Detection of the industrial business cycle using SETAR models

Ferrara, Laurent and Guégan, Dominique

September 2005

Detection of the Industrial Business Cycle using SETAR Models

Laurent Ferrara *and Dominique Guégan †

September 13, 2005

Abstract:
In this paper, we consider a threshold time series model in order to take
into account certain stylized facts of the industrial business cycle, such as
asymmetries in the phases of the cycle. Our aim is to point out some thresh-
olds under (over) which a signal of turning point could be given. First, we
introduce the various threshold models and we discuss both their statisti-
cal theoretical and empirical properties. Especially, we review the classical
techniques to estimate the number of regimes, the threshold, the delay and
the parameters of the model. Then, we apply these models to the Euro-
zone industrial production index to detect, through a dynamic simulation
approach, the dates of peaks and troughs in the business cycle.

Keywords:
Economic cycle, turning point detection, Threshold model, Euro-zone IPI.

JEL Classification:
C22, C51, E32.

---

*Centre d'Observation Economique de la CCIP and E.N.S. Cachan, MORA-IDHE,
CNRS UMR 8533. Adress: COE, 27 Avenue de Friedland, 75382 Paris Cedex 08. E-mail:
lferrara@ccip.fr

†E.N.S. Cachan, Département d'Economie et Gestion, MORA-IDHE, CNRS UMR
8533, Senior Academic Fellow de l'Institut Europlace de finance (IEF). Adress: 61 Av-
enue du President Wilson, 94235, Cachan Cedex, France. E-mail : guegan@ecogest.ens-
cachan.fr

1 Introduction

Recently, we witnessed the development of new tools in business cycle analysis, mainly based on non-linear parametric modeling. Non-linear models have the great advantage to be flexible enough to take into account certain stylized facts of the economic business cycle, such as asymmetries in the phases of the cycle. In this respect, emphasis has been placed on the class of non-linear dynamic models that accommodate the possibility of regime changes. Especially, Markov-Switching models popularized by Hamilton (1989) have been extensively used in business cycle analysis in order to describe the economic fluctuations. Among the huge amount of empirical studies, we can quote the papers of Sichel (1994), Lahiri and Wang (1994), Potter (1995), Chauvet and Piger (2003), Ferrara (2003), Clements and Krolzig (2003) or Anas and Ferrara (2004b) as regards the US economy and the papers of Krolzig (2001, 2004) or Krolzig and Toro (2001) on the Euro-zone economy. Generally, the output of these applications is twofold. The authors aim either at dating the turning points of the cycle or at detecting in real-time the current regime of the economy.

However, dating and detecting the turning points of the cycle are quite different objectives. Dating is an *ex post* exercise for which several parametric and non-parametric methodologies are available. It turns out that simple non-parametric procedures, such as the famous Bry and Boschan (1971) procedure still used by the Dating Committee of the NBER, are more convenient for this kind of work (see Harding and Pagan, 2001, or Anas and Ferrara, 2004a, for a discussion on this issue). Real-time detection refers mainly to short-term economic analysis, which is not an easy task for practitioners. Indeed, several economic indicators are released on a regularly monthly basis, or even on a daily basis as regards the financial sector, adding volatility to the existing volatility and thus leading to an inflation of the available information set. Moreover, the data are often strongly revised and the diverse statistical methods, such as seasonal adjustment or filtering techniques, lead to edge-effects.

Besides the well known Markov-Switching approach, other parametric models have been proposed in the statistical literature to allow for different regimes in business cycle analysis. For instance, probit and logit models have been used by Estrella and Mishkin (1998) to predict US recessions. The threshold autoregressive (TAR) model, proposed first by Tong and Lim (1980), has been used to describe the asymmetry observed in the quartely US real GNP by different authors, such as Tiao and Tsay (1994), Potter (1995) and Proietti (1998) for instance, and using US unemployment monthly data by Hansen (1997). In the TAR model, the transition variable may be either an exogenous variable, such as a leading index for example, or a linear

combination of lagged values of the series. In this latter case, the model is referred to as a self-exciting threshold autoregressive (SETAR) model. This is the main difference with the Markov-Switching model whose generating process varies according to the states of the latent Markov chain. These two approaches are complementary because the notion captured under investigation is not exactly the same. Nevertheless, one of the interest of SETAR processes lies on their predictability, see for instance De Goojier and De Bruin (1999) and Clements and Smith (2001). When dealing with SETAR models, the transition is discrete, but smooth transition is also chosen to study the business cycle by some authors. Then, we get the so-called STAR (smooth transition autoregressive) model, see for instance Terasvirta and Anderson (1992) and van Dijk, Terasvirta and Franses (2002).

In this paper, we focus on the detection of business cycle turning points. Our aim is rather to point out some thresholds under (over) which a signal of turning point could be given. We adopt the SETAR approach because a threshold model seems to be attractive in terms of business cycle analysis. Here, we propose a prospective approach to detect the business cycle as an alternative to other more classical parametric approaches.

This paper is split into two parts. In a first step (section two), we introduce threshold models which allow to capture states in a time series, then, in section three, we specify the method used to estimate the different parameters of the threshold models. In a second step, we apply different threshold models to the Euro-zone industrial production index to detect the dates of peaks and troughs in the business cycle (section four). By using a dynamic simulation approach, we also provide a measure of performance of our model by comparison to a benchmark dating chronology. Lastly, some conclusions and further research directions are proposed.

2   Description of models which capture states

In this section, we specify some of the models which permit to take into account the existence of various states in the data. For sake of simplicity, we describe the models only with two regimes, but they can be easily generalised to more regimes. For a review concerning these kinds of processes, we refer to Tong (1990), Franses and van Dijk (2000) and Guégan (2003).

2.1   Threshold processes

The covariance-stationary process (Y_t) follows a two-regimes threshold autoregressive process, denoted TAR(2,p_1,p_2), if it satisfies the following equa-

tion:

$$Y_t = (1 - I(Z_{t-d} > c))(\phi_{1,0} + \sum_{i=1}^{p_1} \phi_{1,i}Y_{t-i} + \sigma_1\varepsilon_t)$$

$$+ \quad I(Z_{t-d} > c)(\phi_{2,0} + \sum_{i=1}^{p_2} \phi_{2,i}Y_{t-i} + \sigma_2\varepsilon_t), \quad (1)$$

where c is the threshold, $d > 0$ the delay, (ε_t) a standardised white noise process, (Z_t) the transition variable. Here, $I(.)$ is the indicator function such that $I(Z_{t-d} > c) = 1$ if $Z_{t-d} > c$ and zero otherwise. If, $\forall t$, $Z_t = Y_t$, the process is referred to as self-exciting TAR process (SETAR). For a given threshold c and the position of the random variable Z_{t-d} with respect to this threshold c, the process (Y_t) follows here a particular AR(p) model. The model parameters are $\phi_{j,i}$, for $i = 0, \dots, p_j$ and $j = 1, 2$, the standard variance errors σ_1 and σ_2, the threshold c and the delay d. This model has been introduced first by Tong and Lim (1980).

Using some algebraic notations, the model (1), with $p_1 = p_2 = p$, can be rewritten as a regression model. Denote $I_d(c) \equiv I(Z_{t-d} > c)$, $\Phi_1 = [\phi_{1,0}, \cdots, \phi_{1,p}]'$, $\Phi_2 = [\phi_{2,0}, \cdots, \phi_{2,p}]'$ and $\mathbf{Y}'_{\mathbf{t-1}} = [1, Y_{t-1}, \cdots, Y_{t-p}]$, then, we get the following representation from (1):

$$Y_t = (1 - I_d(c))\mathbf{Y}'_{\mathbf{t-1}}\Phi_1 + I_d(c)\mathbf{Y}'_{\mathbf{t-1}}\Phi_2 + ((1 - I_d(c))\sigma_1 + I_d(c)\sigma_2)\varepsilon_t. \quad (2)$$

If we denote π the unconditional stationary distribution of the process (Y_t), to get its analytical form is a non-trivial problem. However an implicit solution is always available if the stationary process (Y_t) can be considered as an ergodic Markov chain over R^n. It is given for any event A, by:

$$\pi(A) = \int_{-\infty}^{\infty} P(A|x)\pi(dx),$$

where π denotes the limiting distribution of (Y_t). For SETAR processes introduced in (1), different numerical solutions have been proposed to solve this problem, see Jones (1978) and Pemberton (1985). In reality, we obtain an approximation of this distribution, computing empirically the percentage of points belonging to the first regime or to the second one. This method gives an estimation of the unconditional probability (π_1 or π_2) to be in a given regime.

On each state, it is possible to propose more complex stationary models like ARMA(p,q) processes, GARCH(p,q) processes (see for instance Za-koian, 1994) or long memory processes (see Dufrénot, Guégan and Peiguin-Feissolle, 2005a and 2005b). Note also that the regimes can be characterized by changes in the variance of the noise process (see Pfann, Schotman and Tchernig, 1996).

2.2 Smooth transition processes

Instead of using a sharp transition between the two states, characterised by the indicator function $I(.)$, we can use a smooth transition. This is the basic idea of the smooth transition autoregressive (STAR) process. In that case, by using the previous notations, the two-regimes STAR(2,p_1,p_2) process (Y_t) follows the recursive scheme:

$$
\begin{aligned}
Y_t &= (1 - G(Z_{t-d}, \gamma, c))(\phi_{1,0} + \sum_{i=1}^{p_1} \phi_{1,i} Y_{t-i} + \sigma_1 \varepsilon_t) \\
&+ G(Z_{t-d}, \gamma, c)(\phi_{2,0} + \sum_{i=1}^{p_2} \phi_{2,i} Y_{t-i} + \sigma_2 \varepsilon_t),
\end{aligned}
\tag{3}
$$

where G is some continuous function, for instance the logistic one:

$$
G(Z_{t-d}, \gamma, c) = \frac{1}{1 + \exp(-\gamma(Z_{t-d} - c))}.
\tag{4}
$$

Note that the transition function G is bounded between 0 and 1. The parameter c can be interpreted as the threshold between the two regimes in the sense that the logistic function changes monotonically from 0 to 1 with respect to the value of the lagged exogeneous or endogenous variable Z_{t-d}. The parameter γ determines the smoothness of the change in the value of the logistic function, and thus, the smoothness of the transition of one regime to the other. As γ becomes very large, the logistic function (4) approaches the indicator function $I(Z_{t-d} > c)$. Consequently, the change of $G(Y_{t-d}, \gamma, c)$ from 0 to 1 becomes instantaneous at $Z_{t-d} = c$. Then we find the TAR model as a particular case of this STAR model. When $\gamma \to 0$, the logistic function approaches a constant (equal to 0.5) and when $\gamma = 0$, the STAR model reduces to a linear AR model. This model has been described by Terasvirta and Anderson (1992). Other generalisations of the STAR process have been proposed, for instance by replacing the logistic function by the exponential function or by using long memory dynamics in each regime (see van Dijk, Franses and Paap, 2002).

3 Estimation for SETAR models

In the following, we use the SETAR process in order to model the economic business cycle using the Euro-zone industrial production index. We now describe the estimation procedure we use in section four.

The TAR model introduced in the eighties' has not been widely used in applications until recently, primarily because it was hard in practice to identify the threshold variable and to estimate the associated values, and, secondly,

because there was no simple modeling procedure available. Recently, some authors have proposed different ways to bypass this problem. We present in this section a way to estimate the parameters of the SETAR models and we specify some recent literature which permits to implement quickly the procedure we use below.

Here, we assume first that the model available for our purpose is a SETAR $(2,p_1,p_2)$ model described by equation (1), with $\sigma_1 = \sigma_2 = 1$. As noted above, a major difficulty in applying TAR models is the specification of the threshold variable, which plays a key role in the non-linear structure of the model. Since there is only a finite number of choices for the parameters c and d, the best choice can be done using the Akaike Information Criterion (AIC), see Akaike (1974). This procedure has been proposed by Tong and Lim (1980) and is used by a large part of the practitioners dealing with this model.

Now, we assume that we observe a sequence of data $(Y_1, \cdots, Y_n)$ from the model (2). The equation (2) is a regression equation (albeit non-linear in parameters) and an appropriate estimation method is least squares (LS). Under the auxiliary assumption that the noise $(\varepsilon_t)_t$ is a strong Gaussian white noise, the least squares estimation is equivalent to the maximum likelihood estimation. Since the regression equation (2) is non-linear and discontinuous, the easiest method to obtain the LS estimates is to use sequential conditional LS. We will use this approach here. Recall that conditional least squares lead to the minimization of:

$$\sum_{Y_{t-d} \leq c, t=1}^{n} \left(Y_t - \phi_{1,0} - \sum_{i=1}^{p_1} \phi_{1,i} Y_{t-i}\right)^2 + \sum_{Y_{t-d} > c, t=1}^{n} \left(Y_t - \phi_{2,0} - \sum_{i=1}^{p_2} \phi_{2,i} Y_{t-i}\right)^2, \ (5)$$

with respect to $\Phi_1, \Phi_2, c, d, p_1, p_2$. Generally, we first assume that the autoregressive orders p_1 and p_2 are known.

Recall that Chan (1993) proves, under geometric ergodicity and some other regularity conditions for the process (2), that the LS parameters estimates of this process have good properties. The threshold parameter is consistent, tends to the true value at rate n and, suitably normalized follows asymptotically a Compound Poisson process. The other parameters of the model are $n^{-1/2}$ consistent and are asymptotically Normally distributed. The limitation of the theory of Chan (1993) concerns the construction of confidence intervals for the threshold c. Indeed, if we denote $\hat{c}$ the LS estimate for c, Chan (1993) finds that $(\hat{c} - c)$ converges in distribution to a functional of a Compound Poisson process and unfortunately, this representation depends upon a host of nuisance parameters, including the marginal distribution of (Y_t) and all the regression coefficients. Hence, this theory does not yield a

6

practical method to construct confidence intervals. Some discussions and extensions to this work can be found in Hansen (2000), Clements and Smith (2001) and Gonzalo and Pitarakis (2002), for instance.

In practice, to determine the parameters c, d, p_1, p_2, we need to assume the existence of a maximal possible order P of the two subregimes and a greatest possible delay D. The threshold parameter c is estimated by using a grid-search procedure. The grid points $(c_1, \ldots, c_s)$ are obtained using the quantiles of the sample under investigation. We use equally spaced quantiles from the 10 (percent) quantiles and ending at the 90 (percent) quantiles. Now, for each fixed pair (d, c_i), $0 < d < D$, $i = 1, \cdots s$, the appropriate TAR model is identified. The AIC criterion is used for selection of the orders p_1 and p_2. In this context, it becomes:

$$AIC(p_1, p_2, d, c) = \ln(\frac{1}{n} \sum \hat{\varepsilon}_t^2) + 2 \frac{p_1 + p_2 + 2}{n}, \qquad (6)$$

where $\hat{\varepsilon}_t$ denotes the residuals when we use the appropriate model for each pair (d, c_i) from the LS approach.

Finally the model with the parameters p_1^*, p_2^*, d^* and c^* that minimize the AIC criterion can be chosen. Since for different d there are different numbers of values that can be used for estimation, the following adjustment should be done. With $n_d = \max(d, P)$ it is:

$$AIC(p_1^*, p_2^*, d^*, c^*) = \min_{p_1, p_2, d, c} \frac{1}{n - n_d} AIC(p_1, p_2, d, c). \qquad (7)$$

Different algorithms have been proposed to improve the properties of the estimates and the speed of the algorithms and we suggest the reader to consult them. It is possible to use a Bayesian approach based on Gibbs sampler proposed by Tiao and Tsay (1994), see also Potter (1999); graphical procedures classifying the observations without knowing the threshold variable to estimate the parameters, see Chen (1995); numerical approaches, see Coakley, Fuertes and Pérez (2003) or a Markov Chain Monte-Carlo approach developed in particular by So and Chen (2003).

In this paper, we consider the Tsay test (1989) to justify the use of SETAR models. Hansen (1997) considers another approach based on a likelihood ratio statistic and a Lagrange Multiplier test has been also proposed by Proietti (1998). To our knowledge, there is no test available to decide between SETAR models and Markov-Switching models.

4   Empirical results

In this section, our aim is to apply a SETAR model to the Euro-zone Industrial Production Index in order to detect the low phases of the industrial

business cycle referred to as the industrial recessions. The application is done in two steps: first we try to find the best SETAR model following the method proposed in section 3 based on the AIC criterion. Secondly we use the model to detect the periods of expansion and recession. By comparing the results to reference recession dates, we can assess the ability of the model to reproduce the industrial business cycle features.

4.1 Data description

The analysis is carried out on the IPI series considered in the paper of Anas *et al.* (2003). This series is a proxy of the monthly aggregate Euro-zone IPI for the 12 countries, beginning in January 1970 and ending in December 2002. The data are working day adjusted and seasonally adjusted by using the Tramo-SEATS model-based methodology, implemented in the Demetra software, which used a Wiener-Kolmogorov filter (see for instance Maravall and Planas, 1999). Moreover, the irregular part including outliers has been removed by using the same methodology. It is noteworthy that there is still a debate among statisticians about the impact of filtering methodologies on the timing of peaks and troughs in business cycle analysis (see for instance
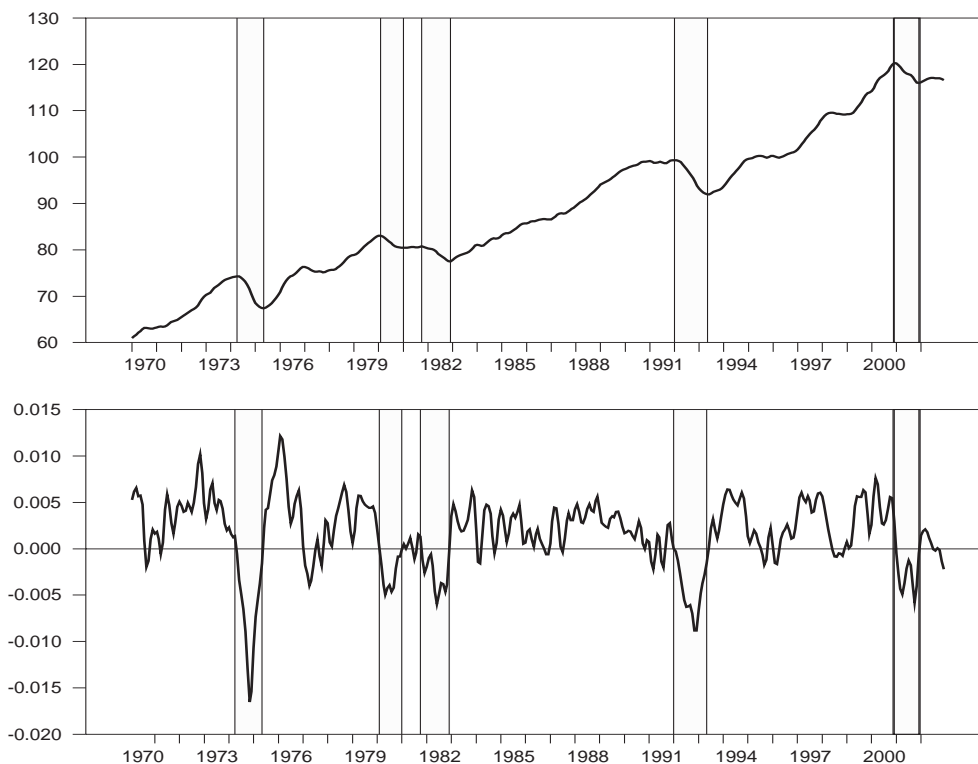

Figure 1: Euro12 IPI (top) and its monthly growth rate (bottom), as well as the reference industrial recession periods (shaded areas), from January 1970 to December 2002.
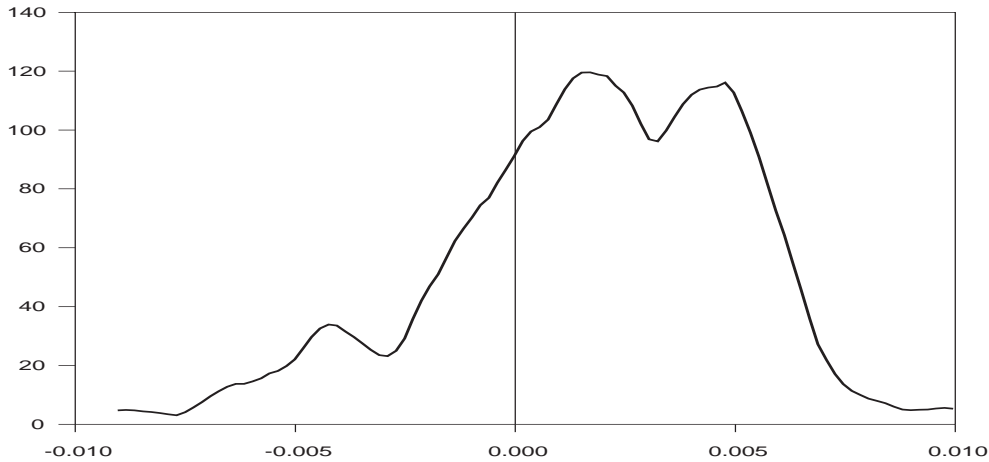
8

Figure 2: Empirical unconditional distribution of the IPI growth rate, from January 1970 to December 2002.

Lahiri *et al.*, 2004).

The original series (X_t) is presented in figure 1 as well as its monthly growth rate (Y_t) defined, for all t, by: $Y_t = \log(X_t) - \log(X_{t-1})$. In figure 1, the shaded areas represent the reference industrial recession dates. Several authors have proposed a turning point chronology for the Euro-zone industrial business cycle, by using different statistical techniques and economic arguments. For example, we refer to Anas *et al.* (2003), who propose a classical NBER-based non-parametric approach, and to Artis *et al.* (2003), Krolzig (2004) or Anas and Ferrara (2004b) who apply parametric Markov-Switching models. Generally, the industrial recession dates are more or less similar. In fact, it turns out that the Euro-zone experienced five industrial recessions: in 1974-75 and 1980-81 due to the first and second oil shocks, in 1981-82, in 1992-93, due to the American recession and the Gulf war, and lastly in 2000-2001 because of the global economic slowdown caused itself by the US recession from March 2001 to November 2001. It is noteworthy that, contrary to a common belief among economists, the Asian crisis in 1997-98 has not caused an industrial recession in the whole Euro-zone, but only a slowdown of the production. Finally, we retain as a benchmark for our study the dates proposed by Anas *et al.* (2003) and summarized in the first column in table 4.

To ensure stationarity, we are going to deal with the monthly industrial growth rate (Y_t). The unconditional empirical distribution of the IPI growth rate computed by using a non-parametric kernel estimate (with the Epanechnikov kernel) is presented in figure 2. There is a clear evidence of three peaks in the estimated distribution. The lowest peak is due to the negative growth

rates during industrial recessions. The intermediate peak seems to be caused by periods of low, but positive, growth rates, experienced for example during the eighties, while the peak corresponding to the highest value is related to periods of fast growth. It is noteworthy that, from 1970 to 2002, periods of low growth rates seem to appear more frequently than periods of high growth rates. Moreover, this empirical distribution is cleary asymmetric (skewness equal to -0.9315) and with heavy tails (excess kurtosis equal to 2.4850). Consequently, the unconditional Gaussian assumption is strongly rejected by a Jarque-Bera test.

4.2 Whole sample modelling

In this subsection we fit various SETAR models to the industrial growth rate series (Y_t), that is, we model the growth of the Euro-zone industry. We consider first a two-regime model, the transition variable being successively the lagged series and the lagged differenced series. Then, we consider a multiple regime model by mixing the conditions on these previous series. For each model, we compare the estimated regimes with the reference recession phases in order to assess the ability of the model to reproduce business cycle features.

4.2.1 Model 1

The first SETAR model uses the lagged series Y_{t-d} as transition variable. Thus, we model the growth of the industrial production according to the regimes of the lagged growth. The delay d and the threshold c are estimated by using the methodology presented in the previous section. However, the autoregressive lag p has to be determined *a priori*. We proceed by using a descendent stepwise approach by considering first $p = 12$. For all estimated models, it turns out that the parameters corresponding to a lag greater than three are statistically not significant by the usual Student test. Therefore, we impose the choice $p = 3$ for all the models. The Tsay (1989) test with $p = 3$ rejects the null of linearity for $d = 1$ and for $4 \leq d \leq 12$, at the usual risk $\alpha = 0.05$, implying thus the presence of two regimes. We get the following estimates for c and d : $\hat{c} = -0.0024$ and $\hat{d} = 1$. We note from table 1 that, in the high regime, the persistence is stronger, because the parameters corresponding to $p = 2$ and $p = 3$ in the low regime are not statistically significant and have been therefore cancelled, and the variance is smaller, which are expected results in business cycle analysis. The full estimated model is as follows (estimates and their standard errors are given in table 1):

$$Y_t = (-0.0049 + 0.8103Y_{t-1})(1 - I_{[Y_{t-1} > -0.0024]})$$
$$+ (0.0025 + 1.3950Y_{t-1} - 0.8742Y_{t-2} + 0.3318Y_{t-3})I_{[Y_{t-1} > -0.0024]} + \varepsilon_t.$$

The empirical unconditional probabilities of being in each regime are $\pi_1 = 0.11$ and $\pi_2 = 0.89$, which is consistent with the usual probabilities of being in recession and expansion in business cycle analysis. As regards the estimated recession dates, we get them by assuming that the low regime matches with the recession regime. The results are presented in figure 3 (top graph) and table 4 along with the two other dating chronologies stemming from the models described below. By comparison with the reference dating chronology, we can observe that the results are basically identical, except that we get a supplementary of recession in 1977, lasting only three months. If we had to establish a dating chronology, this period would not be retained as an industrial recession insofar as its duration is too short in comparison with the minimum duration of a business cycle phase, which generally of six months. However in this paper, to avoid non-persistent signals, we adopt the censoring rule saying that a signal must stay at least three months to be recognized as an estimated recession phase. Thus, this supplementary recession in 1977 is interpreted as a false signal of recession. In the remaining of this paper, a recession phase detected by the model but not present in the reference chronology is interpreted as a false signal of recession. Regarding the last industrial recession, the model estimates a recession period cut into two parts. This can be interpreted as a false signal of recovery. We also note that the other estimated industrial recessions are shorter, especially the 1982 recession but we get a first signal of recession in January 1982 which was not persistent. Otherwise, this model does not provide any other false signal for industrial recession.

	Low regime $[Y_{t-1} \leq -0.0024]$	High regime $[Y_{t-1} > -0.0024]$
$\hat{\phi}_0$	-0.0049 (0.0016)	0.0025 (0.0004)
$\hat{\phi}_1$	0.8103 (0.0931)	1.3950 (0.0513)
$\hat{\phi}_2$		-0.8742 (0.0782)
$\hat{\phi}_3$		0.3318 (0.0513)
$\hat{\sigma}_\varepsilon$	0.0021	0.0012

Table 1: Estimates and standard errors for model described in 4.2.1.

4.2.2 Model 2

The second SETAR model uses the differenced lagged series as transition variable, that is we try to model the growth of the industrial production according the regimes of its acceleration. We note this series Z_{t-d}, defined such as $\forall t$, $Z_{t-d} = Y_{t-1} - Y_{t-d}$. Actually, this series can be considered as a proxy of the acceleration of the IPI over $d-1$ months. It is interesting to investigate how the growth rate is related to the acceleration through a non-linear relationship. The Tsay test (1989) with $p = 3$ rejects the null of linearity for $4 \leq d \leq 13$, at the usual risk $\alpha = 0.05$, implying thus the presence of two regimes. It turns out that the delay d corresponding to the minimum AIC is equal to $d = 10$. That is, the acceleration over nine months seems to be the most significant. The estimated model is given by the following equation (estimates and their standard errors are given in table 2):

$$
\begin{aligned}
Y_t &= (-0.0051 + 0.8100 Y_{t-1})(1 - I_{[Z_{t-10} > -0.0061]}) \\
&+ (0.0024 + 1.7444 Y_{t-1} - 1.3796 Y_{t-2} + 0.5567 Y_{t-3}) I_{[Z_{t-10} > -0.0061]} + \varepsilon_t.
\end{aligned}
$$

Here again, we observe that the persistence is stronger in the higher regime while the variance is smaller and the empirical unconditional probabilities of being in each regime are exactly equal to the previous ones. The estimated industrial recession dates, presented in figure 3 (middle graph) and table 4, slighty differ from the previous estimates. Indeed, we get another false signal of industrial recession in 1998 due to the impact of the Asian crisis. Moreover, we note that the 1977 recession lasts seven months, but the 1982 recession is only of two months. Therefore, by considering the censoring rule adopted above, this model does not recognize this period as a recession. We also note that a non-persistent signal of recession is given in September 1995. Thus, by comparison with the reference dating chronology, this model provides two false signals of recession and misses the 1982 recession.

	Low regime $[Z_{t-10} \leq -0.0061]$	High regime $[Z_{t-10} > -0.0061]$
$\hat{\phi_0}$	-0.0051 (0.0018)	0.0023 (0.0006)
$\hat{\phi_1}$	0.8100 (0.0929)	1.7540 (0.0453)
$\hat{\phi_2}$		-1.3940 (0.0738)
$\hat{\phi_3}$		0.5630 (0.0454)
$\hat{\sigma_\varepsilon}$	0.0023	0.0009

Table 2: Estimates and standard errors for model described in 4.2.2.

Consequently, this model underperforms the previous one in detecting the industrial recession phases. This may be due to the fact that the acceleration, although computed over 9 months, appears to be too volatile.

4.2.3 Model 3

Lastly, the idea which appears to be natural is to combine the two previous SETAR models in a single model with two transition variables : the lagged growth rate and the acceleration. Therefore, the model possesses four regimes and two thresholds c_1 and c_2 have to be estimated. The estimated model which minimizes the AIC is given by the following equations (estimates and their standard errors are given in table 3) :

- Regime 1: if $Y_{t-1} < -0.00148$ and $Z_{t-10} < -0.00076$, then

$$Y_t = -0.0041 + 0.8273 Y_{t-1} + \varepsilon_t^1,$$

- Regime 2: if $Y_{t-1} < -0.00148$ and $Z_{t-10} \geq -0.00076$

$$Y_t = -0.0017 - 0.0934 Y_{t-1} + \varepsilon_t^2,$$

- Regime 3: if $Y_{t-1} \geq -0.00148$ and $Z_{t-10} < -0.00076$

$$Y_t = 0.0010 + 0.6520 Y_{t-1} + \varepsilon_t^3,$$

- Regime 4: if $Y_{t-1} \geq -0.00148$ and $Z_{t-10} \geq -0.00076$

$$Y_t = 0.0036 + 1.3005 Y_{t-1} - 0.7883 Y_{t-2} + 0.3321 Y_{t-3} + \varepsilon_t^4.$$

	Regime 1 $[Y_{t-1} < -0.0015]$ $[Z_{t-10} < -0.0008]$	Regime 2 $[Y_{t-1} < -0.0015]$ $[Z_{t-10} \geq -0.0008]$	Regime 3 $[Y_{t-1} \geq -0.0015]$ $[Z_{t-10} < -0.0008]$	Regime 4 $[Y_{t-1} \geq -0.0015]$ $[Z_{t-10} \geq -0.0008]$
$\hat{\phi}_0$	-0.0041 (0.0016)	-0.0018 (0.0011)	0.0010 (0.0005)	0.0036 (0.0005)
$\hat{\phi}_1$	0.8273 (0.0784)	-0.0934 (0.5282)	0.6520 (0.0749)	1.3005 (0.0660)
$\hat{\phi}_2$				-0.7883 (0.0974)
$\hat{\phi}_3$				0.3321 (0.0662)
$\hat{\sigma}_\varepsilon$	0.0021	0.0028	0.0016	0.0012

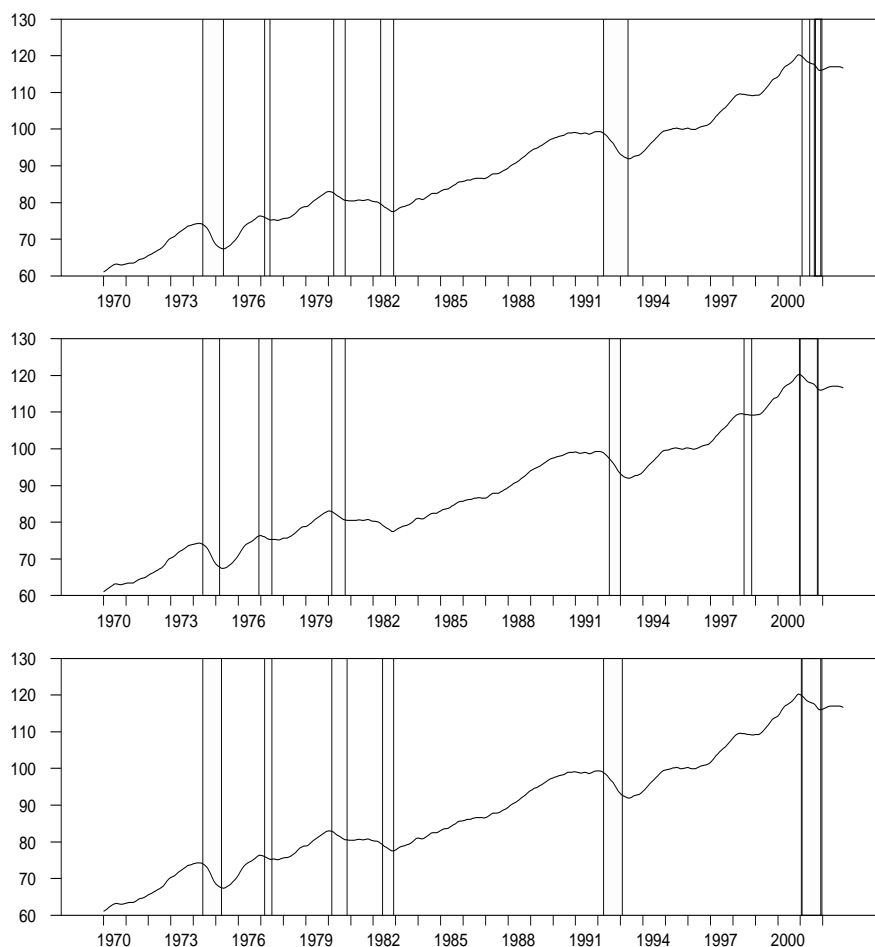Table 3: Estimates and standard errors for model described in 4.2.3.

Figure 3: Industrial recession dates estimated by the 2-regime SETAR with lagged variable as transition variable (top graph), by the 2-regime SETAR with differenced lagged variable as transition variable (middle graph) and by the 4-regime SETAR (bottom graph).

The two thresholds are estimated by using a double loop, but the delays of the model are fixed *a priori* according the two previous estimated models. Both estimated thresholds are negative but very close to zero. The first regime has an empirical unconditional probability of 0.15 and should be considered at a first sight as a period of recession because the estimated recession dates match the reference recession dates. However, the second regime is also meaningful. Indeed, this second regime possesses an unconditional probability of 0.02: only 7 observations over 385 belong to this state. This is the reason why estimates and their standard errors in this regime should be taken with caution. Although the frequency of this second regime is very low, this regime is persistent and appears in clusters. In fact, this regime is very interesting because it corresponds to the end of a recession

14

phase when the economy is accelerating again. This regime was detected twice: at the end of the 1974-75 recession and at the end of the 1992-93 recession. Thus, the sum of regime 1 and regime 2 corresponds to the industrial recession phase. The third regime can be considered as a slowdown of the industrial production, that is the industry is below its trend growth rate without being in recession. Lastly, when the series is in the high regime, we can deduce that the industrial growth rate is over its trend growth rate. Actually, regime 3 and regime 4 correspond to the high phase of the industrial business cycle. It appears that only three regimes would be sufficient to describe the industrial business cycle. However, we decide to keep four states because it gives a deeper understanding of the industrial business cycle features. As regards the dating results, the model provides almost the same results than the first model, the last recession period being not cut into two parts (see figure 3, bottom graph, and table 4). However, this model presents some non-persistent signals of recession.

After this whole sample analysis, we retain the third SETAR model with four regimes for the dynamic analysis, because it provides the more accurate description of the industrial business cycle.

4.3 Dynamic analysis

To be useful for short-term economic analysis, an economic indicator requires at least two qualities: it must be reliable and must provide a readable signal as soon as possible. Thus, there is a well known trade-off between advance and reliability for the economic indicators. By using the previous

	Reference	Model 1	Model 2	Model 3
Peak	m4 1974	m6 1974	m6 1974	m6 1974
Trough	m5 1975	m5 1975	m3 1975	m6 1975
Peak	-	m3 1977	m12 1976	m3 1977
Trough	-	m6 1977	m7 1977	m7 1977
Peak	m2 1980	m4 1980	m3 1980	m3 1980
Trough	m1 1981	m10 1980	m10 1980	m11 1980
Peak	m10 1981	m5 1982	m6 1982	m6 1982
Trough	m12 1982	m12 1982	m8 1982	m12 1982
Peak	m1 1992	m4 1992	m7 1992	m4 1992
Trough	m5 1993	m5 1993	m1 1993	m6 1993
Peak	-	-	m7 1998	-
Trough	-	-	m11 1998	-
Peak	m12 2000	m2 2001	m1 2001	m2 2001
Trough	m12 2001	m12 2001	m10 2001	m12 2001

Table 4: Reference and estimated dating chronologies stemming from the 3 considered SETAR models.

4-regime SETAR model, we assess if it is possible to have a clear and timely signal for the turning points of the industrial business cycle in a dynamic analysis.

In this part, we consider the previous IPI series from January 1970 to December 1999, and we add progressively a monthly data until December 2002. For each step, we re-estimate the model and we classify the observations into one of the four regimes. Thus, by using the conclusions of the whole-sample analysis, if the observations fall into regime 1 or regime 2, we can conclude that the industry is in a recession phase. We are aware that a *true* real-time analysis should be done by using historically released data (see for instance Chauvet and Piger, 2003) in order to take the revisions and the edge-effects of the statistical treatments of the raw data into account. However, such series are very difficult to find in economic data bases.

The dynamically estimated recession period is presented in figure 4. We observe this period matches with the 2001 recession period estimated in the whole-sample analysis. This fact points out the stability of the model. Indeed, we detect a peak in the business cycle in February 2001 and a trough in December 2001. However, it must be noted that a false signal of a change in regime is emitted in August 2001 but lasts only one month. Knowing that a signal must be persistent to be reliable, we have to propose an *ad hoc* real-time decision rule. Thus, it is advocated to wait at least two months before sending a signal of a change in regime. We also note that the exit of the recession is very fast, because the observations go directly from regime 1 in December 2001 to regime 4 in January 2002. Moreover, we observe that the December 2002 observation falls into regime 3.
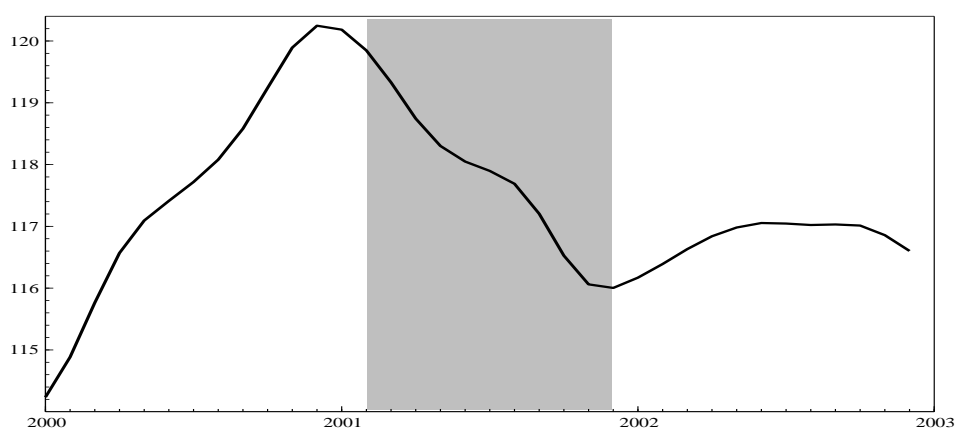

Figure 4: Euro12 IPI and the dinamically estimated recession period (shaded area), from January 2000 to December 2002.

16

Conclusion

This paper is an exploratory analysis of the ability of SETAR models to reproduce the business cycle stylised facts. The results are promising. It appears that these non-linear models allow to identify the turning points of the Euro-zone industrial business cycle and can thus be useful for real-time detection. However, a *true* real-time analysis should be extended by using historically released data, as used in the recent paper of Chauvet and Piger (2003) as regards the US GDP and employment. This *true* real-time analysis would also allow to check the robustness of the model over time. Especially, the stability of the estimated thresholds could be interesting to investigate. If the unstability is effective, an innovative time-varying SETAR could be introduced. Unfortunately, such data are not systematically stored in data bases and are therefore very difficult to get, especially as regards the Euro-zone. As another example of application, consumer and business surveys seem to be good candidates for real-time analysis through SETAR models because they are timely released and are not generally revised. Last, it could be interesting to compare the detections made by these threshold models with analogous detections made by other models, such as Markov-Switching or logit models, based on the same data set, and to compare their goodness of fit through a simulation study.

Acknowledgments

References

[1] Akaike, H. (1974), "A New Look of Statistical Model Identification", *IEEE Transactions on Automatic Control*, 19, 716-722.

[2] Anas, J., Billio, M., Ferrara, L., and LoDuca, M. (2003), "A Turning Point Chronology for the Euro-zone Classical and Growth Cycles", Eurostat Working Paper, presented at the 4th Eurostat Colloquium on Modern Tools for Business Cycle Analysis, Luxembourg, October 2003.

[3] Anas, J., and Ferrara, L. (2004a), "A Comparative Assessment of Parametric and Non-Parametric Turning Points Methods: The Case of the Euro-zone Economy", in *Monographs of Official Statistics : Statistical Methods and Business Cycle Analysis of the Euro zone*, G.L. Mazzi and G. Savio (eds.), Eurostat, 86-121.

[4] Anas, J., and Ferrara, L. (2004b), "Detecting cyclical turning points: The ABCD approach and two probabilistic indicators", *Journal of Business Cycle Measurement and Analysis*, 1, 2, 1-36.

[5] Artis, M., Krolzig, H.M., and Toro, J. (2003), "The European Business Cycle", *Oxford Economic Papers*, 56, 1-44.

[6] Bry, G., and Boschan, C. (1971), *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, NBER, New York.

[7] Chan, K.S. (1993), "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model", *Annals of Statistics*, 21, 520-533.

[8] Chen, R. (1995), "Threshold Variable Selection in Open-Loop Threshold Autoregressive Models", *Journal of Time Series Analysis*, 16, 461-481.

[9] Chauvet, M., and Piger, J.M. (2003), "Identifying Business Cycle Turning Points in Real Time", *Review of the Federal Reserve Bank of St. Louis*, March/April, 47-61.

[10] Clements, M.P., and Smith, J. (2001), "Evaluating Forecasts from SETAR Models of Exchange Rates", *Journal of International Money and Finance*, 20, 133-148.

[11] Clements, M.P., and Krolzig, H.M. (2003), "Business Cycle Asymmetries: Characterization and Testing Based on Markov-Switching Autoregressions", *Journal of Business and Economic Statistics*, 21, 1, 196-211.

[12] Coakley, J., Fuertes, A.M., and Pérez, M.T. (2003), "Numerical Issues in Threshold Autoregressive Modeling of Time Series", *Journal of Economic Dynamics and Control*, forthcoming.

[13] van Dijk, D., Franses, P.H., and Paap, R. (2002), "A Nonlinear Long Memory Model to US Unemployment", *Journal of Econometrics*, 110, 135-165.

[14] van Dijk, D., Terasvirta, T., and Franses, P.H. (2002), "Smooth Transition Autoregressive Models - A Survey of Recent Developments", *Econometric Reviews*, 21, 1-47.

[15] Dufrénot, G., Guégan, D., and Peguin-Feissolle, A. (2005 a), "Long memory dynamics in a SETAR model: Applications to stock markets", *Journal of International Financial markets, Institutions and Money*, 15, 5.

[16] Dufrénot, G., Guégan, D., and Peguin-Feissolle, A. (2005 b), "Modelling squares returns using a SETAR model with long memory dynamics", *Economics Letters*, 86, 237 - 243.

[17] Estrella, A., and Mishkin, F.S. (1998), "Predicting US Recessions: Financial Variables as Leading Indicators", *Review of Economics and Statistics*, 80, 45-61.

[18] Ferrara, L. (2003), "A Three-Regime Real-Time Indicator for the US Economy", *Economics Letters*, 81, 3, 373 - 378.

[19] Franses, P.H. and D. van Dijk (2000), *Non-Linear Time Series Models in Empirical Finance*, Cambridge University Press, Cambridge.

[20] Gonzalo, J., and Pitarakis, J.Y. (2002), "Estimation and Model Selection Based Inference in Single and Multiple Threshold Models", *Journal of Econometrics*, 110, 319 - 352.

[21] de Goojier, J.G., and de Bruin, P.T. (1999), "On Forecasting SETAR Processes", *Statistics and Probability Letters*, 37, 7 - 14.

[22] Guégan, D. (2003), "Point de Vue Personnel sur le Problème de Contagion en Economie et l'Intéraction entre Cycle Réel et Cycle Financier", Note de Recherche MORA-IDHE 06-2003, Ecole Normale Supérieure, Cachan, France.

[23] Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle", *Econometrica*, 57, 357-384.

[24] Hansen, B.E. (1997), "Inference in TAR Models", *Studies in Nonlinear Dynamics and Econometrics*, 2, 1-14.

[25] Hansen, B.E. (2000), "Sample splitting and threshold estimation", *Econometrica*, 68, 575-603.

[26] Harding, D., and Pagan, A. (2001), "A comparison of two business cycle dating methods", unpublished manuscript, University of Melbourne.

[27] Jones, D.A. (1978), "Nonlinear Autoregressive Processes", *Proceedings of the Royal Society*, A, 360, 71-95.

[28] Krolzig, H.M. (2001), "Markov-Switching Procedures for Dating the Euro-zone Business Cycle", *Quarterly Journal of Economic Research*, 3, 339-351.

[29] Krolzig, H.M. (2004), "Constructing Turning Point Chronologies with Markov-Switching Vector Autoregressive Models: the Euro-zone Business Cycle", in *Monographs of Official Statistics : Statistical Methods and Business Cycle Analysis of the Euro zone*, G.L. Mazzi and G. Savio (eds.), Eurostat, 147-190.

[30] Krolzig, H.M., and Toro, J. (2001), "Classical and Modern Business Cycle Measurement: The European Case", Discussion Paper in Economics 60, University of Oxford.

[31] Lahiri, K., and Wang, J.G. (1994), "Predicting Cyclical Turning Points with Leading Index in a Markov-Switching Model", *Journal of Forecasting*, 13, 245-263.

[32] Lahiri, K., Yao, W., and Young, P. (2004), "Cycles in the Transportation Sector and the Aggregate Economy", *Transportation Research Record*, National Academies, 103-111.

[33] Maravall, A. and Planas, C.(1999), "Estimation Error and the Specification of Unobserved Component Models", *Journal of Econometrics*, 92, 325-353.

[34] Pemberton, J. (1985), "Contributions to the Theory of Nonlinear Time Series Models", unpublished Ph.D. Thesis, University of Manchester.

[35] Pfann, G.A., Schotman, P.C., and Tchernig, R. (1996), "Nonlinear Interest Rate Dynamics and Implication for the Term Structure", *Journal of Econometrics*, 74, 149 - 176.

[36] Potter, S.M. (1995), "A Nonlinear Approach to US GNP", *Journal of Applied Econometrics*, 10, 109-125.

[37] Potter, S.M. (1999), "Nonlinear Time Series Modelling: An Introduction", *Journal of Economic Surveys*, 13, 505-528.

[38] Proietti, T. (1998), "Characterizing Asymmetries in Business Cycles Using Smooth-Transition Structural Times Series Models", *Studies in Nonlinear Dynamics and Econometrics*, 3, 141-156.

[39] Sichel, D.E. (1994), "Inventories and the three phases of the business cycles", *Journal of Business and Economic Statistics*, 12, 269 - 277.

[40] So, M.K.P., and Chen, C.W.S. (2003), "Subset Threshold Autoregression", *Journal of Forecasting*, 22, 49-66.

[41] Terasvirta, T., and Anderson, H.M. (1992), "Characterising Nonlinearities in Business Cycles using Smooth Transition Autoregressive Models", *Journal of Applied Econometrics*, 7, S119 - S136.

[42] Tiao, G.C., and Tsay, R.S. (1994), "Some Advances in Non-Linear and Adaptive Modelling in Time-Series", *Journal of Forecasting*, 13, 109 - 131.

[43] Tong, H., and Lim, K.S. (1980), "Threshold Autoregression, Limit Cycles and Cyclical Data", *Journal of the Royal Statistical Society*, B, 42, 245-292.

[44] Tong, H. (1990), *Non-linear Time Series: A Dynamical Approach*, Oxford Scientific Publications, Oxford.

[45] Tsay, R.S. (1989), "Testing and Modeling Threshold Autoregressive Processes", *Journal of the American Statistical Association*, 84, 231-240.

[46] Zakoian, J.M. (1994), "Threshold Heteroskedastic Models", *Journal of Economic Dynamics and Control*, 18, 931-955.