_1}{(1 - \bar{\xi}_1 - \bar{\xi}_2)}\mu_1^{*3} + \frac{\bar{\xi}_2}{(1 - \bar{\xi}_1 - \bar{\xi}_2)}\mu_2^{*3}.$$

We now derive conditions for the presence of steepness which is based on the skewness of the differenced series.

Proposition 2. *An MSM(M)-AR(p) process is* **non-steep** *if the size of the jumps, $\mu_j - \mu_i$, satisfies the following condition:*

$$\sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \left(\bar{\xi}_i p_{ij} - \bar{\xi}_j p_{ji}\right) \left[\mu_j - \mu_i\right]^3 = 0. \tag{5}$$

Symmetry of the matrix of transition parameters (which is stronger than the definition of sharpness) is sufficient but not necessary for non-steepness. A proof of this proposition appears in the appendix.

In contrast to deepness, the condition for steepness depends not only on the ergodic probabilities, $\bar{\xi}_j$, but also directly on the transition parameters.

Example 3. *In an MSM(2)-AR(p) process, condition (5) gives:*

$$\mathsf{E}\left[\Delta\mu_t^3\right] = \left(\bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21}\right) \left[\mu_2 - \mu_1\right]^3 .$$

Example 4. *For an MSM(3)-AR(p) process we get:*

$$\begin{aligned} \mathsf{E}\left[\Delta\mu_t^3\right] &= \sum_{i=1}^{2} \sum_{j=i+1}^{3} \left(\bar{\xi}_i p_{ij} - \bar{\xi}_j p_{ji}\right) \left[\mu_j - \mu_i\right]^3 \\ &= \left(\bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21}\right) \left[\mu_2 - \mu_1\right]^3 + \left(\bar{\xi}_1 p_{13} - \bar{\xi}_3 p_{31}\right) \left[\mu_3 - \mu_1\right]^3 + \left(\bar{\xi}_2 p_{23} - \bar{\xi}_3 p_{32}\right) \left[\mu_3 - \mu_2\right]^3 \end{aligned}$$

While this is a complicated expression, the concept of steepness can be made operational by using the sufficient condition, that is, the symmetry of the matrix of transition parameters, which implies non-steepness. This is stronger than the property of non-sharpness (see, for example, expression (3)).

The concepts of sharpness/steepness and deepness are illustrated in figures 1 and 2. Figure 1 gives representative time-paths for μ_t and $\Delta\mu_t$ for an MS(3) process, drawn so that the duration of each regime equals its expectation, and also shows the densities of μ_t and $\Delta\mu_t$, with gaussian curves super-imposed. The top row corresponds to the non-deep, non-steep and non-sharp (strictly, symmetric transition probabilities) case. There is no skewness in either the μ_t or $\Delta\mu_t$. Deepness (row 2) shows up in negative skewness in μ_t, and steepness of expansions (row 3) in positive skewness of $\Delta\mu_t$. Row 4 shows the two together. Figure 2 depicts deepness in the MS(2) model of Hamilton (1989), and deepness and steepness (of expansions) in the MS(3) model of Clements and Krolzig (1998) (both models are discussed further in section 6).

We close this section with a corollary characterizing the two-regime MS-AR model, which shows the impossibility of the MS(2) exhibiting steepness, and the equivalence of the concepts of deepness and sharpness.

Corollary 1. *A two-regime Markov-switching model is always non-steep. Non-sharpness implies non-deepness and vice versa.*

Non-steepness is evident from **Example 3**. Since $\bar{\xi}_1/\bar{\xi}_2 = p_{21}/p_{12}$, we have that $\bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21} = 0$ and hence $\mathsf{E}[\Delta\mu_t^3] = 0$. Further, non-sharpness (symmetric transition probabilities, $p_{12} = p_{21}$) implies non-deepness, $\mathsf{E}[(y_t - \mu_y)^3] = 0$, and *vice versa*. In particular, both concepts imply that the regime-conditional means μ_1 and μ_2 are equidistant to the unconditional mean μ_y.

4 Testing

As the number of regimes remains unchanged under all three hypotheses, standard asymptotics can be invoked. Wald tests of the hypotheses are computationally attractive, since the model does not have to be estimated under the null. In general terms, we consider Wald (W) tests of the hypothesis:

$$H_0 : \phi(\lambda) = \mathbf{0} \quad \text{vs.} \quad H_1 : \phi(\lambda) \neq \mathbf{0},$$

where $\phi : \mathbb{R}^n \to \mathbb{R}^r$ is a continuously differentiable function with rank r, $r = \mathrm{rk}\left(\frac{\partial\phi(\lambda)}{\partial\lambda'}\right) \leq \dim\lambda$. As the p_{ij} are restricted to the $[0,1]$ interval, the tests are formulated on the logits $\pi_{ij} = \log\left(\frac{p_{ij}}{1-p_{ij}}\right)$ which avoids problems if one or more of the p_{ij} is close to the border. It is worth noting that if $\frac{1}{T}\left(\tilde{\pi}_{ij} - \pi_{ij}\right) \xrightarrow{d} \mathsf{N}(0, \sigma^2_{\pi_{ij}})$, then $\frac{1}{T}\left(\tilde{p}_{ij} - p_{ij}\right) \xrightarrow{d} \mathsf{N}(0, p_{ij}^2(1-p_{ij})^2\sigma^2_{\pi_{ij}})$ as $p_{ij} = \dfrac{e^{\pi_{ij}}}{1+e^{\pi_{ij}}}$. If one of the transition parameters is estimated to lie on the border, $p_{ij} \in \{0,1\}$, then the parameter is taken as being fixed and eliminated from the parameter vector λ.

Let $\tilde{\lambda}$ denote the unconstrained MLE of $\lambda = (\mu_1, \ldots, \mu_M; \alpha_1, \ldots, \alpha_p, \sigma^2; \boldsymbol{\pi})$, and $\hat{\lambda}$ the restricted MLE under the null. Then the Wald test statistic W is based on the unconstrained estimator $\tilde{\lambda}$, which is asymptotically normal:

$$\sqrt{T}\left(\tilde{\lambda} - \lambda\right) \xrightarrow{d} \mathsf{N}\left(\mathbf{0}, \Sigma_{\tilde{\lambda}}\right),$$

where, for the MLE, $\Sigma_{\tilde{\lambda}} = \Im_a^{-1}$ is the inverse of the asymptotic information matrix. This can be calculated numerically. It follows that $\phi(\tilde{\lambda})$ is also normal for large samples:

$$\sqrt{T}[\phi(\tilde{\lambda}) - \phi(\lambda)] \xrightarrow{d} \mathsf{N}\left(\mathbf{0}, \left.\frac{\partial\phi(\lambda)}{\partial\lambda'}\right|_{\tilde{\lambda}} \Sigma_{\tilde{\lambda}} \left.\frac{\partial\phi(\lambda)'}{\partial\lambda}\right|_{\tilde{\lambda}}\right).$$

Thus, if $H_0 : \phi(\lambda) = \mathbf{0}$ is true and the variance–covariance matrix is invertible,

$$T\phi(\tilde{\lambda})' \left[\left.\frac{\partial\phi(\lambda)}{\partial\lambda'}\right|_{\tilde{\lambda}} \tilde{\Sigma}_{\tilde{\lambda}} \left.\frac{\partial\phi(\lambda)'}{\partial\lambda}\right|_{\tilde{\lambda}}\right]^{-1} \phi(\tilde{\lambda}) \xrightarrow{d} \chi^2(r),$$

where $\tilde{\Sigma}_{\tilde{\lambda}}$ is a consistent estimator of $\Sigma_{\tilde{\lambda}}$.

4.1 Deepness

The Wald test for the null of non-deepness is based on:

$$\phi_D(\lambda) = \phi_D(\boldsymbol{\mu};\cdot) := \sum_{m=1}^{M} \bar{\xi}_m(\mu_m - \mu_y)^3$$

where the $\bar{\xi}_m$ and μ_y are taken as fixed. Thus $\frac{\partial\phi}{\partial\lambda_i} = 3\bar{\xi}_m(\mu_m - \mu_y)^2$ for $\lambda_i = \mu_m$, $m = 1, \ldots, M$, and $\frac{\partial\phi}{\partial\lambda_i} = 0$ for $\lambda_i \in [\alpha_1, \ldots, \alpha_p, \sigma^2; \boldsymbol{\pi}]$.

Example 5. *Thus for $M = 2$, the null of non-deepness is tested by:*

$$T\left[\sum_{m=1}^{2} \bar{\xi}_m(\widetilde{\mu}_m - \mu_y)^3\right]^2 \left[\left[\begin{array}{cc} 3\bar{\xi}_1(\widetilde{\mu}_1 - \mu_y)^2 & 3\bar{\xi}_2(\widetilde{\mu}_2 - \mu_y)^2 \end{array}\right] \tilde{\Sigma}_{\tilde{\lambda}_D} \left[\begin{array}{c} 3\bar{\xi}_1(\widetilde{\mu}_1 - \mu_y)^2 \\ 3\bar{\xi}_2(\widetilde{\mu}_2 - \mu_y)^2 \end{array}\right]\right]^{-1} \xrightarrow{d} \chi^2(1),$$

where $\tilde{\lambda}_D = [\widetilde{\mu}_1\ \widetilde{\mu}_2]'$.

4.2 Steepness

A Wald test for the null of non-steepness can be based on:

$$\phi_S(\lambda) = \phi_S(\boldsymbol{\mu};\cdot) := \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \left(\bar{\xi}_i p_{ij} - \bar{\xi}_j p_{ji}\right) \left[\mu_j - \mu_i\right]^3$$

where the $\bar{\xi}_m$, p_{ij} and μ_y again are taken as fixed. Thus the test only concerns the vector of mean parameters:

$$\nabla\boldsymbol{\mu} = \left[\begin{array}{c} \mu_2 - \mu_1 \\ \vdots \\ \mu_M - \mu_1 \\ \vdots \\ \mu_M - \mu_{M-1} \end{array}\right] = Q\boldsymbol{\mu}, \text{ with } Q = \frac{\partial\nabla\boldsymbol{\mu}}{\partial\boldsymbol{\mu}'} = \left[\begin{array}{cccc} -1 & 1 & & 0 \\ \vdots & & \ddots & \\ -1 & 0 & & 1 \\ 0 & -1 & 1 & \\ & -1 & 0 & 1 \\ & & \ddots & \\ 0 & & -1 & 1 \end{array}\right], \boldsymbol{\mu} = \left[\begin{array}{ccc} \mu_1 & \cdots & \mu_M \end{array}\right]'.$$

Thus $\frac{\partial\phi}{\partial\boldsymbol{\mu}'} = \frac{\partial\phi}{\partial\nabla\boldsymbol{\mu}'}\frac{\partial\nabla\boldsymbol{\mu}}{\partial\boldsymbol{\mu}'}$ with $\frac{\partial\phi}{\partial\nabla\boldsymbol{\mu}_m} = 3\left(\bar{\xi}_i p_{ij} - \bar{\xi}_j p_{ji}\right)\left[\mu_j - \mu_i\right]^2$ and $\frac{\partial\phi}{\partial\lambda_i} = 0$ otherwise. The Wald test statistic has the form:

$$\phi(\tilde{\lambda})' \left[\frac{\partial\phi}{\partial\nabla\boldsymbol{\mu}'} Q \left(\tfrac{1}{T}\tilde{\Sigma}_{\tilde{\mu}}\right) Q' \frac{\partial\phi'}{\partial\nabla\boldsymbol{\mu}}\right]^{-1} \phi(\tilde{\lambda}) \xrightarrow{d} \chi^2(1).$$

In the case of a three-state Markov chain, for example:

$$Q = \left[\begin{array}{ccc} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{array}\right] \text{ and } \frac{\partial\phi'}{\partial\nabla\boldsymbol{\mu}} = \left[\begin{array}{c} 3\left(\bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21}\right)\left[\mu_2 - \mu_1\right]^2 \\ 3\left(\bar{\xi}_1 p_{13} - \bar{\xi}_3 p_{31}\right)\left[\mu_3 - \mu_1\right]^2 \\ 3\left(\bar{\xi}_2 p_{23} - \bar{\xi}_3 p_{32}\right)\left[\mu_3 - \mu_2\right]^2 \end{array}\right].$$

4.3 Sharpness

The null of non-sharpness can be expressed as:

$$\phi_{TP}(\lambda) = \phi_{TP}(\boldsymbol{\pi};\cdot) := \Phi\boldsymbol{\pi},$$

where the matrix Φ is defined such that $p_{m1} = p_{mM}$ and $p_{1m} = p_{Mm}$, for all $m \neq 1, M$, and $p_{1M} = p_{M1}$. Let the π_{ij} be collected to the matrix Π:

$$\Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{M1} \\ \vdots & \ddots & \vdots \\ \pi_{1M} & \cdots & \pi_{MM} \end{bmatrix}$$

the matrix of logit transition probabilities. Then the vector $\boldsymbol{\pi}$ is given by $\mathrm{vecd}(\Pi)$, defined as $\mathrm{vec}(\Pi)$ with the diagonal elements π_{ij} excluded. In the case of a three-state Markov chain, for example, we have that:

$$\boldsymbol{\pi} = (\pi_{12}, \pi_{13}, \pi_{21}, \pi_{23}, \pi_{31}, \pi_{32})' \quad \text{and} \quad \Phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{bmatrix}.$$

For linear restrictions the relevant Wald statistic can be expressed as:

$$W_{TP} = T(\Phi\tilde{\lambda} - \varphi)' \left[\Phi\ \tilde{\Sigma}_{\tilde{\lambda}}\ \Phi'\right]^{-1} (\Phi\tilde{\lambda} - \varphi).$$

Thus under the null of symmetric transition probabilities the Wald statistic has the form:

$$W_{TP} = \tilde{\pi}'\Phi' \left[\Phi \left(\tfrac{1}{T}\tilde{\Sigma}_{\tilde{\pi}}\right) \Phi'\right]^{-1} \Phi\tilde{\pi}.$$

5 Extension to Markov switching intercept models

The MSI(M)-AR(p) model is characterised by MS in the **I***ntercept*, rather than MS in the M*ean*:

$$y_t = \mu(s_t) + \sum_{j=1}^{p} \alpha_j y_{t-j} + u_t, \tag{6}$$

where $u_t \sim \mathsf{NID}(0, \sigma^2)$ and $s_t \in \{1, \ldots, M\}$ is generated by a Markov chain.

As before for the MSM-AR process, the MSI(M)-AR(p) process can be written as the sum of two independent processes:

$$y_t - \mu_y = \mu_t + z_t \tag{7}$$

where μ_y is the unconditional mean of y_t, $\mu_y = \alpha^{-1}(1)\sum_{m=1}^{M} \bar{\xi}_m \mu_m$, where $\alpha(L) = 1 - \alpha_1 L - \ldots - \alpha_p L^p$, so that $\mathsf{E}[\mu_t - \mu_y] = 0$. $\{z_t\}$ is a gaussian process, $\alpha(L)z_t = u_t$, and $\mathsf{E}[z_t] = 0$, so that μ_t represents the contribution of the Markov chain, and $\mathsf{E}[\mu_t] = 0$. To derive an expression for μ_t, first rewrite (6) as:

$$\alpha(L)(y_t - \mu_y) = \nu_t + u_t. \tag{8}$$

where ν_t is defined as

$$\nu_t = \mu(s_t) - \bar{\mu} = \sum_{m=1}^{M} \mu_m \left(\xi_{mt} - \bar{\xi}_m\right),$$

and $\bar{\mu} = \alpha^{-1}(1)\,\mu_y$. In the case of a two-regime model we have that $\nu_t = (\mu_1 - \mu_2)\zeta_t$ with $\zeta_t = \xi_{1t} - \bar{\xi}_1$, which equals $1 - \bar{\xi}_1$ if the regime is 1 and $-\bar{\xi}_1$ otherwise. As (7) has to be equivalent to (6), the following expression for μ_t is obtained:

$$\mu_t = \alpha^{-1}(L)\,\nu_t. \tag{9}$$

Thus in contrast to the MSM-AR model, considered so far, where a shift in regime causes a once-and-for-all jump in the level of the observed time series, the MSI-AR model implies a smooth transition in the level of the process after a shift in regime.

Tests for asymmetries in MSI(M)-AR(p) models can be based on ν_t, which can be seen to be equivalent to the μ_t in MSM(M)-AR(p) models. Wald tests for deepness and steepness can be easily constructed by applying the procedures developed in section 4 to parametric tests for the skewness of ν_t and $\Delta\nu_t$, respectively.

A potential problem arises when the roots of $\alpha(L)$ are close to the unit circle, and in the extreme, for the first-order polynomial, $\alpha(L) = 1 - \alpha L$, $\alpha = 1$. Then $\nu_t = \Delta\mu_t$, and testing ν_t for deepness leads to conclusions for the deepness of Δy_t (rather than y_t). In other words applying the conditions derived for deepness in the MSM-AR model to ν_t provides a test of steepness of the MSI-AR model. In our examples, the roots of $\alpha(L)$ are a long way from unity as we are modelling first differences, and these exhibit little dependence relative to models in levels. Furthermore, the extreme case of a unit root implies the data have not been differenced a sufficient number of times prior to modelling.

6 Empirical Illustrations

We apply the parametric tests discussed in section 4 to the MS(2)-AR(4) model of output growth of Hamilton (1989) for the period 1953-1984, to the MS(3)-AR(4) model of Clements and Krolzig (1998) for a number of sample periods, and to a number of bivariate models of post-war US output and employment motivated by Krolzig and Toro (1998). The latter illustrates the extension of the techniques to multiple time series. The outcomes of the tests for asymmetries are compared with non-parametric tests of skewness.

6.1 Empirical MS models of US output growth

MS-AR models have been used in contemporary empirical macroeconomics to capture certain features of the business cycle, but the formal testing of asymmetries has been largely confined to non-parametric approaches. The seminal paper by Hamilton (1989) fit a fourth-order autoregression ($p = 4$) to the quarterly percentage change in US real GNP from 1953 to 1984:

$$\Delta y_t - \mu(s_t) = \alpha_1\left(\Delta y_{t-1} - \mu(s_{t-1})\right) + \ldots + \alpha_4\left(\Delta y_{t-p} - \mu(s_{t-4})\right) + \epsilon_t, \tag{10}$$

where $\epsilon_t \sim \mathsf{NID}(0, \sigma^2)$ and the conditional mean $\mu(s_t)$ switches between two states, 'expansion' and 'contraction':

$$\mu(s_t) = \begin{cases} \mu_1 < 0 & \text{if } s_t = 1 \text{ ('contraction' or 'recession')} \\ \mu_2 > 0 & \text{if } s_t = 2 \text{ ('expansion' or 'boom')} \end{cases}$$

with the variance of the disturbance term, $\sigma^2(s_t) = \sigma^2$, assumed the same in both regimes. This is an MS*M*ean model, with the autoregressive parameters and disturbances independent of the state s_t.

The maximization of the likelihood function of an MS-AR model entails an iterative estimation technique to obtain estimates of the parameters of the autoregression and the transition probabilities governing the Markov chain of the unobserved states: see Hamilton (1990) for an Expectation Maximization

(EM) algorithm for this class of model, and Krolzig (1997) for an overview of alternative numerical techniques for the maximum likelihood estimation these of models.

The results for the original Hamilton model and sample period are recorded in Table 1. The tests for skewness indicate significant negative skewness in Δy_t at the 5% level. The outcome of the tests on the MS-AR model is similar. There is evidence of sharpness at the 10% level.

6.1.1 MSIH(3)-AR(4) Model

Sichel (1994) shows that post-War business cycles have typically consisted of three phases: contraction, followed by high-growth recovery, and then a period of moderate growth. To capture this in a parametric model, we consider the three-state MS model of Clements and Krolzig (1998)[2], where there is a shifting intercept term and a heteroscedastic error term (denoted as an MSIH(3)-AR(4) model — where the H flags the heteroscedastic error term, and 3 and 4 refer to the number of regimes and autoregressive lags, respectively):

$$y_t = \mu(s_t) + \sum_{k=1}^{4} \alpha_k y_{t-k} + \epsilon_t, \tag{11}$$

where $\epsilon_t \sim \mathsf{NID}(\sigma^2(s_t))$ and $s_t \in \{1, 2, 3\}$ is generated by a Markov chain.

Figure 3 and table 2 (reproduced from Clements and Krolzig, 1998) summarize the business-cycle characteristics of this model. The figure depicts the filtered and smoothed probabilities of the 'high growth' regime 3 and the contractionary regime 1 (the middle regime 2 probabilities are not shown). The expansion and contraction episodes produced by the three-regime model correspond fairly closely to the NBER classifications of business-cycle turning points. In contrast to the two-regime model, all three regimes are reasonably persistent.

While Hess and Iwata (1997b) find their three-state MS model estimated for 1949-92 fails to generate contractions of sufficient duration or depth, their estimated p_{11} is only 0.1267, while the lowest value in the MSIH models recorded in table 2 is over 0.78, which directly translates into a longer duration of the recession regime, and so we conjecture that the MSIH model may not have this shortcoming.

The tests for asymmetries in MSIH(3)-AR(4) models are recorded in tables 3 to 6 for various historical periods. For the Hamilton sample period, and an extended period that includes the second half of the eighties, the non-parametric skewness and model-based tests both indicate steepness asymmetries (see tables 3 and 4). Both approaches indicate steepness of expansions, with the MS-AR model test permitting rejection of the null at the 1% level. Moreover, there is clear evidence of asymmetric turning points (or sharpness), which results from a rejection of $p_{21} = p_{23}$, i.e., that moving from moderate to low growth is equally as likely as moving from moderate to high growth. The three-state model permits rejection of the non-sharpness hypotheses at a higher confidence level than the two-state model.

For the later sample period (see tables 5 and 6) the MS models continue to reject non-steepness at the 5% level, in contrast to the non-parametric models that now flag deepness rather than steepness. The major change in inference using the parametric tests is that there is no evidence of sharpness in the later periods.

[2] A number of authors, including Hess and Iwata (1995), Boldin (1996) and Clements and Krolzig (1998), have found that the 2-regime MS model does not yield a particularly good representation of the business cycle when fitted to periods outside that in Hamilton (1989). For example, Clements and Krolzig (1998) find an average duration of contraction (1) of 2–3 quarters for the period 1947–90, and of less than 2 quarters for 1959–96.

6.2 Krolzig and Toro's model of US output and employment

The Krolzig and Toro (1998) model of post-war US employment and output data is a cointegrated vector autoregressive Markov-switching process, where some parameters are changing according to the phase of the business and employment cycle. Employment and output are found to have a common cyclical component, and the long run dynamics are characterized by a proportional cointegrating vector between employment and output, with a trend included as a proxy for technological progress and capital accumulation.

More formally, the long-run relationship between output, y_t, and employment, n_t, is given by the cointegration vector $\beta' = (1 : -1)$ and the regime-dependent deviation from the trend in per-capita output $\mu(s_t) = \mathsf{E}[y_{t-1} - n_{t-1} - \gamma t]$. Then each regime m is associated with a particular attractor (μ_m, δ^*_m) given by the equilibrium growth rate δ^*_m and the equilibrium mean μ_m:

$$\begin{bmatrix} a_{11}(\mathsf{L}) & a_{12}(\mathsf{L}) \\ a_{21}(\mathsf{L}) & a_{22}(\mathsf{L}) \end{bmatrix} \begin{bmatrix} \Delta y_t - \delta^*(s_t) \\ \Delta n_t - \delta^*(s_t) \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} (y_{t-1} - n_{t-1} - \mu(s_t) - \gamma t) + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \quad (12)$$

where $u_t | s_t \sim \mathsf{NID}(\mathbf{0}, \Sigma(s_t))$. Thus the regime-dependent drift term $\delta^*(s_t)$ is the equilibrium growth rate, and shifts in the $\delta^*(s_t)$ map out changes in the business cycle state (e.g., expansion, contraction). The equilibrium mean $\mu(s_t)$ gives the state-dependent equilibrium level of labour productivity: shifts in $\mu(s_t)$ reflect changes in equilibrium per-capita output. As in the univariate MS models, the unobservable regime variable s_t is governed by a Markov chain with a finite number of states defined by the transition probabilities p_{ij}.

Short-run and long-run dynamics are jointly estimated in a Markov-switching vector-equilibrium-correction (MS-VECM) model with three regimes representing recession, growth and high growth. The ML estimation results for a first-order model estimated using data from 1969m2 to 1997m1 are presented in table 7, for the model parameterised as:

$$\begin{bmatrix} 1 - a_{11}L & -a_{12}L \\ -a_{21}L & 1 - a_{22}L \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \Delta n_t \end{bmatrix} = \nu(s_t) + \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} (y_{t-1} - n_{t-1} - \gamma t) + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}, \quad (13)$$

with $\nu(s_t) = \delta(s_t) - \alpha\mu(s_t)$.

Tests for asymmetries in the model of the vector process are given in table 7, and fail to detect any asymmetries. As noted in the introduction, we conjecture that this is because the lagged output terms in the MS-VECM model largely mop up the asymmetries in the employment series (and similarly for the lagged employment terms in the output equation)[3]. Table 9 provides some evidence in support of that conjecture. There we record the results for testing for asymmetries in a series of models that restrict the interaction between the variables permitted in the autoregressive or gaussian part of the model. The top panel relates to models with a homoscedastic error across regimes, the bottom allows regime-dependent error variances.

The bottom left column is the preferred model of Krolzig and Toro (1998), and is shown here as well as in table 7 to aid comparison of the results. The precise specification of each model is given in the notes following the table. Consider the top panel, and compare the VECM(1) (the homoscedastic analogue of Krolzig and Toro's preferred model) with the DVAR(0). The latter excludes any dependence between the series that may arise from the autoregressive part of the model. Then, we find not only evidence of

[3]Clements and Smith (1998) find that the non-linear terms in the Pesaran and Potter (1997) 'ceiling and floor' model of US output growth become less important when lagged unemployment rate terms are added to the model; a similar phenomenon to that alluded to here.

steepness in both output and employment changes (as for the VECM(1)) but also deepness in employment changes (at the 1% level) and in output changes at the 12% level.

Also apparent from the table is that systems models that allow for heteroscedasticity exhibit less evidence of asymmetries than those where homoscedasticity is imposed. A potential drawback with the testing procedures that we have set out is that no allowance is made for asymmetries that may emanate from, or be captured by, regime-dependent error variances. The tests are designed to detect asymmetries in the Markov chain component, μ_t, and do not rule out asymmetries that arise from the regime-dependent variances of the shocks. Figure 4 illustrates the data densities for variants of an MSMH(2)-AR(0) model:

$$y_t = \mu(s_t) + \epsilon_t, \tag{14}$$

where $\epsilon_t \sim \mathsf{NID}(\sigma^2(s_t))$ and $s_t \in \{1, 2\}$. Each panel plots the density of y_t generated by (14), and the density conditional on being in a regime. In the first panel, $\mu_1 = -\mu_2 = -1.5$, $\sigma_1^2 = \sigma_2^2 = 1$, and $prob(s_t = 1) = 0.5$. From the propositions stated in section 3.2, and in particular Corollary 1, it is apparent that the values of $\mu(s_t)$ and $\overline{\xi}_1$ satisfy the conditions for μ_t, and thus y_t, to exhibit non-deepness (non-steepness is a property of the model, and non-deepness implies non-sharpness in this model). The density exhibits skewness in the top right panel because the condition for non-deepness is not satisfied by $\mu_1 = -\mu_2 = -1.5$ and $prob(s_t = 1) = 0.3$. In the bottom left panel the condition for non-deepness is satisfied, because $\mu_1 = -\mu_2 = -1.5$ and $prob(s_t = 1) = 0.5$, but nonetheless heteroscedasticity, $\sigma_1^2 = 1$ and $\sigma_2^2 = 2$, induces skewness in μ_t. The final panel is akin to the top right but now with heteroscedastic errors. In the bottom left panel, then, the contribution of the Markov process is symmetric but the unequal variances result in asymmetry.

That regime-dependent variances can affect the skewness of the observed variables in practice is apparent for the MSIH(3)-AR(4) models starting in 1948 (see tables 3 and 4), where the observed growth rates display negative skewness (deepness of contractions) but the NonDeeepness test, though not significant, indicates positive skewness. In these models the third regime (high-growth) is associated with the Korea boom in 1951-1952, and induces positive skewness of the hidden Markov chain. However, the variance is much higher in regime 1 (recession), so that the observed variable is overall negatively skewed (but not significantly).

7 Conclusions

We have set out the parametric restrictions on MS-AR models for the series generated by those models to exhibit neither deepness, steepness or sharpness business-cycle asymmetries. For the popular two-state model first proposed by Hamilton (1989) we have shown that deepness implies sharpness and vice versa, and that the model (at least with gaussian disturbances) can not generate steepness. For three-state models, which arguably afford a better characterisation of the business cycle, the three concepts are distinct. We have shown how the parameter restrictions can be applied as Wald tests, and to illustrate, report the results of testing for asymmetries in Hamilton's original model, in three-state models, and in bivariate models of output and employment.

The univariate three-state output models generally indicate steepness and, on the earlier sample periods, sharpness, in that the probability of moving from moderate growth to recession is significantly larger than that of moving to high growth.

As expected, the well-specified systems models exhibited little indication of asymmetries, suggesting the latter result to a large extent from omitted variables. Consistent with this, if we exclude linear dependence between output and employment changes, as in the MS-DVAR(0) (with homoscedastic errors), then there is evidence of steepness and deepness asymmetries in both output and employment changes.

A comparison of the results of our tests with the non-parametric outcomes suggests our tests have reasonable power to detect asymmetries. In some cases the ML estimates of the MS-AR model parameters would appear to indicate asymmetries but the tests do not reject the null. For example, the matrix of the estimated transition probabilities may indicate asymmetries, but if the elements of the matrix are imprecisely measured due to there being few observations for some regimes the tests may lack power.

We cautioned about the potential dependence of the results on de-trending by first-differencing, although note that nothing in our approach precludes other methods of de-trending, and of the fact that no account is taken of possible asymmetries accounted for by regime-dependent heteroscedasticity.

References

Beaudry, P., and Koop, G. (1993). Do recessions permanently affect output. *Journal of Monetary Economics*, **31**, 149–163.

Boldin, M. D. (1996). A check on the robustness of Hamilton's Markov switching model approach to the economic analysis of the Business Cycle. *Studies in Nonlinear Dynamics and Econometrics*, **1**, 35–46.

Burns, A. F., and Mitchell, W. C. (1946). *Measuring Business Cycles*. New York: NBER.

Clements, M. P., and Krolzig, H.-M. (1998). A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP. *Econometrics Journal*, **1**, C47–75.

Clements, M. P., and Smith, J. (1998). Evaluating the forecast densities of linear and non-linear models: Applications to output growth and unemployment. mimeo, Department of Economics, University of Warwick.

Diebold, F. X., and Rudebusch, G. D. (1996). Measuring business cycles: A modern perspective. *The Review of Economics and Statistics*, **78**, 67–77.

Diebold, F. X., Rudebusch, G. D., and Sichel, D. E. (1993). Further evidence on business cycle duration dependence. In Stock, J., and Watson, M. (eds.), *Business Cycles, Indicators, and Forecasting*, pp. 255–280: Chicago: University of Chicago Press and NBER.

Doornik, J. A. (1996). *Object-Oriented Matrix Programming using Ox*. London: International Thomson Business Press and Oxford: http://www.nuff.ox.ac.uk/Users/Doornik/.

Falk, B. (1986). Further evidence on the asymmetric behaviour of economic time series over the business cycle. *Journal of Political Economy*, **94**, 1096–1109.

Filardo, A. J. (1994). Business–cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics*, **12**, 299–308.

Filardo, A. J., and Gordon, S. F. (1998). Business cycle durations. *Journal of Econometrics*, **85**, 99–123.

Gordon, S. (1997). Stochastic trends, deterministic trends, and business cycle turning points. *Journal of Applied Econometrics*, **12**, 411–434.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357–384.

Hamilton, J. D., and Lin, G. (1996). Stock market volatility and the business cycle. *Journal of Applied Econometrics*, **11**, 573–593.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, **45**, 39–70.

Hess, G. D., and Iwata, S. (1995). *Measuring Business Cycle Features*: University of Kansas, Research Papers in Theoretical and Applied Economics No. 1995-6.

Hess, G. D., and Iwata, S. (1997a). Asymmetric persistence in GDP? A deeper look at depth. *Journal of Monetary Economics*, **40**, 535–554.

Hess, G. D., and Iwata, S. (1997b). Measuring and comparing business-cycle features. *Journal of Business and Economic Statistics*, **15**, 432–444.

Hicks, H. (1950). *A contribution to the theory of the trade cycle*. Oxford: Clarendon Press.

Krolzig, H.-M. (1997). *Markov Switching Vector Autoregressions: Modelling, Statistical Inference and Application to Business Cycle Analysis*: Lecture Notes in Economics and Mathematical Systems, 454. Springer-Verlag, Berlin.

Krolzig, H.-M., and Sensier, M. (1998). A disaggregated Markov-Switching model of the Business Cycle in UK manufacturing. mimeo, Institute of Economics and Statistics, University of Oxford.

Krolzig, H.-M., and Toro, J. (1998). A new approach to the analysis of shocks and the cycle in a model of output and employment. Discussion paper, Institute of Economics and Statistics, University of Oxford.

McQueen, G., and Thorley, S. (1993). Asymmetric business cycle turning points. *Journal of Monetary Economics*, **31**, 341–362.

Mitchell, W. C. (1927). *Business Cycles: The Problem and its Setting*. New York: National Bureau of Economic Research.

Neftci, S. N. (1984). Are economic time series asymmetric over the business cycle?. *Journal of Political Economy*, **92**, 307–328.

Pesaran, M. H., and Potter, S. M. (1997). A floor and ceiling model of US Output. *Journal of Economic Dynamics and Control*, **21**, 661–695.

Ravn, M. O., and Sola, M. (1995). Stylized facts and regime changes: Are prices procyclical?. *Journal of Monetary Economics*, **36**, 497–526.

Sichel, D. E. (1989). Are business cycles asymmetric? A correction. *Journal of Political Economy*, **97**, 1255–1260.

Sichel, D. E. (1991). Business cycle duration dependence: A parametric approach. *The Review of Economics and Statistics*, **73**, 254–260.

Sichel, D. E. (1993). Business cycle asymmetry. *Economic Inquiry*, **31**, 224–236.

Sichel, D. E. (1994). Inventories and the three phases of the business cycle. *Journal of Business and Economic Statistics*, **12**, 269–277.

Speight, A. E. H., and McMillan, D. G. (1998). Testing for asymmetries in UK macroeconomic time series. *Scottish Journal of Political Economy*, **45**, 158–170.

8 Appendix

Proposition 1. *An MSM(M)-AR(p) process is* **non-deep** *iff:*

$$\sum_{m=1}^{M}\bar{\xi}_m\mu_m^{*3}=\sum_{m=1}^{M-1}\bar{\xi}_m\mu_m^{*3}+\left(1-\sum_{m=1}^{M-1}\bar{\xi}_m\right)\mu_M^{*3}=0 \tag{15}$$

with $\mu_m^*=\mu_m-\mu_y=\sum_{i\neq m}(\mu_m-\mu_i)\,\bar{\xi}_i$ *and where* $\bar{\xi}_m$ *is the unconditional probability of regime* m.

Proof. MSM(M)-AR(p) processes can be rewritten as the sum of two independent processes: $y_t-\mu_y=\mu_t+z_t$. μ_y is the unconditional mean of y_t:

$$\mu_y=\mathsf{E}\left[y_t\right]=\sum_{m=1}^{M}\bar{\xi}_m\mu_m,$$

where $\bar{\xi}_m$ is the unconditional probability of regime m. Both z_t and μ_t are zero mean, $\mathsf{E}[\mu_t]=\mathsf{E}[z_t]=0$. While the process $z_t=\sum_{j=1}^{p}\alpha_j z_{t-j}+u_t$ is gaussian and hence symmetric, the other component, μ_t, is potentially asymmetric, and represents the contribution of the Markov chain:

$$\mu_t=\sum_{m=1}^{M}\xi_{mt}\left(\mu_m-\mu_y\right)=\sum_{m=1}^{M}\xi_{mt}\mu_m^*=\mu_M^*+\sum_{m=1}^{M-1}\xi_{mt}\left(\mu_m^*-\mu_M^*\right)$$

with $\mu_m^*=\mu_m-\mu_y$ and $\xi_{mt}=1$ if the regime is m at period t, and is 0 otherwise.

Thus the k-th moment of μ_t is given by:

$$\mathsf{E}\left[\mu_t^k\right]=\sum_{m=1}^{M}\bar{\xi}_m\left(\mu_m-\mu_y\right)^k=\sum_{m=1}^{M}\bar{\xi}_m\mu_m^{*k}.$$

Using the adding-up restriction, $\sum_{m=1}^{M}\bar{\xi}_m=1$, we have:

$$\mathsf{E}\left[\mu_t^k\right]=\sum_{m=1}^{M-1}\bar{\xi}_m\mu_m^{*k}+\left(1-\sum_{m=1}^{M-1}\bar{\xi}_m\right)\mu_M^{*k}=\mu_M^{*k}+\sum_{m=1}^{M-1}\bar{\xi}_m\left(\mu_m^{*k}-\mu_M^{*k}\right).$$

□

Proposition 2. *An MSM(M)-AR(p) process is* **non-steep** *if the size of the jumps,* $\mu_j-\mu_i$*, satisfies the following condition:*

$$\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\left(\bar{\xi}_i p_{ij}-\bar{\xi}_j p_{ji}\right)\left[\mu_j-\mu_i\right]^3=0. \tag{16}$$

Proof. Write $\mu_t=\mathbf{M}\xi_t$, where $\mathbf{M}=[\mu_1\cdots\mu_M]$ and $\xi_t=[\xi_{1t}\cdots\xi_{Mt}]'$. $\xi_{mt}=1$ if the period t regime is m, and zero otherwise. Then $\Delta\mu_t=\mu_t-\mu_{t-1}=\mathbf{M}\boldsymbol{\Delta}\xi_t=\mathbf{M}\xi_t-\mathbf{M}\xi_{t-1}$. Clearly, $\mathsf{E}[\Delta\mu_t]=0$. We now introduce $\nabla\mathbf{M}=\left[\mathbf{M}'\otimes 1_M-1_M\otimes\mathbf{M}'\right]'$ and $\xi_t^{(2)}=\xi_t\otimes\xi_{t-1}$, such that:

$$\Delta\mu_t=\nabla\mathbf{M}\xi_t^{(2)}=\sum_{i=1}^{M}\sum_{j=1}^{M}\xi_{i,t-1}\xi_{j,t}\left[\mu_j-\mu_i\right].$$

Using that $\mu_j-\mu_i=0$ for $i=j$, we can simplify to:

$$\Delta\mu_t=\sum_{i=1}^{M}\sum_{j\neq i}\xi_{i,t-1}\xi_{j,t}\left[\mu_j-\mu_i\right].$$

The third moment is then given by:

$$\begin{aligned} \mathsf{E}\left[\Delta\mu_t^3\right] &= \sum_{i=1}^{M}\sum_{j\neq i}\bar{\xi}_i p_{ij}\left[\mu_j-\mu_i\right]^3 \\ &= \sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\left\{\left(\bar{\xi}_i p_{ij}-\bar{\xi}_j p_{ji}\right)\left[\mu_j-\mu_i\right]^3\right\}, \end{aligned}$$

where the last line uses $[\mu_j - \mu_i]^3 = -[\mu_i - \mu_j]^3$.

Symmetry of the matrix of transition parameters (which is stronger than the definition of sharpness) is sufficient for non-steepness as it implies that, for all $i, j = 1, \ldots, M$, we have that $\bar{\xi}_i p_{ij} - \bar{\xi}_j p_{ji} = 0$.

□

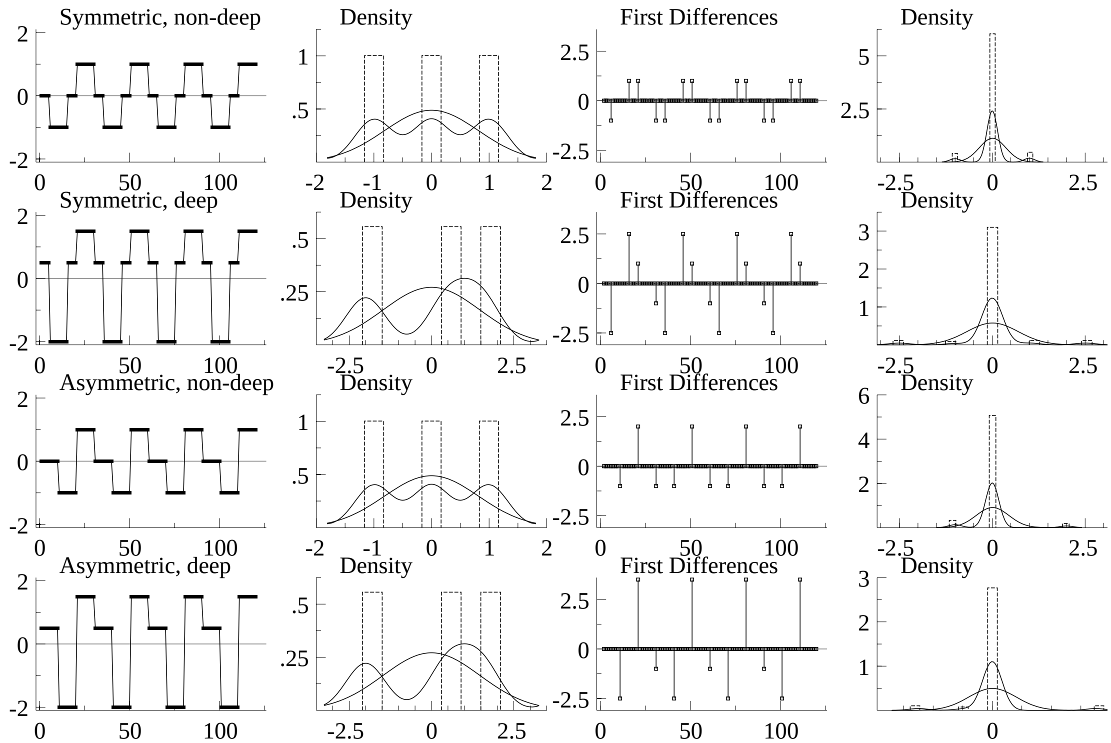

Figure 1 Asymmetries in MS(3)-AR(p) Models.

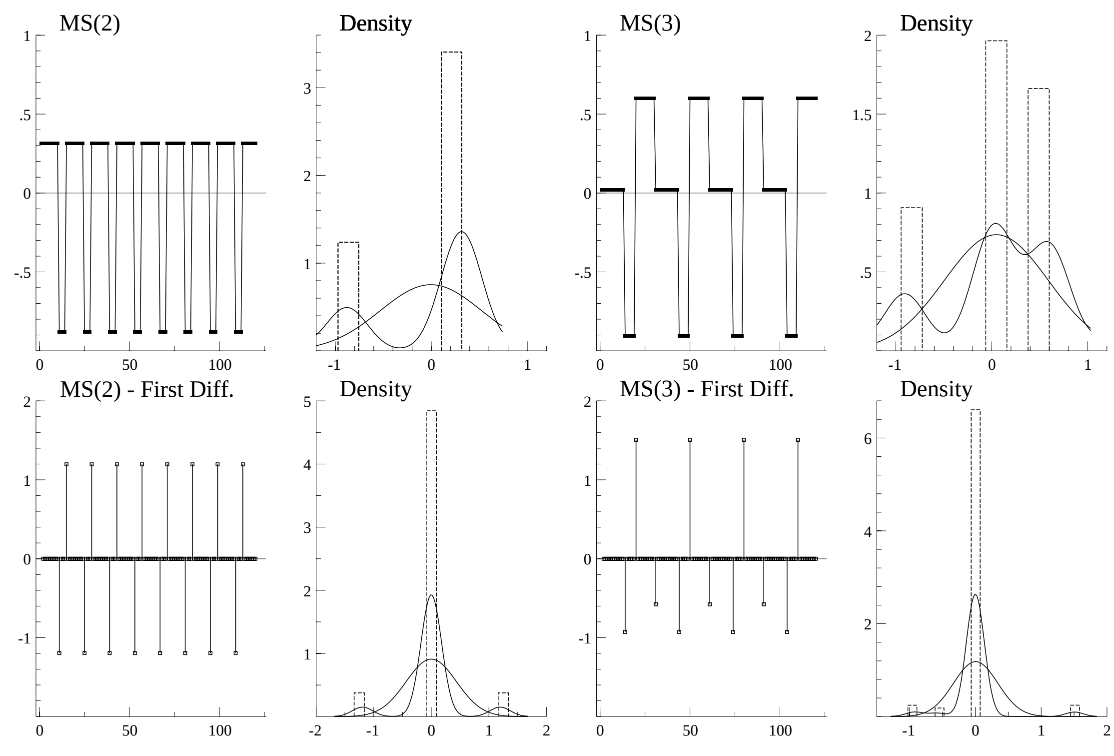

Figure 2 Business Cycle Asymmetries in Hamilton’s model and in a three-regime MS-AR Model.

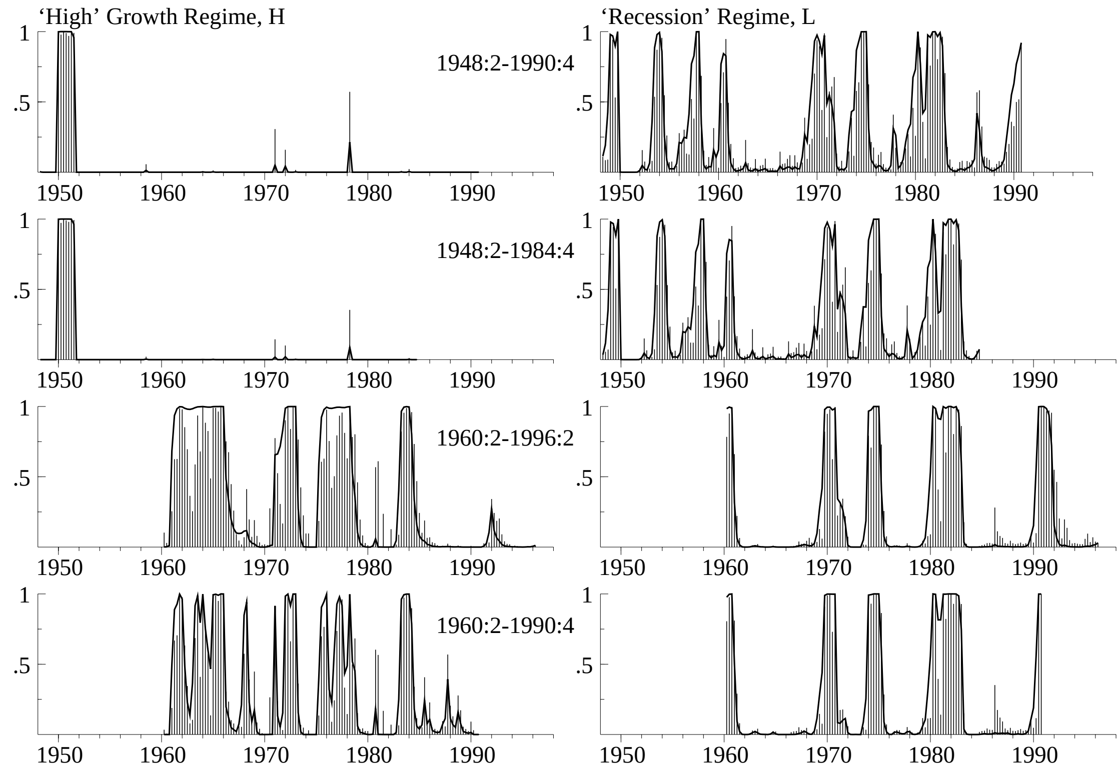

Figure 3 MSIH(3)-AR(4) model smoothed and filtered probabilities of the 'extreme' regimes, H, L.

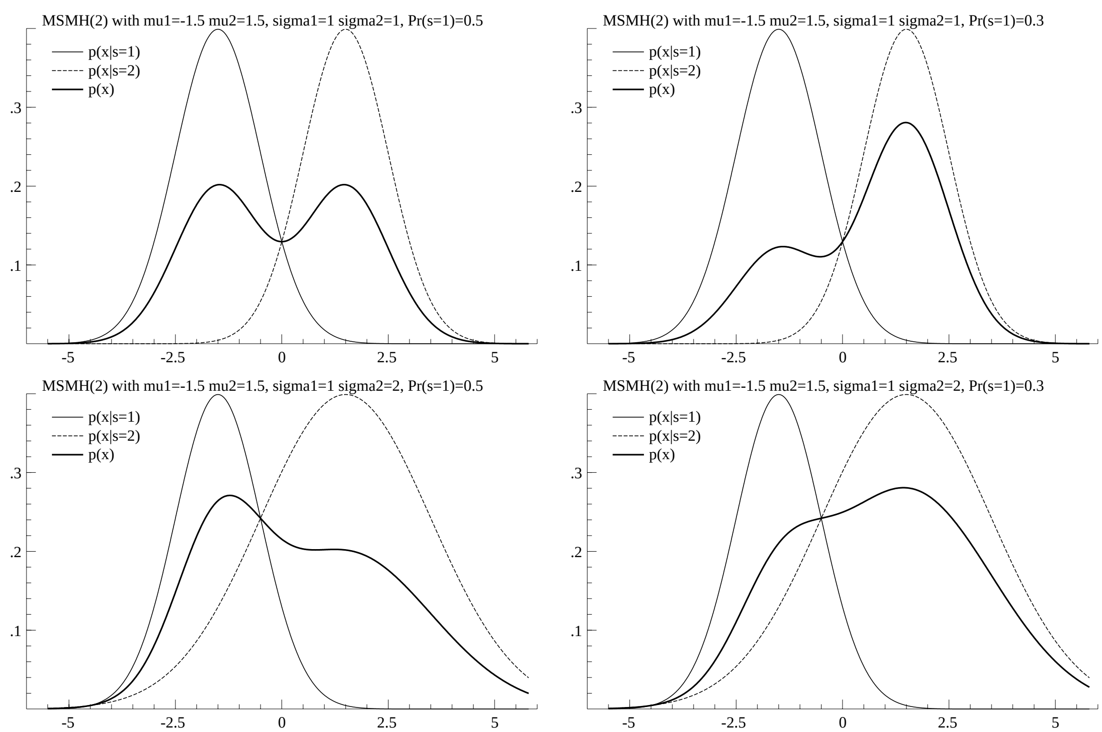

Figure 4 Asymmetries due to regime-dependent heteroscedasticity.

Table 1 Tests for asymmetries in the MSM(2)-AR(4) (Hamilton), 1952:2 – 1984:4.

NonSharpness test:		Chi(1) =	2.9682	[0.0849] *
NonDeepness test:	-0.3148†	Chi(1) =	1.6615	[0.1974]
Skewness (Δy):	-0.4900	Chi(1) =	5.2025	[0.0226] **
Skewness ($\Delta^2 y$):	-0.0015	Chi(1) =	0.0000	[0.9945]

Note: Non-steepness is a property of the model: see Corollary 1.
Here and in subsequent tables * indicates significance at 10% level and ** significance at the 5% level.
† is the value of $\phi(\lambda)$, to flag positive or negative skewness. In subsequent tables the corresponding quantity for the test of NonSteepness is also recorded.

Table 2 MSIH(3)-AR(4) Models.

Sample	48:2-84:4	48:2-90:4	60:2-90:4	60:2-96:2
Mean μ_3	-0.169	-0.081	-0.029	-0.050
Mean μ_2	1.463	1.413	0.904	0.838
Mean μ_1	3.447	3.430	1.463	1.406
α_1	-0.110	-0.102	0.001	0.016
α_2	0.110	0.109	0.021	0.022
α_3	-0.169	-0.172	-0.133	-0.100
α_4	-0.194	-0.191	-0.091	-0.098
σ_1^2	0.848	0.816	0.869	0.796
σ_2^2	0.534	0.496	0.118	0.115
σ_3^2	0.017	0.018	0.402	0.406
Trans.prob p_{12}	0.205	0.187	0.000	0.021
Trans.prob p_{13}	0.024	0.021	0.132	0.128
Trans.prob p_{21}	0.097	0.093	0.104	0.075
Trans.prob p_{23}	0.000	0.000	0.007	0.000
Trans.prob p_{31}	0.000	0.000	0.000	0.000
Trans.prob p_{32}	0.162	0.162	0.087	0.091
Uncond.prob.1	0.285	0.298	0.264	0.231
Uncond.prob.2	0.673	0.663	0.334	0.447
Uncond.prob.3	0.042	0.038	0.402	0.322
Duration 1	4.372	4.814	7.563	6.745
Duration 2	10.318	10.701	9.576	13.089
Duration 3	6.154	6.145	11.544	10.946
Observations	147	171	123	145

Table 3 Tests for asymmetries in the MSIH(3)-AR(4), 1948:2 – 1984:4 .

NonSharpness test:		Chi(3) =	34141.9811	[0.0000] **
$p_{12} = p_{32}$:		Chi(1) =	0.0626	[0.8025]
$p_{13} = p_{31}$:		Chi(1) =	0.0393	[0.8429]
$p_{21} = p_{23}$:		Chi(1) =	34036.1776	[0.0000] **
NonDeepness test:	0.0353	Chi(1) =	0.0082	[0.9277]
NonSteepness test:	0.2376	Chi(1) =	15.9515	[0.0001] **
Skewness (Δy):	-0.1858	Chi(1) =	0.8403	[0.3593]
Skewness ($\Delta^2 y$):	0.3510	Chi(1) =	2.9980	[0.0834] *

Note: As p_{31} and p_{23} are close to zero, the matrix of second derivatives used for the calculation of parameter covariance is singular and the generalized inverse has been used, which explains the magnitude of the test statistics for non-sharpness.

Table 4 Tests for asymmetries in the MSIH(3)-AR(4), 1948:2 – 1990:4 .

NonSharpness test:		Chi(3) =	31828.7577	[0.0000] **
$p_{12} = p_{32}$:		Chi(1) =	0.0225	[0.8808]
$p_{13} = p_{31}$:		Chi(1) =	0.0417	[0.8382]
$p_{21} = p_{23}$:		Chi(1) =	31701.5198	[0.0000] **
NonDeepness test:	0.1271	Chi(1) =	0.1820	[0.6696]
NonSteepness test:	0.1966	Chi(1) =	17.7305	[0.0000] **
Skewness (Δy):	-0.1587	Chi(1) =	0.7055	[0.4010]
Skewness ($\Delta^2 y$):	0.3451	Chi(1) =	3.3736	[0.0662] *

Note: As p_{31} and p_{23} are close to zero, the matrix of second derivatives used for the calculation of parameter covariance is singular and the generalized inverse has been used, which explains the magnitude of the test statistics for non-sharpness.

Table 5 Tests for asymmetries in the MSIH(3)-AR(4), 1960:2 – 1990:4 .

NonSharpness test:		Chi(3) =	0.1474	[0.9856]
$p_{12} = p_{32}$:		Chi(1) =	0.1211	[0.7278]
$p_{13} = p_{31}$:		Chi(1) =	0.0217	[0.8830]
$p_{21} = p_{23}$:		Chi(1) =	0.0043	[0.9478]
NonDeepness test:	-0.1214	Chi(1) =	0.5352	[0.4644]
NonSteepness test:	0.0814	Chi(1) =	4.3970	[0.0360] **
Skewness (Δy):	-0.6775	Chi(1) =	9.3336	[0.0022] **
Skewness ($\Delta^2 y$):	0.2215	Chi(1) =	0.9972	[0.3180]

Table 6 Tests for asymmetries in the MSIH(3)-AR(4), 1960:2 – 1996:2 .

NonSharpness test:		Chi(3) =	0.7539	[0.8605]
$p_{12} = p_{32}$:		Chi(1) =	0.7271	[0.3938]
$p_{13} = p_{31}$:		Chi(1) =	0.0225	[0.8809]
$p_{21} = p_{23}$:		Chi(1) =	0.0045	[0.9462]
NonDeepness test:	-0.0839	Chi(1) =	0.4681	[0.4939]
NonSteepness test:	0.0649	Chi(1) =	4.2702	[0.0388] *
Skewness (Δy):	-0.6412	Chi(1) =	9.8659	[0.0017] **
Skewness ($\Delta^2 y$):	0.1874	Chi(1) =	0.8426	[0.3587]

Table 7 Krolzig and Toro US output and employment model: Summary Table.

	MSI(3)-VECM(1)	
	Δy_t	Δn_t
Regime-dependent intercepts		
ν_1	-0.2119	-0.0522
	.2769	*.1115*
ν_2	0.6203	0.2368
	.1143	*.0428*
ν_3	1.1914	0.4256
	.2022	*.0807*
Autoregressive coefficients		
Δy_{t-1}	-0.0019	0.0295
	.0960	*.0355*
Δn_{t-1}	0.1111	0.5357
	.1692	*.0598*
Adjustment coefficients		
α	-0.0038	0.0666
	.0645	*.0225*
Regime 1: Variance		
Δy	1.0861	0.3979
Δn	.8468	0.2033
Regime 2: Variance		
Δy	0.1826	0.0229
Δn	.3940	0.0184
Regime 3: Variance		
Δy	0.6055	0.1461
Δn	.6403	0.0860
	Erg.Prob	Duration
Regime 1	0.2050	5.898
Regime 2	0.5131	19.472
Regime 3	0.2819	10.205

Table 8 Tests for asymmetries in the MSI(3)-VECM(1), 1960:3 – 1997:1.

NonSharpness test:		Chi(3) =	1.9298	[0.5871]	
$p_{12} = p_{32}$:		Chi(1) =	0.2307	[0.6310]	
$p_{13} = p_{31}$:		Chi(1) =	1.7608	[0.1845]	
$p_{21} = p_{23}$:		Chi(1) =	0.0853	[0.7702]	
NonDeepness test (Δy):	-0.0588	Chi(1) =	0.1786	[0.6726]	
NonSteepness test (Δy):	0.0377	Chi(1) =	2.3960	[0.1216]	
NonDeepness test (Δn):	-0.0020	Chi(1) =	0.1125	[0.7373]	
NonSteepness test (Δn):	0.0014	Chi(1) =	1.4424	[0.2298]	
Skewness (Δy):	-0.4624	Chi(1) =	4.9178	[0.0266]	**
Skewness ($\Delta^2 y$):	0.1477	Chi(1) =	0.5016	[0.4788]	
Skewness (Δn):	-1.0625	Chi(1) =	25.9637	[0.0000]	**
Skewness ($\Delta^2 n$):	0.5329	Chi(1) =	6.5324	[0.0106]	**

Table 9 A Comparison of Business Cycle Asymmetries in the Bivariate System, 1960:2 – 1996:2.

MSI(3)	VECM(1)	VECM(1,0)	VECM(0)	DVAR(1)	DVAR(0)
NonSharpness test:	3.1710 [0.3660]	5.1372 [0.1620]	1.4544 [0.6928]	3.5032 [0.3203]	0.9484 [0.8137]
$p_{12} = p_{32}$:	2.9419 [0.0863]*	4.5876 [0.0322]**	0.3596 [0.5487]	3.4565 [0.0630]*	0.6177 [0.4319]
$p_{13} = p_{31}$:	0.1658 [0.6839]	0.2692 [0.6039]	0.8916 [0.3450]	0.0119 [0.9132]	0.0004 [0.9834]
$p_{21} = p_{23}$:	0.0057 [0.9399]	0.7279 [0.3936]	0.0440 [0.8338]	0.0061 [0.9377]	0.2467 [0.6194]
NonDeepness test ($\Delta\, n$):	0.7355 [0.3911]	0.2010 [0.6539]	0.8405 [0.3593]	1.7913 [0.1808]	12.2866 [0.0005]**
NonSteepness test ($\Delta\, n$):	9.5860 [0.0020]**	5.0429 [0.0247]**	14.6834 [0.0001]**	9.2695 [0.0023]**	28.0373 [0.0000]**
NonDeepness test ($\Delta\, y$):	0.5242 [0.4691]	0.0786 [0.7791]	0.2104 [0.6465]	0.4122 [0.5209]	2.4090 [0.1206]
NonSteepness test ($\Delta\, y$):	9.8849 [0.0017]**	3.4285 [0.0641]*	0.7490 [0.3868]	9.3314 [0.0023]**	2.9357 [0.0866]*
MSIH(3)	VECM(1)	VECM(1,0)	VECM(0)	DVAR(1)	DVAR(0)
NonSharpness test:	1.9299 [0.5871]	1.9387 [0.5852]	0.8690 [0.8329]	1.8665 [0.6006]	0.9790 [0.8063]
$p_{12} = p_{32}$:	0.2314 [0.6305]	0.2568 [0.6124]	0.6576 [0.4174]	0.3543 [0.5517]	0.5805 [0.4461]
$p_{13} = p_{31}$:	1.7598 [0.1847]	1.7612 [0.1845]	0.0308 [0.8606]	1.6774 [0.1953]	0.0922 [0.7614]
$p_{21} = p_{23}$:	0.0862 [0.7690]	0.0729 [0.7871]	0.3962 [0.5290]	0.0820 [0.7746]	0.2722 [0.6019]
NonDeepness test ($\Delta\, n$):	0.1125 [0.7373]	0.1097 [0.7405]	1.0449 [0.3067]	0.5046 [0.4775]	0.8304 [0.3622]
NonSteepness test ($\Delta\, n$):	1.4424 [0.2298]	1.9147 [0.1664]	12.6926 [0.0004]**	1.6250 [0.2024]	10.9591 [0.0009]**
NonDeepness test ($\Delta\, y$):	0.1786 [0.6726]	0.1916 [0.6616]	0.0695 [0.7921]	0.2673 [0.6052]	0.0082 [0.9277]
NonSteepness test ($\Delta\, y$):	2.3960 [0.1216]	2.8544 [0.0911]*	3.0369 [0.0814]*	2.5050 [0.1135]	1.0399 [0.3078]

Note: VECM(1,0) denotes a VECM(1) with $a_{12} = a_{21} = 0$. The VECM(0) imposes the further restrictions $a_{11} = a_{22} = 0$. The DVAR(1) is the VECM(1) with $\alpha = 0$, and the DVAR(0) simply relates the first difference of each variable to a regime-dependent intercept.