DYNAMIC ASYMMETRIES IN U.S. UNEMPLOYMENT

GARY KOOP *
Department of Economics, University of Edinburgh,
Edinburgh, EH8 9JY, U.K.
E-mail: G.Koop@ed.ac.uk

SIMON M. POTTER †
Department of Economics, University of California, Los Angeles,
Box 951477.
Los Angeles, CA 90095-1477, U.S.A.
E-mail: potter@ovett.sscnet.ucla.edu

February 1998

ABSTRACT

We examine dynamic asymmetries in U.S unemployment using nonlinear time series models and Bayesian methods. We find strong statistical evidence in favor of a two-regime threshold autoregressive model. Empirical results indicate that, once we take into account both parameter and model uncertainty, there are economically interesting asymmetries in the unemployment rate. One finding of particular interest is that shocks which lower the unemployment rate tend to have a smaller effect than shocks which raise the unemployment rate. This finding is consistent with unemployment rises being sudden and falls gradual.

Keywords: Nonlinearity, Threshold Autoregression, Bayesian, Unemployment.
JEL: C11,C22,C52,E24.

*Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

†Financial support from the NSF under grant SES 9211726 and the Center for Computable Economics at UCLA is gratefully acknowledged.

1 Introduction

The vast majority of reduced form empirical work in macroeconomics uses linear models (for example, the entire VAR literature initiated by Sims (1980)) despite the fact that theoretical arguments for linearity rarely exist. This dominance is undoubtedly due to the convenience of linear forms: they are easy to work with and their properties are well-understood. In contrast, some recent work that emphasizes the restrictiveness of linearity (see, among many others, Hamilton (1989), Beaudry and Koop (1993) and Pesaran and Potter (1997)) has illustrated by applications to U.S. GNP that there are enormous potential benefits from working with nonlinear macroeconomic models.

However, it is probably fair to state that the general message coming out of this empirical literature is that, although there is some evidence in favor of the hypothesis that economic time series contain nonlinearities, the evidence is not overwhelming nor is it precise on the exact form or meaning of the nonlinearity. In the face of this mixed message, most economists remain unconvinced about the usefulness of considering nonlinearity in empirical specifications. For some, this reluctance to consider alternatives to linear models is due to the perceived weakness of the statistical evidence. Others accept the statistical evidence, but argue that statistically significant results are fragile due to data mining. Still others argue that statistical significance does not imply economic significance.

The present paper is a step towards developing a nonlinear modeling strategy that attempts to convert these three types of skeptics. In particular, we focus on the univariate properties of a time series, the unemployment rate, that captures some of the most important features of the business cycle. The unemployment rate has several properties that an empirical model should account for. Two properties are direct observations:

1. The series is bounded between zero and one. This implies that the unemployment rate cannot exhibit global unit root behavior.

2. As can be seen in Figure 1a and as is commonly believed unemployment dynamics are characterized by fast rises and slower falls.[1] In particular, Neftci (1984) in an influential

[1] We use the civilian male (over 20 years old) unemployment rate (calculated using Citibase LHMU and

article found statistical evidence in favor of this belief using a novel nonparametric technique.[2]

A third observation forms the main motivation for our paper. Unemployment rates are highly persistent. For example, in Figure 1a note the long decline in rates in the 1960s compared to the higher rates of the 1980s. An obvious explanation of this persistence is adjustment costs in the labor market. We focus on whether this persistence is adequately captured by an empirical model that propagates positive and negative shocks symmetrically.

There are many theoretical reasons and some empirical evidence for believing that shocks to the labor market will propagate asymmetrically in the unemployment rate. From the literature on search and matching models for the labor market there are a number of models that imply asymmetric adjustment costs. For example, Mortenson and Pissarides (1993), motivated by measures of asymmetry in job creation and destruction, develop a model where job creation takes longer than job destruction because job creation requires firms to search for good matches. In a similar vein, many search and matching models have been constructed where there is an externality similar to that described in Diamond (1982) which produces more matches at times when economic activity is higher.[3] In general, any model where adjustment costs vary over the cycle will imply asymmetries in the response of the unemployment rate to shocks. For example, the various hysteresis theories of unemployment (e.g.. Blanchard and Summers, 1986) typically focus on the persistence of shocks to unemployment without emphasizing asymmetries, but can be interpreted or extended to imply asymmetries. For instance, hysteresis arguments often hinge on the differential behavior of insiders and outsiders in wage bargaining; the idea being that insiders bargain partly out of concern for their jobs so that the current level of employment is also the equilibrium level. One only needs to add

LHMC). We made this choice to minimize the effect of structural changes on the labor market on the analysis. This monthly series starts in 1948. In our empirical section, we follow Hansen (1997) and use data starting in 1959 through 1996:7. Data from 1996:8 through 1997:6 is used in a forecasting exercise.

[2] His finding of significant asymmetries was questioned by Sichel (1989), who found an error in Neftci's calculations. However, Sichel also argued that Neftci's test had very low power. Rothman (1991) added to the debate by using a modified version of Neftci's tests and finding marginally statistically significant asymmetries in unemployment. Brock and Sayers (1988) found considerable evidence of nonlinearity in unemployment using the nonparametric BDS test.

[3] Burgess (1992) and Storer (1994) examine similar models empirically.

transaction or adjustment costs to this model to obtain a plausible story of asymmetries; viz. small shocks will have no permanent effect on employment since it does not pay to renegotiate contracts, but large shocks will have a permanent effect.

In order to capture asymmetric propagation of shocks, linear models are inadequate. Thus, nonlinear models are required to assess the validity of these descriptions of labor market dynamics. One approach would be to construct structural models of the labor market. A number of authors have attempted to do this (for example, the collection of papers on asymmetries in labor market dynamics in Van Ours *et al*, 1993). However, in order to obtain a form that allows estimation at the aggregate level, a number of simplifying assumptions are made that have left many unconvinced (see the review by Rogerson, 1995). Empirical work at the microeconomic level by Davis and Haltiwanger (1990) has been influential and led to some promising approaches that hold out the hope of developing structural models at the microeconomic level that can be consistently aggregated to produce implications for unemployment. The empirical approach of this paper complements previous analyses by using a relatively simple reduced form nonlinear model to quantify the extent of dynamic asymmetries without committing to their source.

Unlike reduced form linear models, an immediate problem arises with nonlinear models: it is easy to say that a nonlinear model is required, but much harder to choose one model out of the myriad of possibilities. One modeling choice seems uncontroversial to us. Instead of analyzing the unemployment data directly we work with a logistic transformation. We will show that this simple nonlinear transformation captures the first two properties of the unemployment rate described above even when a linear model (for the transformed data) is used.

Within the class of possible nonlinear models for the transformed data, we concentrate on the threshold autoregressive model.[4] Threshold autoregressions (TAR hereafter) are the most widely used class of models in the nonlinear time series literature and are extensively described by their originator Howell Tong (1990). They work by splitting the time series (endogenously) into different regimes. Within each regime the time series is assumed to be described by a

[4]Previous nonlinear time series models of unemployment have been examined by Ham and Sayers(1992) Rothman (1992), Johnson and McCelland (1995) and in particular Hansen (1997).

linear model. The main effort in estimation of TAR is in deciding upon definitions for the different regimes and their number. We examine three possible ways of defining regimes based on: the previous level of unemployment, previous changes in unemployment and averages of previous changes in unemployment and allow for one, two or three regimes.

In terms of techniques used to analyze the TAR we depart from most of the previous literature by using Bayesian methods. Bayesian methods allow us to address directly the concerns of the three types of skeptics of nonlinear modeling. Firstly, as shown in Koop and Potter (1998) in testing nonlinearity versus linearity, Bayesian tests are more conservative (in favor of the linear *status quo*) than classical tests because they include an Occam's razor type penalty for unnecessary complexity. Secondly, Bayesian methods allow one to jointly compare the evidence for a range of linear and nonlinear models rather than being limited to pairwise comparisons using classical methods. Further, instead of focusing on one particular representation of possible nonlinearities (i.e., one choice for the type of threshold structure), we can use the posterior probabilities of each model (including the linear model) to produce an 'average' measure of the amount of nonlinearity present. Thus, we are able to measure asymmetries in the dynamics of the unemployment rate allowing for both parameter and model uncertainty.

Empirical results from this 'weighted average' model indicate that important nonlinearities exist in the US unemployment rate and these nonlinearities offer a deeper understanding of US labor force dynamics than linear models. For instance, impulse responses indicate that, when the unemployment rate is falling quickly, positive shocks (which would tend to increase unemployment) have a much larger effect (in absolute value) than negative shocks. Such a finding is consistent with much of the theoretical literature cited above. They also complement the nonlinear empirical literature involving real output (see Beaudry and Koop, 1993 and Potter 1995) that finds recessionary shocks are less persistent than expansionary shocks. This finding could not be produced by using a reduced form linear model.

The remainder of the paper is organized as follows: Section 2 discusses the data transformation and threshold models. Section 3 describes our empirical techniques. Section 4 presents empirical results, including a Monte Carlo study, impulse response analysis and forecasting exercise. Section 5 is a brief conclusion. There are three appendices. Appendix A gives details on our computational techniques. Appendix B contains some results on the sensitivity for model selection results to the prior. Appendix C gives information on the properties of the nonlinear model with the highest posterior probability and presents a classical analysis of the same model.

2 NONLINEAR MODELS FOR UNEMPLOYMENT

2.1 Transforming the Unemployment Data

Since the unemployment rate lies between zero and one it is bounded. If we were to work directly with this series, the assumption of a symmetric error process would be inappropriate. Furthermore, the bounded nature of the unemployment rate guarantees 'bounded behavior', a feature often overlooked by those who test for a unit root in this series. Statisticians have often stressed the importance of developing a model with reasonable limiting behavior. For instance, Cox and Hinkley (1974, p.6) argue that "even though this limiting behavior may be far from the region directly covered by the data, it will often be wise to use a family of models consistent with the limiting behavior."

Statistical analysis of bounded variables is quite difficult. For instance, to ensure that fitted and forecast values of a series lie in $[0, 1]$ a linear regression must have an error structure which is bounded in a complicated way. Wallis (1987) recommends working with transformations of bounded series; in particular, the logistic transformation. We follow this suggestion and work with $Y_t = \ln(\frac{U_t}{1-U_t})$, where U_t is the unemployment rate.

The transformed series is now unbounded and we assume it is reasonable to take the errors as conditionally Gaussian. Figure 1b shows the transformed and untransformed time series, both normalized to have zero mean and unit variance. The transformation does have some effect, particularly near peaks and troughs in the series. Even the linear model estimated on the transformed unemployment data will imply nonlinearities when converted back to the original form of the data.

An issue that arises for unemployment data is whether to impose a stationarity condition on the models being investigated. In the case of the untransformed series it is clear that one needs to do this to ensure bounded behavior. For the transformed data it is less obvious whether to impose stationarity or not. In this paper we have decided not to explicitly restrict the models being examined to the stationary regime in the model selection phase of the analysis. In terms of the untransformed series this means we allow for the possibility of long term structural shifts in the level of unemployment.

2.2 TAR Models

Let $\{Y_t : t = 0, 1, \ldots\}$ be a time series and let J_t be an index random variable taking values in the set $\{1, 2, \ldots, K\}$. Then a threshold autoregression is defined by:

$$Y_t = \alpha^{\{J_t\}} + \phi^{\{J_t\}}(L)Y_{t-1} + \sigma^{\{J_t\}}V_t, \tag{1}$$

where V_t is an IID sequence of standard normal random variables; and for $J_t = j$, $\alpha^{\{j\}}$ is a constant, σ^j are regime specific scalings of the innovation and $\phi^{\{j\}}(L)$ is a finite order polynomial in the lag operator L.

In the present application, the most general model we consider is a three-regime TAR (3TAR). We consider a number of possible choices for the index variable amongst the class of indicator functions applied to past values of the time series. We label the time series of past values, X_t. In particular, we have

$$Y_t = \begin{cases} \alpha_1 + \phi_1(L)Y_{t-1} + \sigma_1 V_t & \text{if } X_t < r_1, \\ \alpha_2 + \phi_2(L)Y_{t-1} + \sigma_2 V_t & \text{if } r_1 \leq X_t < r_2, \\ \alpha_3 + \phi_3(L)Y_{t-1} + \sigma_3 V_t & \text{if } X_t > r_2, \end{cases} \tag{2}$$

where $\phi_j(L)$ is a polynomial of order p_j in the lag operator. This model is parameterized in terms of $\beta_j = (\alpha_j, \phi_{j1}, \ldots, \phi_{jp_j})'$, $\beta = (\beta_1', \beta_2', \beta_3')'$, $\sigma = (\sigma_1, \sigma_2, \sigma_3)'$, $\gamma = (r_1, r_2, d)'$, $p = (p_1, p_2, p_3)'$ and $\psi = (\beta', \gamma', p')$.

We consider three definitions of X_t:

1. $X_t = \Delta Y_{t-d},\ d = 1, \ldots, p.$

2. $X_t = Y_{t-d},\ d = 1, \ldots, p$

3. $X_t = \frac{Y_{t-1} - Y_{t-d-1}}{d}$, $d = 1, \ldots, p$.

With the first and third type of index variable specification, the model allows for the dynamics of the unemployment rate to differ in times when it has risen rapidly, fallen rapidly, or changed little in recent periods. The difference between them is that the d-th difference index is less sensitive to monthly changes that are not permanent. The second type of threshold nonlinearity is included to capture two possibilities. Firstly, the level of the unemployment rate may affect the dynamics even after the logistic transformation for the reasons discussed in the introduction. Secondly, if the dominant feature of the sample turns out to be secular change then this choice of X_t will be very similar to the choice of a model of structural change.

We simplify the analysis by assuming that the order of the autoregression is the same across all regimes.[5] In our empirical study of monthly unemployment data, we follow Hansen (1997) and set the maximum autoregressive lag at 13.[6] We also consider two regime TARs (2TARs) and homoskedastic versions of TARs as separate classes of models. This gives us 13 separate classes of models to examine:

1. Linear: $\alpha_1 = \alpha_2 = \alpha_3$, $\phi_1 = \phi_2 = \phi_3$, $\sigma_1 = \sigma_2 = \sigma_3$, and the definition of X_t is irrelevant.

2. 2TAR1HOM: $\alpha_2 = \alpha_3, \phi_2 = \phi_3, \sigma_1 = \sigma_2 = \sigma_3$, $X_t = \Delta Y_{t-d}, d = 1, \ldots, p$.

3. 2TAR1HET: $\alpha_2 = \alpha_3, \phi_2 = \phi_3, \sigma_2 = \sigma_3, X_t = \Delta Y_{t-d}, d = 1, \ldots, p$.

4. 2TAR2HOM: $\alpha_2 = \alpha_3, \phi_2 = \phi_3, \sigma_1 = \sigma_2 = \sigma_3$, $X_t = Y_{t-d}, d = 1, \ldots, p$.

5. 2TAR2HET: $\alpha_2 = \alpha_3, \phi_2 = \phi_3, \sigma_2 = \sigma_3, X_t = Y_{t-d}, d = 1, \ldots, p$.

6. 2TAR3HOM: $\alpha_2 = \alpha_3, \phi_2 = \phi_3, \sigma_1 = \sigma_2 = \sigma_3$, $X_t = \frac{Y_{t-1} - Y_{t-d-1}}{d}, d = 1, \ldots, p$.

7. 2TAR3HET: $\alpha_2 = \alpha_3, \phi_2 = \phi_3, \sigma_2 = \sigma_3, X_t = \frac{Y_{t-1} - Y_{t-d-1}}{d}, d = 1, \ldots, p$.

8. 3TAR1HOM: $\sigma_1 = \sigma_2 = \sigma_3$, $X_t = \Delta Y_{t-d}, d = 1, \ldots, p$.

[5] This removes a degree of freedom from the nonlinear model that could lower the posterior probability of nonlinearity.

[6] Hansen uses differenced data and sets $p = 12$, implying $p = 13$ for levels data.

9. 3TAR1HET: $X_t = \Delta Y_{t-d}, d = 1, \ldots, p$.

10. 3TAR2HOM: $\sigma_1 = \sigma_2 = \sigma_3$, $X_t = Y_{t-d}, d = 1, \ldots, p$.

11. 3TAR2HET: $X_t = Y_{t-d}, d = 1, \ldots, p$.

12. 3TAR3HOM: $\sigma_1 = \sigma_2 = \sigma_3$,$X_t = \frac{Y_{t-1} - Y_{t-d-1}}{d}, d = 1, \ldots, p$.

13. 3TAR3HET: $X_t = \frac{Y_{t-1} - Y_{t-d-1}}{d}, d = 1, \ldots, p$.

Our choice of models and the logistic transformation gives a wide range of possible likelihood functions through which we view the unemployment rate data. A major problem for any empirical analysis once one leaves the world of linear models is to decide which of the infinite types of nonlinearity to adopt. The theories discussed in the first section of our paper suggest that unemployment dynamics may differ across regimes, and that these regimes should reflect the tightness of the labor market or the state of the business cycle in some sense. But theory offers us little guidance as to the number of regimes. That is, it does not say whether there should be two regimes reflecting "good times" and "bad times" or, as Pesaran and Potter (1997) find for GDP, three regimes reflecting "good", "bad" and "normal" times. Theory also tells us little about what should trigger regime switches. That is, it does not say whether regimes should be defined according to levels or rates of change of lagged dependant variables. Furthermore, theory provides little help in timing and lag selection issues. We feel that our approach covers virtually all reasonable possibilities that are in accord with theoretical possibilities. One advantage of our Bayesian approach is that we can easily work with large numbers of models. In our case, we have 13 classes and consider $p = 1, 2, ..., 13$ for a total of 165 models.[7] Classical econometric approaches in the literature typically select a small number of models and work with one preferred model. For instance, Hansen (1997) considers only homoskedastic 2TAR models with one choice for p and two threshold definitions.

An alternative class of potentially useful models is discussed in Hamilton (1989). This class is similar in spirit to the TAR. However, simple versions of these so-called Markov switching

[7] That is, $13 \times 13 - 4 = 165$, where the subtractions arise since, for $p = 1$, JTAR1 and and JTAR3 models are identical for J=2,3.

models assume only two regimes and that the probability of switching between regimes depends only on which regime was applicable in the last period. In a serious empirical exercise, it is necessary to allow for the possibility of more regimes and to allow the switching probability to have longer memory. Such extensions are trivial in TAR models, but much more difficult in Markov switching models. Furthermore, from either a Bayesian or a classical econometric perspective, TAR models are easier to work with (see Koop and Potter, 1998, and Hansen, 1992, 1996).

3 BAYESIAN TECHNIQUES

3.1 Bayesian Analysis of Threshold Models

Since the index variable is constructed from the location of observable lags of the time series, the TAR specifications can be estimated and evaluated using classical statistical methods as described in Tong (1990). However, the unconventional likelihood function of TAR models causes many problems for classical statistical inference.[8] It is very difficult to obtain a tractable form for the sampling distribution of the threshold estimates since they converge at a faster rate than the square root of the sample size. Further, because of their superior speed of convergence one is unable to relate uncertainty in the true location of the thresholds to the sampling uncertainty in other parameters in the model (see Chan, 1993). Asymptotically, one can ignore the uncertainty over the delay and thresholds in measuring the uncertainty over the other parameters. In practice, experience suggests that in finite samples there is substantial covariance between the thresholds and other parameter estimates. Further, the finite sample likelihood concentrated with respect to the threshold parameters is very irregular as one would expect for a threshold model and usually has multiple peaks in the case of more than two regimes (see Pesaran and Potter, 1997, Figure 2.b).

These features of the likelihood function of TAR models make it very hard to extract good measures of uncertainty for parameter estimates using classical statistical techniques.

[8]The likelihood function is discontinuous in the thresholds and also contains numerous flats as the threshold value changes but the underlying index variable does not.

For example, popular techniques such as the bootstrap will not capture the multiple peaks in the sample likelihood function since they are centered at the maximum likelihood estimator and only provide information about the uncertainty at this point.[9] In particular, classical methods have major problems in using the model for out of sample forecasting or generating dynamic properties since one is unable to accurately capture the uncertainty surrounding regime classification.[10] Bayesian methods, on the other hand, average over the entire parameter space and the irregular properties of the likelihood function do not hinder the derivation of the posterior. In other words, they provide exact small sample results for a model where asymptotic approximations are guaranteed to be poor.

Other models such as smooth transition autoregressions (Chan and Tong, 1985) or Markov switching impose a considerably greater burden in estimation than the TAR class of models. Furthermore, nonlinear time series (by definition) suffer generically from the problem of likelihood functions with multiple peaks in finite samples.

Bayesian techniques for analyzing simple TAR models are given in Geweke and Terui (1993). These authors derive the posterior density of the parameters and advocate the use of Monte Carlo integration for drawing posterior and predictive inferences. In Appendix A we describe the form for the posterior in our model using informative priors, discuss calculation of marginal likelihoods and features of the posterior and advocate the use of analytical methods.

3.2 Bayesian Model Selection

Our main reason for the adoption of Bayesian methods is that they allow us to explicitly include measures of uncertainty over model type. In the Bayesian approach, no one model has to capture the true data generating process. Instead of choosing just one model, we can weight features of interest (such as impulse responses) from different models by their posterior model probabilities. Leamer (1978) argued that such a weighted average of the properties of different models is particularly attractive in disciplines like economics where theoretical considerations might not give sharp views on which specification is best. The classical approach typically

[9] Sims and Zha (1994) make a similar point in the case of the possibility of unit roots.

[10] The methods discussed in Hansen (1997) provide classical confidence intervals for threshold estimates conditional on the delay lag and do not supply a joint distribution for all the estimates.

presents results based on one peak of the one likelihood function. The Bayesian approach used in this paper presents results based on the whole likelihood function weighted by the prior for several different types of likelihood functions.

Posterior model probabilities can be calculated based on Bayes factors comparing the various models. The Bayes factor for comparing model 1 to model 2 parameterized by θ, η respectively is given by:

$$B_{12} = \frac{\int \ell(\theta) b(\theta) d\theta}{\int \ell(\eta) b(\eta) d\eta},$$

where $\ell(\cdot)$ is the likelihood function and $b(\cdot)$ is the prior belief for the parameters of the models.

The Bayes factor gives the posterior odds in favor of model 1 over model 2 when the prior odds are unity. In the types of models we examine in this paper the marginal likelihoods can be found analytically (see Appendix A) and, hence, the computational burden is not high.

Koop and Potter (1998) discuss the advantages of a Bayesian approach for model comparison in the context of testing for nonlinearities in economic time series. Classical tests for nonlinearity typically run up against the so-called Davies' problem: nuisance parameters that are present under the alternative are unidentified under the null.[11] Andrews and Ploberger (1994) show that 'optimal' solutions to Davies' problem are in the form of average exponentials of standard test statistics (e.g. the LM statistic) where the average is taken over the unidentified nuisance parameters. However, since these parameters are unidentified under the null, there is nothing to pin down the distribution used in the averaging. Thus, the classical approach to model selection is based on a subjectively chosen weighting scheme. The Bayes factor, too involves integrating out the nuisance parameters. However, the integration is performed with respect the posterior distribution of the nuisance parameters.

In addition, Bayes factors include an automatic penalty for more complex models. This Occam's razor property is of great use in nonlinear model selection given the risks of overparameterization with such models. The intuition behind this property of Bayes factors is quite simple in the case of linear versus nonlinear models. Linear models are capable of explaining a smaller range of types of time series data than nonlinear models. If the time series data is approximately linear, the linear model will tend to have a higher marginal likelihood than the

[11]In our case, γ is unidentified for the linear model.

nonlinear model which places considerable weight on time series data that are far from linear.

A major difficulty for many with a Bayesian analysis is the need to specify a prior distribution. Geweke and Terui (1993) carry out a Bayesian analysis of threshold autoregressive models using a 'noninformative prior'.[12] However, for our purposes we require an informative prior. It is well known in the Bayesian literature that in a nested testing situation, improper priors (i.e., priors which do not integrate to one) typically result in Bayes factors containing little information from the observed sample. In general, if priors are too flat and noninformative relative to the data, the restricted (in our case linear) model will always be selected, even if the unrestricted (nonlinear) model is the 'correct' one. The reason for this is that from the point of view of Occam's razor, the nonlinear model under a very flat prior is able to explain an enormous variety of data sets containing nonlinearity. Hence, the marginal likelihood of the observed sample will be low even if it contains important nonlinearities. Alternatively very tight priors on the parameters centered at the same values for the linear and nonlinear models lead to Bayes factors equal to 1 since *a priori* the two models are assumed to be close and the sample information cannot change this too much.

4 Empirical Results

4.1 Prior Selection

We use a prior that attempts to capture some of the theoretical considerations discussed above. We start by considering possible values of the thresholds. For the 3TARs we assume that the prior support of the two thresholds is non-intersecting. In particular r_1 is assumed to take on negative values only and r_2 positive values only.[13]

We take $b(\gamma) = b(r_1)b(r_2)b(d)$ where $b(r_1)$ and $b(r_2)$ are uniform (for the transformed data) over $[\tilde{r}_1, 0]$ and $[0, \tilde{r}_2]$, respectively. The hyperparameters $\tilde{r}_1 < 0$ and $\tilde{r}_2 > 0$ are chosen to

[12] The issue of noninformative prior selection in time series is controversial (see Phillips (1991) plus accompanying discussion). In order to keep the computational burden reasonable we do not attempt to take these issues into account. Since we are not directly interested in testing for 'unit roots' we do not think this is a major issue.

[13] In the case of the 3TAR2 model negative and positive should be interpreted relative to mean adjusted data.

ensure that at least 15 per cent of the observations lie in each of the upper and lower regimes. For the 2TAR, $b(r_1)$ is chosen as above, except that it is no longer restricted to be negative. These priors ensure that our model is not merely fitting a small number of outliers. For both the 3TAR and the 2TAR, $b(d)$ is assumed to be uniform over the integers $1, 2, \ldots, p$.

As discussed in Appendix A, conditional on γ, the heteroskedastic JTARs (where J=2,3) break down into J different regression models, each using a different subset of the data. For each of these J models we use a standard Normal-inverted Gamma prior for the regression coefficients and error variance (see Judge, Griffiths, Hill, Lutkepohl and Lee, 1985, pp. 106-110). Formally, we let $b(\beta_j, \sigma_j^2)$ have a Normal-inverted Gamma distribution. We assume prior independence across regimes and between γ and the other parameters so that the prior can be written as:

$$b(\phi, \sigma, \gamma) = b(\gamma) \prod_{j=1}^{J} b(\beta_j, \sigma_j^2).$$

The homoskedastic JTAR's can, through the use of appropriate dummy variables, be written in the form of a linear regression model, conditional on γ. The prior for these models can be written as $b(\beta, \sigma_1^2, \gamma) = b(\gamma) b(\beta, \sigma_1^2)$ where $b(\gamma)$ is described above and $b(\beta, \sigma_1^2)$ is Normal-inverted Gamma. For the linear autoregressive model, we let $b(\beta_1, \sigma_1^2)$ be Normal-inverted Gamma.

It remains to specify the hyperparameters of the Normal-inverted Gamma priors. Precise details are given in Appendix A. The results presented in this section are based on a prior that is reasonably flat. Furthermore, the various models have basically the same prior. The relative flatness of the prior tends to stack the odds in favor of the linear model as we noted above.

4.2 Performance of Bayes Factors in Artificial Data Sets

Before examining the properties of U.S. unemployment, it is useful to briefly examine the properties of our Bayesian methodology using simulated data. To this end, we artificially simulated 100 data sets, each of length 200, from each of six data generating processes. These six DGPs are selected to represent a wide variety of behavior, but are not intended to be exhaustive. We focus on linear models (i.e. to investigate "size" issues) and 3TAR's (i.e. to investigate "power" issues). We do not consider 2TAR DGPs on the grounds that, if our methods work well for less parsimonious 3TARs, they will likely also work well for 2TARs. All of the DGPs are restricted versions of (2) where intercepts are always set to zero (i.e. $\alpha_1 = \alpha_2 = \alpha_3 = 0$) and, with one exception, $p_1 = p_2 = p_3 = 1$. The exception is DGP_2 which is an AR(2) model. The five DGPss are given by:

1. DGP_1: $\phi_{11} = \phi_{21} = \phi_{31} = 0.8, \sigma_1 = \sigma_2 = \sigma_3 = 1.0$.

2. DGP_2: $\phi_{11} = \phi_{21} = \phi_{31} = 1.2, \phi_{12} = \phi_{22} = \phi_{32} = -0.6, \sigma_1 = \sigma_2 = \sigma_3 = 1$.

3. DGP_3: $\phi_{11} = 0.2, \phi_{21} = 0.8, \phi_{31} = -0.5, \sigma_1 = \sigma_2 = \sigma_3 = 1, r_1 = -0.5, r_2 = 0.5, X_t = Y_{t-1}$.

4. DGP_4: same as DGP_3, except $\sigma_1 = 0.5, \sigma_2 = 1.0, \sigma_3 = 1.5$.

5. DGP_5: same as DGP_4, except $\phi_{11} = \phi_{21} = \phi_{31} = 0.8$.

6. DGP_6: same as DGP_3, except $X_t = \Delta Y_{t-1}$.

In other words, we simulate from a linear AR(1), AR(2), a homoskedastic 3TAR, a heteroskedastic 3TAR, a linear heteroskedastic model and another homoskedastic 3TAR. The latter 3TAR has thresholds triggered by lagged differences, while the former use lagged levels. The prior is identical to that used with the unemployment data, except that the prior mean of σ_j^2 is set to 1.0. We set the maximal lag length to 2.

Table 1 presents posterior model probabilities averaged across the 100 simulated data sets. The correct DGP is indicated by a *. It can be seen that the Bayesian methodology tends to allocate most of the posterior model probability in virtually every case. Results for DGP_1 and DGP_2 indicate that, if the DGP is truly linear, then our Bayesian methodology indicates this. A comparison of results for these two DGPs also indicates that the methodology also works well in finding the correct lag length.

DGP_3 is a nonlinear homoskedastic model with one lag with threshold defined by lagged levels, and our methods allocate, on average, over thirty per cent of the posterior model probability to this model. However, substantial weight is allocated to the linear and to the corresponding 2TAR model. This finding is undoubtedly due to the large prior standard deviations on the autoregressive coefficients. As discussed previously, Bayes factors have a strong reward for parsimony built in, especially if the prior is relatively noninformative. We view this property in a positive light: only if there is overwhelming evidence against the linear status quo will our methods indicate nonlinearity.

The next two DGPs exhibit heteroskedasticity. The posterior model probabilities indicate this strongly and in all simulated data sets. As before, the corresponding 2TAR's receive some posterior model probability indicating the Occam's razor property of Bayes' factors, especially where relatively flat priors are used.

DGP_6 was introduced to see if our Bayesian methodology is good at selecting the correct threshold definition. Results strongly indicate that it is.

Overall, we find these Monte Carlo results to be strongly supportive of the theoretical properties discussed in the previous section of this paper. Posterior model probabilities based on Bayes factors do seem to perform well in repeated samples and contain a strong reward for parsimony.

4.3 Model Selection Results

We calculate posterior probabilities for the various models using equal prior probabilities for each model.[14] The results are presented in Table 2. Since none of the 3TAR models receive probability greater than 0.00001, we do not include these models in the table. Furthermore, since no model with $p > 8$ receives appreciable posterior model probability, we do not present results for high lag lengths individually. The poor performance of these classes of models with a large number of parameters is undoubtedly due to the reward for parsimony built into Bayes factors.

It is clear that the 2TAR3HET models (i.e. two regime heteroskedastic TAR with regimes

[14] In particular, since we have 165 models each receives prior probability equal to 1/165. If the reader wishes to choose different prior model probabilities, he/she can simply reweight the numbers in the tables. For instance, if one wishes to double the prior weight attached to each of the linear models, then the model probabilities in the "Linear" row of Table 2 can be doubled and the other rows downweighted appropriately.

defined by average differences of the dependant variable) receive the majority of the posterior model probability. Note that it is the average change, rather than the level, of the unemployment rate which defines the two regimes in the model. The preferred model has $p = 5$, and for all model classes this tends to be the preferred lag length. If we average across values of p then the posterior probability of the 2TAR3HET is 0.802, indicating this model in particular and nonlinearity in general is strongly supported in this data set. The remainder of the probability is allocated to the 2TAR3HOM models (0.143) and the linear models (0.055). The prior sensitivity in Appendix B suggests these results are robust to changes in the prior. It is worth noting that much of the nonlinearity we find seems to be of a heteroskedastic nature (i.e. nonlinear heteroskedastic models tend to outperform nonlinear homoskedastic models). Appendix C contains a brief discussion of parameter estimates and a comparison with the classical analysis of Hansen (1997). Suffice it to note here that our results are in line with those obtained by a classical econometric study of this data set.

4.4 Generalized Impulse Responses

Potter (1997) presents a thorough discussion of impulse response analysis in nonlinear univariate models. Koop (1996) extends the discussion by allowing for parameter uncertainty using Bayesian methods. The generalized impulse response (GI) developed in these papers measures the effect on a time series of a shock relative to a suitably-defined base case. These papers note that, unlike in linear models, impulse responses depend both on the initial conditions when the shocks hits and the other shocks in the model. To motivate this, consider a simple AR(1) model with autoregressive coefficient ϕ. Any realization of this series can be written as:

$$Y_{t+n} = \phi^{n+1} Y_{t-1} + \sum_{i=0}^{n} \phi^i V_{t+n-i}.$$

This expression depends on Y_{t-1} (the initial condition or history, which we will denote in the general case ω_{t-1}), the shocks $V_t, \ldots, V_{t+n}$, and the parameter, ϕ. In a linear model, when we take the difference between a perturbed and base case, many simplifications arise. However, with nonlinear models we have to take into account the effect of initial conditions and future shocks and their interaction with the estimated parameters. Many different types of generalized impulse responses can be defined depending on the treatment of these three factors. In this paper, we always integrate out the parameters with respect to the posterior distribution and all shocks expect for V_t in the perturbed case. We condition on values for V_t and the initial condition ω_{t-1}. defined by the particular specification of the 3TAR or linear model drawn from the posterior.

Formally, we define the generalized impulse response at horizon n as:

$$GI(n, V_t = \delta_1 \sigma X_t, \Delta U_{t-i} = \delta_2 : i = 1, \ldots, p, U_{t-p-1} = \delta_3) =$$

$$E[U_{t+n} | V_t = \delta_1 \sigma(X_t), \Delta U_{t-i} = \delta_2 : i = 1, \ldots, p, U_{t-p-1} = \delta_3]$$

-

$$E[U_{t+n} | \Delta U_{t-i} = \delta_2 : i = 1, \ldots, p, U_{t-p-1} = \delta_3].$$

The expectations in the previous expression are taken over all parameters in the model as well as the errors which are not specified as conditioning arguments (i.e. $V_{t+1}, .., V_{t+n}$ for the first and $V_t, .., V_{t+n}$ for the second expectation). Note that we consider the effect of a shock of size $\delta_1 \sigma(X_t)$, which hits after a period which unemployment has been growing by δ_2 per cent for each of the last p years starting from a level of δ_3. That is, δ_2 and δ_3 define the initial condition when the shock hits. The standard deviation of the shock in the HET models will depend on the regime. Hence, we make σ a function of X_t. In principal, we could integrate out δ_2 and δ_3 (see Koop, 1996), but it is often revealing to consider the effect of shocks which hit at various points on the business cycle. Note also that we are calculating the effect of shocks to the unemployment rate (U_t), not to Y_t, the logistically transformed series.

We consider shocks equal to twice the standard deviation (from a particular draw of the posterior) of the innovation for each model. That is, $\delta_1 = \pm 2$. Hence, we are considering the effects of shocks that are large, but not unreasonably so. For the history, we try three different setups corresponding to a shock hitting when the unemployment rate is

1. **Fast Decrease** i.e., falling rapidly,

2. **No Change** i.e., constant

3. **Fast Increase** i.e., rising rapidly.

In particular, we choose initial conditions based on data considerations. δ_3 is set to the mean unemployment rate (4.72%) for all setups and δ_2 is selected based on the average absolute value of the monthly change in the unemployment rate (i.e. 0.15%). That is, $\delta_2 = -0.15\%, 0.0\%, 0.15\%$ correspond to fast decrease, no change and fast increase setups. Since we report the impulse responses for the original unemployment series (i.e.. U_t) a value of 1.00 implies a 1 point rise in the level of the unemployment rate relative to the base case (i.e., the second conditional expectation in the definition of the GI). Recall that, because of the logistic transformation, even the linear model will show some nonlinearities in its GI. As described in Koop (1996) or Koop, Pesaran and Potter (1996), GIs can be calculated in a simple manner using simulation methods. That is, we can simulate from the posterior for the parameters and the distribution of the errors to construct simulated draws of Y_{t+n}, which can be transformed to provide draws of U_{t+n}. Averages of these latter draws will converge to the expectations in the formula for the GI.

We concentrate on GI averaged across all models.[15] The weights in this average are given by the posterior model probabilities given in Table 2. Appendix C presents results for the preferred model (2TAR3HET with $p = 5$) and the most popular linear model ($p = 5$) separately. We present the results in various forms. Table 3 contains all the information on the responses for the 24 month horizon. In order to help describe the various dynamic asymmetries we also present the information in graphical form. It is worth stressing that we have six GI's, coming from two shocks and three histories. We start by grouping the GI by type of shock (i.e. positive/negative) in Figures 2a and 2b. We also normalize so the initial effect of the shock is unity for all histories. In Figures 2c-2e we plot the effect of the two different shocks for each history (i.e. fast decrease/no change/fast increase). The GI's in these latter figures are not normalized. To aid in comparison, we take the absolute value of the GI for the negative shock.

Three main findings are evident in the GI's. First, the classic hump-shaped pattern often

[15] In practice, to save computer time we average across all models which receive posterior model probability of 1% or more.

seen in the impulse response functions of business cycle variables is still clearly present. Secondly, Figure 2 shows clearly that the effects of positive and negative shocks are asymmetric. In particular, positive shocks (which raise unemployment) tend to have a larger effect than negative shocks (which lower unemployment). This is consistent with the commonly observed fast rise/slow fall behavior of the unemployment rate. Thirdly, the effect of shocks differs markedly over the histories. In particular, shocks which hit during times when the unemployment rate is decreasing have a larger effect than those which hit when the unemployment rate is stable. Shocks which hit when the unemployment rate is increasing have least effect. This finding probably relates to the hysteresis of the unemployment rate. That is, if unemployment is more sluggish during bad times (i.e. when it is increasing), then shocks should have little effect in bad times. This is exactly what we observe in Figure 2d.

The GIs reflect all the models considered, but the 2TAR3HET with $p = 5$ receives most of the weight in the averaging process. In fact, the overall GI's look quite similar to those for the preferred model (see Appendix C). However, the GIs coming from the linear model are not that different from the 2TAR. The asymmetry between positive and negative shocks is more pronounced for the nonlinear model, but the linear model exhibits the same type of history dependence as the nonlinear ones. This latter finding is due to the logistic transformation. In the present application, this transformation of the series has a role at least as important as the nonlinear modelling in uncovering important regularities in the unemployment rate.[16]

Overall, there seem to be several interesting dynamic asymmetries in the U.S. unemployment rate. From one point of view, we are merely discovering stylized facts about this variable that are commonly known by macroeconomists. However, it is worth emphasizing that such stylized facts could not be found using linear methods and an untransformed unemployment rate. Furthermore, we are able to exactly quantify the degree of asymmetry in a manner that is difficult to do using alternative methodologies. The GIs presented here relate to the mean of the series rather than the variance. The fact that the heteroskedastic models receive so

[16]Note that our transformation is not imposing asymmetries in the model in the same way that a linear model with untransformed data would impose symmetry. It would be possible for our TAR models to remove the effect of the transformation. That is, it is possible for the nonlinear model with transformed data to look like the linear model with untransformed data.

much support indicates that there may also be interesting nonlinearities in the variance. The GI methodology can be used to investigate the effect of shocks on the variance of the series. However, to do so here would be beyond the scope of the present paper.

4.5 Forecasting Exercise

Throughout our empirical work, we use data from 1959:1 through 1996:7, even though data through 1997:6 is currently available. This is partly done to aid in comparison with the results of Hansen (1997). However, the withholding of data from 1996:8 through 1997:6 allows us to carry out a forecasting exercise. As for the GIs, we present results averaged over all models.[17] Figure 3a contains our results. If we let U_t^* for t=1996:8,...,1997:6 denote the predictive random variables, then Figure 3a is plotted using $E(U_t^*)$ and $VAR(U_t^*)$. Note that the actual value of unemployment does tend to lie with the band defined by the predictive mean +/- two standard deviations, although our model is tending to predict too high unemployment rates. This is seen most clearly in Figure 3b where predictive means for several models are presented. Even though 2TAR3HET with p=5 does tend to predict the best, all models tend to be predicting poorly. If we examine the actual data, note that $U_{1996:7} = 4.62\%$ while $U_{1996:8} = 4.23\%$. In other words, there is a sizeable drop in unemployment just as we begin our out-of-sample period.

Even though all models forecast somewhat poorly, it is worth stressing that the 2TAR3HET with p=5 does outperform the linear AR(5) model. This is noted most clearly by considering the expectation of predictive sum of squares errors:

$$PSSE = E(\sum_{i=1996:8}^{1997:6} (U_t^* - U_t)^2),$$

where the expectation is with respect to the posterior of the parameters and models. $PSSE = 7.941$ for the 2TAR3HET with p=5 while $PSSE = 8.231$ for the AR(5). This provides additional evidence in support of the nonlinear specification.

[17]Note that variances cannot simply be averaged across models since they are not moments. Rather we average first and second moments across models and then use these to build up an overall variance.

5 Conclusions

In this paper, we have advocated the use of nonlinear methods and suitable transformation of the variable of interest for macroeconomic modeling, arguing that linear models using untransformed data are too restrictive to uncover important features of many economic time series. Unfortunately, with nonlinear models we run into problems of model choice and possible over-parameterization. As a way of getting around these problems, we use multiple-regime TAR models which have several advantages. Firstly, they are based in economic theory in the sense that they can exhibit the types of dynamic asymmetries that theoretical labor models lead us to expect. Secondly, even though they are very flexible, they are reasonably parsimonious. Thirdly, they are computationally easy to work with.

We use Bayesian methods to estimate our models since these allow for the recovery of small-sample properties in a case where we would expect classical asymptotic approximations to be very poor. Furthermore, classical model selection methods are plagued by Davies' problem, which is not a problem for Bayesian methods. The use of such methods allows us to avoid the selection of one single model. Rather we consider 165 different models and present results which average across all these models using posterior model probabilities.

Empirical results indicate that there are both statistically and economically significant nonlinearities in the unemployment rate. For instance, impulse responses indicate that positive shocks (which would tend to increase unemployment) have a much larger effect (in absolute value) than negative shocks.

Appendix A

Geweke and Terui (1993) carry out a Bayesian analysis of a two-regime TAR model using a noninformative prior using Monte Carlo integration. The techniques used in this paper are extensions of their techniques to incorporate informative priors and calculate marginal likelihoods. Given that our formulae are closely related to theirs, we do not provide exact details here. Rather we describe the steps necessary to extend the Geweke and Terui results. Throughout we will refer to the JTAR model, where $J = 2, 3$ indicates the appropriate number of regimes.

We first specify the hyperparameters of the Normal-inverted Gamma priors for β_j and σ_j^2. For all models we set the prior mean of β_j ($j = 1, ..., J$ for JTAR, $j = 1$ for AR) equal to zero. The prior degrees of freedom for σ_j^2 is set to 3, which is the smallest (and hence least informative) value consistent with the existence of the first two prior moments of β_j. The prior mean of σ_j^2 is set to 45. This value was set after looking at the variance of the transformed data. The prior covariance matrix for β_j is assumed to be diagonal. The prior variance is set to 16 for the intercept term(s) and 1 for the first autoregressive coefficient (i.e. var(ϕ_{j1})=1). For autoregressive coefficients for lags greater than 1, we use a prior which gradually tightens as lag length increases (i.e. var(ϕ_{ji})=.81var($\phi_{j,i-1}$) for $i = 2, .., p$). This prior tightening is common in the Bayesian VAR literature (see, e.g. Hamilton, 1994, chapter 12.2). Given the bounded nature of the underlying data, we believe this prior is sensible and, if anything, errs on the side of being too flat and noninformative. All these values are the same for all models.

The posterior, $p(\psi|\mathbf{data})$, can easily be calculated once one notes that, conditional on γ, the JTAR breaks into J independent linear regression models, and results for the linear regression model are well-known (Judge, Griffiths, Hill, Lutkepohl and Lee, 1985). Given the Normal-inverted Gamma form for the prior, it follows immediately that $p(\beta_j, \sigma_j^2|\gamma, \mathbf{data})$ has a Normal-inverted Gamma form. Further, there are only a finite number of ways of breaking the data into three subsets, and the delay parameter d also takes on a finite number of values. Hence, γ can be treated as having a discrete distribution with each point of the probability mass function involving the integration of the product of normal-inverted gamma densities. This latter integration has a known form. Marginal likelihoods can then be found by summing the posterior mass function of γ weighted by the width of the thresholds interval where the split of the data remains constant.

The posterior for the homoskedastic JTAR is even simpler since, conditional on γ, it can be written in the form of a linear regression model, using appropriate dummy variables. $p(\gamma|\mathbf{data})$ can be calculated as described above so that unconditional results can easily be obtained. For the linear model, we can draw directly on standard results (e.g.. Judge *et al* 1985).

In order to obtain features of the posterior distribution of ψ we use Monte Carlo simulation techniques. In generating draws of $\beta_j, j = 1, .., 3$ we impose the usual stationarity restrictions on the polynomials for the outer regimes $\phi_1(L)$ and $\phi_J(L)$ (see Tong (1990) for a discussion of stationarity conditions in nonlinear models). We do not adjust the posterior distribution of $p(\gamma|\mathbf{data})$ for the stationarity condition so there is an element of approximation in the analysis. The stationarity condition is enforced by rejecting draws of β that do not satisfy it. Since these were small in number we do not think the approximation error is large.

The GI are calculated by generating 1000 paths of length 24 from each initial condition and for fixed parameters. The resulting paths are transformed back to the original units of the unemployment data and averaged to approximate the conditional expectations (see Koop, Pesaran and Potter (1996) for more details). This exercise is repeated for 10,000 draws from the posterior and an overall average taken.

Appendix B: Prior Sensitivity Analysis

In this appendix we examine the sensitivity of the posterior model probabilities to changes in the prior distribution. As is well known posterior model probabilities are more sensitive to the prior than features of the posterior distribution of the particular models. In addition i the similarity of the GIs from various models implies the effect of changes in the prior on the model weighted GI will be small. We know from Koop and Potter (1998) that as the variance of the prior on the regression coefficients becomes large, the more parsimonious linear model will receive more support (assuming homoskedasticity). This property is illustrated in Table 4 where we double the prior standard deviation for the regression coefficients. This sizeable change in the prior has relatively little effect on our results. As we would expect, there is some tendency for more probability to be allocated to parsimonious models (i.e. either linear models or models with lower lag lengths). But, overall the main thrust of our empirical results are unchanged.

We also examine separately the effect of halving the prior mean of the variance of the innovations keeping the degrees of freedom fixed at 3 (see Table 5). This relatively sizeable change in prior also has little effect on our empirical results. For the sake of brevity, we do not present results for other priors. But other priors that are similar to the ones we consider yield similar results. It is only if we use completely flat priors, or priors which are very tight in unreasonable areas of the parameter space that our basic empirical conclusions are altered.

Appendix C: Posteriors and GI from Two Models

In this appendix we present properties of the posteriors and GIs for the preferred model (2TAR3HET with $p = 5$) and the linear model which receives most support ($p = 5$). Tables 6, 7 and Figure C1 present results for the 2TARHET3 model while Tables 8, 9 and Figure C2 present results for the linear model. Note that we are working with transformed data so that it is difficult to interpret the estimates of r_1. Beyond what has already been said in the body of the paper, the GIs for the individual models contain no surprises.

It is worthwhile to briefly compare our Bayesian results with classical results. Bayesian posterior means are roughly similar to MLEs, but asymptotic standard errors are quite different. This is undoubtedly due to the fact that the Bayesian measure takes into account uncertainty in the estimation of d and r_1, whereas classical standard errors condition on the MLEs for these parameters. Posterior means of d and r_1 are hard to interpret since the former variable takes on only integer values while the interpretation of the latter depends on the value of d. Hence, Figures C3a and b plot the entire posterior p.d.f.'s for these parameters. The posterior for r_1 is plotted conditional on the most likely value for d (i.e. $d = 5$). It can be seen that the data strongly support a value for r_1 which is near, but not at, the lowest acceptable threshold value. The only noteable difference between Bayesian posteriors and MLEs is for d. We have investigated this difference closely and it does not seem to be due to the Bayesian prior. Rather, it appears to be due to the way the data enters the relevant criteria. Both likelihood function and the Bayesian marginal likelihood depend on the sum of squared errors in the two regimes (i.e. SSE_1 and SSE_2). The likelihood depends on $\text{SSE}_1+\text{SSE}_2$, but the marginal likelihood depends on SSE_1 and SSE_2 in a more complicated way (roughly, $\text{SSE}_1^a\text{SSE}_2^b$ where a and b depend on the number of observations in each regime). In practice, it seems these linear and multiplicative functions can be somewhat different.

Figure C3b is similar to Figure 3 in Hansen (1997) and is calculated using the 2TAR3HET specification. This figure can be used to construct a classical confidence interval for r_1. A detailed explanation of the procedure is given in Hansen (1997) and the reader is referred there for precise details. Briefly, consider the likelihood ratio statistic for testing H_0: $r_1 = r^*$ against an unrestricted alternative. This depends on the residual variance under H_0. Hansen derives critical values for such a test (see his Table 1). Our figure C3c plots the residual variance for every possible r^*. The horizontal line is the 95% critical value from Hansen's test. Values for r_1 where the residual variance lies below this horizontal line are in the confidence interval. Note that the resulting confidence interval will be disconnected. However, it exhibits a similar pattern to the Bayesian posterior, i.e. it indicates the threshold is near the lower bound.

It is difficult to compare Bayesian model selection procedures with classical testing procedures since they are, by their nature, very different. However, it is worth noting that Hansen (1997) carries out a classical econometric analysis of this data. He first differences the data (although classical tests do not find a unit root once a 2TAR is used), rather than transforming the data logistically, considers only p=12 and does not explicitly consider heteroskedasticity. His classical test procedures lead him to select a 2TAR with regimes defined by long differences of the series. In other words, despite great differences in analysis, he ends up with a preferred model similar to the one which receives most support in our Bayesian analysis.

References

[1] Andrews, D. and Ploberger, W. (1994). "Optimal tests when a nuisance parameter is present only under the alternative," *Econometrica,* 62, 1383-1414.

[2] Beaudry, P. and G. Koop (1993). "Do recessions permanently change output?" *Journal of Monetary Economics,* 31, 149-164.

[3] Blanchard, Olivier and Lawrence Summers (1986). "Hysteresis and the European unemployment problem," in *NBER Macroeconomics Annual 1986,* Stanley Fischer (editor), The MIT Press: Cambridge, Mass.

[4] Brock, W.A., and Sayers, C. (1988) "Is the business cycle characterized by deterministic chaos," *Journal of Monetary Economics,* 22, 71-90.

[5] Burgess, S. (1992),"Nonlinear dynamics in a structural model of unemployment," *Journal of Applied Econometrics* 7, S101-S118.

[6] Chan, K.S. (1993). "Consistency and limiting distribution of the least squares estimator of a threshold autoregression model," *Annals of Statistics,* 31, 520-533.

[7] Chan, K. and H. Tong (1985)."On estimating thresholds in autoregressive models," *Journal of Time Series Analysis,* 7, 178-90.

[8] Cox, D. and D. Hinkley (1974). *Theoretical Statistics,* Chapman and Hall, London.

[9] Davis, S., and Haltiwanger, J. (1990). "Gross job creation and destruction: microeconomic evidence and macroeconomic implications," in Blanchard, O.J. and Fischer, S., eds *NBER Macroeconomics Annual*, Cambridge, MIT Press, 123-186.

[10] Diamond, P., (1982), " Aggregate demand management in a search equilibria," *Journal of Political Economy,* 881-894.

[11] Geweke, J. and N. Terui (1993). "Bayesian threshold autoregressive models for nonlinear time series," *Journal of Times Series Analysis,* 14, 441-454.

[12] Ham, M., and Sayers, C., (1992).“ Testing for nonlinearities in U.S. unemployment by sector,” manuscript, University of Virginia and University of Houston.

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

[14] Hamilton, J.D. (1989). “A new approach to the economic analysis of nonstationary time series and the business cycle,” *Econometrica,* 57, 357-384.

[15] Hansen, B. (1997). “Inference in TAR models,” *Studies in Nonlinear Dynamics and Econometrics,* forthcoming.

[16] Hansen, B. (1996). “Inference when a nuisance parameter is not identified under the null hypothesis,” *Econometrica*, 64, 413-440.

[17] Hansen, B. (1992). “The likelihood ratio test under non-standard conditions,” *Journal of Applied Econometrics*, 7S, 61-82.

[18] Johnson, D., and McCelland, R., (1995), “ A general test for independence,” manuscript, Bureau of Labor Statistics.

[19] Judge, G., W. Griffiths, R.C. Hill, H. Lutkepohl, and T.-C. Lee (1985). *The Theory and Practice of Econometrics,* Second edition, John Wiley and Sons, Inc., New York.

[20] Koop, G. (1996). “Parameter uncertainty and impulse response analysis,” *Journal of Econometrics,* 72, 135-149.

[21] Koop, G., M.H. Pesaran and S. Potter (1996). “Impulse response analysis in nonlinear multivariate models,” *Journal of Econometrics,* 74, 77-118.

[22] Koop, G., and Potter, S. (1998). “Bayes factors and nonlinearity: evidence from economic time series,” *Journal of Econometrics*, forthcoming.

[23] Leamer, E.E. (1978).*Specification Searches* John Wiley, New York.

[24] Mortensen, D., and Pissarides, C., (1993) in Van Ours et al edited, *Labor Demand and Equilibrium Wage Formation.*

[25] Neftci, S. (1984). "Are economic time series asymmetric over the business cycle?" *Journal of Political Economy,* 92, 307-328.

[26] Van Ours, J.C., Pfann, G., and Ridder, G., editors (1993). *Labor Demand and Equilibrium Wage Formation,* North Holland New York.

[27] Pesaran, M.H. and S. Potter (1997). "A floor and ceiling Model of U.S. Output," *Journal of Economic Dynamics and Control*, 21, 661-698.

[28] Phillips, P.C.B. (1991). "To criticize the critics: An objective Bayesian analysis of stochastic trends," *Journal of Applied Econometrics,* 6, 333-364.

[29] Potter, S.M., (1995a). "A nonlinear approach to US GNP," *Journal of Applied Econometrics,* 10, 109-126.

[30] Potter, S.M., (1997). "Nonlinear impulse response functions, " *mimeo* Department of Economics, UCLA.

[31] Rogerson, R., (1995)"Review of Ours et al" *Journal of Economic Literature,* XXXIII, 255-256.

[32] Rothman, P. (1991). "Further evidence on the asymmetric behavior of unemployment rates over the business cycle," *Journal of Macroeconomics,* 13, 291-298.

[33] Rothman, P., (1992). " Forecasting asymmetric unemployment rates," Department of Economics, University of Delaware, WP # 92-03.

[34] Sichel, D., (1989).*Three essays in business cycle asymmetry,* unpublished Ph.D thesis, Department of Economics, Princeton University.

[35] Sims, C. (1980). "Macroeconomics and reality," *Econometrica,* 48, 1-48.

[36] Sims, C.A. and Zha, T. (1994). "Error bands for impulse responses," manuscript.

[37] Storer, P. (1994). "Unemployment dynamics and labour market tightness: An empirical evaluation of matching function models," *Journal of Applied Econometrics,* 9, 389-420.

[38] Tong, H. (1990). *Non-linear Time Series Models,* Clarendon Press, Oxford.

[39] Wallis, K. (1987). "Time series analysis of bounded economic variables," *Journal of Time Series Analysis,* 8, 115-123.

Table 1: Average Posterior Model Probabilities

Model	DGP_1	DGP_2	DGP_3	DGP_4	DGP_5	DGP_6
Linear(p=1)	0.7794*	0.0000	0.2985	0.0000	0.0001	0.0457
Linear(p=2)	0.1065	0.9868*	0.0568	0.0000	0.0000	0.0085
2TAR1HOM(p=1)	0.0136	0.0002	0.0549	0.0000	0.0000	0.0107
2TAR1HOM(p=2)	0.0006	0.0023	0.0008	0.0000	0.0000	0.0107
2TAR1HET(p=1)	0.0113	0.0000	0.0233	0.0000	0.0001	0.0048
2TAR1HET(p=2)	0.0049	0.0015	0.0010	0.0000	0.0000	0.0000
2TAR2HOM(p=1)	0.0513	0.0000	0.1499	0.0000	0.0000	0.0331
2TAR2HOM(p=2)	0.0022	0.0059	0.0009	0.0000	0.0000	0.0008
2TAR2HET(p=1)	0.0188	0.0000	0.0382	0.2585	0.3465	0.0130
2TAR2HET(p=2)	0.0006	0.0015	0.0003	0.0000	0.0000	0.0002
2TAR3HOM(p=1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2TAR3HOM(p=2)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2TAR3HET(p=1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2TAR3HET(p=2)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3TAR1HOM(p=1)	0.0003	0.0001	0.0108	0.0000	0.0000	0.7809*
3TAR1HOM(p=2)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3TAR1HET(p=1)	0.0006	0.0000	0.0014	0.0000	0.0000	0.0000
3TAR1HET(p=2)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3TAR2HOM(p=1)	0.0074	0.0000	0.3035*	0.0000	0.0000	0.0197
3TAR2HOM(p=2)	0.0000	0.0002	0.0002	0.0000	0.0000	0.0002
3TAR2HET(p=1)	0.0024	0.0000	0.0593	0.7415*	0.6533*	0.0025
3TAR2HET(p=2)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3TAR3HOM(p=1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3TAR3HOM(p=2)	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
3TAR3HET(p=1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
3TAR3HET(p=2)	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000

Table 2: Posterior Model Probabilities

p	Linear	2TAR1HOM	2TAR1HET	2TAR2HOM	2TAR2HET	2TAR3HOM	2TAR3HET
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.017	0.020
4	0.001	0.000	0.000	0.000	0.000	0.049	0.132
5	0.035	0.000	0.000	0.000	0.000	0.067	0.618
6	0.016	0.000	0.000	0.000	0.000	0.006	0.031
7	0.002	0.001	0.000	0.000	0.001	0.003	0.001
≥ 8	0.001	0.000	0.000	0.000	0.000	0.001	0.000

Table 3: Generalized Impulse Responses: Average over all Models

History	**Fast Decrease**		**Fast Increase**		**No Change**	
Horizon \ Shock	+2	-2	+2	-2	+2	-2
0	0.3636	-0.3450	0.3249	-0.3094	0.3551	-0.3369
1	0.3912	-0.3701	0.2875	-0.2304	0.3731	-0.3437
2	0.4617	-0.4329	0.3038	-0.2341	0.4280	-0.3691
3	0.5682	-0.5208	0.3480	-0.2343	0.5102	-0.4009
4	0.6356	-0.5776	0.3860	-0.2745	0.5550	-0.4294
5	0.6669	-0.6006	0.3840	-0.2582	0.5654	-0.4236
6	0.7017	-0.6246	0.3953	-0.2621	0.5798	-0.4225
7	0.7220	-0.6366	0.4027	-0.2614	0.5835	-0.4195
8	0.7271	-0.6353	0.4075	-0.2682	0.5764	-0.4129
9	0.7232	-0.6265	0.4054	-0.2632	0.5636	-0.4011
10	0.7133	-0.6132	0.4046	-0.2635	0.5485	-0.3900
11	0.6962	-0.5943	0.4016	-0.2611	0.5297	-0.3776
12	0.6740	-0.5719	0.3973	-0.2602	0.5084	-0.3645
13	0.6489	-0.5473	0.3914	-0.2563	0.4864	-0.3508
14	0.6211	-0.5214	0.3852	-0.2538	0.4641	-0.3374
15	0.5917	-0.4945	0.3784	-0.2499	0.4413	-0.3242
16	0.5619	-0.4674	0.3712	-0.2463	0.4191	-0.3114
17	0.5318	-0.4408	0.3636	-0.2419	0.3976	-0.2991
18	0.5022	-0.4150	0.3558	-0.2378	0.3772	-0.2873
19	0.4733	-0.3902	0.3481	-0.2333	0.3577	-0.2760
20	0.4455	-0.3666	0.3403	-0.2288	0.3395	-0.2654
21	0.4190	-0.3444	0.3324	-0.2241	0.3224	-0.2556
22	0.3941	-0.3236	0.3246	-0.2195	0.3066	-0.2463
23	0.3708	-0.3041	0.3170	-0.2148	0.2919	-0.2376
24	0.3487	-0.2861	0.3095	-0.2102	0.2784	-0.2294

Table 4: Prior Standard Deviation of β Doubled

p	Linear	2TAR1HOM	2TAR1HET	2TAR2HOM	2TAR2HET	2TAR3HOM	2TAR3HET
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.001
3	0.000	0.000	0.000	0.000	0.000	0.167	0.119
4	0.006	0.000	0.000	0.000	0.000	0.089	0.172
5	0.144	0.000	0.000	0.000	0.000	0.037	0.221
6	0.032	0.000	0.000	0.000	0.000	0.000	0.000
7	0.002	0.000	0.000	0.000	0.000	0.000	0.000
≥ 8	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 5: Prior Mean of σ_j^2 Halved

p	Linear	2TAR1HOM	2TAR1HET	2TAR2HOM	2TAR2HET	2TAR3HOM	2TAR3HET
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.087	0.046
4	0.003	0.000	0.000	0.000	0.000	0.094	0.154
5	0.109	0.000	0.000	0.000	0.001	0.074	0.378
6	0.035	0.000	0.000	0.000	0.000	0.004	0.010
7	0.004	0.000	0.000	0.000	0.000	0.000	0.001
≥ 8	0.001	0.000	0.000	0.000	0.000	0.000	0.000

Table 6: Bayesian and ML Properties of Parameters: 2TAR3HET with p=5

Parameter	Posterior Mean	Posterior SD	MLE	Asymptotic SE
α_1	-0.8292	0.8468	5.1081	2.1507
ϕ_{11}	0.7678	0.1333	1.2443	0.2855
ϕ_{12}	0.1816	0.1594	0.5098	0.1783
ϕ_{13}	-0.0140	0.1584	-0.6995	0.2992
ϕ_{14}	0.2004	0.1369	0.0213	0.1351
ϕ_{15}	-0.1589	0.1278	-0.1191	0.1389
σ_1^2	12.4926	—	12.5937	—
α_2	-0.2252	0.2028	-0.4457	0.1998
ϕ_{21}	1.0615	0.0516	1.1607	0.0523
ϕ_{22}	0.1090	0.0721	0.0468	0.0692
ϕ_{23}	0.0813	0.0709	-0.0224	0.0744
ϕ_{24}	-0.1115	0.0684	-0.0579	0.0688
ϕ_{25}	-0.1506	0.0481	-0.1328	0.0464
σ_2^2	14.9226	—	13.5903	—
d	4.9580	—	2.0000	—
r_1	-2.1761	—	-2.9562	—

Table 7: Generalized Impulse Responses: 2TAR3HET with p=5

History	**Fast Decrease**		**Fast Increase**		**No Change**	
Horizon \ Shock	+2	-2	+2	-2	+2	-2
0	0.3623	-0.3437	0.3229	-0.3076	0.3551	-0.3369
1	0.3868	-0.3660	0.2798	-0.2274	0.3717	-0.3517
2	0.4543	-0.4264	0.2931	-0.2266	0.4256	-0.3725
3	0.5594	-0.5178	0.3333	-0.2256	0.5079	-0.4063
4	0.6345	-0.5812	0.3802	-0.2766	0.5599	-0.4401
5	0.6665	-0.6053	0.3761	-0.2558	0.5695	-0.4313
6	0.7047	-0.6325	0.3889	-0.2615	0.5852	-0.4300
7	0.7303	-0.6481	0.3995	-0.2619	0.5913	-0.4272
8	0.7382	-0.6483	0.4069	-0.2717	0.5844	-0.4204
9	0.7369	-0.6405	0.4064	-0.2670	0.5715	-0.4079
10	0.7292	-0.6279	0.4079	-0.2689	0.5565	-0.3964
11	0.7133	-0.6089	0.4068	-0.2676	0.5370	-0.3836
12	0.6914	-0.5858	0.4043	-0.2682	0.5148	-0.3700
13	0.6661	-0.5601	0.3998	-0.2651	0.4918	-0.3557
14	0.6375	-0.5329	0.3949	-0.2636	0.4684	-0.3418
15	0.6069	-0.5043	0.3893	-0.2604	0.4443	-0.3281
16	0.5755	-0.4753	0.3831	-0.2576	0.4208	-0.3148
17	0.5436	-0.4467	0.3764	-0.2537	0.3980	-0.3020
18	0.5120	-0.4189	0.3694	-0.2502	0.3764	-0.2898
19	0.4812	-0.3921	0.3623	-0.2461	0.3558	-0.2782
20	0.4513	-0.3667	0.3550	-0.2419	0.3366	-0.2672
21	0.4229	-0.3427	0.3476	-0.2375	0.3186	-0.2572
22	0.3961	-0.3202	0.3402	-0.2331	0.3021	-0.2476
23	0.3709	-0.2992	0.3328	-0.2285	0.2868	-0.2387
24	0.3472	-0.2798	0.3255	-0.2240	0.2727	-0.2304

Table 8: Bayesian and ML Properties of Parameters: Linear model with p=5

Parameter	Posterior Mean	Posterior SD	MLE	SE
α_1	0.0144	0.1829	0.0146	0.1852
ϕ_1	1.0063	0.0471	1.0113	0.0475
ϕ_2	0.1535	0.0660	0.1489	0.0668
ϕ_3	0.0230	0.0639	0.0222	0.0647
ϕ_4	-0.0395	0.0630	-0.0385	0.0640
ϕ_5	-0.1547	0.0450	-0.1552	0.0456
σ^2	14.9323	0.0000	15.0241	0.0000

Table 9: Generalized Impulse Responses: Linear model with p=5

History	**Fast Decrease**		**Fast Increase**		**No Change**	
Horizon \ Shock	+2	-2	+2	-2	+2	-2
0	0.3956	-0.3755	0.3215	-0.3048	0.3568	-0.3385
1	0.4334	-0.4118	0.2886	-0.2736	0.3545	-0.3363
2	0.5415	-0.5109	0.3059	-0.2876	0.4088	-0.3848
3	0.6636	-0.6210	0.3310	-0.3083	0.4718	-0.4402
4	0.7745	-0.7195	0.3549	-0.3277	0.5286	-0.4892
5	0.8155	-0.7568	0.3487	-0.3215	0.5385	-0.4976
6	0.8636	-0.7995	0.3519	-0.3235	0.5572	-0.5134
7	0.8924	-0.8249	0.3526	-0.3237	0.5673	-0.5217
8	0.9023	-0.8337	0.3503	-0.3213	0.5686	-0.5226
9	0.8969	-0.8289	0.3450	-0.3167	0.5626	-0.5171
10	0.8858	-0.8189	0.3403	-0.3126	0.5549	-0.5102
11	0.8660	-0.8011	0.3342	-0.3073	0.5433	-0.4999
12	0.8407	-0.7782	0.3271	-0.3012	0.5290	-0.4873
13	0.8120	-0.7524	0.3193	-0.2945	0.5133	-0.4732
14	0.7813	-0.7246	0.3111	-0.2874	0.4966	-0.4585
15	0.7497	-0.6959	0.3024	-0.2797	0.4791	-0.4428
16	0.7178	-0.6669	0.2934	-0.2718	0.4613	-0.4269
17	0.6864	-0.6383	0.2842	-0.2637	0.4436	-0.4110
18	0.6560	-0.6104	0.2749	-0.2553	0.4261	-0.3952
19	0.6268	-0.5835	0.2656	-0.2470	0.4092	-0.3798
20	0.5987	-0.5577	0.2565	-0.2388	0.3927	-0.3649
21	0.5724	-0.5335	0.2476	-0.2307	0.3769	-0.3505
22	0.5473	-0.5105	0.2388	-0.2227	0.3620	-0.3369
23	0.5240	-0.4889	0.2304	-0.2150	0.3477	-0.3239
24	0.5020	-0.4686	0.2224	-0.2076	0.3343	-0.3115