Tests for business cycle asymmetries are developed for Markov-switching autoregressive models. The tests of deepness, steepness, and sharpness are Wald statistics, which have standard asymptotics. For the standard two-regime model of expansions and contractions, deepness is shown to imply sharpness (and vice versa), whereas the process is always nonsteep. Two and three-state models of U.S. GNP growth are used to illustrate the approach, along with models of U.S. investment and consumption growth. The robustness of the tests to model misspecification, and the effects of regime-dependent heteroscedasticity, are investigated.

KEY WORDS: Deepness; Regime-switching; Steepness and sharpness; Wald 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.
—W. C. 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) defined an asymmetric cycle as “one in which some phase of the cycle is different from the mirror image of the opposite phase.” McQueen and Thorley (1993, pp. 342–343) and Sichel (1993, pp. 225–226) discussed the importance, from both theoretical and empirical viewpoints, of establishing whether there are asymmetries in the business cycle. The finding of asymmetry is compatible with a number of business cycle models, but would rule out linear models with symmetric errors. Sichel noted that models of asymmetric price adjustment can generate deepness (defined later), and referenced work of De Long and Summers (1998) and Ball and Mankiw (1994). More recently, the “output-gap” literature (see, e.g., Laxton, Meredith, and Rose 1995; Clark, Laxton, and Rose 1996) suggested that more of the adjustment follows. A stationary time series \( \{x_t\} \) is assumed to have been generated by an AR(\( p \)) process, providing parametric tests as alternatives to the NP tests typically used in the literature. Our tests are able to detect asymmetries in the propagation mechanisms of shocks, or first-moment asymmetries, whereas NP tests are unable to discriminate between first-moment asymmetries and asymmetries in the shocks. Since the work of Hamilton (1989), the MS-AR model class has been used extensively in the empirical macroeconomics literature to analyze business cycle phenomena, and availability of good estimation and inferential procedures makes it an obvious choice for the development of parametric tests of asymmetry.

The basic MS-AR model at the center of our analysis is as follows. A stationary time series \( \{x_t\} \) is assumed to have been generated by an AR(\( p \)) with \( M \) MS regimes in the mean of the process, which we label an MSM(\( M \))-AR(\( p \)) process,

\[
x_t - \mu(s_t) = \sum_{k=1}^{p} \alpha_k (x_{t-k} - \mu(s_{t-k})) + u_t,
\]

\[
u_t | s_t \sim \text{NID}(0, \sigma^2). \tag{1}
\]

We can order the regimes by the magnitude of \( \mu \) such that \( \mu_1 < \cdots < \mu_M \). The Markov chain is ergodic, irreducible, and there does not exist an absorbing state, that is, \( \tilde{\xi}_m \in (0,1) \) for all \( m = 1, \ldots, M \), where \( \tilde{\xi}_m \) is the ergodic or unconditional probability of regime \( m \). The transition probabilities
are time-invariant,

$$p_{ij} = \Pr(s_{t+1} = j \, | \, s_t = i),$$

$$\sum_{j=1}^M p_{ij} = 1, \quad \forall i, j \in \{1, \ldots, M\}, \quad (2)$$

so that the probability of a switch between regimes $i$ and $j$ does not depend on how long the process has been in regime $i$. A number of authors (including Diebold et al. 1993, Filardo 1994, Filardo and Gordon 1998) have extended the Hamilton (1989) model to allow for time-varying transition probabilities. Generally there appears to be positive duration dependence in contractions in the U.S. post-World War II period, so that the probability of moving out of recession increases with the duration of recession. However, nonconstant transition probabilities would complicate derivation of the SDS tests.

The MS-AR framework can be readily extended to multivariate settings (see, e.g., Ravn and Sola 1995; Diebold and Rudebusch 1996; Hamilton and Lin 1996; Krolzig 1997; Krolzig and Sensier 2000; Krolzig and Toru 1998), which is a distinct advantage given that business cycles were originally viewed as consisting of comovements in many economic variables (see, e.g., Burns and Mitchell 1946). In (1), $x_t$ can be a vector of variables. The extension of our tests to multivariate settings, and models that include regime-dependent heteroscedasticity and switches in intercepts, are discussed in Section 4.

Along with constructing tests of asymmetries, a related goal is to establish precisely which types of asymmetries that MS-AR models are capable of generating in principle, given the widespread popularity of these models in applied research and the apparent confusion on this point in the literature. For example, Sichel (1993, p. 232, footnote 19) stated that the Hamilton (1989) two-state model implies steepness in U.S. GNP. In fact, steepness (as defined formally later) cannot arise in such a model. Recently, Hess and Iwata (1997b) have investigated by simulation whether empirically estimated models, including MS-AR models, are able to replicate the “fundamental business cycle features” of observed durations and amplitudes of contractions and expansions. Our focus is different, because we ask whether in principle MS-AR models can generate certain asymmetric features, in addition to whether empirically estimated models have these features.

Finally, because measures of economic activity exhibit secular increases, the asymmetries relate to the detrended log of output. For example, Speight and McMillan (1998) considered the detrended component ($y_t$) of the variable $x_t$, where $x_t = y_t - \tau_t$. Here $\tau_t$ is a nonstationary trend component, and $x_t$ is stationary, possibly consisting of cycle and noise components. Throughout the article, $x_t$ refers to the detrended series. We assume that the nonstationarity can be removed by differencing, that is, $x_t = \Delta y_t$. Trend elimination by differencing is natural in our setup, because the MS-AR model is typically estimated on the first difference of the log of output. An exception is the approach of Lam (1990), which allows for a general AR process in the level of the log of output, rather than imposing a unit root. However, none of the propositions on asymmetries in MS-AR models that follow, nor the testing procedures, requires this method of detrending, and all remain valid whichever method is used. All that we require is that a MS-AR model can be estimated for the detrended series, however obtained. The sensitivity of the findings of asymmetries to the method of trend elimination requires further research. Gordon (1997) showed that in general the model of the short-run fluctuations in output may depend on the treatment of the trend component.

The article is organized as follows. Section 2 reviews the literature on testing for business cycle asymmetries and shows how (the absence of) SDS asymmetries can be mapped into parameter restrictions of MS-AR models, paying particular attention to the empirically relevant two- and threeregime models. Then, Section 3 derives the Wald tests of SDS hypotheses. Wald tests obviate the need to estimate the restricted (null) MS-AR model and are attractive for that reason. Section 4 shows how the basic testing approach can be extended in a number of directions. Section 5 uses Monte Carlo simulations to investigate the small-sample properties of the tests, their performance in the presence of heteroscedasticity, and their robustness to model misspecification. Section 6 sets out the empirical illustrations, and Section 7 concludes.

2. BUSINESS CYCLE ASYMMETRIES

2.1 A Brief Review of the Literature on Business Cycle Asymmetries

2.1.1 Steepness and Deepness. Sichel (1993) distinguished 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 (or is shallower than) that of peaks. The top left panel of Figure 1 depicts a schematic business cycle that is nondeep and nonsteepest. The second panel in the first column shows deepness of troughs (but nonsteepness), the third panel shows steepness of expansions (but nondeepness), and the last panel shows both properties.

A number of ways of testing for steepness and deepness have been proposed in the literature. Neftci (1984) proposed a test of whether there are longer runs of increases than decreases in a series, indicating that the length of expansions exceeds that of contractions, so that contractions are necessarily steeper than expansions. He defined an indicator variable $I_{t-1} = 1$ if $x_t > 0$ (expansion) and $I_{t-1} = -1$ if $x_t \leq 0$ (recession). Suppose that $I_t$ can be represented by a second-order Markov process; then $p_{22} > p_{11} \quad (where \quad p_{22} = \Pr(I_t = 1 \, | \, I_{t-1} = 1, \, I_{t-2} = 1) \, and \, p_{11} = \Pr(I_t = -1 \, | \, I_{t-1} = -1, \, I_{t-2} = -1))$ implies a form of cyclical asymmetry because the length of expansions exceeds that of contractions. A possible 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. Using this approach, Neftci found evidence of steepness in postwar U.S. unemployment during contractions. In contrast, Falk (1986) failed to find evidence of steepness in other U.S. 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. Rothman (1991) found evidence of asymmetry in the quarterly unemployment rate series using a modified version of Neftci’s test, and Sichel (1989) found strong evidence of asymmetry in annual unemployment, for which measurement error is presumably less important. Luukkonen and Teräsvirta (1991) noted that self-exciting threshold AR models (see, e.g., Tong and Lim 1980; Tong 1995) and smooth-transition AR models (see, e.g., Luukkonen, Saikkonen, and Teräsvirta 1988; Teräsvirta and Anderson 1992) may imply cyclical asymmetry in this sense, in that the probabilities of remaining in the regimes, once entered, may not be equal due to different dynamic structures.

Sichel (1993) suggested a test of steepness based on the coefficient of skewness calculated for the detrended series \( \{x_t\} \). Deepness of contractions will show up as negative skewness, because it implies that the average deviation of observations below the mean will exceed that of observations above the mean. Steepness (of expansions) implies positive skewness in the first difference of the detrended series, \( \{\Delta x_t\} \); increases should be larger, though less frequent, than decreases. Figure 1 illustrates this. The second column depicts the histograms (and densities) for \( \{\Delta x_t\} \) corresponding to the schematic business cycles in the first column. The densities are symmetric for the first and second rows because the business cycles are nonsteep; those in the third and fourth rows exhibit positive skewness because of the steepness of expansions.

On the basis of these tests, deepness is found to characterize quarterly postwar U.S. unemployment and industrial production, with weaker evidence for GNP, whereas only unemployment (of the three) appears to exhibit steepness. We also report these NP tests based on the coefficients of skewness in our empirical work. In addition, the concepts of nondeepness and nonsteepness are used to construct parametric tests based on the MS-AR model in (1), which we turn to after discussing the third notion of asymmetry.

2.1.2 Sharpness. Sharpness, or TP asymmetry, as introduced by McQueen and Thorley (1993), result if, for example, troughs are “sharp” and peaks more “rounded.” McQueen and Thorley presented two tests. The first test is based on the magnitude of growth rate changes around National Bureau of Economic Research–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 found the null of equal TP 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 distinguished between contraction, moderate, and high (recovery) states. The hypothesis of Hicks (1950) that troughs are sharper than peaks corresponds to the probability of jumping from the contraction to high-growth state exceeding the probability of jumping directly from high growth to contraction. “Complete” TP symmetry requires that these switches be equally likely, that switches to moderate growth from contraction and from high growth be equally likely, and that movements to high growth and contraction from moderate growth be also equally likely. McQueen and Thorley again found evidence of sharpness asymmetry for postwar unemployment and industrial production, but the susceptibility of the test to noise became evident when they considered prewar industrial production and postwar agricultural unemployment. In both cases, quarterly volatility in the series interrupted runs of ones and threes, reducing the number of sharp TPs and the power of the test. Their second approach can be implemented directly in a MS-AR model.

2.2 Formal Definitions of Asymmetries

For clarity, we formally define the concepts of steepness, deepness, and sharpness.

**Definition 1:** Deepness (Sichel 1993). The process \( \{ x_t \} \) is said to be nondeep (nontall) iff \( x_t \) is not skewed,

\[
E[(x_t - \mu_x)^3] = 0.
\]

Analogously, we can define steepness as skewness of the differences.

**Definition 2:** Steepness (Sichel 1993). The process \( \{ x_t \} \) is said to be nonsteep iff \( \Delta x_t \) is not skewed,

\[
E[\Delta x_t^3] = 0.
\]

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

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

\[
P_{m3} = P_{m2} \quad \text{and} \quad P_{1m} = P_{Mm},
\]

for all \( m \neq 1, M \), and \( P_{1M} = P_{M1} \).

In a two-regime model, for example, nonsharpness implies that \( p_{12} = p_{21} \). In a three-regime model, it requires \( p_{13} = p_{31} \) as well as \( p_{12} = p_{23} \) and \( p_{21} = p_{32} \). When \( M = 4 \), the following restrictions on the matrix of transition probabilities are required to hold for nonsharpness,

\[
P = \begin{bmatrix}
1 - a - b - c & a & b & c \\
0 & d & * & * \\
0 & e & * & * \\
c & a & b & 1 - a - b - c
\end{bmatrix},
\]

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

According to Definition 1, deepness implies skewness. Using the properties of the MS-AR defined in (1) and (2), the following necessary and sufficient moment condition results.

**Proposition 1.** An MSM(1)-AR(p) process is nondeep iff

\[
\sum_{m=1}^{M} \xi_m \mu_m^3 = \sum_{m=1}^{M-1} \xi_m \mu_m^3 + \left(1 - \sum_{m=1}^{M-1} \xi_m \right) \mu_M^3 = 0 \quad (4)
\]

with \( \mu_m^3 = \mu_m - \mu_x = \sum_{i \neq m} (\mu_x - \mu_i) \xi_i \), where \( \xi_m \) is the unconditional probability of regime \( m \) and \( \mu_x \) is the unconditional mean of \( x_t \).

The expression (4) is a complicated third-order polynomial in the regime-dependent parameters of the process, \( \mu_1, \ldots, \mu_M \), and the unconditional regime probabilities, \( \xi_1, \ldots, \xi_{M-1} \), which are nonlinear functions of the transition parameters \( p_{ij} \). Equation (4) is derived from the condition that the \( k \)th moment of \( \mu_x \) (with \( k = 3 \)) equals 0, where \( \mu_x \) is the Markov chain component of the process (see the Appendix),

\[
E[\mu_x^k] = \sum_{m=1}^{M} \xi_m \left( \mu_m - \sum_{i=1}^{M-1} \xi_i \mu_i \right)^k = \sum_{m=1}^{M} \xi_m \left[ \mu_m - \mu_x - \sum_{i=1}^{M-1} (\mu_i - \mu_x) \xi_i \right]^k = \sum_{m=1}^{M} \xi_m \left[ (\mu_m - \mu_x) - \sum_{i=1}^{M-1} (\mu_i - \mu_x) \xi_i \right]^k + \left(1 - \sum_{m=1}^{M-1} \xi_m \right) \left[ - \sum_{i=1}^{M-1} (\mu_i - \mu_x) \xi_i \right]^k.
\]

For \( M = 2 \), the problem becomes more tractable analytically.

**Example 1.** Consider the case of the two-regime MSM(2)-AR(p) process. Invoking Proposition 1, the skewness of the Markov chain is given by

\[
E[\mu_x^3] = \sum_{m=1}^{2} \xi_m \mu_m^3 = \xi_1 \mu_1^3 + (1 - \xi_1) \mu_2^3,
\]

where \( \xi_1 = p_{11}/(p_{11} + p_{12}) \) is the unconditional probability of regime one, \( \mu_1^3 = \mu_1 - \mu_x = (1 - \xi_1)(\mu_1 - \mu_x) \) and \( \mu_2^3 = \mu_2 - \mu_x = (-\xi_1)(\mu_2 - \mu_x) \). Substituting for \( \xi_1, \mu_1, \) and \( \mu_2 \) we obtain

\[
E[\mu_x^3] = \xi_1 (1 - \xi_1)[1 - 2\xi_1] \left( \mu_1 - \mu_x \right)^3.
\]

Because the MS model implies that \( \mu_1 \neq \mu_2 \) and \( \xi_1 \in (0, 1) \), it is apparent that nondeepness, \( E[\mu_x^3] = 0 \), requires that \( \xi_1 = .5 \). Hence the matrix of transition probabilities must be symmetric, \( p_{12} = p_{21} \). This also implies that the regime-conditional means \( \mu_1 \) and \( \mu_2 \) are equidistant to the unconditional mean \( \mu_x \).
Hence, in the case of two regimes, nondeepness can be tested by testing the hypothesis \( p_{12} = p_{21} \). This is equivalent to the test of nonsharpness. For processes with \( M > 2 \), we propose testing for nondeepness based on the \( \mu^*_m \) conditional on \( \mu_\ast \) 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 \( \mu_\ast \) is given by

\[
E[\mu_\ast^3] = \sum_{m=1}^{3} \bar{\xi}_m \mu_m^3 = \bar{\xi}_1 \mu_1^3 + \bar{\xi}_2 \mu_2^3 + (1 - \bar{\xi}_1 - \bar{\xi}_2) \mu_3^3,
\]

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

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

Nondeepness, \( E[\mu_\ast^3] = 0 \), 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 based on the skewness of the differentiated series.

**Proposition 2.** An MSM(M)-AR(p) process is nonsteep if the size of the jumps, \( \mu_\ast - \mu_i \), satisfies the following condition:

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

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

In contrast with deepness, the condition for steepness depends not only on the ergodic probabilities, \( \bar{\xi}_i \), but also directly on the transition parameters.

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

\[
E[\Delta \mu^3] = (\bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21}) (\mu_2 - \mu_1)^3 = 0.
\]

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

\[
E[\Delta \mu^3] = \sum_{i=1}^{2} \sum_{j=i+1}^{3} (\bar{\xi}_i p_{ij} - \bar{\xi}_j p_{ji}) (\mu_j - \mu_i)^3
\]

\[
= (\bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21}) (\mu_2 - \mu_1)^3
\]

\[
+ (\bar{\xi}_1 p_{13} - \bar{\xi}_3 p_{31}) (\mu_3 - \mu_1)^3
\]

\[
+ (\bar{\xi}_2 p_{23} - \bar{\xi}_3 p_{32}) (\mu_3 - \mu_2)^3.
\]

Although 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 nondeepness. As noted earlier, this is stronger than the property of nonsharpness.

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

**Corollary 1.** A two-regime MS model is always nonsteep. Nonsharpness implies nondeepness and vice versa.

Nonsteepness is evident from Example 3. Because \( \bar{\xi}_1/\bar{\xi}_2 = p_{21}/p_{12} \), we have that \( \bar{\xi}_1 p_{12} - \bar{\xi}_2 p_{21} = 0 \) and hence \( E[\Delta \mu^3] = 0 \). Further, nonsharpness (symmetric transition probabilities, \( p_{12} = p_{21} \)) implies nondeepness, \( E[(x_i - \mu_\ast)^3] = 0 \), and vice versa. In particular, both concepts imply that the regimeconditional means \( \mu_1 \) and \( \mu_2 \) are equidistant to the unconditional mean \( \mu_\ast \).

### 3. Parametric Tests Based on the Markov-Switching Autoregressive Model

Testing the MS-AR model against a linear null (or three regimes versus two regimes) is complicated due to the presence of unidentified nuisance parameters under the null of linearity (i.e., the transition probabilities) and because the scores associated with parameters of interest under the alternative may be identically 0 under the null. These issues have been looked at by a number of authors (see, e.g., Hansen 1992, 1996), but are not of direct interest to us here, because the number of regimes remains unchanged under all three asymmetry hypotheses, so that standard asymptotics can be invoked. But we note that in practice they may complicate identification of the appropriate model on which to carry out the asymmetry tests.

Wald tests of the asymmetry hypotheses are computationally attractive, because the model does not have to be estimated under the null. In general terms, we consider Wald (W) tests of the hypothesis

\[ H_0: \phi(\lambda) = 0 \quad \text{versus} \quad H_1: \phi(\lambda) \neq 0, \]

where \( \phi: \mathbb{R}^n \to \mathbb{R}^r \) is a continuously differentiable function with rank \( r, r = \text{rk}(\frac{\partial \phi(\lambda)}{\partial \lambda}) \leq \text{dim} \lambda \). Because the \( p_{ij} \) are restricted to the \([0, 1] \) interval, the tests are formulated on the logits \( \pi_{ij} = \log(\frac{p_{ij}}{1-p_{ij}}) \), which avoids problems if one or more of the \( p_{ij} \) is close to the border. It is worth noting that if \( \frac{1}{\lambda} (\pi_{ij} - \pi_{ij}) \xrightarrow{d} \mathcal{N}(0, \sigma_{ij}^2) \), then \( \frac{1}{\lambda} (\bar{p}_{ij} - p_{ij}) \xrightarrow{d} \mathcal{N}(0, p_{ij}^2 (1 - p_{ij})^2 \sigma_{ij}^2) \) as \( p_{ij} \xrightarrow{d} \frac{1}{1 + e^{\pi_{ij}}} \). If one of the transition parameters is estimated to lie on the border, \( p_{ij} \in (0, 1) \), then the parameter is taken as fixed and is eliminated from the parameter vector \( \lambda \).

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

\[
\sqrt{T} (\hat{\lambda} - \lambda) \xrightarrow{d} \mathcal{N}(0, \Sigma),
\]
where, for the MLE, $\Sigma_\lambda = \hat{\Sigma}_a^{-1}$ is the inverse of the asymptotic information matrix. This can be calculated numerically.

It follows that $\phi(\lambda)$ is also normal for large samples,

$$\sqrt{T}[\phi(\lambda) - \phi(\lambda)] \overset{d}{\rightarrow} \mathcal{N}\left(0, \frac{\partial \phi(\lambda)}{\partial \lambda'} \Sigma_\lambda \frac{\partial \phi(\lambda)'}{\partial \lambda} \right).$$

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

$$T(\hat{\phi}(\lambda)) \left[ \frac{\partial \phi(\lambda)}{\partial \lambda'} \right] \left[ \Sigma_\lambda \frac{\partial \phi(\lambda)'}{\partial \lambda} \right]^{-1} \phi(\lambda) \overset{d}{\rightarrow} \chi^2(r),$$

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

### 3.1 Deepness

The Wald test for the null of nondeppness is based on

$$\phi_D(\lambda) = \phi_D(\pi, \mu; \cdot) = \sum_{m=1}^{M} \tilde{\xi}_m (\mu_m - \mu),$$

where $\tilde{\xi}_m(\pi)$ is the ergodic probability of regime $m$ and $\mu_m(\pi, \mu) = \sum_{m=1}^{M} \tilde{\xi}_m \mu_m$ is the unconditional mean of $x$. As $\frac{\partial \phi}{\partial \lambda} = 0$ for $\lambda_1 \in \{\alpha_1, \ldots, \alpha_p, \sigma^2, \pi\}$ the Wald test statistic for nondeppness is given by

$$T \phi_D(\lambda) = \left[ \tilde{\Sigma}_\lambda \Sigma_{\mu} \tilde{\Sigma}_{\mu'} \right]^{-1} \phi_D(\lambda).$$

**Example 5.** For $M = 2$, the null of nondeppness is based on

$$\phi(\pi_{12}, \pi_{21}, \mu_1, \mu_2) = \phi(\tilde{\xi}_1(\pi_{12}, \pi_{21}), \mu_1, \mu_2)$$

$$= \tilde{\xi}_1 (\mu_1 - \xi_1) \mu_1 - (1 - \xi_1) \mu_2)^3$$

$$+ (1 - \xi_1) (\mu_2 - \xi_2) \mu_1 - (1 - \xi_2) \mu_2)^3$$

Differentiating with respect to $\xi_1, \mu_1,$ and $\mu_2$ gives

$$\frac{\partial \phi}{\partial \xi_1} = -(\mu_2 - \mu_1)^3(1 - 6 \xi_1 (1 - \xi_1)),$$

$$\frac{\partial \phi}{\partial \mu_1} = -3 \xi_1 (\mu_2 - \mu_1)^2 (2 \xi_1 - 1) (1 - \xi_1),$$

and

$$\frac{\partial \phi}{\partial \mu_2} = 3 \xi_1 (\mu_2 - \mu_1)^2 (2 \xi_1 - 1) (1 - \xi_1).$$

Using that $\xi_1 = \frac{p_{11}}{p_{12} + p_{21}}$ and $\pi_{ij} = \log(p_{ij}) - \log(1 - p_{ij})$, we have

$$\frac{\partial \phi}{\partial \pi_{21}} = \frac{\partial \phi}{\partial \xi_1} \frac{\partial \xi_1}{\partial \pi_{21}}$$

$$= -(\mu_2 - \mu_1)^3(1 - 6 \xi_1 (1 - \xi_1)) \xi_1 (1 - \xi_1)(1 - \pi_{21})$$

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

$$\nabla \mu = \begin{bmatrix} \mu_2 - \mu_1 \\ \vdots \\ \mu_M - \mu_{M-1} \end{bmatrix} = Q \mu,$$

with

$$Q = \frac{\partial \nabla \mu}{\partial \mu} = \begin{bmatrix} -1 & 1 & 0 \\ \\ \vdots & \ddots & \ddots \\ 0 & -1 & 1 \end{bmatrix}$$

and

$$\frac{\partial \phi}{\partial \pi_{12}} = \frac{\partial \phi}{\partial \xi_1} \frac{\partial \xi_1}{\partial \pi_{12}}$$

$$= (\mu_2 - \mu_1)^3(1 - 6 \xi_1 (1 - \xi_1))(1 - \pi_{12}).$$

Because this test is difficult to implement for $M > 2$, for models with more than two regimes we use a version of the deepness test with $\xi_m$ and $\mu_j$ taken as fixed. This Wald test for the null of nondeppness is based on

$$\phi_{D2}(\lambda) = \phi_{D2}(\mu; \cdot) = \sum_{m=1}^{M} \tilde{\xi}_m (\mu_m - \mu)^3$$

where $\tilde{\phi} = 3 \tilde{\xi}_m (\mu_m - \mu)^2$ for $\lambda_1 = \mu_m, m = 1, \ldots, M$, and $\tilde{\phi} = 0$ for $\lambda_1 \in \{\alpha_1, \ldots, \alpha_p, \sigma^2, \pi\}$.

**Example 6.** For $M = 3$, the null of nondeppness is tested by $\phi_{D2}(\lambda) = 0$ and so has the form

$$T \left[ \sum_{m=1}^{3} \tilde{\xi}_m (\tilde{\mu}_m - \mu)^3 \right]$$

$$= \tilde{\Sigma}_\mu \left[ \tilde{\xi}_1 (\tilde{\mu}_1 - \mu)^2 \tilde{\xi}_2 (\tilde{\mu}_2 - \mu)^2 \tilde{\xi}_3 (\tilde{\mu}_3 - \mu)^2 \right]^{-1}$$

$$\times \tilde{\Sigma}_\lambda \left[ \tilde{\xi}_1 (\tilde{\mu}_1 - \mu)^2 \tilde{\xi}_2 (\tilde{\mu}_2 - \mu)^2 \tilde{\xi}_3 (\tilde{\mu}_3 - \mu)^2 \right]$$

3.2 Steepness

A Wald test for the null of nonsteepness can be based on

$$\phi_S(\lambda) = \phi_S(\mu; \cdot) = \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} (\tilde{\xi}_i p_{ij} - \tilde{\xi}_j p_{ij}) (\mu_j - \mu_i),$$

where $\tilde{\xi}_m, p_{ij},$ and $\mu_j$ again are taken as fixed. Thus the test concerns only the vector of mean parameters,
and
\[ \mathbf{\mu} = [\mu_1 \cdots \mu_M]^\prime. \]
Thus \( \frac{\partial \phi}{\partial \mathbf{\mu}} = \frac{\partial \phi}{\partial \mathbf{\mu}} \frac{\partial \mathbf{\mu}}{\partial \mathbf{\mu}} \) with \( \frac{\partial \phi}{\partial \mathbf{\mu}} = 3(\tilde{\xi}_i p_{ij} - \tilde{\xi}_j p_{ji})(\mu_j - \mu_i)^2 \) and \( \frac{\partial \phi}{\partial \mathbf{\lambda}} = 0 \) otherwise. The Wald test statistic has the form
\[ \phi(\lambda) \left[ \frac{\partial \phi}{\partial \mathbf{\mu}} \right] \left[ Q\left( \frac{1}{T} \sum \right) \frac{\partial \phi}{\partial \mathbf{\mu}} \right]^{-1} \phi(\lambda)^\prime = \chi^2(1). \]
In the case of a three-state Markov chain, for example,
\[ Q = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \]
and
\[ \frac{\partial \phi}{\partial \mathbf{\mu}} = \begin{bmatrix} 3(\tilde{\xi}_1 p_{12} - \tilde{\xi}_2 p_{21})(\mu_2 - \mu_1)^2 \\ 3(\tilde{\xi}_1 p_{13} - \tilde{\xi}_3 p_{31})(\mu_3 - \mu_1)^2 \\ 3(\tilde{\xi}_2 p_{23} - \tilde{\xi}_3 p_{32})(\mu_3 - \mu_2)^2 \end{bmatrix}. \]

### 3.3 Sharpness

The null of nonsharpness can be expressed as
\[ \phi_{FP}(\lambda) = \phi_{FP}(\pi; \cdot) := \Phi \pi, \]
where the matrix \( \Phi \) 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 \( \pi_{ij} \) be collected to the matrix \( \Pi \),
\[ \Pi = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1M} \\ \vdots & \ddots & \vdots \\ \pi_{M1} & \cdots & \pi_{MM} \end{bmatrix}, \]
the matrix of logit transition probabilities. Then the vector \( \pi \) is given by \( \text{vec}(\Pi) \), defined as \( \text{vec}(\Pi) \) with the diagonal elements \( \pi_{jj} \) excluded. In the case of a three-state Markov chain, for example, we have that
\[ \pi = (\pi_{12}, \pi_{13}, \pi_{21}, \pi_{23}, \pi_{31}, \pi_{32})^\prime \]
and
\[ \Phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{bmatrix}. \]
For linear restrictions, the relevant Wald statistic can be expressed as
\[ W_{FP} = \sum_{i=1}^{M} \alpha_i x_i + \eta_i, \]
where \( \eta_i \) is defined as
\[ \nu_i = \mu(s_i) - \tilde{\mu} = \sum_{m=1}^{M} \mu_m (\xi_{mi} - \tilde{\xi}_m) \]
and \( \tilde{\mu} = \alpha^{-1}(1) \mu_i \). In the case of a two-regime model, we have \( \nu_i = (\mu_i - \mu) \), with \( \tilde{\eta}_i = \xi_{1i} - \xi_{2i} \), which equals \( 1 - \tilde{\xi}_i \) if the regime is 1 and \( -\tilde{\xi}_i \) otherwise. Because (9) must be equivalent to (8), the following expression for \( \mu_i \) is obtained:
\[ \mu_i = \alpha^{-1}(L) \nu_i. \]

### 4. EXTENSIONS TO THE TESTING FRAMEWORK

In this section we outline three extensions to the basic framework for testing for asymmetries in MS-AR models. We deal with models in which the intercept, rather than the mean switches between regimes and models that display regime-dependent heteroscedasticity and consider multivariate settings.

#### 4.1 Switching Intercepts

The MS(M)-AR(p) model is characterized by switching in the intercept, rather than the mean (MSM-AR),
\[ x_t = \mu(s_t) + \sum_{j=1}^{p} \alpha_j x_{t-j} + u_t, \]
where \( u_t \sim NID(0, \sigma^2) \) and \( s_t \in \{1, \ldots, M\} \) is generated by a Markov chain.

As in the MSM-AR process, the MS(M)-AR(p) process can be written as the sum of two independent processes,
\[ x_t = \mu_i + \xi_t, \]
where \( \mu_i = \alpha^{-1}(1) \sum_{m=1}^{M} \xi_m \mu_m \), \( \alpha(L) = 1 - \alpha_1 L - \cdots - \alpha_p L^p \), so \( \alpha(1) = 1 - \sum_{m=1}^{p} \alpha_m \) and \( \mathbb{E}[\mu_i - \mu] = 0 \). Here \( \{\xi_t\} \) is a Gaussian process, \( \alpha(L) \xi_t = u_t \), and \( \mathbb{E}[\xi_t] = 0 \), so \( \mu_i \) represents the contribution of the Markov chain and \( \mathbb{E}[\mu_i] = 0 \). To derive an expression for \( \alpha(L) \), first rewrite (8) as
\[ \alpha(L)(x_t - \mu_i) = \nu_i + u_t, \]
where \( \nu_i \) is defined as
\[ \nu_i = \mu(s_i) - \tilde{\mu} = \sum_{m=1}^{M} \mu_m (\xi_{mi} - \tilde{\xi}_m) \]
and \( \tilde{\mu} = \alpha^{-1}(1) \mu_i \). In the case of a two-regime model, we have \( \nu_i = (\mu_i - \mu) \), with \( \tilde{\eta}_i = \xi_{1i} - \xi_{2i} \), which equals \( 1 - \tilde{\xi}_i \) if the regime is 1 and \( -\tilde{\xi}_i \) otherwise. Because (9) must be equivalent to (8), the following expression for \( \mu_i \) is obtained:
\[ \mu_i = \alpha^{-1}(L) \nu_i. \]

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

Tests for asymmetries in MS(M)-AR(p) models can be based on \( \nu_i \), which can be seen to be equivalent to the \( \mu_i \) in MSM(M)-AR(p) models. Wald tests for deepness and steepness can be easily constructed by applying the procedures developed in Section 3 to parametric tests for the skewness of \( \nu_i \) and \( \Delta \nu_i \).

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 \( \nu_i = \Delta \mu_i \), and testing \( \nu_i \) for deepness leads to conclusions for the deepness of \( \Delta x_i \) (rather than \( x_i \)). In other words, applying the conditions derived for
deepness in the MSM-AR model to \( \nu \), provides a test of steepness of the MSI-AR model. In our examples, the roots of \( \alpha(L) \) are a long way from unity because we are modeling first differences, and these exhibit little dependence relative to models in levels. Furthermore, the extreme case of a unit root implies that the data have not been differenced a sufficient number of times before modeling.

4.2 Regime-Dependent Heteroscedasticity

In the original Hamilton model, the variance of the disturbance term does not depend on the regime. However, regime-dependent heteroscedasticity is often manifest when the model is applied to financial data and also perhaps (albeit to a lesser extent) to macroeconomic data. For example, Koop, Pesaran, and Potter (1996) found evidence of regime-dependent error variances in a nonlinear model of output growth and unemployment rate changes. Therefore, the assumption that \( u_s \sim \text{NID}(0, \sigma^2) \) in (1) may need to be replaced by \( u_s \sim \text{NID}(0, \sigma^2(s)) \), where in a two-regime model, for example, \( \sigma^2(s) = \sigma^2_1 \) when \( s = 1 \) and \( \sigma^2(s) = \sigma^2_2 \) when \( s = 2 \). In such a model, asymmetries in the observed variable can arise either from asymmetries in the model’s propagation mechanism or from asymmetries in the innovations. Failure to allow for heteroscedasticity in the MS-AR model when it is present in the data may affect the properties of the SDS tests, as shown in Section 5. Luukkonen and Teräsvirta (1991) tested for cyclical asymmetry by testing whether a linear AR model can be rejected in favor of a smooth-transition AR, and are concerned that asymmetry may result simply because of regime heteroscedasticity. Consequently, these authors also tested for AR conditional heteroscedasticity. The SDS tests can be calculated within a model that explicitly allows for heteroscedasticity.

Our tests are designed to detect asymmetries in the model’s propagation mechanism, whereas the NP tests are unable to discriminate between the two sources of asymmetry.

4.3 Multivariate Models

The tests that we have outlined apply equally to a vector process with a single state variable. Models of this sort arise when the variables share a common cyclical component, as in the MS-VAR model of postwar U.S. employment and output of Krolzig and Toro (1998) or the dynamic-factor model with regime switching of Diebold and Rudebusch (1996). Note that the test of sharpness is intrinsically a system-based test, because it evaluates the transition probabilities of the common latent regime variable. In contrast, the tests for deepness and steepness (for \( M > 2 \)) focus on variable-specific asymmetries. It is therefore possible to test each variable for asymmetry.

In some instances, it may be appropriate to allow more than one state variable, as in the bivariate model of stock returns and output growth of Hamilton and Lin (1996), where each variable responds to a specific state variable. The procedures developed in Section 3 can then be applied in the same way, using the regime means and ergodic and transition probabilities relevant for each pairing of variable and state variable.

5. PROPERTIES OF TESTING PROCEDURES

In this section we explore by Monte Carlo (a) the size and power properties of our parametric tests relative to NP tests, (b) the impact on our testing procedures of ignoring regime-dependent heteroscedasticity, and (c) their properties under model misspecification, by which we mean applying the MS-AR model-based tests when the process was generated by an alternative model.

5.1 Size and Power

Table 1 reports the empirical sizes and powers of the tests for a Monte Carlo (based on 1,000 replications) for two data-generating processes (DGPs) based on

\[
x_t = \mu(s_t) + \varepsilon_t, \quad \text{where} \quad \varepsilon_t \sim \text{NID}(0, \sigma^2) \quad \text{and} \quad s_t \in \{1, 2\}.
\]

For the first process, labeled “Symmetric MSM” in Table 1, \( \mu_1 = -\mu_2 = -1.5, \sigma^2 = 1 \), and \( p_{11} = p_{22} = .85 \). Thus, from the propositions stated in Section 2.3, it is apparent that the values of \( \mu(s_t) \) and \( \xi_t \) satisfy the conditions for \( \mu_1 \), and thus \( x_t \), to exhibit nondeepness. (Nonsteepness is a property of the model, and nondeepness implies nonsharpness in this model.) The parameter values for the second process, labeled “Asymmetric MSM,” are the same except that \( p_{11} = .65 \), and the conditions for nondeepness do not hold.

Because the DGP consists of two regimes, there are no entries for our steepness test (MS:Steepness) in Table 1; such processes can not exhibit steepness. The NP test rejection frequencies for steepness (NP:Steepness) are close to the nominal sizes for all of the DGPs. The test for sharpness has size close to nominal even for \( T = 100 \), and power of nearly 60% at a 5% size. The parametric deepness test (MS:Deepness) is correctly sized asymptotically and only a little too large for \( T = 100 \). Moreover, it has good power for the asymmetric process for \( T = 100 \).

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Symmetric MSM</th>
<th>Asymmetric MSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>T = 100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS:Sharpness</td>
<td>.100 , .055 , .012</td>
<td>.696 , .585 , .280</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>.162 , .098 , .029</td>
<td>.587 , .440 , .195</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.212 , .153 , .084</td>
<td>.762 , .691 , .525</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>.138 , .081 , .020</td>
<td>.112 , .061 , .021</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>0 , 0 , 0</td>
<td>0 , 0 , 0</td>
</tr>
<tr>
<td>T = 1,000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS:Sharpness</td>
<td>.087 , .045 , .010</td>
<td>1.000 , 1.000 , 1.000</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>.154 , .088 , .027</td>
<td>1.000 , .999 , .999</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.098 , .051 , .016</td>
<td>1.000 , 1.000 , 1.000</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>.167 , .100 , .030</td>
<td>.116 , .063 , .015</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>0 , 0 , 0</td>
<td>0 , 0 , 0</td>
</tr>
</tbody>
</table>

NOTE: This table records the rejection frequencies of tests for asymmetry at three nominal sizes (0.10, 0.05, 0.01) for a “symmetric” MS-AR DGP and an “asymmetric” MS-AR DGP. The former is nonsharp, nondeep, and nonsteep. The latter is deep. The rejection frequencies are empirical sizes for the symmetric DGP and powers for the asymmetric DGP. MS:Deepness refers to the MS-AR parametric test of deepness; NP:Deepness, to the nonparametric test of deepness; and so on.
(a) $\mu_1 = -1.5, \mu_2 = 1.5, \sigma_1 = 1, \sigma_2 = 1, \Pr(s=1) = .5$

(b) $\mu_1 = -1.5, \mu_2 = 1.5, \sigma_1 = 1, \sigma_2 = 1, \Pr(s=1) = .3$

(c) $\mu_1 = -1.5, \mu_2 = 1.5, \sigma_1 = 1, \sigma_2 = 2, \Pr(s=1) = .5$

(d) $\mu_1 = -1.5, \mu_2 = 1.5, \sigma_1 = 1, \sigma_2 = 2, \Pr(s=1) = .3$

Figure 2. Asymmetries Due to Regime-Dependent Heteroscedasticity. The figure depicts densities (conditional on the regime and unconditional) of $x_t$ constructed for two-regime MS-AR models \{\ldots, \Pr(s=1); \ldots, \Pr(s=2)\; ; \ldots \Pr(x)\}. In (a), the MS-AR model satisfies the conditions for a symmetric propagation mechanism, and the process is homoscedastic. The unconditional density (the solid line) is symmetric. In (b), depicting a homoscedastic process with an asymmetric propagation mechanism, the unconditional density is skewed. The process in the (c) has a symmetric propagation mechanism but has regime-dependent heteroscedasticity. The unconditional density is skewed. The process in (d) has asymmetric innovations and an asymmetric propagation mechanism.

5.2 Robustness Under Heteroscedasticity

Our tests are designed to detect asymmetries in the Markov chain component, $\mu_t$, which we have termed “first-moment asymmetry,” whereas the NP tests would be expected to reject the null of asymmetry in the presence of regime-dependent variances of the shocks (“heteroscedasticity”). The data-generating processes chosen to explore these issues is a simple extension of (12) to allow for heteroscedasticity, $x_t = \mu(s) + \epsilon_t$, where $\epsilon_t \sim \text{NID}(0, \sigma^2(s))$ and $s \in \{1, 2\}$. (13)

To illustrate, Figure 2 plots the density of $x_t$ generated by (13) and the density conditional on being in a regime. In (a), $\mu_1 = -\mu_2 = -1.5, \sigma_1^2 = \sigma_2^2 = 1$, and $\Pr(s = 1) = .5$. From the propositions stated in Section 2.3, it is apparent that the values of $\mu(s)$ and $\tilde{\xi}_1$ satisfy the conditions for $\mu_t$, and thus $x_t$, to exhibit nondeepness. (Nonsteepestness is a property of the model.) The density exhibits skewness in (b) because the condition for nondeepness is not satisfied by $\mu_1 = -\mu_2 = -1.5$ and $\Pr(s = 1) = .3$. In (c) the condition for nondeepness is satisfied, because $\mu_1 = -\mu_2 = -1.5$ and $\Pr(s = 1) = .5$, but nonetheless heteroscedasticity, $\sigma_1^2 = 1$ and $\sigma_2^2 = 2$, induces skewness in $x_t$. So the contribution of the Markov process is symmetric, but the unequal variances result in asymmetry in the marginal distribution of $x_t$. Figure (d) is akin to (b) but with heteroscedastic errors.

Table 2 reports the properties of the testing procedures for heteroscedastic MS-AR processes, when the estimated model allows for regime-dependent error variances. The NP test rejection frequencies for steepness (NP:Steepness) are close to the nominal sizes for all of the DGPs, so the presence of heteroscedasticity does not inflate the size of the test. The test for sharpness has size close to nominal even for $T = 100$. The power approximately halves when the DGP is heteroscedastic (compare the “Asymmetric MSMH” columns of Table 2 with the “Asymmetric MSMH” columns of Table 1), but as the entries for $T = 1,000$ confirm, this is a small-sample effect. The parametric steepness test (MS:Deepness) is correctly sized asymptotically and only a little too large for $T = 100$. Moreover, it has good power for the asymmetric DGP for $T = 100$. In contrast, the NP (NP:Deepness) test is less powerful, has a size approaching 1 asymptotically for the symmetric MSMH, and has power approximately equal to size for a 5% test for the asymmetric MSMH when $T = 1,000$.

The second aspect the Monte Carlo explores is the effect of using homoscedastic models when the DGP is heteroscedastic. Table 2 shows that the sizes of the parametric sharpness and steepness tests are a little inflated for $T = 100$ and are 25%–30% for a nominal size of 5% for the large sample. Finally, the failure to model the heteroscedasticity reduces the power of the asymmetry tests for the DGP considered.
5.3 Robustness Under Model Misspecification

Alternative regime-switching models have been proposed in the literature to explain business cycle phenomena. Because these models may in some instances provide a better characterization of the data than the MS-AR model, a relevant question is how well the parametric tests perform when the model on which they are based (i.e., the MS-AR model) is not the model that generated the data. In this section we report Monte Carlo simulations designed to investigate the properties of the proposed testing procedures under model misspecification. Results are recorded in Table 3 for a self-exciting threshold AR (SETAR) DGP and a smooth-transition AR (STAR) DGP. SETAR models have been estimated for U.S. output growth by a number of authors, including Tiao and Tsay (1994), Potter (1995), Pesaran and Potter (1997), and Clements and Smith (2000). A statistical analysis of the SETAR model was provided by Tong (1995), and of the STAR model by Luukkonen et al. (1998) and Teräsvirta and Anderson (1992).

Table 2. Empirical Size and Power of Tests of Asymmetries When There are Heteroscedastic Disturbances

<table>
<thead>
<tr>
<th>Sample size</th>
<th>MS-AR models matching the MS-AR DGPs</th>
<th>Homoscedastic MS-AR models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MS-AR DGP</td>
<td>Asymmetric MS-AR DGP</td>
</tr>
<tr>
<td></td>
<td>Symmetric</td>
<td>Asymmetric</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T = 100</td>
<td>.10</td>
<td>.05</td>
</tr>
<tr>
<td>MS:Sharpeness</td>
<td>.099</td>
<td>.111</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>.361</td>
<td>.043</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.173</td>
<td>.493</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>.119</td>
<td>.096</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>T = 1,000</td>
<td>.073</td>
<td>.096</td>
</tr>
<tr>
<td>MS:Sharpeness</td>
<td>.75</td>
<td>.120</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>.986</td>
<td>.100</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.078</td>
<td>.006</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>.141</td>
<td>.085</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

NOTE: The table records the rejection frequencies of tests for asymmetry at three nominal sizes, $a = .10, .05, .01$, for two DGPs: a heteroscedastic MS-AR with a symmetric propagation mechanism, labeled "symmetric MSMH," and a heteroscedastic MS-AR with an asymmetric propagation mechanism, labeled "asymmetric MSMH," which is deep. For both DGPs, $d_1 = 1$, $d_2 = 2$. For the symmetric MSMH, $\mu_1 = -1.5$, $\mu_2 = 1.5$, and $\rho_1 = \rho_2 = 85$; for the asymmetric MSMH, $\mu_1 = -1.5$, $\mu_2 = 1.5$, $\rho_1 = 65$, and $\rho_2 = 85$. For each DGP, the table records the results of testing for asymmetries within MS-AR models that allow for the heteroscedasticity (the left half of the table) and in MS-AR models that impose equal error variances across the two regimes (the right half of the table). The rejection frequencies are empirical sizes for the symmetric MS-AR, and powers for the asymmetric MS-AR.

MS:Deepness refers to the MS-AR parametric test of deepness; NP:Deepness, to the nonparametric test of deepness; and so on.

Table 3. Empirical Size and Power: SETAR and LSTAR DGPs

<table>
<thead>
<tr>
<th>Sample size</th>
<th>MS-AR model applied to data from SETAR and LSTAR DGPs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symmetric SETAR</td>
</tr>
<tr>
<td></td>
<td>.10</td>
</tr>
<tr>
<td>T = 100</td>
<td>.116</td>
</tr>
<tr>
<td>MS:Sharpeness</td>
<td>.398</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>.306</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.057</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>0</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>0</td>
</tr>
<tr>
<td>T = 1,000</td>
<td>.077</td>
</tr>
<tr>
<td>MS:Sharpeness</td>
<td>.410</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>.113</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.073</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>0</td>
</tr>
</tbody>
</table>

NOTE: The table records the rejection frequencies of tests for asymmetry at three nominal sizes, $a = .10, .05, .01$, for four DGPs: symmetric and asymmetric SETAR models and symmetric and asymmetric STAR models. The $c$ and $\gamma$ DGP parameter values are as follows: for the symmetric SETAR, $c = 0$; for the asymmetric SETAR, $c = -75$; for the symmetric STAR, $c = 0$, $\gamma = 10$; and for the asymmetric STAR, $c = -75$, $\gamma = 10$.

MS:Deepness refers to the MS-AR parametric test of deepness; NP:Deepness, to the nonparametric test of deepness; and so on.

The regime-generating process in both models is not assumed to be exogenous, but rather directly linked to a transition variable. For the SETAR model, the transition variable is a lag of the endogenous variable, say, $x_{t-d}$,

$$x_t = \left( \mu_1 + \sum_{i=1}^{p} \alpha_i x_{t-i} \right) \left( 1 - I(x_{t-d} < c) \right)$$

$$+ \left( \mu_2 + \sum_{i=1}^{p} \alpha_i x_{t-i} \right) I(x_{t-d} \geq c) + \epsilon_t,$$

(14)

where $\epsilon_t \sim \text{IID}(0, \sigma^2)$, $I(x_{t-d} < c) = 1$ if $x_{t-d} > c$ and 0 otherwise and $c$ is the threshold at which the switching between regimes occurs.

In the STAR model, the weight attached to the regimes depends on the realization of an exogenous or lagged endogenous variable $z_t$, so that the transition between regimes
is “smooth,”

\[ x_t = \left( \mu_1 + \sum_{i=1}^{p} \alpha_{1i} x_{t-i} \right) (1 - G(z_t; \gamma, c)) \]

\[ + \left( \mu_2 + \sum_{i=1}^{p} \alpha_{2i} x_{t-i} \right) G(z_t; \gamma, c) + \varepsilon_t, \tag{15} \]

where \( \varepsilon_t \sim \text{IID}(0, \sigma^2) \) and the transition function \( G(z_t; \gamma, c) \) is a continuous function, usually bounded between 0 and 1. We consider the LSTAR model, where the transition function is given by

\[ G(z_t; \gamma, c) = \frac{1}{1 + \exp[-\gamma(z_t - c)]}, \]

where \( \gamma \) is the smoothness parameter and the transition variable \( z_t \) is taken to be the lagged endogenous variable \( z_t = x_{t-d} \). For \( \gamma > 0 \), as \( z_t \to -\infty \), \( G(\cdot) \to 0 \), and for \( z_t \to \infty \), \( G(\cdot) \to 1 \).

To simplify matters, in our simulations we set \( \alpha_{1i} = \alpha_{2i} = 0 \) for all \( i, d = 1 \), \( \sigma^2 = 1 \), \( \mu_1 = -\mu_2 = -1.5 \), and vary \( c \) to give symmetric and asymmetric models. The results in Table 3 demonstrate the excellent performance of the MS model-based sharpness test. Even for the small sample size of \( T = 100 \), it has approximately the correct size under model misspecification. It also has good power in small samples for both the asymmetric SETAR and LSTAR DGPs. In contrast, the parametric and NP nondeepness tests work less well in small samples, but for large samples the parametric nondeepness test (MS:Deepness) is only slightly oversized and is clearly better behaved than the NP one.

Overall, the Monte Carlo results support the use of the proposed tests for business cycle asymmetries. The tests (a) behave as expected for correctly specified MS models, (b) are robust against skewness due to heteroscedasticity, and have been found to be reliable tests for business cycle asymmetries even if the underlying DGP is different from the empirical model. The tests proposed in this article enhance the role of the MS model as a flexible tool for empirical modelling.

### 6. EMPIRICAL ILLUSTRATIONS

The SDS tests are illustrated on a number of datasets. Specifically, we apply the parametric tests discussed in Section 3 to the MS(2)-AR(4) model of output growth of Hamilton (1989) for his original sample period of 1953–1984, to the MS(3)-AR(4) model of Clements and Krolzig (1998) on more recent data, and to various models of U.S. investment and consumption growth. In each case, the outcomes of the tests for asymmetries are compared with NP tests of skewness.

#### 6.1 The Hamilton Model of U.S. 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 NP approaches. The seminal article by Hamilton (1989) fit a fourth-order autoregression \( (p = 4) \) to the quarterly percentage change in U.S. real GNP, \( x_t \), from 1953 to 1984,

\[ x_t - \mu(s_t) = \alpha_1 (x_{t-1} - \mu(s_{t-1})) + \cdots + \alpha_4 (x_{t-p} - \mu(s_{t-p})) + \varepsilon_t, \tag{16} \]

where \( \varepsilon_t \sim \text{NID}(0, \sigma^2) \) 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 to be the same in both regimes. This is an MS mean model, with the autoregressive parameters and disturbances independent of the state \( s_t \).

Maximization of the likelihood function of an MS-AR model entails an iterative estimation technique to obtain estimates of the parameters of the AR and the transition probabilities governing the Markov chain of the unobserved states. Hamilton (1990) presented an expectation maximization (EM) algorithm for this class of model, and Krolzig (1997) gave an overview of alternative numerical techniques for maximum likelihood estimation of these models.

The results of testing for asymmetry based on the original Hamilton model and dataset (1952:2–1984:4) are recorded in Table 4. The NP test for skewness indicates significant negative skewness in output growth (i.e., deepness of contractions) at the 5% level. Our parametric test of nondeepness also indicates negative skewness, but is significant only at the 20% level. There is evidence of sharpness at the 10% level, with the probability of switching from contraction to expansion exceeding the probability of movement in the reverse direction.

#### 6.2 The Three-Regime Heteroscedastic Model of U.S. Output Growth

Sichel (1994) argued that postwar business cycles typically consist 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-AR model of Clements and Krolzig (1998). This model consists of a shifting intercept term and a heteroscedastic error term and is denoted an MSIH(3)-AR(4) model, where H flags the

<table>
<thead>
<tr>
<th>Test</th>
<th>Sign</th>
<th>Test statistic value</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS:Sharpness</td>
<td>–</td>
<td>2.77</td>
<td>[.09]†</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>–</td>
<td>1.74</td>
<td>[.19]</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>–</td>
<td>5.20</td>
<td>[.02]†</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>–</td>
<td>0</td>
<td>[.99]</td>
</tr>
</tbody>
</table>

NOTE: The NP and MS test statistics are \( \chi^2(1) \) under the null of symmetry. A positive (negative) value of “Sign” flags positive (negative) skewness. Nonsteepness is a property of the two-regime model; see Corollary 1.

†Significance at the 10% level.

†Significance at the 5% level.
heteroscedastic error term and 3 and 4 refer to the number of regimes and AR lags. It is written as

\[ x_t = \mu(s_t) + \sum_{k=1}^{4} \alpha_k x_{t-k} + \epsilon_t, \quad (17) \]

where \( \epsilon_t \sim \text{NID}(\sigma^2(s_t)) \) and \( s_t \in \{1, 2, 3\} \) is generated by a Markov chain. The three-state model is preferred over the two-state model, because the latter does not always yield a good representation of the business cycle when fitted to periods outside that of Hamilton (1989) (see, e.g., Boldin 1996; Clements and Krolzig 1998). Clements and Krolzig (1998) found an average duration of contraction of 2–3 quarters when the two-state model is fit to the period 1947–1990, and of less than 2 quarters when it is fit for 1959–1996.

Figure 3 and Table 5 (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.

Although Hess and Iwata (1997b) found that their three-state MS-AR model estimated for 1949–1992 fails to generate contractions of sufficient duration or depth, their estimated \( p_{11} \) is only .1267, whereas the lowest value in the MSIH models recorded in Table 5 is 0.78, which directly translates into a longer duration of the recession regime. Thus 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 6 and 7 for two historical periods. For the first sample period (1948–1990), the NP skewness and model-based tests both indicate steepness of expansions, with

![Figure 3. MSIH(3)-AR(4) Model of U.S. Output Growth: Smoothed (Solid Lines) and Filtered (Dashed Blocks) Probabilities of the High-Growth H (left column) and “Recession” L Regime (Right column). Each row refers to a different historical period.](image)

Table 5. MSIH(3)-AR(4) Models of U.S. Output Growth

<table>
<thead>
<tr>
<th>Sample</th>
<th>48:2–90:4</th>
<th>60:2–96:2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ( \mu_1 )</td>
<td>-.08</td>
<td>-.05</td>
</tr>
<tr>
<td>Mean ( \mu_2 )</td>
<td>1.41</td>
<td>.84</td>
</tr>
<tr>
<td>Mean ( \mu_3 )</td>
<td>3.43</td>
<td>1.41</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>-.10</td>
<td>.02</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-.11</td>
<td>.02</td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>-.17</td>
<td>-.10</td>
</tr>
<tr>
<td>( \alpha_4 )</td>
<td>-.19</td>
<td>-.10</td>
</tr>
<tr>
<td>( \sigma^2_1 )</td>
<td>.82</td>
<td>.80</td>
</tr>
<tr>
<td>( \sigma^2_2 )</td>
<td>.50</td>
<td>.12</td>
</tr>
<tr>
<td>( \sigma^2_3 )</td>
<td>.02</td>
<td>.41</td>
</tr>
<tr>
<td>( \rho_{12} )</td>
<td>.19</td>
<td>.02</td>
</tr>
<tr>
<td>( \rho_{13} )</td>
<td>.02</td>
<td>.13</td>
</tr>
<tr>
<td>( \rho_{23} )</td>
<td>.09</td>
<td>.08</td>
</tr>
<tr>
<td>( \rho_{11} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \rho_{22} )</td>
<td>.16</td>
<td>.09</td>
</tr>
<tr>
<td>( \rho_{33} )</td>
<td>.30</td>
<td>.23</td>
</tr>
<tr>
<td>( \rho_{23} )</td>
<td>.66</td>
<td>.45</td>
</tr>
<tr>
<td>( \rho_{32} )</td>
<td>.04</td>
<td>.32</td>
</tr>
<tr>
<td>Duration 1</td>
<td>4.81</td>
<td>6.75</td>
</tr>
<tr>
<td>Duration 2</td>
<td>10.70</td>
<td>13.09</td>
</tr>
<tr>
<td>Duration 3</td>
<td>6.15</td>
<td>10.95</td>
</tr>
<tr>
<td>Observations</td>
<td>171</td>
<td>145</td>
</tr>
</tbody>
</table>

**NOTE:** \( \mu_i, \sigma^2_i, \rho_{ij} \) denote the intercept, disturbance variance, and ergodic probability of regime \( i \). The \( \alpha_s \) are the autoregressive parameters, which are constant across regimes, and the \( p_{ij} \) are the transition probabilities.
the MS-AR model test permitting rejection of the null at the 1% level. Moreover, there is clear evidence of asymmetric TPs (or sharpness), which results from a rejection of $p_{21} = p_{23}$, because moving from moderate to low growth is more likely than moving from moderate to high growth. The three-state model permits rejection of the nonsharpness hypotheses at a higher confidence level than does the two-state model. Table 7 illustrates how heteroscedasticity can affect the skewness of the unconditional distribution of $x_t$, as shown in Section 5.2. The observed growth rates ($x_t$) display negative skewness (deepness of contractions), but the nondeepness test (though not significant) indicates positive skewness. The positive skewness in the hidden Markov chain emanates from the high-growth third regime and is partly associated with the 1951–1952 period. However, the variance is much higher in regime 1 (recession), so that the observed variable is overall negatively skewed (but not significantly).

For the later sample period (shown in Table 7), the MS model test continues to reject nonsteepness at the 5% level, in contrast to the NP test, which now flags deepness of recessions rather than steepness of expansions. The major change in inference using parametric tests is that there is no evidence of sharpness in the later period.


<table>
<thead>
<tr>
<th>Test</th>
<th>Sign</th>
<th>Test statistic value</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS:Sharpness</td>
<td>+</td>
<td>31.828</td>
<td>[0]</td>
</tr>
<tr>
<td>$p_{12} = p_{13}$</td>
<td>.02</td>
<td>[.88]</td>
<td></td>
</tr>
<tr>
<td>$p_{21} = p_{23}$</td>
<td>.04</td>
<td>[.84]</td>
<td></td>
</tr>
<tr>
<td>$p_{21} = p_{23}$</td>
<td>31.701</td>
<td>[.0]</td>
<td></td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>+</td>
<td>.18</td>
<td>[.70]</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>+</td>
<td>17.73</td>
<td>[0]</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>−</td>
<td>.71</td>
<td>[.40]</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>+</td>
<td>3.37</td>
<td>[.07]</td>
</tr>
</tbody>
</table>

NOTE: The NP and MS test statistics are $\chi^2(1)$ under the null of symmetry. A positive (negative) value of “Sign” flags positive (negative) skewness. Because $p_{21}$ and $p_{23}$ are close to 0, 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 nonsharpness.

The three-state model permits rejection of the nonsharpness hypotheses at a higher confidence level than does the two-state model. Table 7 illustrates how heteroscedasticity can affect the skewness of the unconditional distribution of $x_t$, as shown in Section 5.2. The observed growth rates ($x_t$) display negative skewness (deepness of contractions), but the nondeepness test (though not significant) indicates positive skewness. The positive skewness in the hidden Markov chain emanates from the high-growth third regime and is partly associated with the 1951–1952 period. However, the variance is much higher in regime 1 (recession), so that the observed variable is overall negatively skewed (but not significantly).

For the later sample period (shown in Table 7), the MS model test continues to reject nonsteepness at the 5% level, in contrast to the NP test, which now flags deepness of recessions rather than steepness of expansions. The major change in inference using parametric tests is that there is no evidence of sharpness in the later period.


<table>
<thead>
<tr>
<th>Test</th>
<th>Sign</th>
<th>Test statistic value</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MS:Sharpness</td>
<td>+</td>
<td>.75</td>
<td>[.86]</td>
</tr>
<tr>
<td>$p_{12} = p_{13}$</td>
<td>.73</td>
<td>[.39]</td>
<td></td>
</tr>
<tr>
<td>$p_{21} = p_{23}$</td>
<td>.02</td>
<td>[.88]</td>
<td></td>
</tr>
<tr>
<td>$p_{21} = p_{23}$</td>
<td>0</td>
<td>[.95]</td>
<td></td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>+</td>
<td>.47</td>
<td>[.49]</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>+</td>
<td>4.27</td>
<td>[.04]</td>
</tr>
<tr>
<td>NP:Deepness</td>
<td>−</td>
<td>9.87</td>
<td>[0]</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>+</td>
<td>.84</td>
<td>[.36]</td>
</tr>
</tbody>
</table>

NOTE: The NP and MS test statistics are $\chi^2(1)$ under the null of symmetry. A positive (negative) value of “Sign” flags positive (negative) skewness. Because $p_{21}$ and $p_{23}$ are close to 0, 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 nonsharpness.

6.3 Models of U.S. Investment and Consumption Growth

To further illustrate the method of testing for asymmetries, we apply the tests to U.S. investment and consumption growth using a number of MS models. These models contain either two or three regimes and either allow the error variance to depend on the regime or restrict it to be heteroskedastic. In all cases we consider models without lags, so that MSI and MSM models are equivalent.

The first four panels of Figure 4 depict the recession regime probabilities for investment growth. The allocation of observations to the recession regime is more dependent on whether or not the errors are allowed to be heteroskedastic than on whether there are two or three regimes. The main difference between the two- and three-regime models with heteroskedastic errors is that there is some evidence of a recession in the investment series around 1990 in the former. Table 8 shows that the model-based steepness tests (MS nonsteepness) reject the null at the 5% level in both the homoscedastic and heteroskedastic three-regime models and indicate steepness of expansions, whereas the NP tests suggest “tallness” of expansions.

The last four panels in Figure 4 give the recession probabilities for consumption growth. Here the estimates of the “recession” regime for the two- and three-regime heteroskedastic models are quite different, and from Table 8 the homoscedastic model indicates tallness and steepness of expansions, in line with the NP tests, whereas both features are absent in the heteroskedastic model.

These examples suggest a number of points. The results of testing for asymmetries based on parametric models may be sensitive to the model specification used. Specifically, it is likely to matter whether the model allows for heteroskedastic errors. The findings here confirm the Monte Carlo results in Section 5.2. The regime categorization in models that allow heteroskedastic disturbances will reflect shifts in both the mean and the variance of the series, and so will not necessarily coincide with that in homoskedastic models if, for example, the shifts in mean and variance are not in line. Thus it is important to adequately capture the business cycle features of the series; we argued in Section 6.2, following Sichel (1994), that for modeling U.S. output growth, a three-regime model with heteroskedastic errors appears to be required. In the case of the investment and consumption series, a closer examination of the individual models would reveal which one is the most appropriate; the results in Table 8 are simply illustrative.

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, nor 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) cannot generate steepness. For three-state models, which arguably afford a better characterization of the business cycle, the three concepts are distinct. We showed how
Figure 4. MS Models of U.S. Investment and Consumption Growth. Shown are the estimated smoothed probabilities with which each observation falls in the recession regime for a variety of models, for investment growth (top two rows) and consumption growth (bottom two rows).

Table 8. Tests for Asymmetries Using MS-AR Models of U.S. Investment and Consumption Growth

<table>
<thead>
<tr>
<th></th>
<th>Investment</th>
<th>Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MSM(2)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS:Sharpness</td>
<td>.02 [.89]</td>
<td>7.67 [.01] *</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>3.68 [.06]</td>
<td>4.25 [.04] †</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSMH(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS:Sharpness</td>
<td>2.18 [.14]</td>
<td>.26 [.61]</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.12 [.73]</td>
<td>.32 [.57]</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>MSM(3)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS:Sharpness</td>
<td>1.57 [.67]</td>
<td>3.83 [.28]</td>
</tr>
<tr>
<td>( \rho_{32} = \rho_{03} )</td>
<td>.33 [.56]</td>
<td>.03 [.86]</td>
</tr>
<tr>
<td>( \rho_{33} = \rho_{31} )</td>
<td>1.13 [.29]</td>
<td>3.09 [.08]</td>
</tr>
<tr>
<td>( \rho_{21} = \rho_{23} )</td>
<td>.15 [.70]</td>
<td>.74 [.39]</td>
</tr>
<tr>
<td>MS:Deepness</td>
<td>.79 [.37]</td>
<td>6.79 [.01] *</td>
</tr>
<tr>
<td>MS:Steepness</td>
<td>7.61 [.01] *</td>
<td>9.15 [.0] *</td>
</tr>
<tr>
<td><strong>MSMH(3)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MS:Sharpness</td>
<td>1.04 [.79]</td>
<td>1.73 [.63]</td>
</tr>
<tr>
<td>( \rho_{32} = \rho_{03} )</td>
<td>1.01 [.32]</td>
<td>.12 [.73]</td>
</tr>
<tr>
<td>( \rho_{33} = \rho_{31} )</td>
<td>.01 [.94]</td>
<td>1.69 [.19]</td>
</tr>
<tr>
<td>( \rho_{21} = \rho_{23} )</td>
<td>.01 [.93]</td>
<td>.0 [94]</td>
</tr>
<tr>
<td>MS Nondeepness</td>
<td>1.35 [.25]</td>
<td>.13 [.71]</td>
</tr>
<tr>
<td>MS Nonsteepness</td>
<td>4.77 [.029]</td>
<td>.44 [.51]</td>
</tr>
</tbody>
</table>

Nonparametric tests

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NP:Deepness</td>
<td>14.32 [0] *</td>
</tr>
<tr>
<td>NP:Steepness</td>
<td>.13 [.72]</td>
</tr>
</tbody>
</table>

**NOTE:** The data are from the FRED database (http://www.stls.frb.org/fred/data/gdp.html) and cover the period 1960–1999:2. Investment growth is the first difference of the log of investment (FRED database mnemonic GDPIC92), and consumption growth is the first difference of the logarithm of consumers’ expenditure (mnemonic PCEDC92).

* Significant at the 1% level
† Significant at the 5% level

The parameter restrictions can be applied as Wald tests, and to illustrate, reported the results of testing for asymmetries in Hamilton’s original model of U.S. output growth and in two- and three-state models of U.S. investment and consumption growth. The tests detect first-moment asymmetries and are not affected by regime-dependent heteroscedasticity, provided that this is modeled.

A comparison of the empirical results for our tests with the NP outcomes suggests that our tests have reasonable power to detect asymmetries. This was confirmed by a Monte Carlo study showing that our tests have good size and power properties and perform well relative to the NP tests. The latter are adversely affected by regime-dependent heteroscedasticity and can give misleading inferences concerning first-moment asymmetries. Moreover, our tests work reasonably well when the data are generated from other classes of regime-switching models.

**ACKNOWLEDGMENTS**

Financial support from the U.K. Economic and Social Research Council under grant L138251009 is gratefully acknowledged by both authors. All the computations reported herein were carried with the MSVAR class for Ox: (see Krolzig 1998). Helpful comments were received from two anonymous referees, as well as seminar audiences at the Bank of England, Cambridge, the 1999 Society of Economic Dynamics Conference at CRENOS, Sardinia, the 1999 Meeting of the European Economic Association, Santiago de Compostela, Exeter University, and Nuffield College. Comments from David Hendry and Neil Shephard were especially helpful.
APPENDIX: DERIVATION OF PARAMETRIC RESTRICTIONS FOR NONDEEP AND NONSTEEP PROCESSES

Proposition A.1.
An MSM($M$)-AR($p$) process is nondeep iff
\[
\sum_{m=1}^{M} \bar{\xi}_m \mu_m^3 = \sum_{m=1}^{M} \bar{\xi}_m \mu_m^3 + \left(1 - \sum_{m=1}^{M} \bar{\xi}_m\right) \mu_M^3 = 0 \tag{A.1}
\]
with \(\mu_m^* = \mu_m - \mu_x = \sum_{i=1}^{p} (\mu_m - \mu_x) \hat{\xi}_i\) 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, \(x_t - \mu_x = \mu_i + z_t\). Here \(\mu_x\) is the unconditional mean of \(x_t\),
\[
\mu_x = E[x_t] = \sum_{m=1}^{M} \bar{\xi}_m \mu_m,
\]
and both \(z_t\) and \(\mu_t\) have zero mean, \(E[\mu_t] = E[z_t] = 0\). The process \(z_t = \sum_{i=1}^{p} \alpha_i z_{t-i} + u_t\) is Gaussian and hence symmetric. The component \(\mu_t\) is potentially asymmetric and represents the contribution of the Markov chain,
\[
\mu_t = \sum_{m=1}^{M} \bar{\xi}_m \mu_m = \sum_{m=1}^{M} \bar{\xi}_m \mu_m^* + \sum_{m=1}^{M} \bar{\xi}_m (\mu_m^* - \mu_m^*),
\]
with \(\mu_m^* = \mu_m - \mu_x\) and \(\bar{\xi}_m = 1\) if the regime is \(m\) at period \(t\) and is \(0\) otherwise.

Thus the \(k\)th moment of \(\mu_t\) is given by
\[
E[\mu_t^k] = \sum_{m=1}^{M} \bar{\xi}_m (\mu_m - \mu_x)^k = \sum_{m=1}^{M} \bar{\xi}_m \mu_m^k ,
\]
where \(\bar{\xi}_m = E[\xi_m]\) and \(E[\xi_m^k] = \bar{\xi}_m\) \(\forall k\).

Using the adding-up restriction, \(\sum_{m=1}^{M} \bar{\xi}_m = 1\), we have
\[
E[\mu_t^k] = \sum_{m=1}^{M} \bar{\xi}_m \mu_m^k + \left(1 - \sum_{m=1}^{M} \bar{\xi}_m\right) \mu^k_M
\]
\[
= \mu^k_M + \sum_{m=1}^{M} \bar{\xi}_m (\mu_m^k - \mu^k_M).
\]

Proposition A.2.
An MSM($M$)-AR($p$) process is nonsteep if the size of the jumps, \(\mu_j - \mu_i\), satisfies the following condition:
\[
\sum_{j=1}^{M-1} \sum_{i=1}^{M} (\xi_i p_{ij} - \xi_j p_{ij}) [\mu_j - \mu_i] = 0. \tag{A.2}
\]

Proof. Write \(\mu_t = M \xi_t\), where \(M = [\mu_1 \cdots \mu_M]\) and \(\xi_t = [\xi_{1t} \cdots \xi_{Mt}]\). \(\xi_{it} = 1\) if the period \(t\) regime is \(m\), and \(0\) otherwise. Then \(\Delta \mu_t = \mu_t - \mu_{t-1} = M \Delta \xi_t = M \xi_t - M \xi_{t-1}\). Clearly, \(E[\Delta \mu_t] = 0\). We now introduce \(\nabla M = [M' \otimes I_M - I_M \otimes M']\) and \(\xi^{(2)}_t = \xi_t \otimes \xi_{t-1}\), such that
\[
\Delta \mu_t = \nabla M \xi^{(2)}_t = \sum_{i=1}^{M} \sum_{j=1}^{M} \xi_{i,t-1} \xi_{j,t} [\mu_j - \mu_i].
\]

Using that \(\mu_j - \mu_i = 0\) for \(i = j\), we can simplify to
\[
\Delta \mu_t = \sum_{i=1}^{M} \sum_{j=1}^{M} \xi_{i,t-1} \xi_{j,t} [\mu_j - \mu_i].
\]

The third moment is then given by
\[
E[\Delta \mu_t^3] = \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{k=1}^{M} \{(\xi_{i,t-1} \xi_{j,t}) [\mu_j - \mu_i] [\mu_j - \mu_k]\}
\]
\[
= \sum_{i=1}^{M} \sum_{j=1}^{M} \{(\xi_{i,t-1} \xi_{j,t}) [\mu_j - \mu_i]^3\},
\]
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 nonsteepness, because it implies that for all \(i, j = 1, \ldots, M\), we have that \(\xi_{i,t-1} \xi_{j,t} - \xi_{j,t} \xi_{i,t} = 0\).

[Received August 1999. Revised November 2001.]

REFERENCES


