Time series models of GDP: a reappraisal.

Marchese, Malvina

London School of Economics

2010

Time series models of real GDP.

Over the past 25 years time-series models have been used increasingly in the empirical analysis of real GDP fluctuations.

Initially the emphasis of this literature has been on linear time series models, however more recently a number of studies have found important evidences of asymmetries in real GDP business cycle. Empirical analysis of such asymmetries is possible only in the context of non-linear models and consequently recent studies are based on the assumption that first difference of the log of GDP follow a non-linear stationary process rather than a linear one.

A variety of non-linear specifications have been proposed to model the asymmetric behaviour of real GDP: different parametric and non-parametric non-linear models have been employed to render insight into the differing dynamics over the business cycle. Within the non-parametric line of research regime-switching models have been particularly popular: Markov-switching autoregressive models (MS), self exciting threshold autoregressive models (SETAR) and smooth transition autoregressive models (STAR) have been applied most frequently.

All the above models are developed to capture a specific source of asymmetry in the dynamic behaviour of real GDP over the cycle.

Initially the literature has focused on defining concept of such as " deepness", "steepness" and "sharpness" of a cycle and in modelling phase-dependent properties of the series and it is now well understood that recessions and expansions obey to different dynamics.

Although the two-regime characterisations proposed by this literature have played an important role in non-linear modelling of real GDP, current research proposes growing evidence that at least three regimes are required to represent adequately business cycle movements, with expansions typically consisting of a period of rapid recovery followed by one of slower growth.

This approach claims that two-regimes characterisations fail to emphasize a distinctive feature of real GDP business cycle:output growth tends to be relatively strong following recessions. The need to model this so called " post-recession bounce-back effect " in the level of aggregate output is at the base of recent research in empirical business cycle literature.

Time -series models or real GDP attempt to explore which is the "best" statistical model of the series.

The choice of a specific model to represent the behaviour of real GPD over the cycle has striking economic implications since it implies a very different representation of the permanent effect of recessions on the level of aggregate output and leads the way to different economic policy measures to cope with recessions.

There is a very important underlying debate on whether the effect of shocks on output is asymmetric, permanent or both. Linear time-series models implicitly impose symmetry on the measure of shock persistence whereas different non-linear time series models originate different measures of shock persistence and do not assume symmetry in the persistence of positive and negative shocks.

It is therefore of considerable interest to inquire whether these models that are designed to capture persistence or asymmetric character of the real GDP series perform well, and to develop a model comparison diagnostic to compare the different statistical representations of the series that they imply.

The conventional practise of model evaluation essentially focuses on the overall fit of the model or its ability to predict future outcomes and does not pay much attention to the model-induced stochastic processes of business-cycle features per se.

This practise is partly motivated by the serious difficulties that arise in making formal inference in highly non-linear state-switching statistical representations,such as Markov-switching models.

The aim of this project is to provide a model-diagnostic device and to compare different popular linear and non-linear time series models of real GDP on the basis of the consideration that business-cycle features themselves should motivate a good metric for judging a macroeconomic time-series model. This approach is based on the previous work of Harding and Pagan (2002) and King and Plosser(1994).Like conventional residual-based diagnostic tests, this device does not rely on asymptotic distributional theory but ,in contrast to them, it offers a direct interpretation for the economist 's ultimate focus of inquiry- the business -cycle features of the data.

All time-series real GDP models propose specific data generating processes (dgp) that generate specific patterns of expansions and contractions.

The model-diagnostic will focus on the ability of these models to reproduce the duration and the amplitude of business-cycle phases.

Accordingly I will focus on the ability of a model's dgp to produce expansions and contractions whose average duration is similar to those observed in the actual data and on its ability to produce expansions and contractions whose average amplitude is similar to those found in the data. A model that performs poorly in these checks would be missing relevant stylised facts of actual business cycle and different linear and non-linear model might be compared on the basis of this evaluation criteria.

In section one I illustrate the different linear and non-linear approaches to modelling the real GDP series, and the specific concerns of each approach. I then focus on six popular models that are widely employed in the literature : two linear ARMA models, and four non-linear models:a two-regime SETAR model , a multistage SETAR model, a two-regime Markov-switching model and a three-regime Markov switching model .Each of these models is explained in some details.

In the section two I explore the economic implications of each of the models in term of measures of shock persistence and economic stability .In these different implications lies the economic motivation of my project: different models impose very different patterns of GDP recovery from a shock and therefore suggests different policy responses to cope with them.It thus appear of great economic relevance to compare them .Limitations of inferential procedures for Markov-switching models are then explained.

In section three I introduce a model-comparison device, based on a set of definitions of contractions and expansions alternative to the NBER 's dating practise. Stylized facts of business cycle are derived using those definitions.

In section four I estimate the models presented in section one and present the results of a Monte Carlo simulation calibrated to U.S real GDP for those models.

In section five I explain the project's limitations and draw conclusions.

1-Time series models of real GDP.

Linear time-series models.

A number of studies have sought to characterize the nature of the long term trend in real GDP and its relation to the business cycle within a linear time-series framework.

Researchers such as Beveridge and Nelson (1981), Nelson and Plosser (1982) and Campbell and Mankiw (1987) explored this question using ARIMA models or ARMA processes around a deterministic trend.

Other, such as Harvey (1985), Watson and Clark (1986) based their analysis on linear unobserved components models. A third approach employs the co-integrated specification of Engle and Granger (1987).

All these approaches are based on the assumption that first differences of the log of real GDP follow a linear stationary process, that is in all the above studies, optimal forecasts of variables are assumed to be linear functions of their lagged values.

Linear ARMA models.

Campbell and Mankiw ARIMA (2,1,2) model with drift is perhaps the most popular linear ARMA model of real GDP in the literature.

Departing from the most standard approach to GDP modelling in the prevailing literature of the time,Campbell and Mankiw stress how detrending data would bias a priori any measure of shock persistence on output fluctuations.They argue that detrendig would force the resulting series to be trend reverting , so that today's innovation would not have ultimate effect on output at an infinite horizon and propose instead to difference the series of log real GDP. The differenced series ,the growth rate of real GDP,appears stationary ,allowing them to invoke asymptotic distribution theory.

The change in log GDP is modelled as a stationary ARMA process,that is

$\phi(L)\Delta Y_t = \vartheta(L)\varepsilon_t$

where

$$\varphi(L) = 1 - \phi_1 L - \varphi_2 L^2 - .... - \phi_p L^p$$

and

$$\vartheta(L) = 1 + \theta_1 L + \vartheta_2 L^2 + ... + \vartheta_q L^q$$

To select the optimal parametrisation of the ARMA model Campbell and Mankiw rely on the Akaike criterion,according to which the parametrisation with the maximum likelihood after imposing a penalty for the number of parameters should be chosen .In particular the criterion tells to maximize

-2lnL-2k where L is the likelihood and k+p=q is the number of parameters.

They find that the Akaike criterion selects the ARIMA (2,1,2) model over the white noise ,the AR(1) and AR(2) and MA(1) models.

However,although they prefer the ARIMA (2,1,2) model,they too advocate the ARIMA (1,1,0) model as a more parsimonious representation and I will therefore report the distributions of the business cycle features generated by both these models in the Monte Carlo simulation.

Finally to estimate the model they employ an exact maximum likelihood estimation method which uses a Kalman filter to build up the log likelihood function of the model as a sum of conditional log likelihoods.

I will explore the consequences of assuming that the growth rate of real GDP follows a liner stationary process in the next section where ,having introduced a variety of non-linear models,I will be able to focus on the comparative differences in term of prediction of shock permanence on the level of real GDP.

Non-linear time-series models.

The choice to model the growth rate of real GDP as a linear stationary process implicitly rules out the existence of sources of asymmetry in business cycle dynamics by imposing symmetry.

However the notion of business-cycle asymmetry has been around for quite some time in economic theory. In particular the idea of inherently different dynamics in expansions and recessions has a long history in business cycle theory, dating back at least to Keynes who observed that

> " The substitution of a downward for an upward tendency often takes place suddenly and violently,whereas there is,as a rule,no sharp turning point when an upward is substituted for a downward tendency"

Recent advances in econometrics have allowed this idea to be formally modelled and tested and there has now accumulated abundant evidence that departures from linearity are an important feature of many key macro series.

Starting with the documentation works of Neftci (1984) and Sichel (1987),the bispectral analysis of Hinch and Patterson (1985) ,the ARCH-M model of Engle and Robins (1987) and the chaos model of Brock and Sayers (1988) a variety of parametrizations for non-linear dynamics have been proposed and formally tested.

The statistical procedures that have been employed can be divided into two main categories:non-parametric and parametric techniques.

Non parametric techniques have been used by Neftci(1984), Sichel (1989) and Thorley (1993) who, among many others, test from asymmetry between expansions and contractions by using Markov chain methods to examine whether the transition probabilities from one regime to the other differ. However this studies have failed to provide very compelling evidence for asymmetry in the behaviour of real GDP, for example Sichel cannot reject symmetry when examining US real GDP and Thorley rejects linearity only marginally.

Alternatively parametric non- linear time- series models have been employed to examine a variety of sources of asymmetry in the business cycle and within this line of research regime-switching models have been particularly popular. Typically, these models consist of a system of linear models of which,at each point in time,only one or a linear combination of the models is active to describe the behaviour of a time series,where the activity depends on the regime at that particular moment.

Within the class of regime switching models,two main categories can be distinguished, depending on whether the regimes are determined exogenously ,by an unobservable state variable,or endogenously, by a direct observable variable.

The most prominent member of the first class of models is the Markov-switching autoregressive model (MS),which has been applied for the first time to model business cycle asymmetry by Hamilton in his seminal paper of 1989.

From the second class of models the more frequently applied have been the (self exciting) autoregressive models (SETAR) ,by Potter (1996) , Beaudry and Koop (1993) and the smooth-transition autoregressive model (STAR) ,by Filardo (1994) , Terasvirta(1995) and Skalin (1996).

Much of the initial research on non-linear parametric models has focused on portraying the short, violent nature or recession relative to expansions and has lead to a two-phase characterisations of the business cycle ,i.e. expansions and contractions.

However more recently this two-phase view has been questioned in favour of a three stage characterisation,that suggests the possibility of three business-cycle phases:contractions,high growth recovery that immediately follow troughs of the cycle, and subsequent moderate growth phases.

The need for a three regimes representations of business cycle movements has arisen from a growing evidence on the existence of a phase of high recovery typical of post recession dynamics that can't be adequately modelled with two stage models.

Sichel (1994) observes that real GDP tends to grow faster immediately following a trough than in the rest of the expansion phase.

Wynne and Balke (1992) and Emery and Koenig(1992) present additional evidences for this "bounce-back effect" in the level of aggregate output.

These works highlight the existence of a relatively stronger phase of output growth following recession and of a correlation between each recession and the strength of the subsequent recovery.

The existence of a striking asymmetry in the correlations between succeeding phases of the business cycle had been already pointed out in 1993 by Friedman who noticed that

> "the amplitude of a contraction is strongly correlated with the succeeding expansion although the amplitude of an expansion is uncorrelated with the amplitude of the succeeding contraction"

but no attempt to formally model it had been done till quite recently .

Recent research in non- linear parametric models has focused on extending a variety of regime -switching models to allow for the existence of a bounce back effect.

Morley ,Kim and Piger (2005) augment Hamilton's original model with a " bounce-back" term that is scaled by the length of each recession and can generate faster growth in the quarters immediately following a recession.

Tiao and Tsay (1994) extend Potter model and estimate a multiple regime SETAR model in which one regime is a high growth phase following economic contraction.

Van Dick and Franses (2001) explore how STAR models can be modified to allow for more than two regimes and propose a multiple regime STAR model (MRSTAR) to describe the behaviour of postwar U.S real GDP.

Beaudry and Koop (1999) propose a non-linear ARMA model in which dynamics change when an observed indicator variable exceeds a given threshold.

All these models allow for the existence of three business cycle phases ,in some cases,like in Tiao and Tsay's SETAR model, by explicitly adding furthers regimes to the original model , in others ,like in Kim and Piger's MS model by adding a specific term in a two stage model that can account for the bounce-back effect.

Among all the parametric non-linear models that have been proposed in the literature I will present results from the Monte- Carlo simulation for two business cycle two- phases models : the Markov-switching model of Hamilton (1989) and the SETAR model of Potter(1991).These are the models that have mostly influenced the subsequent research and the most conceptually important.

It will be evident that they both perform well in reproducing basic business cycle features,better than linear ARIMA models to some extent, but Hamilton's moldel fails to take into account the bounce-back effect in the level of aggregate output.

Within the class of the three business cycle phases I will report results for Tiao and Tsay's multiple regime SETAR model (1994), and for Kim and Piger's (2005) augmented MS model.

I will now proceed to illustrate the basic features and estimation procedures of the above four models in some details.

Markov-switching models.

Model instability is sometimes defined as a switch in a regression equation from one sub sample period (a regime) to another .

Economists have long recognised the possibility that parameters may not be constant through time but rather that structural shifts may occur, dividing the period into distinct regimes with different parameters values. In the regression model context, Quandt (1972) studied the case of independent switches in regime.

In many cases,however, researchers may have little information on the dates at which the parameters change and thus need to make inference about the turning points as well as on the significance of the parameters shift.

There are many models that address the question of modelling a switch in the data series with an unknown turning point, some of which assume that the probability of a switch depends upon which regime is in effect. Goldfeld and Quandt 's(1973) model was the first to explicitly allow for such a dependence by introducing Markov switching.

A Markov-switching model is thus a model where the transition probability, that is the probability of the switch, is assumed to follow a Markov process, and so depends on the state of the model .

Markov-switching autoregressive models address the question of modelling state-dependent structural changes in dependent data.

These models however pose serious computational difficulties in estimation and testing in the classical approach due to the potentially very large number of evaluations of the likelihood function required. The filtering proposed by Hamilton ,that I will shortly illustrate below, makes maximum likelihood feasible but still the degree of approximation in any particular case in unknown.

Hamilton's markov-switching model

Hamilton's state-dependent Markov-switching model can be viewed as an extension of Goldefeld and Quandt's (1973) model to the important case of structural changes in the parameters of an autoregressive process .The turning point is treated as a structural event that is inherent in the data-generating process .

Hamilton allows the mean of the growth of real GDP to be evolving according to a two-state Markov-switching process ,thus allowing the dynamics of recession to be qualitatively distinct from those of expansions .

Growth in real GDP is modelled as an AR(4) process:

$$(\Delta y_t - \mu_{st}) = \varphi_1(\Delta y_{t-1} - \mu_{st-1}) + \phi_2(\Delta y_{t-2} - \mu_{st-2}) + ... + \phi_4(\Delta y_{t-4} - \mu_{st-4}) + e_t$$

$$e_t \sim N(0, \sigma^2)$$

$$\mu_{st} = \mu_0(1 - S_t) + \mu_1 S_t$$

$$\Pr(S_t = 1 \mid S_{t-1} = 1) = p$$
$$\Pr(S_t = 0 \mid S_{t-1} = 0) = q$$

where roots of $\phi(L) = (1 - \phi_1 L - .... - \phi_4 L^4) = 0$
lie outside the unit circle and y is the log of real GDP.

Hamilton's estimation approach is to solve for the actual marginal likelihood function for GDP ,maximize this likelihood function with respect to populations parameters ,and then use these parameters and the data to draw the optimal statistical inference about the unobserved regimes .Estimation of the state variables is therefore conditional on maximum likelihood estimate of the parameters.

If S_t were known a prori,the above would be nothing more than a dummy variable model. Calculation of the density of y_t given past information (ψ_{t-1}) and the loglikelihood would be straightforward ,but in writing the density of y_t , S_t and S_{t-1} are need ,which are unobserved .

To solve this problem Hamilton implements a filtering algorithm ,that accepts as input the joint conditional probability

Pr($S_{t-1} = s_t, S_{t-2} = s_{t-2}, ..., S_{t-r} = s_{t-r} \mid y_{t-1}, y_{t-2,....,y_{-r+1}}$)

and has as output

Pr($S_t = s_t, S_{t-1} = s_{t-1}, ..., S_{t-r+1} = s_{t-r+1} \mid y_t, y_{t-1}, ..., y_{-r+1}$)

along which,as a by-product, the conditional likelihood of y_t :

f($y_t \mid y_{t-1}, y_{t-2}, ..., y_{-r+1}$)

Another by-product of the filtering is inference about the state s_t based on currently available information,

Pr($S_t = s_t \mid y_t, y_{t-1}, ..., y_{-r+1}) = \sum_{st=0}^{1} \sum_{st-1=0}^{1} .... \sum_{st-r+1=0}^{1}$ Pr($S_t = s_t, S_{t-1} = s_{t-1}, ..., S_{t-r+1} = s_{t-1+r} \mid y_t, y_{t-1}, ..., y_{-r+1}$)

For details of the filtering please refer to Kim and Nelson(1999).

Hamilton applied the above model to US real GDP for the sample period :1952:2-1984:4 and obtained the following estimates(standard errors are in parentheses)

p	0.9008	(0.0443)
q	0.7606	(0.1206)
φ_1	0.0898	(0.1981)
φ_2	-0.0186	(0.2082)
φ_3	-0.1743	(0.1381)
φ_4	-0.0839	(0.1248)
σ	0.7962	(0.0858)
μ_0	-0.2132	(0.2613)
μ_1	1.1283	(0.1596)

Log likelihood -175.24

A very interesting feature of this model is that the specific inferences about the historical incidence of growth states generated by the filter and the smoother correspond extremely closely to conventional NBER dating of business cycles. Hamilton reports an alternative dating of U.S business cycles peaks and troughs as determined from full sample smoother and shows that there is a strong correspondence between the smoothed probability of being in a contractionary regime and the NBER recession dates.

Hamilton's model is a two regime model that portrays the short ,violent nature of recessions relative to expansions and offers an alternative to linear representation of the data such as the ARIMA discussed above. However it fails to take into account the post bounce-back effect in the level of aggregate output as will be clear from the Monte Carlo results. Kim and Piger's augmented MS models deal with this limitation by introducing a a ' bounce-back ' term .

Kim and Piger's MS bounce-back augmented model.

Kim and Piger (2005) extend Hamilton's MS model to take into account the high growth post recession recovery phase in economic activity,while maintaining endogenously estimated business cycles regimes.

Their model implies an expansion , a recession and a recovery phase and takes into account the strong link between each recession and the strength of the subsequent recovery. The 'bounce-back ' term is directly related to the underlying recessionary regimes and is therefore, endogenously estimated .

The model is given as follow:

$$\phi(L)(\Delta y_t - \mu_0 - \mu_1 S_t - \lambda \sum_{j=1}^{m} S_{t-j}) = \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma^2)$$

where the lag operator $\phi(L)$ is kth order with roots outside the unit circle, Δy_t is the first difference of log real GDP , and S_t is an unobserved markov-switching state variables that takes on discrete value of 0 or 1 according to transition probabilities Pr($S_t = 0 \mid S_{t-1} = 0) = q$

and Pr($S_t = 1 \mid S_{t-1} = 1) = p$

S_t corresponds to a contractionary regime .

The key variable of the model is the summation term $\sum_{j=1}^{m} S_{t-j}$

This term implies a bounce back effect if $\lambda \succ 0$, while Hamilton's model is obtained if λ =0.If $\lambda \succ 0$ $\sum S_{t-j}$ implies that growth will be above average for the first m periods of an expansionary regime.

The model is estimated with data of log of quarterly Us real GDP over the period 1947:1 to 2003:1via maximum likelihood ,using the filter presented in Hamilton .

The lag length k for the autoregressive polynomial and the length m of the bounce-back term are selected using the Schwartz information criterion ,that maximizes -2lnL -2klnT, where L is the likelihood and T is the number of observations.

They consider upper bounds of k=4 and m=9 . For the autoregressive polynomial they find that k=0 suggesting that the non linear dynamics in the model are sufficient to capture most of all the serial correlation of the data .For the bounce-back effect they find that m=6 quarters.

The estimates they obtain are reported below

Parameter	Estimate	standard error
μ_0	0.836	0.064
μ_1	-2.055	0.232
$\mu_0 + \mu_1$	-1.219	0.229
λ	0.319	0.050
q	0.957	0.017
p	0.695	0.101
σ	0.768	0.042
loglikelihood	-288.088	

SETAR models.

Self-exciting threshold autoregressive models (SETAR) are special cases of nonlinear models with a single index restriction.

Let Δy_t represents the observed univariate time series, in our case the first difference of real GDP, and let Z_t be an unobserved time series .Let H_t denote the single index which is assumed to be a continuous map from the history of (Δy_t Z_t) to the line. Let F(.) be a function form the line to the unit interval with at most a finite number of discontinuities .Then a univariate first order Single-index Generalised Multivariate Autoregressive model (SIGMA) would be:

$$\Delta y_t = \alpha_1 + \alpha_2 F(H_t) + \langle \varphi_1 + \phi_2 F(H_t) \rangle \Delta y_{t-1} + \langle \varphi_1 + \varphi_2 F(H_t) \rangle V_t$$

As a special case if F(H)=$\bar{\text{I}}(\Delta y_{t-d} \succ r)$ then the model above becomes a SETAR (1,d,r) model, where $\bar{\text{I}}$(A) is the indicator function equal to one if the event A occurs and zero otherwise , and d is known as the delay parameter and r as the threshold parameter.

The model is called self-exciting autoregressive model (TAR) because it uses its own lagged value Δy_{t-d} as the threshold variable

$$\Delta y_t = \Phi_0^j + \textstyle\sum_{i=1}^{p} \Phi_i^j \Delta y_{t-i} + \varepsilon_t^j$$

$r_{j-1} \preceq \Delta y_{t-d} \preceq r_j$

where j=1,.......,k and d is a positive integer. The thresholds are $-\infty = r_0 \prec r_1 \prec ..... \prec r_k = \infty$;

for each j,($\varepsilon_t{}^j$) is a sequence of martingale differences satisfying

$$\text{E}(\varepsilon_t^j \mid F_{t-1}) = 0$$

with F_{t-1} the σ field generated by $(\varepsilon_{t-1}^j \mid i = 1, 2, ... : j = 1, ..., k)$.

Such a process partitions the one dimensional Euclidean space into k regimes and follows a linear AR model in each regime.

The overall process Δy_t is non-linear when there are at least two regimes with different linear models.

In contrast to Markov-switching models ,in SETAR models the nonlinearity is defined by the direct observable history of the time series. This greatly simplifies the estimation and gives to the SETAR model much greater flexibility in fitting the observed data .

A first step in modelling a TAR model is the specification of the threshold variable and the threshold values ,which play a key role in the non-linear nature of the model. In the case of a SETAR model the specification amounts to the selection of the delay parameter d and the values of r_i. Lim (1980) uses the Akaike information criterion to select d after choosing all the other parameter,Tong (1990) suggests a grid search method for the estimation of d and r_i

Potter and Tiao and Tsay implement yet another procedure that I illustrate below.

Once the delay parameters and the threshold values have been selected , and thus the model has been identified,estimation of a SETAR model can simply be done by least squares. Two techniques are available:

one can split the data into two groups and run a least square regression for each regime separately,thus the estimated residual variance for each regime will be different,or one can run a single regression with indicator functions given by the single index multiplying the lags of the time series,thus the estimated residual variance is restricted to be constant across regimes .Both methods give consistent estimates of the intercept and slope coefficient in each regimes conditional on the correct choice of r_i and d (Tong 1990).

Potter's SETAR model of real GDP.

Potter (1991) estimates a SETAR model of real GDP growth using the series of quarterly data on US real GDP from 1948:3 to 1990:4 ,for a total of 170 observations and seasonally adjusted data.

To specify the delay parameter and the threshold values the following method is implemented:

1) First a linear AR(5) model is estimated .

Since the coefficients of the AR3 ,AR4 and AR5 term of the linear model are not significantly different form zero at a 5% level and have similar values in both regimes ,the linear model is re estimated with this term restricted to be zero and the following AR(2) model is obtained:

$\Delta y_t = 0.0041 + 0.33\Delta y_{t-1} + 0.13\Delta y_{t-2} + \varepsilon_t$

with standard errors 0.001,0.075and 0.076

residual standard deviation : 0.00986

2)Then a threshold non-linearity test based on arranged AR(2) autoregressions with possible threshold d $\in$ (1,2,.........6) is performed. For each value of d, the data are arranged according to the order of Δy_t.

Then predictive residuals from the arranged autoregression are regressed against the predictor variables ,giving rise to an asymptotic F test for indipendency between the residuals and the predictors which would be consistent with linearity.

The results are reported below:

d	1	2	3	4	5	6
F test	0.37	3.16	2.55	2.65	1.70	1.80
pvalue	0.778	0.026	0.058	0.051	0.169	0.150

from the table it can be seen that the linear model hypothesis seems untenable and that d=2 is reasonable for the series as the corresponding p value is the smallest. Therefore the delay parameter is set d=2.

To determine the number of regimes and the threshold values r_i Potter plots the sequential t ratio of a AR(2) estimate against the threshold variable Δy_t

in an arranged autoregression of order 2 with d=2. He suggests that major changes in the slope of the t-ratios indicate regime partition. He concludes that the data can be partitioned into two regimes with a threshold at $\Delta y_{t-2} = 0$.

A SETAR (2) model of real GDP growth is thus specified as :

$\Delta y_t = \phi_0^{(1)} + \phi_1^{(1)} \Delta y_{t-1} + \phi_2^{(1)} \Delta y_{t-2} + \varepsilon_{i,t}$
if $\Delta y_{t-2} \leq 0$

$\Delta y_t = \phi_0^{(2)} + \phi_1^{(2)} \Delta y_{t-1} + \phi_2^{(2)} \Delta y_{t-2} + \varepsilon_{2,t}$
if $\Delta y_{t-2} \succ 0$

The following least square estimates for the model are obtained (standard errors are in parenthesis)

j	$\phi_0^{(j)}$	$\phi_1^{(j)}$	$\phi_2^{(j)}$	σ_j
1	-0.0039(0.0033)	0.44(0.18)	-0.79(0.33)	0.0120
2	0.0038(0.0014)	0.31(0.08)	0.20(0.11)	0.0087

standard error of the regression 0.95597
AIC= -4.89

The model thus portraits two regimes: a contraction (regime 1) and an expansion (regime2) . Perhaps the most striking feature of this model is that, by treating a negative growth in GDP as a contraction and a positive growth as an expansion , it shows that the economy behaves differently after a contraction and an expansion. For example, the AR polynomial of the first regime has a pair of complex roots ,indicating some cyclical behaviour of GDP after a contraction. On the other hand the AR polynomial of the second regime has two real roots,showing that the economy tends to decay experientially to some mean level after an expansion.

Tiao and Tsay's multiple regime SETAR model.

Tiao and Tsay (1994) point the attention to a very interesting founding of the above SETAR model: the very large negative coefficient of the AR2 term in the contractionary regime. In this regime the lag 2 value must be negative implying a positive effect on growth when multiplied by the negative coefficient.

To gain further insight into this features of the data they refine Potter's SETAR model by incorporating the relative size of Δy_{t-1}

with respect to Δy_{t-2} ,the threshold variable of the model. Thus the SETAR(2) model above is extended to a four regime model:

regime1: $\Delta y_{t-1} \preceq \Delta y_{t-2} \leq 0$.

this regime denotes a recession period in which the economy changed form contraction to an even worse one .

regime2: $\Delta y_{t-1} \succ \Delta y_{t-2}, \Delta y_{t-2} \preceq 0$

here the economy was in a contraction but improving

regime3:$\Delta y_{t-1} \leq \Delta y_{t-2}, \Delta y_{t-2} \succ 0$

the regime corresponds to a period in which the economy was declining but reasonable

regime4:$\Delta y_{t-1} \succ \Delta y_{t-2} \succ 0$

this is an expansion period in which the economy was reasonable and became stronger

They use seasonally adjusted quarterly observations of US real GDP from 1947:1 to 1990:4 and fit the following four stages SETAR model of real GDP growth:

regime 1: $\Delta y_t = -0.015 - 1.076\Delta y_{t-1} + \varepsilon_{1,t}$
regime 2: $\Delta y_t = -0.006 + 0.630\Delta y_{t-1} - 0.756\Delta y_{t-2} + \varepsilon_{2,t}$
regime3: $\Delta y_t = 0.006 + 0.438\Delta y_{t-1} + \varepsilon_{3,t}$
regime4: $\Delta y_t = 0.004 + 0.443\Delta y_{t-1} + \varepsilon_{4,t}$

All the coefficients in the model are statistically significant. The residual standard deviations are σ_1=0.0062, σ_2 =0.0132, σ_3=0.0094, σ_4 =0.0082. the numbers of observations in each regime are 6,31,79,58.

In this multistage model it is interesting to notice the negative explosive nature of of the regression function in regime one,indicating that the economy usually recovers quickly from the recession period.

Furthermore regime 2 appears to be consistent with the existence of a bounce-back effect in the level of aggregate output, this seems to be indicated by the large negative coefficient of the AR2 term. The regression function in this regime tends to be positive,suggesting that the economy is more likely to grow continuously out of a recession once a recovery has started.

2-Paper motivation.

Linear ARMA models , Markov-switching models and SETAR models offer alternative statistical representations of the real GPD series.

Policy debates depend on whether the GDP series is best characterized by a linear or a non-linear model .Using a linear model with fixed parameters to describe the series may lead to a wrong quantitative assessment of policy effects if structural changes occur during the period of study.

Furthermore different linear and non-linear representations of the series have different economic implications in term of both symmetry and persistence of shocks on the level of aggregate output. Macroeconomic stabilisation policies depend ,both in term of their quantitative assessment and timing of the intervention,on the estimates of shock persistence that are used.

I will now explore the economic implications of the above models in term of shock persistence making use of impulse response functions.

Estimating shock persistence

In the linear ARMA model literature there seems to be considerable agreement that post-war output fluctuations are highly persistent .For example,at horizons that are typically associated with a downturn(e.g. 8 quarters) ,this literature almost never finds significant evidence of dampening. Therefore it is generally agreed that explaining why recession have such a long impact is a necessary requirements of any macroecomic theory.

However this literature imposes symmetry on the measure of persistence,that is a positive and a negative shock to output are restricted to have identical impulse responses.

To see this consider a linear ARMA model

$$\phi(L)\Delta y_t = \vartheta(L)\varepsilon_t$$

and rearrange this equation to arrive at the moving average representation ,or impulse response function for Δy_t :

$$\Delta y_t = \phi(L)^{-1}\vartheta(L)\varepsilon_t = A(L)\varepsilon_t$$

If it is assumed that the change in log real GDP is stationary then $\sum_{i=0}^{\infty} A_i^2$ is finite,implying that the limit of A_i as i approaches infinity is zero.

The moving average representation for the *level* of y_t can be derived by inverting the difference operator (1-L):

$$y_t = (1-L)^{-1}A(L)\varepsilon_t = B(L)\varepsilon_t$$

where $B_i = \sum_{j=o}^{i} A_j$

y_i needs not to be stationary and thus B_i needs not to approach zero as i approaches infinity .Instead, the limit of B is the infinite sum of A_j coefficients .

The value of B_i for large i is exactly what is to be estimated to measure shock persistence, since it measures the response of y_{t+i} to an innovation at time t.

Clearly this type of time- series representation imposes a symmetric updating rule for forecasting output,that is, both a positive and a negative innovation lead to the same size update for future output.

Within the linear ARMA model literature Camplbell and Mankiw 's ARMA(2,2) model has been widely employed to compute the long run impact of innovations on the level of real GDP.

According to their estimates a one per cent (positive or negative) innovation in real GDP should change one's forecast of real GDP by over one percent over a long horizon,thus implying a strong permanent effect of shocks on the level of aggregate output.

The non-linear model literature questions the imposition of symmetry on shock persistence ,arguing that it can cause severe biases in the characterization of persistence and, for this reason, should be tested and not assumed.

Non-linear models do not impose symmetry on the measure of persistence and they all argue in favour of asymmetry . However the sources of asymmetry in response to shocks and their implications for the measure of persistence vary quite dramatically across models.

Hamilton's MS representation of real GDP provides an alternative perspective with respect to linear models of the measure of the long -run effects of shocks .Hamilton considers the expected difference in the long -run level of output given that the economy is currently in a contractionary regime versus an expansionary regime. The measure is thus evaluated:

$$\lim_{j \to \infty} \{E\left[y_{t+j} \mid S_t = 1, \Omega_{t-1}\right] - E\left[y_{t+j} \mid S_t = 0, \Omega_{t-1}\right]\}$$

where $\Omega_{t-1} = \{S_{t-1} = 0, S_{t-2} = 0, ...; y_{t-1}, y_{t-2}, ...\}$

Hamilton's estimates of this limit yield the results that, if at date t the economy is in a recession (S_t=0) the consequences for the long run future level of real GDP are of about a 3 per cent drop.

On the other hand if the economy is in an expansionary regime the permanent effect of a shock is estimated around 0.66 percent.

Thus the response of output to shocks at different stages of the business cycle is asymmetric and recessions have a large permanent effect on the level of output.

When the same measure of persistence is applied to Kim and Piger's MS augmented model it yields quite different results.

Although the asymmetry of responses is not questioned ,the estimate of the long run effect of recession differ dramatically both form the Hamilton MS model and from the linear ARMA models.

The limit above has for this model the following expression:

$$\Lambda = \mu_1 + m\lambda/(2 - q - p)$$

Using the parameters estimates of Kim and Piger's the estimated value for Λ is -0.412, just under a 0.5% permanent drop in the level of GDP.

Kim and Piger's model thus implies a non linear dynamic of shock response with much smaller long-run effect than are implied both by linear and non linear two-regime MS representations of the data.

It is clear therefore that,not only MS models yield different implications with respect to linear models in term of shock persistence, but also that within the class of MS models two regimes and three regimes characterisations provide quite different measures of shock persistence.

The distinction between two regime and multiple regime representations of the real GDP series does not appear to play such an important role in estimating the long-run effect of shocks with SETAR models.

In fact both Potter's two stage and Tiao and Tsay's multiple stage SETAR models highlight the same two main sources of asymmetry between a contractionary and an expansionary regime:the change in the intercept and in the AR2 coefficient.

In order to illustrate and quantify the extent of the asymmetry in these two models non-linear impulse response functions can be used. A nonlinear impulse response function for a SETAR model is simply:

$$\mathrm{NLIRF}_n(\varepsilon; Y_t, Y_{t-1,.....}) = E\left[Y_{t=n} \mid Y_t = y_t + \varepsilon, Y_{t-1}, ...\right] - E\left[Y_{t+n} \mid Y_t = y_t, Y_{t-1} = y_{t-1}, ...\right]$$

where lower case letters represent realized values and ε is the postulated impulse.

Non linear response functions are obviously functions of the history of the time series and the size and magnitude of the shock. Asymmetric responses in SETAR models occur in two main form:

1)for any specific history the effect of shocks of varying magnitudes and signs is not a simple scaling of a unit shock

2)for the same shock but different histories the response can differ markedly

Looking at impulse response functions of Potter's model for shocks of +1%,-1%,+2%,-2% in various historical periods, interesting conclusions about the effects of such shocks on the level of real GDP can be drawn:

1)if the shocks keep the growth rate positive then the response is very similar to that obtained from a linear model.

2) if the negative unit shock turns the growth rate negative then its effect will be magnified compared to point one by the switch in the intercept term. Magnification also occurs for the positive shocks because the probability of a future contraction decreases. In fact this effect is similar to the abrupt switch between a contraction and expansion in Hamilton's model .However ,for the -2% shock the stabilizing influence of the AR(2) coefficient in the contractionary regime starts to take hold. Thus the implied responses are similar to those of Kim and Piger's MS model The effect after two years of the -2% shock is smaller in absolute value than the-1% shock. As noticed above Hamilton's MS model is unable to capture such an effect since it constrains the probability of movement out of a contraction to be fixed no matter how large is the negative shock

3)If the value of growth perturbed in the starting values is only slightly greater than zero (then for the positive shocks the effects are very similar to those in point 2 above.)

4)if the value of growth perturbed in the starting value is only slightly below zero then there is almost a doubling effects of positive shock compared to point one.The main reason is the switch in the intercept values produced by the perturbation. For negative shocks the stabilizing property is in this case more powerful, with output returning to a trend after 8 quarters.

Thus SETAR models ,as MS models and in opposition to linear ARMA models,imply that the response of output to shocks at different stages of the business cycle is asymmetric.

However,no matter if the SETAR model considered is a two regime or a multi regime , there appear to be a stabilizer effect in these models,that significantly diminishes the effect of negative shock on the level of aggregate output that is implied by two regimes MS models.

This stabilisation mechanism is consistent with evidences on the bounce-back effect ,suggesting that the growth rate in recoveries tends to be higher than the average expansion growth rate and that the magnitude of the recovery is positively correlated with the magnitude of the recession.

The mechanism has similar implications for the persistence of long-run negative shock on the level of output than those of Kim and Piger's MS augmented model, however SETAR models,probably for their piecewise linear nature, predict stronger responses to positive shocks at some stage of the business cycle,which are more similar to those of linear ARMA model .

It is also interesting to notice that SETAR models offer a greater range of dynamic responses than MS models and are far easier to estimate.

In conclusion linear ARMA models, two-regime and three regimes MS models and SETAR models yield different economic implications on whether the effect of shocks on aggregate output is asymmetric, permanent or both. Different models predict different qualitative and quantitative responses to negative and positive shocks and thus suggest alternative economic policies to deal with shocks.

Therefore it is of considerable importance to compare these alternative statistical representations of real GDP and assess which of them can better reproduce the behaviour of the series over the cycle.

Which comparison criterion?

The absence of a body of finite sample theory for non-linear models means that applied research must rely on asymptotic theory for inference.

In order to compare the above linear and non-linear models one would thus like to apply the Cox Non-nested testing methodology, however there are many serious conceptual difficulties in doing this due to the filtering required for the estimation of MS models.

To understand this recall that the primary asymptotic distributional theory requires that in a sufficiently large sample the estimator nears the true parameter vector .For non-linear models it runs roughly as follows:

via a Taylor's expansion the parameter estimates are equal to their true value plus the score evaluated at the true value ,divided by the second derivative matrix evaluated at median points. The likelihood surface is assumed to be approximately quadratic in this region,so the second derivatives are approximately constant (that is, not a function of the parameters).Since the score is zero mean,if it has positive variance , one can apply the central limit theorem ,and conclude that the estimator has an as asymptotic multivariate normal distribution.

There appear to be two keys assumption in this procedure.

First the likelihood surface must be locally quadratic, where locally means that the likelihood surface is approximately quadratic over the region in which the global optimum lies .This condition is violated if for example some parameters are not identified under the null hypothesis,because then the likelihood function is flat with respect to the unidentified parameters at the optimum.

Second the score must have a positive variance. This condition is violated when the score is identically to zero under the null hypothesis,which occurs when the null hypothesis yields a local maximum ,minimum or an inflection point.

Unfortunately , Markov- switching models violate both the two above necessary conditions.

Two nuisance parameters (the transition probabilities) are not identified under the null hypothesis .The null hypothesis also yield a local optimum of the likelihood surface, and higher order derivatives also appear to be zero. This yields a singular information matrix under the null. Being highly non-linear ,the model produces numerous local optima as well.

Standard asymptotic theory appears therefore inapplicable to formally test hypothesis involving MS models.

This lack gives rise to the need of a different *model comparison device* to compare MS models with linear and non linear time -series models of real GDP.

In the following paragraph I propose a simple comparison device based on the consideration that business cycle features themselves could motivate a good metric for comparing different models.

3-Towards a comparison device.

In the literature statistical evaluation of MS models and comparison with non-linear and linear models is mostly done in term of residual based tests and forecasting criteria. Not much attention is paid to the models -induced stochastic process of business cycle per se.

However each time series model ,as a data generating process, generates specific patterns of expansions and contractions.

It appears therefore of considerable interest to address a simple question: how well do alternative popular time-series models reproduce the business cycle features of real GDP? In particular how well do they reproduce two basic stylised facts such as duration and amplitude of business-cycle phases?

The diagnostic check for business cycle duration answer the following question: does a given time series model produce expansions and contractions whose average duration is similar to those observed in the data?

Correspondingly, the diagnostic check for business-cycle amplitude asks whether a time series model produces expansions and contraction whose average amplitude is similar to those observed in the data.

This provides a simple comparison device for the above linear and non-linear model ,enabling one to evaluate *how the above alternative dgp perform at replicating the business cycle features of the data.*

In what follows I explain the set of definitions of contractions and expansions that I have used,then,on the basis of this definitions,I derive stylised facts of business features for US GDP over the period 1970 -2003 ,finally I report results of a Monte Carlo simulation for the above models and draw conclusions on *which model is the most likely to generate the business-cycle features that have been highlighted.*

Definitions

In the literature the NBER chronology for US data is normally regarded as providing the most authentic standard. However a well known problem with NBER 's dating practice is that it is based on an informal judgement about turning points, and it is difficult to reproduce by a formalised algorithm. One direct solution to this problem is to propose a new set of definitions of contractions and expansions.

Harding and Pagan (2002) extend the Bry and Boshan 's algorithm to identify peaks and trough in a given data series,and find that, when this algorithm is applied to US real GDP,it provides a chronology very close to that established by the NBER.

Kim and Piger employ this algorithm for evaluation of their MS augmented model and find that it performs considerably better than linear AR(1) ,AR(2) and MA(1) and Sichel's (1994) three regimes MS model. However they don't compare their model with any linear ARMA model, which is in fact the most popular linear model in the literature, and with any SETAR model. Thus the only non-linear alternative specification that is considered is still another MS model. Instead the purpose of my comparison is to draw conclusions about how well MS models compare to linear ARMA models and to a different non-linear class of models , SETAR models, which are far easier to estimate and with well defined asymptotics.

To identify business cycle turning points I refer to the set of definitions of expansions and contractions introduced by King and Plosser (1994) and add an "average post recession growth" stylised fact to in order to capture the bounce-back effect in the level of aggregate output.

To begin with define a running peak m_t of the time series $\{y_\tau\}_{\tau=0}^{\infty}$ at time t as the current historical maximum of the series $m_t = \max_{0 \leq S \leq t} y_s$.

This is to be distinguished from the NBER reference-cycle peak ,which is the month prior to which the economy is in a recession.

A *contraction* is a whole time period during which the current value of the series is lower than the running peak m_t.

The *duration of a contraction* is equal to h when $y_s \prec m_s$ for $s=t+1, ..., t+h$ and $y_s = m_s$ for s=t and t+h+1.

The *depth of a contraction* ,starting at time t with duration h ,is the maximum difference of m_s and y_s between time period t and t+h.

An *expansion* is any time period not belonging to a contraction. Therefore at any time in a contraction $m_t = y_t$.

The *duration of an expansion* is equal to k when $y_s = m_s$ for s_t =t+1,.....,t+h and $y_s \prec m_s$ for s=t and s=t+h+1.

The *height of an expansion* ,starting at time t with duration k, is defined as the difference of the running peaks between the starting and ending time of the expansion.

The *average post-recession growth* is defined as growth in the four quarters following a business cycle trough.

Data.

The data I employ to compute statististics for the above business cycle features and to estimate the presented models are from the National Bureau of Economic Analysis.

I use non seasonal, quarterly observations of US real GDP from 1947:Q1 to 2003:Q4.

Data are originally in level but the models are estimated in growth rates which are formed as 100 times the log of first difference of the level data.

The reason for which I use US data is that most of the literature on real GDP time-series models concentrates on the US business cycle, thus for result comparison purposes this seems to be the better choice.

Real GDP business-cycle stylised facts.

On the basis of the above definitions, the following stylised facts of US business cycle form 1947 to 2003 have been derived:

number of peaks	number of whole cycles	average post recession growth
133	21	4.96

average depth of contractions	average duration of contractions
1.30%	2.61Q

average height of expansion	average duration of expansion
8.79%	6.7Q

Total of 223 observations.
Durations are measured in quarters.

4-Monte Carlo simulation.

Models estimation.

The estimates of the six models that I have used to calibrate the dgp used in the Monte Carlo are reported below:

ARIMA (1,1,0):
$\Delta\hat{y}_t = 0.0048 + 0.3650\Delta y_{t-1}$
St.errors: 0.009;0.0640
$\sigma_\varepsilon = 0.0095$

ARIMA (2,1,2)

$\Delta\hat{y}_t = 0.068 + 0.501\Delta y_{t-1} - 0.4983\Delta y_{t-2} - 0.2301\varepsilon_{t-1} + 0.6557\varepsilon_{t-2}$
St.errors:0,210;0.074;0.1230;o.1172;0.0896
$\sigma_\varepsilon = 0.0098$

Potter 's SETAR

regime 1:$\Delta y_{t-2} \leq 0$
$\Delta\hat{y}_t = -0.0071 + 0.302\Delta y_{t-1} - 0.591\Delta y_{t-2}$
St.errors:0.0025;0.105;0.352
$\sigma_\varepsilon = 0.0121$
53 observations

regime2:$\Delta y_{t-2} \succ 0$
$\Delta\hat{y}_t = 0.0039 + 0.316\Delta y_{t-1} + 0.298\Delta y_{t-2}$
St.errors:0.0014;0.090;0.129
$\sigma_\varepsilon = 0.0098$
170 observations

Tiao and Tsay's multistage SETAR

regime 1: $\Delta \mathrm{y}_{t-1} \leq \Delta y_{t-2} \leq 0$
$\Delta\hat{y}_t = -0.057 - 0.032\Delta y_{t-1}$
st.errors:0.0052;0.309
$\sigma_{\varepsilon 1} = 0.0089$
12 observations

regime2: $\Delta y_{t-1} \succ \Delta y_{t-2}$ and $\Delta y_{t-2} \leq 0$

$\Delta\hat{y}_t = -0.01 + 0.389\Delta y_{t-1} - 0.487\Delta y_{t-2}$
st.errors:0.0028;0.091;0.053
$\sigma_{\varepsilon 2} = 0.0214$
67 observations

regime 3: $\Delta y_{t-1} \leq \Delta y_{t-2}$ and $\Delta y_{t-2} \succ 0$
$\Delta\hat{y}_t = 0.013 + 0.526\Delta y_{t-1}$
st.errors:0.079;0.192
$\sigma_{\varepsilon 3} = 0.0172$
98 observations

regime4:$\Delta y_{t-1} \succ \Delta y_{t-2} \succ 0$
$\Delta\hat{y}_t = 0.002 + 0.503\Delta y_{t-1}$
st.errors: 0.117;0.071
$\sigma_{\varepsilon 4} = 0.0079$
49 observations

I haven't been able to estimates the two MS models presented since I had no computer program for estimating switching-regime models.

For *Kim and Piger's MS augmented model* I rely on their original estimates. This is essentially the reason for which I have chosen to employ data on real US GDP from 1947 to 2003 so that the estimates of the other models that I have derived and the business cycle features are all based on the same sample period .

For *Hamilton's MS model* the estimation problem is not so easily bypassed.

Kim and Nelson in their authoritative book on regime switching models (Kim and Nelson 1999) suggest that when Hamilton's model is extended to include 11 more years of recent observations it fails to provide reasonable parameter estimates and thus it fails to provide reasonable inferences on the probability of a recession or a boom.

They suggest that this might be due to the lack in the model of a mechanism to account for a productivity slowdown in the more recent sample since the model assumes that the average growth rate of output during a boom or a recession is the same over the entire sample.

To account for this possibility they propose a different specification of Hamilton's model including a dummy. The model is thus specified as:

$$(\Delta y_t - \mu_{st}) = \phi_1(\Delta y_{t-1} - \mu_{st-1}) + \phi_2(\Delta y_{t-2} - \mu_{st-2}) + ... + \phi_4(\Delta y_{t-4} - \mu_{st-4}) + \varepsilon_t$$

$$\varepsilon_t \sim i.i.dN(0, \sigma^2)$$

$$\mu_{st} = (\mu_0 + \mu_0^* D_t)(1 - S_t) + (\mu_1 + \mu_1^* D_1)S_t$$

where D_t is a dummy variable set set equal to 1 for the subsample 1983:1 1995:4and 0 for the earlier sample period.

The inclusion of the dummy variable potentially captures a change in the mean growth rates during booms and recessions.

I have used the estimates of Hamilton's MS model they derived to calibrate the dgp for this model in the Monte Carlo.

The sample period on which these estimates are based goes from 1952:1 to 1995:4.

There thus appear to be a gap of 13 years between the sample period for which all the other models are estimated and the one for which the Hamilton's model is estimated. This is of course a limitation of my work,even if ,luckily enough,the missing 13 years of observation for the estimation of the MS model are spread into two different historical periods.

The estimates are reported below, standard errors are in parentheses:

p	0.9113(0.0363)	0.9187(00309)
q	0.7658(0.0857)	0.7668(0.0863)
ϕ_1	0.0496(0.1347)	0.0477(0.1117)
ϕ_2	-0.0495(0.1295)	-0.0422(0.1103)
ϕ_3	-0.2112(0.1129)	-0.2095(0.1008)
ϕ_4	-0.0953(0.1140)	-0.0984(0.0970)
σ	0.6902(0.0505)	0.6939(0.0474)
μ_0	-0.2996(0.1892)	-0.2328(0.1895)
μ_1	1.1479(0.0768)	1.1510(0.0776)
μ_0^*	0.4516(0.3209)	-
μ_1^*	-0.3346(0.1340)	-0.3699(0.1244)
Log likelihood	-212.17	-212.99

μ_0^* is statistically insignificant, μ_1^* is negative and statistically significant .

Results.

In order to evaluate how the alternative models of real GDP estimated above perform at replicating business cycle features I have conducted a Monte Carlo simulation generating 5000 samples of 223 artificial observations for each of the calibrated models.

Below I report the mean values and the standard deviations of the various business cycle features obtained for each model with the corresponding stylised fact of the data. In the next section I use these results to draw conclusions.

model	number of peaks	number of whole cycles
ARIMA(1,1,0)	114.06(12.70)	22.36(3.13)
ARIMA(2,1,2)	110.54(12.15)	23.88(3.23)
Potter SETAR	119.96(12.17)	20.84(3.36)
multiregime SETAR	118.36(13.37)	19.36(2.98)
Hamilton MS	117.47(13.37)	17.75(3.06)
Augmented MS	119.83(10.13)	18.90(2.48)
observed data	133	21

model	average depth of the contraction	average length of the contraction
ARIMA(1,1,0)	1.29(0.48)	3.21(0.86)
ARIMA(2,1,2)	1.23(0.38)	3.09(0.75)
Potter SETAR	1.32(0.48)	2.98(0.93)
multiregime SETAR	1.34(0.52)	2.87(0.97)
Hamilton MS	1.49(0.59)	3.43(0.98)
Augmented MS	1.26(0.40)	2.96(0.80)
observed data	1.30%	2.61Q

model	average height of the expansion	average length of the expansion
ARIMA(1,1,0)	6.98(2.11)	5.90(1.24)
ARIMA(2,1,2)	5.25(1.55)	4.86(0.97)
Potter SETAR	7.93(2.37)	6.9(1.38)
multistage SETAR	7.96(1.48)	6.5(1.85)
Hamilton MS	6.93(2.39)	5.91(1.66)
Augmented MS	7.11(2.39)	6.34(1.45)
observed data	8.79%	6.7Q

model	average post-recession growth
ARIMA (1,1,0)	3.98(0.70)
ARIMA(2,1,2)	2.98(0.87)
Potter SETAR	4.89(0.38)
multistage SETAR	4.87(0.21)
Hamilton MS	3.83(0.69)
Augmented MS	5.25(0.90)
observed data	4.96

5-Conclusions

On the basis of the above results of the Monte Carlo simulation some interesting conclusions can be drawn with respect to the ability of each of the presented time-series models of real GDP to reproduce business-cycle features.

Linear ARMA models perform quite well in reproducing the number of peaks and of whole cycles and the average length and depth of recessions. However they underestimate both the average length and the average height of expansions and they fail to capture the post recession "bounce back "effect that is evident in the data.

Hamilton MS model in fact provides no significant improvement on linear models in reproducing the selected business cycle features.

The augmented MS model of Kim and Piger , Potter's SETAR model and the multistage SETAR model are the three models that perform better in the overall.

It is interesting to notice that Potter's SETAR model and Tiao and Tsay's multistage SETAR model perform quite similarly in this check. It thus appears that the distinction between two regime and multiple regime model ,that is so much stressed in the literature ,is not that relevant within the class of SETAR models ,at least when the ability of reproducing business cycle features is the point of interest.

Instead this same distinction appears very relevant within the class of MS models: in fact the augmented MS model performs considerably better than the two stage one.

Only the two SETAR models and the augmented MS are able to capture adequately the post recession "bounce-back " effect in the level of aggregate output thus over performing linear ARMA models and the two stage MS model.

The most relevant limitation of the Monte Carlo results is the relative shortness of the sample period on which the Hamilton MS model has been estimated. I am not able to say how much this lack has influenced the above results and thus my conclusions.

The project itself has many limitations.

The comparison between non-linear models does not take into account any STAR model,whereas these are quite popular time-series models of real GDP in the current literature.

The project addresses a model comparison question within the classical statistical framework . However MS models are frequently employed in a Bayesian ,Gibbs-sampling ,framework. It would thus be very interesting to address the same question and compare SETAR and MS model in a Gibbs-sampling setting. To my knowledge this has not yet been attempted.

Econometric modelling should always be guided by the principle of parsimony.

Simpler models should be preferred to more complicated models *ceteris paribus.*

In this project the definition of *ceteris paribus* is given as the model ability to reproduce stylized facts of the business cycle. Bearing in mind this specific choice, we can conclude that a simple two stage SETAR model should be preferred to any Markov switching representation of the real GDP series.

The two stage SETAR model performs as well as the multistage SETAR and the augmented MS in reproducing business cycle feature but it's a far easier model to estimate. Furthermore there is a well developed asymptotic theory for SETAR model that allows the econometrician to make robust inference within this class of models.