BANCA D'ITALIA

Temi di discussione

del Servizio Studi

Nonlinear VAR: Some Theory and an Application to US GNP and Unemployment

by Filippo Altissimo and Giovanni Luca Violante

Number 338 - October 1998

NONLINEAR VAR: SOME THEORY AND AN APPLICATION TO US GNP AND UNEMPLOYMENT

by Filippo Altissimo* and Giovanni Luca Violante**

Abstract

A generalization of the endogenous threshold model is developed by extending this class to a multivariate framework and to cases where the feedback acts at multiple lags. The feedback is specified, following Beaudry and Koop, by a variable which measures the depth of recessions. We give conditions for the ergodicity of the model and prove strong consistency of the maximum likelihood estimator, although the objective function is discontinuous in the threshold parameter. The model is applied to a bivariate VAR of output growth and changes in the unemployment rate for the US economy. The nonlinearity is found to be statistically significant only in the unemployment equation and it transmits to GNP through the cross-correlation between the series. We also find that, owing to the nonlinear structure, shocks hitting the economy in downturns have lower persistence than those occurring during expansions. Since this dampening effect is stronger for negative than for positive shocks, the feedback from recessions is found to contribute positively to the long-run growth of the economy and we estimate this contribution to be about 1/6 of the total growth over the sample period. We interpret this result as an empirical validation of those economic theories that model recessions as cleansing times. Finally, we suggest that the state-dependence in persistence is a possible key to interpret the divergence in the measures of persistence existing in the literature.

---

* Banca d'Italia, Servizio Studi.

** University College of London, Department of Economics.

Contents

1. Introduction[1]

Starting from the work by Neftçi (1984), a substantial interest has arisen in nonlinear and asymmetric features of economic time series.[2] While the evidence on GNP and other production series is rather mixed, generally nonlinearity seems to be present in the unemployment rate.

One limitation of this literature is that the analysis is exclusively univariate. In a multivariate framework the shocks interact, producing much richer dynamics, so one might conjecture that extending the analysis to a multivariate *linear* framework is sufficient to capture what in the univariate world needs a nonlinear structure to be well approximated. Also, the nonlinearity might be present only in some series and transmit to other series through their cross-correlation, so the multivariate system would help to identify where the nonlinearity is originally active. This could represent a useful piece of information for theoretical model building and for a more complete assessment of the relevance of nonlinear features in economic time series.

Another large body of literature, following the seminal work of Nelson and Plosser (1982), has grown around the issue of persistence of shocks in macroeconomic time series. A common finding of this literature is that GNP shows some persistence, but the estimates provided by the different authors vary across a wide range.[3] In a nonlinear world the nature of the response of the economy to a given impulse depends on the sign and the size of the shock hitting the economy and on the history of the system. Hence, potentially a much richer analysis of persistence can be performed in order to assess whether the restriction of linearity is truly binding and therefore hides qualitatively interesting properties of the data.

In this paper we explore the interaction between nonlinearities and persistence in a multivariate environment and we show that allowing for the presence of nonlinear features

---

[1] We are grateful to M. Hashem Pesaran for helpful discussion and his encouragement. We thank Valentina Corradi, Frank Diebold, Lutz Killian, Lee Ohanian, Simon Potter and two anonymous referees for their comments and also partecipants in workshops at IGIER, Università ”Ca’ Foscari” and the University of Pennsylvania. We thank Giancarlo Marra for technical assistence. The usual disclaimer applies.

[2] See DeLong and Summers (1986), Falk (1986), Hamilton (1989), Rothman (1991), Brunner (1992), Potter (1995), Sichel (1993) among others.

[3] See Watson (1986), Campbell and Mankiw (1987), Cochrane (1988) and Evans (1989) among others.

in the series can shed light on several issues: first, the spurious presence of nonlinearities due to cross-correlation between the series; second, the characterization of the propagation mechanism of impulses and the measurement of persistence; third, the potential constraints imposed by the linear structure on the data, which prevents us from uncovering interesting dynamics and testing theoretical models which are inherently nonlinear.

To accomplish this task we consider a reduced-form nonlinear bivariate VAR model of changes in US GNP and unemployment rate. The nonlinear structure we introduce in the VAR is an endogenous delay threshold, first formalized by Pesaran and Potter (1994) (hereafter PP). That model is a special case of the more general class of threshold autoregressive models (TAR).[4] TAR models specify different regime dynamics according to a given threshold rule and restrict the process to be of the autoregressive form within each regime.

Even within the relatively restricted class of endogenous delay threshold models, there are a number of ways to define the threshold rule. We followed the specification introduced by Beaudry and Koop (1993) where the regimes are upturns and downturns and the state variable is a measure of the current depth of recession. We believe that this type of nonlinearity is particularly attractive since, in a system with persistence, it captures the link between aggregate fluctuations and long-run growth, a fact that has been stressed by the recent macroeconomic literature. In the past few years some authors (for a survey, see Hall, 1991; Caballero and Hammour, 1994; Aghion and Saint-Paul, 1993) have emphasized the idea that recessions are periods in which a *cleansing process* is activated in the economy. The least productive firms exit the market, whereas more innovative ones survive; moreover, the low opportunity cost of investing resources in production stimulates firms to use resources in order to reduce organizational inefficiencies and restructure their plant. The result of this process is that recessions might have a positive feedback on the future productivity growth of the economy. Our model allows us to examine this issue in a very intuitive way and provides a measure of how much the feedback originating from downturns has contributed to the growth of the economy. Note that linear models would not be able to uncover this interaction between recessions and growth, as they would treat symmetrically both phases of the business cycle.

---

[4] The theory of TAR models has a rather long history, starting with the work of Tong and Lim (1980). An extended survey of the literature on TAR models can be found in Tong (1990).

The choice of the variables used in the VAR is guided by three considerations. First, as mentioned above, these are time series for which some evidence of significant nonlinearity exists in the univariate literature; hence it is of interest to see whether some spurious nonlinearity induced by correlation arises in our bivariate framework. Second, GNP and unemployment are a suitable set of variables to investigate the transmission mechanism of the shocks between the labor market and the goods market, crucial in many policy analysis. Third, linear VARs for this pair of variables have been already estimated in the past.[5] Our methodology is original in respect to this previous literature and reveals new properties of the interaction between the two series that may put the existing results in a different light.

The rest of the paper is organized as follows. The model is introduced in Section 2. Section 3 presents the estimation method and a discussion of ergodicity of the model and consistency of the estimator. The model selection, results of the estimation and testing procedures are contained in Section 4. In Section 5 we characterize and measure the long-run effect of recessions on the economy and in Section 6 this analysis is deepened by looking at the impulse response functions and measures of persistence. Section 7 concludes.

2. The endogenous delay threshold model

2.1 *The threshold principle and the endogenous delay threshold VAR*

The traditional approach to the analysis of a stochastic nonlinear dynamic system is to decompose it into piecewise linear subsystems. This is essentially what is called the *threshold principle*, which means splitting the dynamics into different regimes through a threshold on certain key variables in order to define a local linear approximation of the stochastic process in every regime. The general class of linear threshold autoregressive models (TAR) is based on this principle and uses autoregressive specifications for the local approximation. The canonical form of the TAR, as proposed by Tong and Lim (1980), is

(1) $$x_t = \alpha^{(J_t)} + \Phi^{(J_t)} x_{t-1} + \sigma^{(J_t)} \varepsilon_t,$$

[5] See Evans (1989) and Blanchard and Quah (1989).

where ε_t is a white noise with unit standard deviation and the indicator $\{J_t\}$, $J_t \in [1, ..., M]$, is a stochastic process indicating the dynamic regime which is in effect. The specification (1) is quite flexible and can be specialized both to the Markov switching model introduced in Hamilton (1989) and to the self-exciting TAR (SETAR),[6] according to the way in which $\{J_t\}$ is parametrized. A new specialization of the TAR class is the endogenous delay threshold (EDTAR) model proposed by PP. The EDTAR model is a framework for modeling feedback processes from the past realizations of the system to its current dynamics in a more articulated way than the SETAR. As first proposed the EDTAR model was univariate and it did not allow for feedback effects operating at more than one lag. We will extend that framework to a multivariate setting with multiple lags in the feedback process.[7]

Consider the time series $\{X_t\}$, where X_t is a $(k \times 1)$ vector, whose dynamics are modeled with M feedback channels from past values to the current realization. These feedbacks can be characterized constructing M feedback index functions, F_{mt}, $m \in [1, ..., M]$, defined as

(2) $$F_{mt} = 1(b_{ms}(X_t, ..., X_{t-s}) \in A_m),$$

where $b_{ms} : \Re^{s+1} \times \Re^k \to \Re$, A_m is an interval on the real line which characterizes the threshold rule for the $m-th$ regime, $1(\cdot)$ is the indicator function and τ_m is the maximum memory of the $m-th$ feedback and it is finite. In the original specification of the $EDTAR$ model, τ is not restricted to be finite. The finiteness of τ greatly simplifies the proof of consistency of the maximum likelihood estimator and the requirements for the ergodicity. The feedback index F_{mt} is activated when a particular transformation b_{ms} of some lags of X_t satisfies the conditions implicitly represented by the set A_m. F_{mt} is only an indicator of the activation of the feedback and it is used to switch on and off a feedback variable Y_{mt} which enters the dynamics of the process. The feedback variable is recursively built as

(3) $$Y_{mt} = 1(F_{mt} > 0)\left[\theta Y_{m,t-1} + G_m(X_t, ..., X_{t-l})\right],$$

6 An extensive discussion on the properties of SETAR models and their estimation procedure are described in Tong (1990) and an application to the dynamics of the growth rate of US GNP is given by Potter (1995).

7 In the rest of this section we will follow closely the notation in PP.

where $G_m : \Re^{l+1} \times \Re^k \to \Re$, and θ is a scalar. This representation is very flexible, as the set A_m and the functions b_{ms}, f_m and G_m can be specified according to the nature of the feedback process modeled.

We can now specify the dynamics of X_t as

$$\Phi(L)X_t = \alpha + \sum_{m=1}^{M} \Theta_m(L)Y_{m,t-1} + H_t(\cdot)u_t, \tag{4}$$

with $u_t \sim N(0, I_k)$, and $H_t(\cdot)$ is a general functional form for the conditional variance.

We call the specification in (4) $EDTVAR(p, q_1, ..., q_M)$ where p is the lag order of Φ and $(q_m - 1)$ is the lag order of $\Theta_m(L)$, $m = 1, 2, ..., M$. Equation (4), together with equations (2) and (3) and a specification for the conditional variance, defines the general structure of the model.

To gain some intuition on the formalized model, consider as an example the case in which we want to model an autoregressive process with slope β if in the previous period the process assumed a positive value and with slope α in the other case. In this case the process can be modeled by constructing a single feedback index and a single feedback variable, $i.e.$ $M = 1$. Given that the memory of the feedback is one lag, $i.e.$ $\tau = 1$, then the implied feedback index is

$$F_t = 1(X_t \in \Re^+).$$

Accordingly, the feedback variable is

$$Y_t = 1(F_t = 1)\left[(\beta - \alpha)X_t\right]$$

where θ is identically zero and the function $G(X_t) = (\alpha - \beta)X_t$ and the dynamics of the process can be written as

$$\begin{aligned} X_t &= \alpha X_{t-1} + Y_{t-1} + \varepsilon_t \\ &= \alpha X_{t-1} + 1(F_{t-1} = 1)\left[(\beta - \alpha)X_{t-1}\right] + \varepsilon_t. \end{aligned}$$

Note that this is also a different way of representing a SETAR(1) model.

Finally it is interesting to note how this kind of approach to nonlinear time series has a natural interpretation in terms of a state-space model. The Y_{mt} variables are the states of the system and equations (3) and (4) are, respectively, the state equation and the measurement equation of the model, using the terminology of that literature. Notice, however, that here the state is entirely a function of the observables, while many models that are usefully cast in state-space form contain a latent state variable.

2.2 *A VAR of output and unemployment with feedback from the depth of recessions*

Within the framework of the $EDTVAR$ described above, we focus on a bidimensional model of changes in log of real output and in the unemployment rate,[8] hence $X_t \equiv (\log GNP, U)$. We introduce a nonlinear feedback process in the dynamics of the variables ΔX_t through a unique feedback variable intended to capture the effect of recessions on the economy. Our feedback index F_t is defined as

$$F_t(r) = \sum_{i=0}^{\tau-1}\left[1 - \prod_{j=0}^{i} 1(X_{1t} - X_{1,t-\tau+j} > r)\right], \tag{5}$$

hence this index is zero when the current realization of GNP is higher than each of the past τ realizations augmented by r and is equal to i when the previous peak of X_{1t} occurred at $(t-i)$. The implied feedback variable, CDR, is defined as

$$CDR_t(r) = \begin{cases} 1(F_t > 0)(\Delta X_{1t} - r) & if\ F_{t-1} = 0 \\ 1(F_t > 0)(CDR_{t-1} + \Delta X_{1t}) & otherwise. \end{cases} \tag{6}$$

We can define CDR_t in more compact notation in two alternative ways:

$$\begin{aligned} &CDR_t(r) = X_{1t} - \max(X_{1t}, X_{1,t-1} + r, ..., X_{1,t-\tau} + r), \text{ or} \\ &CDR_t(r) = \min(0, \Delta X_{1t} - r, \Delta X_{1t} + \Delta X_{1,t-1} - r, ..., \Delta X_{1t} + ... + \Delta X_{1,t-\tau+1} - r) \end{aligned} \tag{7}$$

where τ is a finite number and the $\max$ ($\min$, respectively) is taken over the $\tau+1$ terms inside the brackets. The variable CDR measures the current depth of the recession of the

[8] The series are Citibase quarterly data of U.S. GNP at 1982 prices (GNP82) and unemployment rate (LHURR) monthly averaged, from 1952.1 to 1990.4.

economy. It is clear that CDR will be zero as long as output grows at least at a rate r and –once it is switched on through F_t– it will stay activated as long as output is lower than its previous maximum increased by r, that is to say until the economy has recovered the gap of the recession.

The threshold parameter r, which determines the activation of CDR, will be endogenously estimated. Hence our measure of recession is endogenous as well since it depends on the estimate of r;[9] note that our specification is more general than that of Beaudry and Koop (1994) since they exogenously fix r at zero. In the rest of the paper, we will use the terminology of PP and refer to the situation in which the dynamics of the system are not under the effect of any feedback, i.e. $\sum_{i=1}^{q} CDR_{t-i} = 0$, as *corridor* regime, while we will speak of *floor* regime in the case in which at least one of the different q lags of the CDR variable is not null.

We allow the feedback effect to operate also at lags higher than one, up to q, so the dynamics of ΔX_t follow the specification:

(8) $$\Phi(L)\Delta X_t = \alpha + \Theta(L)CDR_{t-1}(r) + H_t^{\frac{1}{2}} u_t \quad \text{with } u_t \overset{iid}{\sim} N(0, I_2),$$

where p is the order of Φ and $(q-1)$ the order of Θ. The conditional variance H_t is modeled as

(9) $$H_t = 1\left(\sum_{i=1}^{q} CDR_{t-i} = 0\right)(\Omega_c - \Omega_f) + \Omega_f,$$

where Ω_c and Ω_f are the conditional covariance matrices when the system is respectively in the corridor regime or in the floor regime. This structure of the conditional variance is also known in the literature as qualitative threshold autoregressive conditional heteroskedasticity (QTARCH).[10] Note that since CDR depends on the threshold parameter r, the latter enters both the conditional mean and the conditional variance.

---

[9] In the next section we contrast our definition of recession as implied by the estimation with the NBER business cycle dates.

[10] See Gourieroux and Monfort (1992) for the theory on QTARCH and French and Sichel (1993) for evidence on asymmetry in the conditional variance of real GNP in different phases of the business cycle.

In conclusion, our model is described by equations (**??**)-(9). To gain some intuition about the structure of the model, it is helpful to observe from (**??**) that CDR can be written as a constant plus the sum of the changes in GNP from the date of the previous peak to the current value, so that by substituting (**??**) into (8) we obtain a linear VAR with time variant parameters and time variant number of lags determined by an endogenous deterministic threshold rule. More specifically, when the economy is under the effect of the feedback, at each period the linear structure of the corridor regime is modified with additional lags and different parameter values.

3. Ergodicity of the model and asymptotics of the ML estimator

The normality assumption on the error term allows us to write the conditional log-likelihood function of the model as:

$$\ell(\delta) \equiv \sum_{t=1}^{T} \ell_t(\delta|\mathcal{F}_t) = -\frac{1}{2}\sum_{t=1}^{T} \ln|H_t| + \tag{10}$$
$$-\tfrac{1}{2}\textstyle\sum_{t=1}^{T}(\Phi(L)\Delta X_t - \alpha - \Theta(L)CDR_{t-1})H_t^{-1}(\Phi(L)\Delta X_t - \alpha - \Theta(L)CDR_{t-1})',$$

where $\delta = \{\alpha, \Phi, \Theta, \Omega_c, \Omega_f, r\}$ is the set of parameters to be estimated and $\mathcal{F}_t$ is the information set at time t.

The numerical maximization of this log-likelihood is not a standard problem, given that the function is discontinuous with respect to the parameter r, the threshold coefficient. It is clear from simple inspection of (9) and (10) that what generates the discontinuity in the likelihood is the conditional variance term, which changes discretely between regimes, while the conditional mean changes smoothly.[11] The estimation method applied here is the two-step procedure suggested by Tong (1990) for the case of $SETAR$ models and also used by PP. In the first step, we generate a finite grid of points over the domain of the threshold parameter r and at each point on the grid we estimate the model by maximizing the likelihood function conditional on a given value of r using a standard hill-climbing algorithm. Then we choose the value of r on the grid for which the likelihood in (10) attains its global maximum. We

[11] The $EDTVAR$ model with homoskedastic disturbances was tested against the one with the $QTARCH$ conditional variance, but we rejected the first.

have worked with a grid of 400 point in the interval $(-.013, +.013)$ [12] and at every point we allow for 100 iterations over the likelihood function with a convergence criteria of 10^{-5} for each parameter.

This procedure enables us to find the maximum of the sample log-likelihood function, but it is not sufficient to ensure that the estimated parameter vector $\hat{\delta}_{MLE}$ is consistent for the true value of the underlying DGP. There are two main problems in proving consistency of the ML estimator. The first is to characterize the conditions under which the model is ergodic. The second is to control for the fact that the number of discontinuity in the log-likelihood function is "small" and to show that the unconditional expectation of the log-likelihood function is itself continuous.

3.1 *Ergodicity*

The presence of the CDR variable induces a nonlinearity in the conditional mean of the process. In a nonlinear environment assessing the ergodicity of a model is a complex matter.

Given some regularity conditions on the disturbances, to prove ergodicity it is sufficient to show that a certain "drifting condition" is satisfied. The latter ensures that, for any initial value, the process is expected to move towards the "centre" of the space in a finite number of steps. Tweedie (1975) stated this condition in its simplest form and Tong (1990) contains an interesting interpretation of Tweedie's condition through the concept of Lyapunov function, which establishes a link between the stability of deterministic dynamical systems and the ergodicity of stochastic systems. In particular, we will follow the approach of Tjøstheim (1990) in using a generalization of Tweedie's original result to characterize the ergodicity of our nonlinear multivariate system.

To characterize this drifting condition the assumption of a finite memory in the feedback process, i.e. a finite τ, is critical. In fact, given a finite τ it is possible to show that the model has a Markovian representation and therefore to apply the well-established theory on stability of Markov chains[13] to our specific case. We exploit the finite τ in order to write the model as a multivariate $SETAR$ with a large number of states and constraints across the parameters

---

[12] This is the largest interval for r_f such that both Ω_c and Ω_f are nonsingular.

[13] See the seminal book by Nummelin (1984).

characterizing the linear dynamics of each different state. In this way we obtain a Markovian representation of the process and we can therefore apply an argument based on the "$N-$step ahead drifting condition" along the lines of Tjøstheim (1990). The formal proof of geometric ergodicity of the $EDTVAR$ is given in Proposition 1 of Appendix 1. Also, note that strict stationarity of the process is obtained from ergodicity, once we assume that the chain is started with initial distribution equal to the invariant distribution.

The following example provides some intuition for this stability result in a simple case. Consider a univariate version of the model in (8) with $\tau = 1$, $p = 2$, $q = 2$ and $r = 0$:

$$\begin{aligned}\Delta X_t = \; & \phi_1 \Delta X_{t-1} + \phi_2 \Delta X_{t-2} + \theta_1 \min(0, \Delta X_{t-1}) \\ & + \theta_2 \min(0, \Delta X_{t-2}) + u_t.\end{aligned}$$

Define $Z_t = \{\Delta X_t, \Delta X_{t-1}\}$ and F_{t-i}, $i = 1, 2$ as the feedback indexes in (5). It follows from that definition that $F_{t-i} = 1$ when $Z_{1,t-i} \leq 0$ and $F_{t-i} = 0$ otherwise. The potential activation of the two lags of CDR_t generates 4 possible states that create a partition $\wp \equiv \{P_0, P_1, P_2, P_3\}$ on the space of Z_{t-1}:

$$\begin{aligned}P_0 &\equiv \{Z_{t-1} : F_{t-1} = 0, F_{t-2} = 0\} \\ P_1 &\equiv \{Z_{t-1} : F_{t-1} = 1, F_{t-2} = 0\} \\ P_2 &\equiv \{Z_{t-1} : F_{t-1} = 0, F_{t-2} = 1\} \\ P_3 &\equiv \{Z_{t-1} : F_{t-1} = 1, F_{t-2} = 1\}\end{aligned}$$

This partition is used to build the following Markovian representation of the model:

$$Z_t = \sum_{k=0}^{3} A_k 1(Z_{t-1} \in P_k) Z_{t-1} + u_t \quad .$$

The (2×2) matrices A_k associated to the four states are:

$$\begin{aligned} A_0 &= \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix}, \qquad A_1 = \begin{bmatrix} \phi_1 + \theta_1 & \phi_2 \\ 1 & 0 \end{bmatrix}, \\ A_2 &= \begin{bmatrix} \phi_1 & \phi_2 + \theta_2 \\ 1 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} \phi_1 + \theta_1 & \phi_2 + \theta_2 \\ 1 & 0 \end{bmatrix}. \end{aligned}$$

The matrix A_0 describes the dynamics in the corridor regime and the other three matrices describe how the corridor regime dynamics are modified through the activation of CDR.

The key condition for ergodicity is that, independently of the initial state, the process moves toward the centre of the space in N (finite) periods. This condition requires some restrictions on the products of the N matrices A which define the dynamics between Z_t and Z_{t+N} (see $Assumption\ A1$, Appendix 1).

3.2 *Consistency and asymptotic distribution*

Having characterized the conditions for the ergodicity of the model, we turn now to the discussion of the consistency of the ML estimator. The key step is to show that the discontinuity in the sample likelihood triggered by the conditional heteroskedasticity of the error term does not invalidate the consistency of the estimator. In Appendix 2, a careful characterization of the properties of the objective function ℓ_t is given and it is shown that the set of realizations of the process that generate discontinuities has measure zero in the space of histories. This fact, with the addition of a set of standard regularity conditions on $\Delta X_t,$ and under the above condition for ergodicity, leads to a proof of strong consistency of the estimator (see Proposition 2, Appendix 2) which is constructed following Andrews (1987).

Finally, assessing the asymptotic distribution of the estimator requires deriving the limiting distribution of the threshold parameters, which is likely to be non standard owing to the discontinuity of the likelihood function. We do not pursue this strategy, but rather we observe that conditional on r, the asymptotic normality follows from standard asymptotic theory. Hence, if the speed of convergence of r is sufficiently quick, then the threshold value can be treated as known in performing inference on the autoregressive parameters and the standard asymptotic theory will hold. This conjecture is based on the result in Chan (1988) where the superconsistency of r is proved for a two-regime $SETAR$ and it is supported by some of our Monte Carlo experiments.

4. Results of estimation and testing

The previous literature on VARs of output and unemployment, such as Blanchard and Quah (1989) and Evans (1989), considers the unemployment rate to be stationary, although it recognizes that the evidence on this point is not unequivocal. In our model the presence of the nonlinearity invalidates the standard asymptotic theory of the unit root test; therefore a different route should be pursued in addressing the issue of stationarity for the series.[14] A necessary condition for the consistency of the estimator is the ergodicity of the system and a sufficient condition for ergodicity is ASSUMPTION A1. When we picked the specification with unemployment in levels, the maximum eigenvalue of the matrix driving the dynamics in the corridor regime was found to be outside the unit circle which invalidates ASSUMPTION A1, whereas when we estimated the model in first differences of both series the condition was satisfied.

In addition to this point, one should consider that working with a VAR with highly persistent series may induce a large small–sample bias of the estimates.[15] These two arguments have led us to the choice of a specification in first differences and in the rest of the paper we will provide some additional evidence on the validity of this assumption, in particular through the long-run behavior of the generalized impulse response functions.

4.1 *Model selection*

The next step in the estimation procedure is the choice of the lag order of the polynomials Φ and Θ. Note that we need to select the lag order also for the CDR variable, allowing in principle the presence of delays in the feedback effect from the recession to the current dynamics of the variables. We start from a maximum lag for ΔX of eight and a maximum lag for CDR of four and for each combination we compute Akaike Information Criterion and Schwartz Information Criterion. Table 1 clearly shows that for any given lag of CDR the model with four lags of ΔX is always preferred. Moreover, it seems that the competing models are the linear model and the models with one and two lags of CDR.

The comparison between models with different order of CDR is not merely a statistical exercise since the $EDTVAR(4,1)$ gives rise to a very different picture on the way the nonlinearity operates in the economy compared with the $EDTVAR(4,2)$. In particular, the nonlinear term is significant only in the GNP equation in the first case and only in the U equation in the second case. The first model would support theories which locate the direct effect of the feedback in the product market, whereas in the second model the feedback affects

---

[14] Although we are aware of their limited value, ADF tests for unit root have been performed and for both series we were not able to reject the null hypothesis of first difference stationarity.

[15] See Nicholls and Pope (1988).

first the labor market dynamics and transmits to output mainly through its correlation with unemployment. Hence, comparing the models with one and two lags of CDR is not only a matter of model selection but it can also give some insights about the propagation of the feedback process and the origin of the nonlinearity found by the literature in both series. This point illustrates the relevance of the two extensions of the $EDTAR$ model that we have proposed.

From the statistical point of view, a non-nested hypothesis testing procedure is needed to compare the two specifications. If we consider the conditional covariance structure of the two models

$$H_{41t} = 1(CDR^{41}_{t-1} = 0)(\Omega_{41c} - \Omega_{41f}) + \Omega_{41f}$$

and

$$H_{42t} = 1(\sum_{i=1}^{2} CDR^{42}_{t-i} = 0)(\Omega_{42c} - \Omega_{42f}) + \Omega_{42f},$$

we find that there is no combination of the parameter values such that the structures of the variance of the error terms in the two models are equivalent. This is due to the different number of lags of the CDR variable inside the indicator function. Note that the non-nested tests are actually two, one of the $EDTVAR(4,1)$ versus the $EDTVAR(4,2)$ and the other of the $EDTVAR(4,2)$ versus the $EDTVAR(4,1)$. A Cox type statistics in its multivariate version as proposed by Pesaran and Deaton (1978) was applied. The proposed statistics for testing $EDTVAR(4,1)$ versus $EDTVAR(4,2)$ is given by

$$\hat{C}_{41,42} \equiv 2\left(\ell_{41}(\hat{\delta}_{41}) - \ell_{42}(\hat{\delta}_{42})\right), \tag{11}$$

where ℓ_{41} and ℓ_{42} are the log-likelihoods of the two models; similarly for the test of $EDTVAR(4,2)$ versus $EDTVAR(4,1)$. The discontinuity of the log-likelihood function prevents us from using the asymptotic results proposed by Pesaran and Deaton and obliges us to resort to resampling techniques in computing the empirical distribution of the two statistics.[16] The computed standardized statistics and bootstrap p-values are reported in Table 4.1.

[16] In Appendix 2, we describe the bootstrap methodology used through out all the paper.

Table 4.1

	Cox Test	p-value
EDTVAR(4,1) VS EDTVAR(4,2)	-6.735	0.002
EDTVAR(4,2) VS EDTVAR(4,1)	1.274	0.392

When the model with one lag is the null hypothesis we are able to reject it strongly; when the model with two lags is the null hypothesis, we found a p-value of 39.2 per cent. Therefore, in the rest of the paper we work with a VAR with two lags in the feedback process. This test suggests that the feedback process from recessions affects the dynamics of the unemployment directly, while it operates on output mainly through the cross-correlation of the series. This point is rather important, since it stresses the danger of misspecification that is implicit in testing nonlinearity on GNP in a univariate framework.

4.2 *Results of the estimation*

The results of the estimation procedure described in the previous section are shown in the following table.[17]

Table 4.2

ESTIMATES FOR $\Delta \ln$ GNP

	parameter	t-value
CONST.	0.0046	2.3127
$\Delta \ln \mathrm{GNP}_{-1}$	0.0583	0.5376
$\Delta \mathrm{U}_{-1}$	-1.0898	-3.0396
CDR_{-1}	-0.1161	-0.5212
$\Delta \ln \mathrm{GNP}_{-2}$	0.1906	1.8203
$\Delta \mathrm{U}_{-2}$	0.5810	1.6534
CDR_{-2}	-0.1214	-0.5566
$\Delta \ln \mathrm{GNP}_{-3}$	0.0225	0.2204
$\Delta \mathrm{U}_{-3}$	0.2086	0.5995
$\Delta \ln \mathrm{GNP}_{-4}$	0.0373	0.3705
$\Delta \mathrm{U}_{-4}$	0.1122	0.4230

ESTIMATES FOR ΔU

	parameter	t-value
CONST.	0.0013	2.5107
$\Delta \ln \mathrm{GNP}_{-1}$	-0.0531	-1.8169
$\Delta \mathrm{U}_{-1}$	0.3924	3.7822
CDR_{-1}	-0.1295	-1.7782
$\Delta \ln \mathrm{GNP}_{-2}$	-0.0803	-2.9213
$\Delta \mathrm{U}_{-2}$	-0.1295	-1.3071
CDR_{-2}	0.1623	2.2916
$\Delta \ln \mathrm{GNP}_{-3}$	-0.0421	-1.5971
$\Delta \mathrm{U}_{-3}$	-0.0911	-0.9554
$\Delta \ln \mathrm{GNP}_{-4}$	-0.0042	-0.1626
$\Delta \mathrm{U}_{-4}$	-0.0865	-1.2269

The t-values are based on the asymptotic standard errors conditional on the estimated value of the threshold parameter. We regard these values as good approximations of the unconditional standard errors, as argued in the previous section.

---

[17] For the estimation, we have initialized the feedback effect to be null.

Table 4.3

Ω_c	
5.1952e-5	-7.2639e-6
-7.2639e-6	2.2988e-6

Ω_f	
1.1805e-4	-3.2413e-5
-3.2413e-5	1.6167e-5

As expected, the estimated variance of the innovations is significantly higher in downturns; hence recessions appear as times of strong volatility and uncertainty (see Table 4.3).

Table 4.4

OTHER PARAMETERS

r	-0.0016
Times in corridor	97
Times in floor	55

Notice that the estimated value of r is very close to zero, which implies that roughly 1/3 of the observations belongs to the floor regime.[18] Figure 1 shows the values for CDR implied by the estimation and the NBER chronology for business cycles from the quarter following the peak to the quarter of the trough. Interestingly, in most cases the CDR variable activates at a point corresponding to the NBER definition of peak and starts decreasing at one corresponding to the trough. Therefore its timing coincides with the "conventional wisdom" about recessions and, in addition, it provides a measure of their depth.

The CDR variable enters the equation of the GNP with a negative sign in both lags, but it is not significant, and the equation of the unemployment with a negative sign in the first lag and with a positive sign in the second.[19] The latter combination of signs offers a very intuitive interpretation of the *direct* effect of the feedback on unemployment. When the system is sliding into a recession, the effect is in the direction of worsening unemployment, since the first lag of CDR has a bigger magnitude than the second. On the contrary, when the economy is in the recovery process, the total effect of the feedback on unemployment is positive.

It is also interesting to note that, when we looked at the deterministic dynamics of the system implied by the estimated parameters, we found that all the steady-states of the

[18] This finding is consistent with the fact that in the past the NBER Business Cycle Dating Committee has dated recessions as if the appropriate cutoff is approximately zero growth in real GNP.

[19] Recall that the variable CDR is negative by definition.

different linear dynamics of the model are located in the corridor regime. Therefore, when the deterministic version of the model is started in the floor regime, it is attracted into the corridor regime, which suggests the existence of a positive feedback effect that contributes moving the economy out of recession. At this point of the analysis, though, the net total effect of the feedback on the economy appears ambiguous; it will be investigated further in the following sections.

4.3 *Testing nonlinearity*

At this point, having picked the best nonlinear model, one has to test the significance of the nonlinearity itself. This test is aimed at understanding whether a specification of the dynamics of ΔX_t with a nonlinear term but also with a possibly induced heteroskedasticity in the error term fits the data better than the best possible linear model, an homoskedastic linear VAR with four lags. The drawback of this testing procedure is that under the null of linearity and homoskedasticity, the threshold parameter r vanishes. This problem is known in the statistical literature as *Davies problem*. One possible solution, proposed by Davies (1977) himself, is to fix the nuisance parameter r and to compute the likelihood ratio test

$$LR(r) = 2\left(\ell_{42}(\alpha, \Phi, \Theta, \Omega_c, \Omega_f | r) - \ell_{40}(\alpha, \Phi, \Omega)\right), \tag{12}$$

where ℓ_{42} is the log-likelihood of the nonlinear model given a value for the nuisance parameter r, while ℓ_{40} is the log-likelihood of the alternative linear model. Davies suggested the statistics

$$SupLR = \sup_r(LR(r)). \tag{13}$$

The distribution of the $SupLR$ is unknown and has to be computed by simulation. Andrews and Ploberger (1994) pointed out that the $SupLR$ test lacks asymptotical optimal properties and proposed the use of two other statistics:

$$\begin{array}{l} ExpLR = E_r \exp(\frac{1}{2}LR(r)) \\ AvgLR = E_r LR(r), \end{array} \tag{14}$$

where the expectation is taken with respect to the nuisance parameter. For computing the critical value of these three test-statistics, we again used bootstrap methods. Table 4.5 reports the bootstrap p-values for the three statistics and shows that all these tests largely rejected the linear model.

Table 4.5

	Test	p-value
SUP-LR	27.704	0.001
AVG-LR	16.293	0.000
EXP-LR	5.0391	0.001

Although we are aware that we picked just one specification of nonlinearity among many possibilities, the evidence against the linear specification is surprisingly strong; even if the true unknown DGP may not be a $EDTVAR$, it seems to be much closer to this nonlinear model than to a linear one. Hence, using a linear VAR in this context may induce misleading conclusions about the general dynamic behavior of the series, the pattern of the impulse responses and the persistence of the shocks and may in the end lead to poor forecasting ability.

5. The long-run effect of recessions

Although the estimation provided useful information on where the nonlinearity originates, it did not make explicit the role of recessions in the dynamics of the system. In particular the sign and the magnitude of the effect of the CDR variable on the long-run growth are still ambiguous at this point of the analysis. The objective of this section is to resolve this ambiguity.

Assuming that the roots of the $\Phi(L)$ polynomial are inside the unit circle, the linear part of the VAR can be inverted to give

$$\Delta X_t = M + B(L)CDR_{t-1} + \Psi(L)\varepsilon_t, \tag{15}$$

where:

$$\varepsilon_t = H_t^{\frac{1}{2}} u_t,$$
$$\Psi(L) = \Phi^{-1}(L),$$
$$M = \Phi^{-1}(1)\alpha,$$
$$B(L) = \Phi^{-1}(L)\Theta(L).$$

Note that $B(L)CDR_{t-1}$ represents the nonlinear feedback which affects the current dynamics of the variables in first differences and cumulates its effect over time on the variables in levels. The idea is to subtract this component from the series and compare the pair of series excluding the cumulation of the CDR polynomial with the actual pair: the difference in the drift of the series can be interpreted as a measure of how recessions have contributed to the long-run behavior of the variables over the sample period.

From Figures 2a and 2b it can be seen that this difference amounts approximately to 0.5 per cent per year for the growth of GNP –which means slightly more than one sixth of its sample average yearly growth rate– and roughly to -0.15 per cent per year for the unemployment rate.[20] We therefore conclude that the contribution of the feedback is positive and large. A useful way to interpret Figure 2a is that if the economy were hit by negative recessionary shocks but did not benefit from the positive feedback (especially strong in the exit from recession), in the subsequent periods output would have grown at a lower rate. Likewise, looking at the actual series for unemployment, it appears at first glance that the unemployment rate rises abruptly at the very beginning of a recession and decreases at a much slower pace when the downturn ends, so it never catches up with the pre-recession level.[21] A closer look at this series shows in fact that as soon as the recovery starts the unemployment rate usually decreases sharply and then keeps declining, but at a much slower rate. When the feedback from recessions is not present, the linear part of the dynamics of unemployment lacks exactly that sharp decrease; hence the corresponding series after each recession remains at a higher level than the actual series. Therefore, although the total effect of recessions on unemployment is negative, since it never recovers its starting level, the feedback goes in the direction of reducing unemployment. In the next section we provide more insight into the reasons why the

---

[20] The two series, with and without the cumulation of the feedback, are started from the same initial point, although, in principle, even the cumulation of the CDR that occurred before the beginning of the sample should be considered.

[21] Blanchard and Summers (1987) used the term *hysteresis* to label this apparent path-dependence of the equilibrium unemployment rate.

nonlinear feedback yields this long-run positive effect on the economy, through the persistence of the impulse response functions.

Finally, as mentioned in the introduction, although ours is not a structural approach, we believe that these results give empirical support to economic theories that model recessions as times of cleansing and reorganizations.

6. Impulse response analysis

6.1 *The generalized impulse response*

In the last section we began to set out the argument that the model offers rich interaction between the nonlinearity represented by the feedback from downturns and the persistence embedded in the fact that our variables are nonstationary on the levels. A complete analysis of this interaction requires the estimation of the empirical impulse response functions (IR) implied by the model.

In a nonlinear environment the IR analysis is a complex matter. The theory of nonlinear IR has been developed in the last few years.[22] As clearly stated in Koop, Pesaran and Potter (1996), the main difference between linear and nonlinear IRs is that the latter are *state* and *shock dependent*. The state dependence means that IRs are sensitive to the history of the system up to the point in which the model is shocked. Shock dependence is a twofold concept. First, if we take the nonlinear IR as function of the shock perturbing the economy, this function is nonlinear and not symmetric around zero. Second, in the definition of IR the treatment of future shocks matters. Whereas in the linear case turning off all the shocks gives the same result as keeping the stochastic structure activated over all the time horizon and averaging out the different futures, in the nonlinear case this equivalence does not hold. Hence it is necessary to explicitly consider the future realizations to evaluate the IR.

Our model being multivariate, we face the additional problem of the *composition dependence* of the disturbance, also encountered in linear VARs. We will not try to exploit the theoretical framework to which we referred in the introduction in order to identify the ”structural” sources of the shocks, but rather we follow a reduced-form approach in which we

---

[22] See Potter (1991), Gallant, Rossi and Tauchen (1993), Koop, Pesaran and Potter (1996).

use the empirical distribution of the residuals to shock both equations at the same time. Note that these two methods are designed to answer different questions. While the first can uncover the propagation pattern of disturbances originating from different specific sources, the second allows us to study the effect of representative economy-wide shocks on the two dependent variables and, in general, has the advantage of generating IRs that are unique, whereas in the structural approach uniqueness is obtained only up to some identification restrictions.

Our IR function is defined as the generalized impulse response in KPP:

$$(16) \qquad IR_{\Delta X}(N, \varepsilon_t = \eta,\ \mathcal{F}_{t-1}) = E(\Delta X_{t+N} | \varepsilon_t = \eta,\ \mathcal{F}_{t-1}) - E(\Delta X_{t+N} |\ \mathcal{F}_{t-1}),$$

where N is the horizon at which the effect of the innovations is examined, η is the vector of shocks hitting the system at time t and $\mathcal{F}_{t-1}$ is a history until time $t-1$.[23] Being an economy-wide shock, η is a draw from the joint distribution of $\{\varepsilon_{y_t}, \varepsilon_{u_t}\}$. Moreover, in the baseline forecast –the second expectation in expression (16)– we do not condition on a null realization of the current shock, but rather we average over all possible realizations. The rationale for this choice comes directly from the meaning of the baseline which is to represent the average behavior of the system.

From the representation in equation (15), and from the definition of IR in (16) we obtain the IR for ΔX implied by our model

$$(17) \qquad IR_{\Delta X}(N, \varepsilon_t = \eta, \mathcal{F}_{t-1}) = \Psi_N \eta + \sum_{j=0}^{N} \{B_j [E(CDR_{t+N-1-j} | \varepsilon_t = \eta, \mathcal{F}_{t-1}) - E(CDR_{t+N-1-j} | \mathcal{F}_{t-1})]\}.$$

The first term in (17) is the linear part of the IR, while the lag polynomial in the difference between the two expectations conditional on different information sets comes from the nonlinearity. As $N \to \infty$, the model in first differences should not show persistence since, given its ergodicity, $\Psi_N \to 0$ and therefore also the coefficients of $B(L)$ quickly go to zero and the two expectations become arbitrarily close, because the effect of the different initial conditions dies off. In levels of the variables, the IR cumulates to

[23] All the relevant information at time $t-1$ is given by CDR_{t-i}, ΔX_{t-j} with $i = 1, 2$ and $j = 1, 2, 3, 4$.

$$(18) \quad IR_X(N, \varepsilon_t = \eta, \mathcal{F}_{t-1}) = \sum_{k=0}^{N} \{\Psi_k \eta + \sum_{j=0}^{k} B_j [E(CDR_{t+k-1-j} | \varepsilon_t = \eta, \mathcal{F}_{t-1}) - E(CDR_{t+k-1-j} | \mathcal{F}_{t-1})]\}.$$

Therefore, the persistence now arises not only from $\Psi(1)$ as in the linear case, but also from the nonlinear structure. In particular, the difference in the realizations of CDR_t in the two expectations composing the IR (due to the non equal initial shock) permanently affects the level of the variables. Intuitively, being at time t in the floor regime or in the corridor regime will make a sharp difference in the final persistence and so will the magnitude of the shock and its sign. We will analyze these issues in the final part of this section, after describing the methodology for computing the empirical IR functions.

6.2 *Computation of impulse response functions*

Our computation of the IRs follows closely the procedure of Koop, Pesaran and Potter (1996) who suggest a resampling technique to numerically integrate the expectations in (18). The three main steps in the implementation of the procedure are:

- choice of the histories, $\mathcal{F}_{t-1}$;
- calibration of the shocks, ε_t;
- treatment of the future.

Regarding histories, we have only used the observed histories, without generating any new one. Altogether, we have 97 sample paths that end up in the corridor regime and 55 leading to the floor regime at time t for $t = 5, 6, ..., 156$. Our main interest is in measuring the asymmetry in persistence across the two regimes; we therefore estimate the IR conditional on the regime by averaging over the histories that end up in the same regime. Moreover, since we have some realization of the floor which is very mild (in the sense that the model at that point is very close to the linear model), we discard all the histories in which either CDR_{t-1} or CDR_{t-2} are below their mean conditional on the system being in floor. Following this strategy we are able to identify 15 histories in which the economy is *undoubtedly* in a phase of recession and this allows us to compare two potentially sharply different situations.

The calibration of the shock is a problem when dealing with multivariate non orthogonalized disturbances since the contemporaneous correlation of the innovations has to be taken into account. As in Gallant, Rossi and Tauchen, we have adopted a graphical method consisting of scattering the residuals for the two time series and, by inspection, determining what could be regarded as a typical shock to the system. These plots suggest that we pick the pair $(.008, -.002)$ respectively for lnGNP and U as a representative positive shock. By changing the sign of this shock and doubling its magnitude we can explore the sign-effect and the size-effect, so altogether we have the four cases in Table 6.1.

Table 6.1

Shocks	ΔlnGNP	ΔU
POS	0.008	-0.002
NEG	-0.008	0.002
PPOS	0.016	-0.004
NNEG	-0.016	0.004

Once the initial shock is chosen from the four cases above, the futures sample paths of the system are generated by bootstrap as discussed in Appendix 3. The maximum horizon of the IR has been set to 28 quarters, which turns out to be sufficient for the long-run behavior to set in. The regime-dependent empirical IR is computed averaging over the realizations of the future conditional on the history and over the histories in each regime, i.e.

$$\hat{I}R_X(N,\eta,R) = \frac{1}{H_R}\sum_{j=1}^{H_R}\left\{\frac{1}{M}\sum_{i=1}^{M}\left[X^i_{t+N}(\eta,\mathcal{F}_j) - X^i_{t+N}(\mathcal{F}_j)\right]\right\}, \tag{19}$$

where $R = \{floor, corridor\}$, H_R indicates the number of histories for floor and corridor, $\mathcal{F}_j$ is the observed $j-th$ history for regime R associated with the realizations until time $t-1$, M is the number of replications. The law of large numbers for i.i.d. random variables ensures convergence of the sample mean in (19) –for each history– to the time invariant expectation[24] characterizing the true IR conditional on the same history, i.e. the right hand side of (18).

[24] Time invariance of the true impulse response function follows from the strict stationarity of our process (ASSUMPTION A2 a)).

6.3 *Results of the impulse response analysis*

The empirical IRs are presented[25] in Figures 3 and 4. Here we analyze the issue of dependence of the IR on the regime of the system, the sign and the size of the shock.

A strongly regime-dependent pattern of persistence stands out of our results: shocks are much *less* persistent when they hit in a recession. This is true independently of the sign and the magnitude of the shock. For example, if we consider the GNP hit by a Neg shock $(-.008, .002)$, we observe that when the system is in the floor regime it responds through an initial negative effect that drives the IR well below its initial point, but after only four quarters the upward pressure deriving from the downturn offsets all the initial drop and quickly brings the IR above its starting point. In the corridor regime this effect is weaker –as it is evident from the less sharp and delayed hump– so the IR levels off below its initial value. An analogous argument holds for unemployment. It is also interesting to note that in both regimes unemployment has a similar abrupt rise in the following 2-3 quarters after the shock; then, depending on the regime, the IR drops quickly or dies out at a higher level. Thus a sharp short-run peak in unemployment following a negative shock is a common feature of both regimes; only its timing differs slightly.

If we turn to positive shocks, the same regime-dependent pattern of persistence is found. It might seem counterintuitive that even in the occurrence of a positive shock the floor regime is associated with a lower rate of persistence. To see why this happens one has to recall that the IR is the difference between the response of the shocked economy and that of the baseline economy. In the floor regime, the shocked economy is pushed out of the recession quickly, while in the baseline economy the feedback is likely to stay activated longer and its positive effect contributes to reduce the difference between the shocked and the baseline systems, hence reducing the persistence.

Another interesting finding is that, independently of the regime, the positive shocks show more persistence than the negative shocks. This result is in line with the univariate model in Beaudry and Koop (1994), although the magnitude of the relative persistence is very different. This clear sign asymmetry is the crucial factor in explaining why in the long run the effect of the feedback from recessions enhances growth. Indeed, although the feedback from downturns

---

[25] In all the Figures, the Impulse Responses have been normalized so that their starting value is always 1.

tend to reduce the persistence of all shocks, this dampening effect is larger for negative shocks, which induces a net positive contribution to the long-run growth of the system.

Note that Gali and Hammour (1991) and Hall (1991) found that a negative aggregate shock has a long-run positive effect on productivity, which they take as a validation of theories that assign a positive role to bad times. Our model predicts that shocks keep their sign even in the long run, but the key role of recessions is to reduce the persistence of negative more strongly than positive shocks. We find this sign-asymmetry view more appealing than the sign-reversion in explaining the positive role of the feedback from recessions on growth.

Figure 4 shows the result of hitting the system with a shock of double magnitude. Here the asymmetry in signs is more visible, especially for the corridor regime: the IR for NNeg has a much steeper hump compared with the PPos case since the disturbance is so negative that the economy falls immediately into a recession; hence the upward push occurs early and is more intense. Interestingly, if we contrast IRs across magnitude of the shocks, the size-asymmetry is rather weak. Indeed, it seems that the only case in which it is relevant is for the negative shock in corridor. In this case, doubling the magnitude induces the system to fall in a deeper recession and substantially increases the probability and the size of the feedback effect.

In relation to the issue of stationarity of the unemployment rate, it is also important to note that the IR computed on the levels does not die off at zero in any of the cases considered. This is an additional piece of evidence in support of our choice of first differencing both series.

One of the objectives of this paper is to explore the type of misspecification which arises in IR analysis from omitting the nonlinearity. For this purpose we compare our IRs with the analogous functions computed for a linear VAR(4) estimated on the same series. Braun and Mittnik (1993) point out that in a linear VAR excessive lag-length truncation has serious consequences in terms of misspecification of the IR. Since the nonlinearity is a function of past lags of ΔX_t, its omission roughly fits into this category, but as additional lags enter the model only when the floor regime is activated, we expect the misspecification to be worse for the floor regime and for negative shocks. This is indeed our finding, as documented in Figure 5. For the PPos case, the linear IR is an average of the two, while for the NNeg case it completely misses the positive effect of CDR and greatly overstates the measure of persistence.

6.4 *The issue of persistence in GNP*

Following the influential paper of Nelson and Plosser (1982), many other authors have tried to measure the persistence of shocks in economic time series. Campbell and Mankiw (1987), using univariate parsimonious ARMA models for GNP, found a persistence coefficient of about 1.5. Cochrane's (1988) nonparametric approach provides estimates between 1.1 and 1.4, according to the window-size selected. Watson's (1986) decomposition of the process into stochastic trend and cycle, using the assumption of orthogonality of the shocks to the two components, gives a measure of persistence between .36 and .57. Lippi and Reichlin (1992) showed that this latter method provides measures that are constrained to be below unity. Evans (1989), using a multivariate approach, yields an estimate between .26 and .55, hence close to Watson's results, although his VAR is not subject to the same criticism. Evans reconciles his finding with the previous literature by arguing that his VAR specification implies a high order ARMA process for the GNP, so that his measure of persistence does include the dampening effect of higher lags, all entering with a negative sign. On the contrary, the low order ARMA models as in Campbell and Mankiw miss this effect and overestimate persistence, producing measures above the random-walk level.

Since our model is nonlinear, we cannot adopt one of the standard methods in the literature, but we can easily generalize the Campbell and Mankiw measure by defining persistence as that number at which the IR levels off, once it is normalized to one at the initial period. Persistence measurements for the log real GNP in our model are summarized in Table 6.2.

Table 6.2

COEFFICIENTS OF PERSISTENCE

Shocks	lnGNP
POS CORR.	1.5
PPOS CORR.	1.5
NEG CORR.	1.3
NNEG CORR.	0.9
POS FLOOR	0.8
PPOS FLOOR	0.9
NEG FLOOR	0.6
NNEG FLOOR	0.5

As already noted in Section 6.3, persistence is fairly asymmetric across regimes. One striking fact is that the coefficient is in the range .5-.9 for the floor regime and between .9 and 1.5 in the corridor regime and these are roughly the two sets of numbers over which the debate in the literature has developed. Our model suggests that there is not a unique optimal lag order and therefore not a unique coefficient of persistence, and can reconcile the two different set of measures in the literature through the nonlinearity. Using Evans' argument, during expansions GNP can be well approximated by a low order ARMA process and the persistence of the shock is high. During downturns, to capture correctly the dynamics of the system, more lags of GNP should enter the specification, and this is done in our model through the CDR variable, with the effect of decreasing the persistence of the innovations.

7. Conclusions

An extension of the $EDTAR$ model to the multivariate framework and to the case in which the feedback operates at multiple lags is proposed in this paper. Although the model presents some difficulties induced by the nonlinear structure and the discontinuity of the log-likelihood function, we could state a set of simple assumptions under which the model is ergodic and we could prove the strong consistency of the ML estimator.

The model is applied to a bivariate VAR of output growth and changes in the unemployment rate for the US economy, where the nonlinearity is introduced through a variable which measures the depth of recessions and defines a *corridor* and a *floor* regime. The two extension we propose turn out to be relevant since, depending on the lag order of the feedback process, the nonlinearity is found to be significant in one equation or the other but not in both. The appropriate testing procedures and the estimation suggest that the relevant model is that where asymmetries are strongly present on the unemployment dynamics and transmit to output through the cross-correlation of the series. This finding points at the danger of misspecification that can heavily affect tests aiming at detecting nonlinearities in the univariate framework.

The generalized impulse response analysis confirms the presence of rich interaction between persistence and nonlinearity. Indeed, it emerges that asymmetries in persistence are clearly present across different regimes and shocks of different sign. Shocks in the floor regime are always less persistent than in the corridor. Negative shocks are less persistent than positive shocks. We argue why these asymmetries explain our finding that the feedback from the downturns has a permanent positive effect on the long run growth process of the economy. Quantitatively, this fact –which we interpret as a cleansing effect of recessions following some recent contributions to macroeconomics– accounts for more than one sixth of the average growth of GNP over our 40-years sample. Moreover, our estimates of regime-dependent persistence for GNP provide a new way of looking at the wide range of persistence measures existing in the literature.

To conclude, we believe we have shown, through our application, that there are cases in which nonlinearities do matter in characterizing the dynamic properties of relevant macroeconomic time series. At the current state of the art, a great amount of theoretical work has been developed on nonlinear time series models. In parallel, there has been a fast decline in the computational time needed to estimate and test these models. Both facts have clearly reduced the advantage of working with linear VAR models. As we have argued, the neglect of nonlinearities may induce wrong inference on the propagation mechanism and persistence measures of the shocks and may hide interesting properties of the data. Therefore, in many cases there is a large net gain in adopting nonlinear frameworks for time series analysis.

The issue of forecasting precision of these models versus their linear counterpart is another important dimension along which the two approaches should be compared. Although this issue has not been explicitly the focus of this paper, it deserves attention and will be object of future work.

Appendix I

In this appendix we will give a Markovian representation of the model in Section 2.2 and we will then use it to state some sufficient conditions for its geometric ergodicity (Proposition 1).

We start with some notation. Let $\{\Delta X_t\}$ be a sequence of $\Re^2$-valued random variables (r.v.'s) that are defined on the complete probability space $(S_{\Delta X}, F_{\Delta X}, \mu_{\Delta X})$ where $\mu_{\Delta X}$ is some $\sigma-$finite measure on $F_{\Delta X}$. Let also $Z_t = \{\Delta X_t, \Delta X_{t-1}, ..., \Delta X_{t-q-\tau+2}\}$ be a $(q+\tau-1)-tuple$ defined on the product space (S_Z, F_Z,μ_Z). Z_t is needed for the vectorization used to construct the Markovian representation of the model.

We can generate q different partitions $\wp^i$, $i = 1, ..., q$, of the space of Z_{t-1} through the feedback indexes F_{t-i} defined in (5), based on the different ways in which the CDR_{t-i} variable can be activated. The partition $\wp^i$ is composed by $\tau + 1$ elements and the generic element $P^i(j) \equiv \{Z_{t-1} : F_{t-i} = j\}$.

Let $\wp^*$ be the joint[26] of the $\wp^i, i = 1, ..., q,$ partitions. This joint partition has maximum dimension $(1 + \tau)^q$ and we can denote its generic element by $P(K) \equiv \cap_{i=1}^q P^i(k_i)$ where $P^i(k_i) \in \wp^i$, $k_i = 0, ..., \tau$ and $K = \{k_1, ..., k_q\}$. Hereafter, $P_0 \equiv \cap_{i=1}^q P^i(0)$ corresponds to that region of the space where the model is in the linear regime, i.e. the region were all the feedback are zeros. The joint partition $\wp$ divides the space of Z_{t-1} into sub-regions $P(K)$ in which the dynamics of the model are modified with respect to those of the linear regime through a matrix $\Lambda(K)$. This matrix can be written as $\Lambda(K) = \sum_{i=1}^q \Lambda^i(k_i)$, where $\Lambda^i(k_i)$ is associated with the element $P^i(k_i) \in \wp^i$ and it is a square matrix of dimension $2(q + \tau - 1)$:

$$\Lambda^i(k_i) = \begin{bmatrix} \underset{(2\times 2(i-1))}{0} & \underset{(2\times 2k_i)}{\overline{\Theta}_i} & \underset{(2\times 2(q+\tau-k_i-i)}{0} \\ \underset{(2(q+\tau-2)\times 2(i-1))}{0} & \underset{(2(q+\tau-2)\times 2k_i)}{0} & \underset{(2(q+\tau-2)\times 2(q+\tau-k_i-i))}{0} \end{bmatrix}$$

where the matrix $\overline{\Theta}_i$ is

$$\overline{\Theta}_i = e \otimes \begin{bmatrix} \Theta_i & 0 \end{bmatrix},$$

[26] Given two partitions $\mathcal{F}_1$and $\mathcal{F}_2$, F^* is an element of the joint partition $\mathcal{F}_1\cup \mathcal{F}_2$ if for some $F_1 \in \mathcal{F}_1$ and for $F_2 \in \mathcal{F}_2$, $F^* \subseteq F_1 \cap F_2$ and there is no other element F' of the joint such that $F' \subseteq F_1 \cap F_2$ and $F^* \subset F'$.

e is a $(1 \times k_i)$ unit vector and Θ_i is the (2×1) vector of coefficients of CDR_{t-i} in (8).

Using Z_t and the defined partition, the model can be rewritten in a vectorized form as

$$
\begin{aligned}
Z_t = \quad & \begin{bmatrix} \alpha \\ 0 \end{bmatrix} + \begin{bmatrix} \Phi_1, ..., \Phi_p & 0 \\ I & 0 \end{bmatrix} Z_{t-1} + \sum_{K=1}^{(1+\tau)^q} \Lambda(K) Z_{t-1} 1(Z_{t-1} \in P(K)) \\
& + r \sum_{i=1}^{q} \Theta_i 1(Z_{t-1} \notin P^i(0)) + U_t.
\end{aligned}
$$

where $U_t' = \{(H_t^{\frac{1}{2}} u_t)', 0, ..., 0\}$.

Even if the partition P has $(1+\tau)^q$ elements, this does not imply that the process Z can move among all elements of the partitions owing to the way the feedback effect has been constructed. Two elements, $P(K)$ and $P(K^*)$, of the partition P will be defined adjoin from K to K^* if $Z_t \in P(K)$ and $Z_{t+1} \in P(K^*)$ which is the case if $k_{i+1}^* = k_i$ for $i = 1, ..., (q-1)$ and $|k_1^* - k_2^*| \leq 1$. Given two elements, $P(K)$ and $P(K^*)$, of the partition P they will be defined as reachable if there is a sequence of adjoin elements of the partition P which allow to move from $P(K)$ to $P(K^*)$. So the subset of the partition P of interest will be the set of element of P which are reachable form P_0.

Given that a Markovian representation of the model exists, the proof of geometric ergodicity can be based upon the theory on stability of Markov chains as in Nummelin (1984) and Tjøstheim (1990). If the process is μ_Z-irreducible and aperiodic, this proof essentially requires verifying the drift conditions for a given power function of Z_t. For a non negative measurable function g and $R > 1$, the $N-$step ahead drift conditions are

$$
\begin{aligned}
&\text{DC1} \quad RE(g(Z_t)|Z_{t-N} = z) < g(z) \text{ on } z \in \kappa^c \\
&\text{DC2} \quad E(g(Z_t)|Z_{t-N} = z) \leq M < \infty \text{ on } z \in \kappa
\end{aligned}
$$

where κ is a small set. In our framework the continuity of the conditional mean function and the following regularity conditions on the error term u_t ensure that every compact set on the space of Z_t is a small set. For the formal definitions of irreducibility, aperiodicity and small set, the reader can refer to Nummelin (1984), definitions 2.2, 2.4 and 2.3 respectively.

The geometric ergodicity of the process is based on the following assumption:

ASSUMPTION A1: $\exists$ an integer $N > 0$ s.t.

$$\max \left\| \prod_{j=1}^{N} \left\{ \begin{bmatrix} \Phi_1, ..., \Phi_p & 0 \\ I & 0 \end{bmatrix} + \Lambda\left(K_j\right) \right\} \right\| < 1,$$

where the max operator is taken over all the possible finite number of combinations of $\Lambda\left(K_j\right)$ such that $P\left(K_j\right)$ for $j = 1, ..., N$ belongs to subset reachable from P_0 and $P\left(K_j\right)$ and $P\left(K_{j+1}\right)$ are adjoin for $j = 1, ..., N-1$, and where $\|\cdot\|$ stands for the euclidean norm.

This assumption is meant to capture that, independently of all its possible paths, the process after N steps always tends to return towards the centre. We are now ready to state:

PROPOSITION 1: If ASSUMPTIONS A1 holds and the marginal pdf of the error term u_t is absolutely continuous and positive on $\Re^2$ then the process ΔX_t is geometrically ergodic;

PROOF: The aperiodicity and μ_Z-irreducibility follow directly from definitions under the assumptions on the marginal pdf of the error term u_t. Define the power function $g\left(\cdot\right)$ as the Euclidean norm $\|\cdot\|$ in the space of Z_t. We can now verify DC1 and DC2. The norm of the constant vector, the r_f term and the covariance of the errors can be bounded by a positive constant C. Furthermore, by Assumption A1, it is easy to see that there is an $\alpha < 1$ s.t. the following inequality holds:

$$E\left(\|Z_t\| \mid Z_{t-N} = z\right) \quad \leq C + \alpha \|z\|.$$

So it is possible to find an $R > 1$ such that $R\alpha < 1$ and to rewrite that previous condition as

$$\begin{aligned} RE\left(\|Z_t\| \mid Z_{t-N} = z\right) \quad & \leq RC + R\alpha \|z\| \\ & \leq \|z\| + RC + \left(\alpha R - 1\right)\|z\| \end{aligned}$$

Let us now define the small set $\kappa \equiv \{z : \|z\| \leq r\}$ and take $r > \frac{RC}{1-R\alpha}$, then :

$$\begin{aligned} R\left(E\|Z_t\| \, |Z_{t-N} = z\right) \quad & < \|z\| \quad \text{on } \kappa^c \text{ and} \\ R\left(E\|Z_t\| \, |Z_{t-N} = z\right) \quad & < M < \infty \quad \text{on } \kappa, \end{aligned}$$

where the second condition derives from the fact that κ is compact and the conditional mean of the process is continuous. Given that the drift conditions hold, the result follows by Proposition 5.21 of Nummelin (1984). ■

Given the geometric ergodicity property, if we also assume that the initial distribution of the process was the invariant distribution, then this implies that ΔX_t is a strictly stationary process.

Appendix II

Under the conditions of Proposition 1 in Appendix 1, we can establish the geometric ergodicity of our $EDTVAR$ model. We are now ready to look at the properties of the ML estimator.

Let us define $W_t \equiv \{\Delta X_t, Z_{t-1}\}$ on the probability space (S_W, F_W, μ_W). W_t is the history of ΔX_t up to its maximum relevant past, i.e. time $(t-q-\tau+1)$, and it will be used to express the likelihood in compact notation. Denote by B_k the Borel $\sigma-$algebra generated by the space $\Re^k$. Recall that the conditional likelihood function for observation ΔX_t is:

$$\ell_t(\delta) \equiv \ell(W_t, \delta) =$$
$$-\frac{1}{2}\left[\ln|H_t| + \left(\Phi(L)\Delta X_t - \alpha - \Theta(L)CDR_t\right)' H_t^{-1}\left(\Phi(L)\Delta X_t - \alpha - \Theta(L)CDR_t\right)\right],$$

where δ is the vector of parameters which lies in a metric space D. The dependence of $\ell(\cdot)$ from W_t reflects that we allow the variable CDR_t to contain up to τ lags of ΔX_t and that the polynomial $\Theta(L)$ has order q. Define also

$$\overline{\ell}(W_t, \delta^*, \rho) \equiv \sup_\delta \left(\ell(W_t, \delta) : d(\delta^*, \delta) < \rho\right)$$
$$\underline{\ell}(W_t, \delta^*, \rho) \equiv \inf_\delta \left(\ell(W_t, \delta) : d(\delta^*, \delta) < \rho\right),$$

where $d(.,.)$ is a metric on D and ρ is a positive number. We shall refer to $\widehat{\delta}_{MLE}$ as the maximizer of the sample likelihood, $\frac{1}{T}\sum_{t=1}^{T}\ell(W_t, \delta)$, on D. For the proof of Proposition 2, we need four Lemmas in which we characterize the continuity of ℓ and the measurability of $\overline{\ell}$ and $\underline{\ell}$.

LEMMA 1A: For any $\delta^* \in D$, $\ell(\cdot, \delta^*)$ is continuous in W_t almost everywhere (a.e.) under μ_W;

PROOF: Take any arbitrary $\delta^* \in D$. By simple inspection it turns out that the function $\ell(\cdot, \delta^*)$ is discontinuous *only if* for some $j = 1, 2, ..., (q+\tau-1)$, $\Delta X_{1,t-j} = r^*$ where r^* is the last element of δ^*. Define now the set $G_j^* \equiv \left\{W_t \in \times_{q+\tau-1}\Re^2 : \Delta X_{1,t-j} = r_f^*\right\}$. It contains the set of histories of ΔX_t which induce a discontinuity of $\ell(\cdot, \delta^*)$ at time $t-j$. This is a $\times_{q+\tau-1}B_2$-measurable set and it has measure zero. Define now $G^* \equiv \cup_{j=1}^{q+\tau-1} G_j^*$. This set is larger than the set of *all* histories up to time t which induce discontinuities of $\ell(\cdot, \delta^*)$. Being a countable union of measure-zero sets, G^* has measure zero and the conclusion follows for the arbitrariness of δ^* and the completeness of (S_W, F_W, μ_W). ■

LEMMA 1B: For any fixed W_t, $\ell(W_t, \cdot)$ has at most one discontinuity with respect to r.

PROOF: Suppose $\ell(W_t, \cdot)$ is discontinuous at r^*. Denote by $B_\rho(r^*)$ an open ball around r^* of radius ρ. Then, for every $r \in B_\rho(r^*)$ the following two conditions are verified

$$\begin{aligned}
&\text{(i) if } r > r^*, \text{ then } \exists i^* \in \{1, 2, ...q\} \text{ such that } CDR_{t-i^*}(r) < 0; \\
&\text{(ii) if } r \leq r^*, \text{ then } CDR_{t-i}(r) = 0, \text{ for } \forall i \in \{1, 2, ...q\}.
\end{aligned}$$

Hence, from the definition of CDR_{t-i^*} and (i) it follows that $\exists j > i^*$ such that $X_{t-i^*} - (X_{t-j} + r) < 0$. So for every $r > r^*$, $CDR_{t-i^*} < 0$ and it is not possible to have a new discontinuity since for any other $r > r^*$ the system at time t will be always in the floor regime. Similarly, for every $r \leq r^*$, $CDR_{t-i} = 0$ and we cannot have a new discontinuity. Given that r^* and ρ are arbitrary, the conclusion follows. ■

REMARK: A simple corollary of Lemma 1b is that, once Δ is equipped with some appropriate $\sigma-$algebra and measure, the set of discontinuities of $\ell(W_t, \cdot)$ has measure zero.

LEMMA 2: $\ell(\cdot, \delta)$ is a F/B-measurable function.

PROOF: It follows from Lemma 1a and the completeness of (S_W, F_W, μ_W). ■

LEMMA 3: For any $\delta^* \in D$, $\overline{\ell}(\cdot, \delta^*, \rho)$ and $\underline{\ell}(\cdot, \delta^*, \rho)$ are F/B-measurable functions.

PROOF: To prove measurability of $\overline{\ell}(\cdot, \delta^*, \rho)$ we will construct a sequence of measurable functions that converges to it. Denote the set $\overline{W}$ of histories such that $\inf_j |\Delta X_{1t-j} - r^*| > \rho$. By construction, the function $\ell(W_t, \delta)$ is jointly continuous on $\overline{W} \times B_\rho(\delta^*)$. This implies that $\overline{\ell}(\cdot, \delta^*, \rho)$ is continuous on $\overline{W}$. Now, define the smooth switching-function $F_{\chi^2}\left(\frac{-\sum_{i=1}^{q} CDR_{t-i}}{\lambda_n}\right)$ as the cumulative distribution of a chi-squared random variable and let $\{\lambda_n\}$ be a sequence of scalars converging to zero. Note that in our definition the variable CDR_t is nonpositive, so if the process is in the corridor regime, i.e. all $CDR_{t-i} = 0$ for $i = 1, ..., q$, then F_{χ^2} will be identically zero. If the process is in the floor regime, i.e. $\exists\ i^*$ such that $CDR_{t-i^*} < 0$, then the function F_{χ^2} will be greater than zero and approach one as

λ_n goes to zero. Now we introduce the following function

$$\begin{aligned} &\ell^{'}(W_t, \delta, \lambda_n) = \\ &-\tfrac{1}{2}\left[\ln|\Omega_c| + (\Phi(L)\Delta X_t - \alpha)^{'}\Omega_c^{-1}(\Phi(L)\Delta X_t - \alpha)\right] \\ &+\left\{-\tfrac{1}{2}\ln\tfrac{|\Omega_f|}{|\Omega_c|} - \tfrac{1}{2}(\Theta(L)CDR_t)^{'}\Omega_f^{-1}(\Theta(L)CDR_t)\right. \\ &-\tfrac{1}{2}(\Phi(L)\Delta X_t - \alpha)^{'}\left(\Omega_f^{-1} - \Omega_c^{-1}\right)(\Phi(L)\Delta X_t - \alpha) \\ &\left.+(\Theta(L)CDR_t)^{'}\Omega_f^{-1}(\Phi(L)\Delta X_t - \alpha)\right\}F_{\chi^2}\left(\tfrac{-\sum_{i=1}^{q}CDR_{t-i}}{\lambda_n}\right). \end{aligned}$$

Being a sum of jointly continuous functions, this function is jointly continuous with respect to (W_t, δ). Moreover, the sequence of functions $\left\{\ell^{'}(W_t, \delta, \lambda_n)\right\}$ converges to $\ell(W_t, \delta)$ as λ_n goes to zero. Denote

$$\overline{\ell}^{'}(W_t, \delta^*, \rho, \lambda_n) \equiv \sup_{\delta}\left(\ell^{'}(W_t, \delta, \lambda_n) : d(\delta^*, \delta) < \rho\right).$$

For the same argument as above, the function $\overline{\ell}^{'}$ is continuous on all $\times_\tau\Re^2$. This allows the following sequence of functions to be constructed

$$\widehat{\ell}(W_t, \delta^*, \rho, \lambda_n) = \begin{cases} \overline{\ell}(W_t, \delta^*, \rho) & \text{if } W_t \in \overline{W} \\ \overline{\ell}^{'}(W_t, \delta^*, \rho, \lambda_n) & \text{if } W_t \in \times_\tau\Re^2/\overline{W}, \end{cases}$$

Since the function $\widehat{\ell}$ is continuous a.e., the sequence converges to $\overline{\ell}(W_t, \delta^*, \rho)$ and δ^* is arbitrary, we have established the measurability of $\overline{\ell}(\cdot, \delta^*, \rho)$. The proof goes through similarly for $\underline{\ell}(\cdot, \delta^*, \rho)$. ■

For the main proposition we also need the following assumptions:

ASSUMPTION A2: D is a compact metric space.

ASSUMPTION A3: $a)$ $\int \sup_{\delta\in D}|\ell(W_t, \delta)|\, h(W_t)d\mu_W \leq M < \infty$
where $h(W_t)$ is the probability density function of W_t.
$b)$ $h(W_t) < \infty$ almost everywhere (a.e.) under μ_W.

ASSUMPTION A4: $E\left(\ell(W_t,\cdot)\right)$ attains its unique global maximum on Δ at $\delta = \delta^{\circ}$.

We are now ready to prove the consistency of $\widehat{\delta}_{MLE}$. We will follow, with minor modifications, the argument in the main Theorem of Andrews (1987).

PROPOSITION 2: If Assumptions A1-A4 hold, then

(a) $E\left(\ell(W_t,\delta)\right)$ is continuous on D;

(b) $\sup_{\delta\in\Delta}\left|\frac{1}{T}\sum_{t=1}^{T}\ell(W_t,\delta) - E\left(\ell(W_t,\delta)\right)\right| \to 0$
as $T\to\infty,$ almost surely (a.s.) under μ_W

(c) $\widehat{\delta}_{MLE} \to \delta^{\circ}$ a.s. under μ_W.

PROOF: For $\forall\delta\in D$, we have that

$$\begin{aligned}
&\lim_{\rho\to 0}\left|\int\left(\overline{\ell}(W_t,\delta,\rho) - \ell(W_t,\delta)\right)h(W_t)d\mu_W\right| \\
&\le \lim_{\rho\to 0}\int\left|\left(\overline{\ell}(W_t,\delta,\rho) - \ell(W_t,\delta)\right)h(W_t)\right|d\mu_W \\
&= \int\lim_{\rho\to 0}\left|\left(\overline{\ell}(W_t,\delta,\rho) - \ell(W_t,\delta)\right)h(W_t)\right|d\mu_W = 0,
\end{aligned}$$

where the first equality holds by the Lebesgue Dominated Convergence Theorem using Assumption A3 a) and the last equality holds by Lemma 1a and Assumption A3 $b)$. The same is true if we replace $\overline{\ell}$ by $\underline{\ell}$. Given this, part (a) follows.

By (a), for any $\varepsilon > 0$, we can choose $\rho\left(\delta\right)$ so that

$$E\left(\ell(W_t,\delta)\right) - \varepsilon \le E\left(\underline{\ell}(W_t,\delta,\rho\left(\delta\right))\right) \le E\left(\overline{\ell}(W_t,\delta,\rho\left(\delta\right))\right) \le E\left(\ell(W_t,\delta)\right) + \varepsilon.$$

The collection of balls $\{B(\delta,\rho\left(\delta\right)) : \delta\in D\}$ is an open cover of the compact set D, hence there is a finite subcover $\{B(\delta_i,\rho\left(\delta_i\right))\}$, $i = 1,...,I$. For any $\delta\in B(\delta_i,\rho\left(\delta_i\right))$, we have

$$\begin{aligned}
&\ell(W_t,\delta) - E\left(\ell(W_t,\delta)\right) \\
&\le \overline{\ell}(W_t,\delta_i,\rho\left(\delta_i\right)) - E\left(\underline{\ell}(W_t,\delta_i,\rho\left(\delta_i\right))\right) \\
&\le \overline{\ell}(W_t,\delta_i,\rho\left(\delta_i\right)) - E\left(\overline{\ell}(W_t,\delta_{i,},\rho\left(\delta_i\right))\right) + 2\varepsilon,
\end{aligned}$$

and

$$\ell(W_t,\delta) - E\ell(W_t,\delta)$$
$$\geq \underline{\ell}(W_t,\delta_i,\rho(\delta_i)) - E\left(\overline{\ell}(W_t,\delta_i,\rho(\delta_i))\right)$$
$$\geq \underline{\ell}(W_t,\delta_i,\rho(\delta_i)) - E\left(\underline{\ell}(W_t,\delta_{i,},\rho(\delta_i))\right) - 2\varepsilon.$$

From A.7 and A.8, averaging over T observations and subsequently taking the minimum and the maximum over i, we obtain:

$$\min_{i\leq I}\frac{1}{T}\sum_{t=1}^{T}\underline{\ell}(W_t,\delta_i,\rho(\delta_i)) - E\left(\underline{\ell}(W_t,\delta_{i,},\rho(\delta_i))\right) - 2\varepsilon$$
$$\leq \frac{1}{T}\sum_{t=1}^{T}\ell(W_t,\delta) - E\left(\ell(W_t,\delta)\right) \leq$$
$$\leq \max_{i\leq I}\frac{1}{T}\sum_{t=1}^{T}\overline{\ell}(W_t,\delta_i,\rho(\delta_i)) - E\left(\overline{\ell}(W_t,\delta_{i,},\rho(\delta_i))\right) + 2\varepsilon.$$

The min and max operator are defined because I is finite. By Proposition 1 and Lemmas 2 and 3, ℓ, $\overline{\ell}$ and $\underline{\ell}$ are stationary and ergodic (see Stout, 1974, p. 182) hence the expectations of these functions do not depend on t. By Assumption A3 $a)$ the first moment of ℓ, $\overline{\ell}$ and $\underline{\ell}$ exist, so we can apply the Strong Law of Large Number for stationary and ergodic sequences to the three terms of A.9. The two bounds converge then to -2ε and 2ε. Since ε does not depend on δ and it is arbitrary, the uniform convergence of the central term follows. This establishes part (b) of the Proposition.

By (a) and Assumption A2 δ° exists and by Assumption A4 it is unique. Part (b) of Proposition 1 ensures that the sequence $\left\{\widehat{\delta}_T\right\}$ will converge to δ°almost surely under μ_W, so part (c) is also proved. ■

Appendix III

Due to the nonstandard asymptotics of the nonlinear model, we make extensive use of bootstrap techniques throughout all the paper. Recall that in our model the variance of the error term follows a qualitative threshold type of process of the form:

$$\varepsilon_t = H_t^{\frac{1}{2}} u_t$$
$$H_t = 1(\textstyle\sum_{i=1}^{q} CDR_{t-i} = 0)(\Omega_c - \Omega_f) + \Omega_f.$$

We adapt the method of resampling proposed by Lamoureux and Lastrapes (1990) for models with conditional heteroskedastic errors. From the estimation procedure we have consistent estimators for $\widehat{\varepsilon}_t$ and for $\widehat{H_t}$ at every t, hence the adjusted-homoskedastic error $\widehat{u}_t$ can be computed as

$$\widehat{u}_t = \widehat{H}_t^{-\frac{1}{2}} \widehat{\varepsilon}_t.$$

For the errors $\{\widehat{u}_t\}$ we tested and could not reject the null hypothesis of no serial correlation. A test of serial correlation has also been performed on the series of errors $\widehat{\varepsilon}_t$ to make sure that there is no dependence left in the residuals, since this would also affect the consistency of the estimates. The test in this case is nonstandard due to the conditional heteroskedasticity and has to be performed following the modification of the Box-Pierce test proposed by Diebold (1987). Again the test does not signal the presence of autocorrelation.

In bootstrapping, we used the following procedure. At every replication, we draw with replacement a new series $\left\{u_t^{(i)}\right\}$ from the homoskedastic and uncorrelated residuals. Given this new sample and assuming the original initial conditions, we generate the new sample $\left\{\Delta X_t^{(i)}\right\}$, using the recursive structure of the model specified in equations (**??**)-(9).

In the non-nested test for the order selection on the CDR variable and in the test for linearity, we use this resampling strategy to compute the empirical distribution of the statistics under the null. For both tests at each replication, we regenerate the data under the null of the test and then compute a new value of the statistics given the bootstrap sample. For each test, 2,000 bootstrap replications were performed.

The same resampling methodology has also been used in the computation of the impulse responses. For each history, we draw $1,000 \times 28$ realizations of the homoskedastic residuals for the baseline model and $1,000 \times 27$ for the shocked model, since the initial shock is fixed at the calibrated values.[27] For each of the $1,000$ replications, a future of length 28 for ΔX and CDR is recursively built both for the shocked and the baseline economies. We then average over the replications to compute the value of the two expectations in the right-hand side of (16), so the IR conditional to the given history is obtained. To induce a negative correlation between the sample estimates of the two expectations in (16) and reduce the experimental variance, we use the same set of random numbers in generating the 27 period futures for the shocked and the baseline model. The regime-dependent impulse response is computed by averaging over all the observed histories in a given regime; hence, in all, we have 97,000 realizations in the corridor and 15,000 in the floor regime.

All computations have been performed using GAUSS 3.1 for UNIX running on IBM RISC 6000.

[27] 28 is the maximum horizon chosen for the IR analysis.

Table 1

MODEL SELECTION CRITERIA FOR THE VAR ORDER

Lag CDR	Lag Linear	AIC	SIC
0	4	**246.14**	**192.19**
0	5	239.37	173.43
0	6	231.85	153.92
0	7	224.49	134.57
0	8	217.92	116.01
1	4	**249.86**	**189.92**
1	5	242.55	170.62
1	6	235.99	152.07
1	7	228.48	132.57
1	8	225.37	117.47
2	4	**250.95**	**185.01**
2	5	244.17	166.24
2	6	237.44	147.52
2	7	230.40	128.49
2	8	225.08	111.19
3	4	**246.89**	174.96
3	5	240.85	156.93
3	6	234.39	138.48
3	7	227.51	119.61
3	8	223.90	104.01
4	4	244.33	166.40
4	5	237.16	147.24
4	6	230.30	128.39
4	7	223.51	109.62
4	8	218.83	92.65

Number of observations: 148

Legenda: AIC, Akaike Information Criterion; SIC, Schwartz Information Criterion.

Figure 1

CURRENT DEPTH OF RECESSION AND NBER DATES

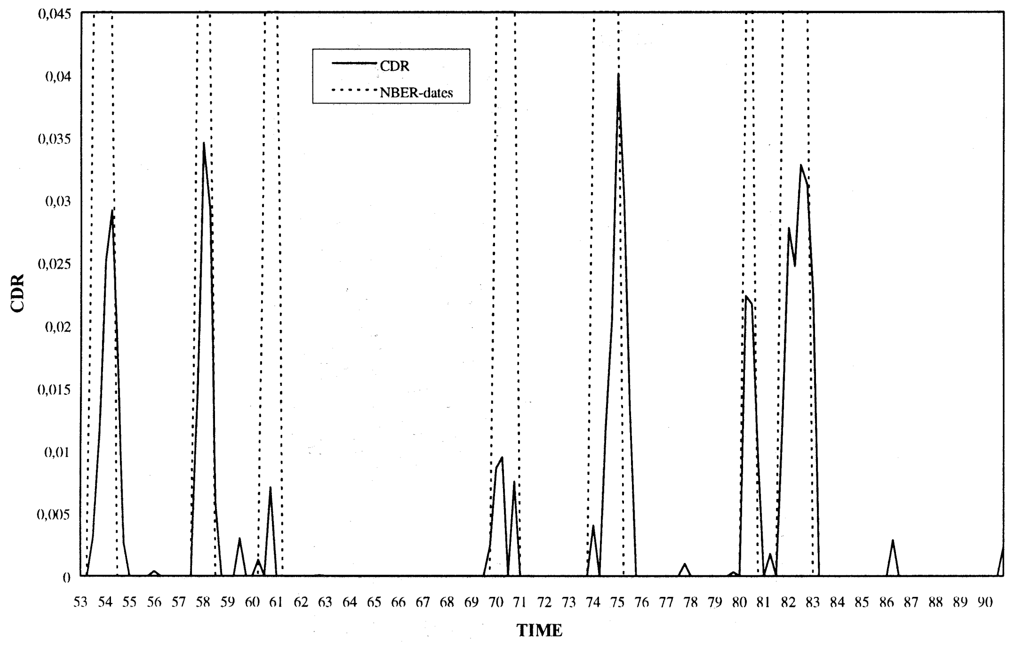


Figure 2a

EFFECT OF RECESSION ON OUTPUT

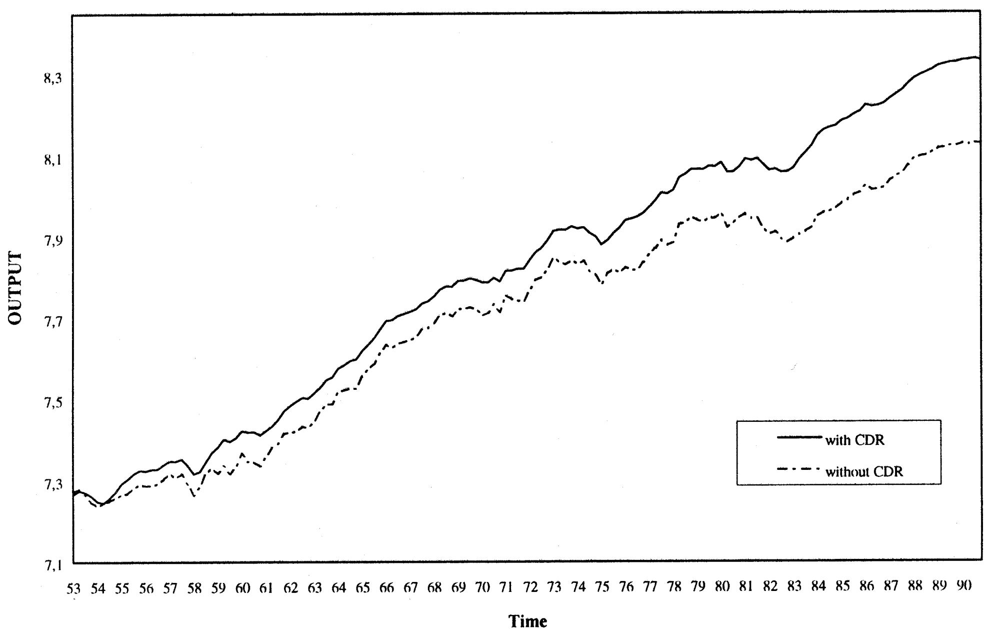


Figure 2b

EFFECT OF RECESSION ON UNEMPLOYMENT

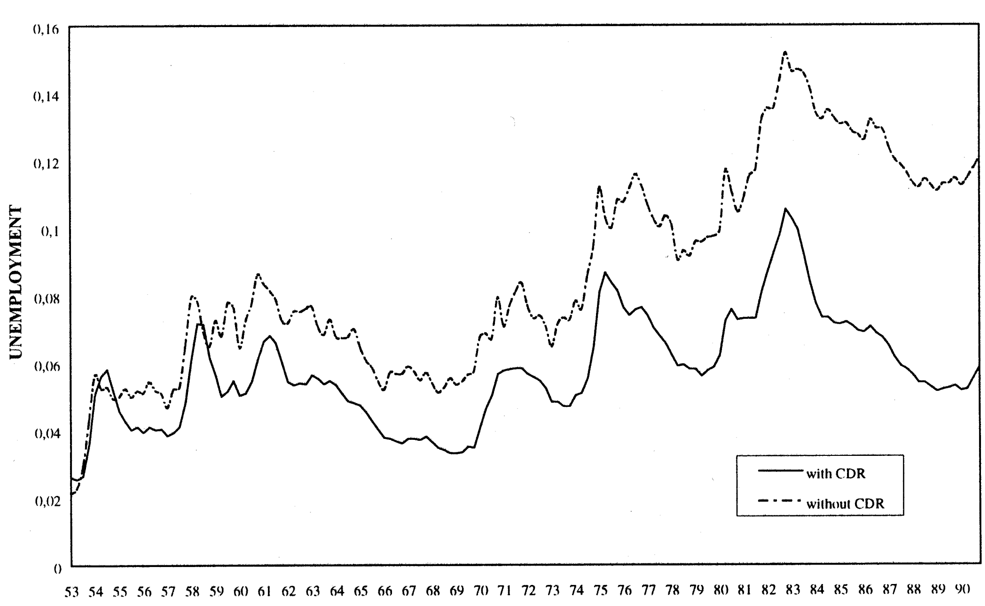
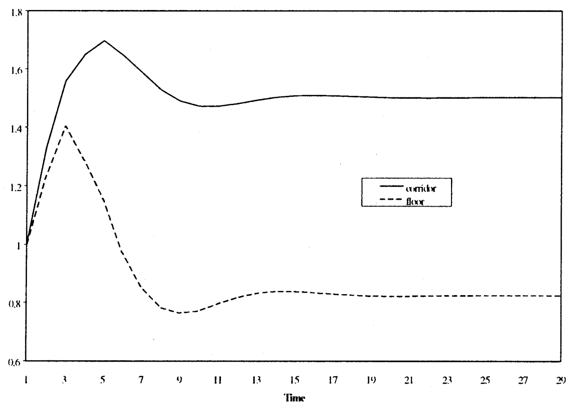
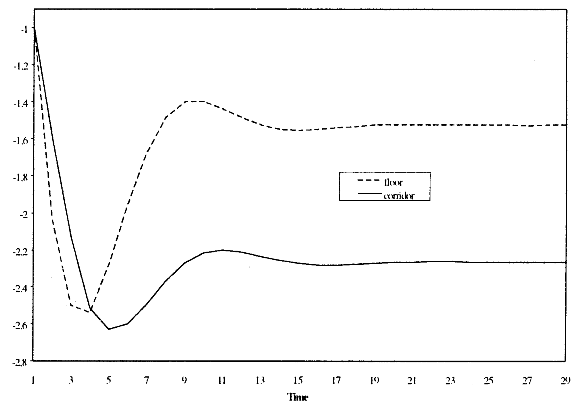


Figure 3

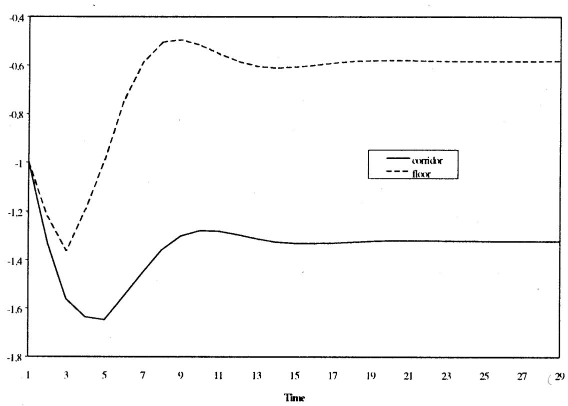
IMPULSE RESPONSE



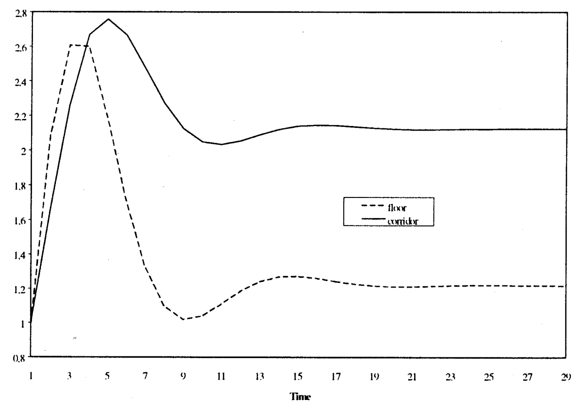
IR of output to pos shock



IR of unemployment to pos shock



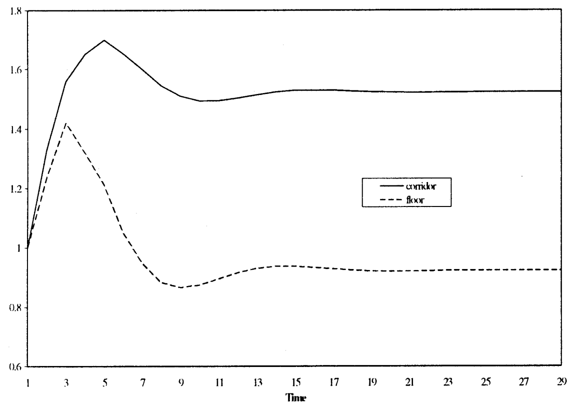
IR of output to neg shock



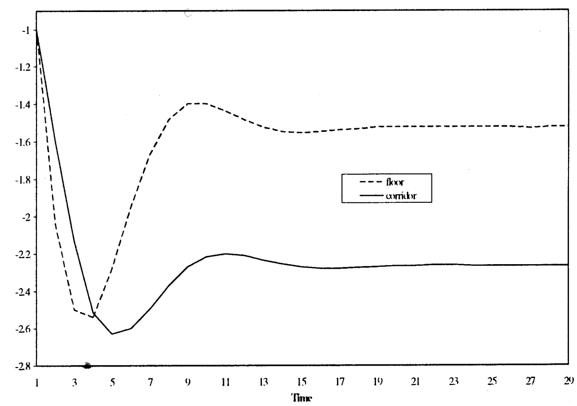
IR of unemployment to neg shock

Figure 4

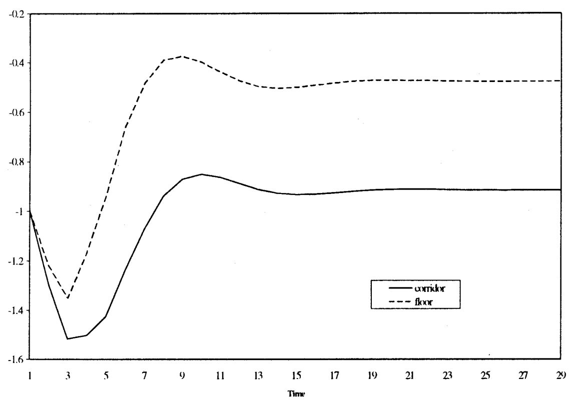
IMPULSE RESPONSE



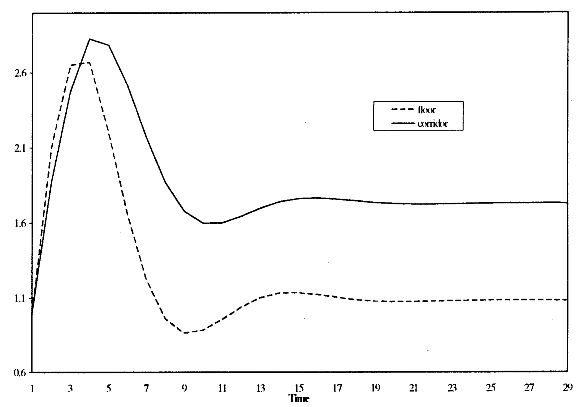
IR of output to ppos shock



IR of unemployment to ppos shock



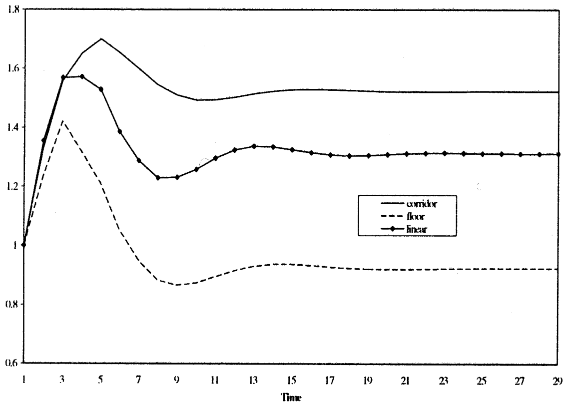
IR of output to nneg shock



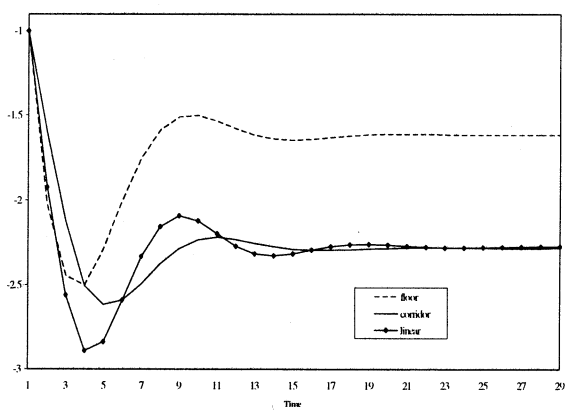
IR of unemployment to nneg shock

Figure 5

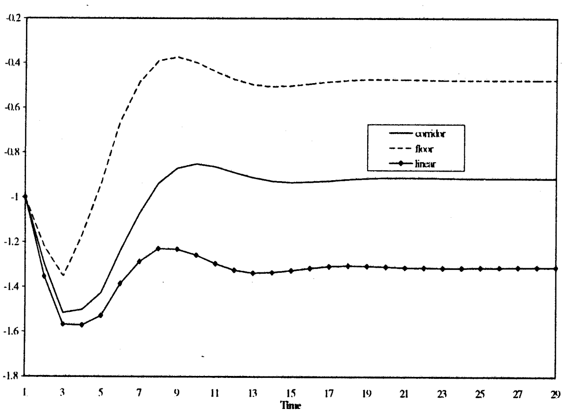
NONLINEAR VS. LINEAR IMPULSE RESPONSE



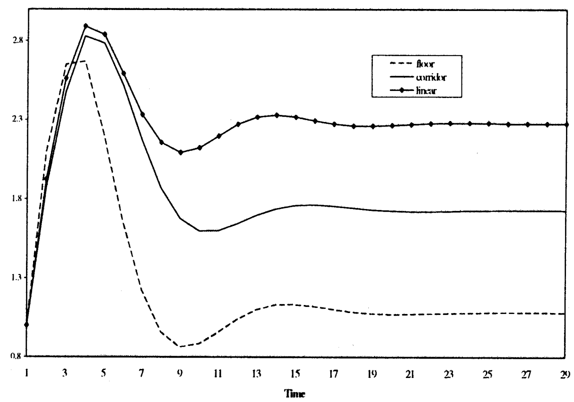
IR of output to ppos shock



IR of unemployment to ppos shock



IR of output to nneg shock



IR of Unemployment to NNeg Shock

References

Aghion, P. and G. Saint-Paul (1993), "Uncovering Some Causal Relationships between Productivity Growth and the Structure of Economic Fluctuations: A Survey," National Bureau of Economic Research, Working Paper, No. 4603.

Andrews, D. W. K. (1987), "Consistency in Nonlinear Econometric Model: A Generic Uniform Law of Large Numbers," *Econometrica*, Vol. 55, pp. 1465-471.

Andrews, D.W.K. and W. Ploberger (1994), "Optimal Tests When a Nuisance Parameter is Present only under the Alternative," *Econometrica*, Vol. 64, pp.1383-414.

Beaudry, P. and G. Koop (1993), "Do Recessions Permanently Affect Output?" *Journal of Monetary Economics*, Vol. 31, pp. 149-63.

Blanchard, O. and D. Quah, (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances," *American Economic Review*, Vol. 79, No. 4, pp. 655-73.

Blanchard, O. and L. Summers (1987), "Hysteresis in Unemployment," *European Economic Review*, Vol. 31, pp. 288-95.

Braun, P. A. and S. Mittnik (1993), "Misspecifications in Vector Autoregression and Their Effect on Impulse Responses and Variance Decompositions," *Journal of Econometrics*, Vol. 59, pp. 319-41.

Brunner, A. (1992), "Conditional Asymmetries in Real GNP: A Seminonparametric Approach," *Journal of Business and Economic Statistics*, Vol. 10, pp. 65-72.

Caballero, R. and M. Hammour (1994), "The Cleansing Effect of Recessions," *American Economic Review*, Vol. 84, pp. 1350-68.

Campbell, J. and G. Mankiw (1987), "Permanent and Transitory Components in Macroeconomic Time Series," *American Economic Review: Papers and Proceedings*, Vol. 77, pp. 111-17.

Chan, K.S. (1988), "Consistency and Limiting Distribution of Least Square Estimator in a Threshold Autoregressive Model," University of Chicago, Department of Statistics, Technical Report, No. 245.

Cochrane, J. H. (1988), "How Big is the Random Walk in GNP?" *Journal of Political Economy*, Vol. 5, pp. 893-920.

Davies, R. B. (1977), "Hypothesis Testing When a Nuisance Parameter is Present only under the Alternative," *Biometrika*, Vol. 64, pp. 247-54.

DeLong, J. B., and L. Summers (1986), "Are Business Cycles Symmetrical?" in R. J. Gordon (ed.), *The American Business Cycle: Continuity and Change*, Chicago, University of Chicago Press.

Diebold, F. X. (1987), “Testing for Serial Correlation in the Presence of ARCH,” in American Statistical Association (ed.), *Proceedings of the Business and Economic Statistics Section*, Washington DC.

Evans, G. W. (1989), “Output and Unemployment Dynamics in the US: 1950-1988,” *Journal of Applied Econometrics*, Vol. 4, pp. 213-37.

Falk, B. (1986), ”Further Evidence on the Asymmetric Behavior of Economic Time Series over the Business Cycle,” *Journal of Political Economy*, Vol. 94, pp. 1096-109.

French, M. W. and D. Sichel (1993), “Cyclical Patterns in the Variance of Economic Activity,” *Journal of Business and Economic Statistics*, Vol. 11, pp. 113-19.

Gali, J. and M. Hammour (1991), “Long-Run Effects of Business Cycles,” Columbia University, mimeo.

Gallant, R.A., P.E. Rossi and G. Tauchen (1993), “Nonlinear Dynamic Structures,” *Econometrica*, Vol. 61, pp. 871-907

Gourieroux, C. and A. Monfort (1992), “Qualitative Threshold ARCH Models,” *Journal of Econometrics*, Vol. 52, pp. 159-99.

Hall, R. E. (1991), “Recessions as Reorganizations,” in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*, Cambridge, MIT Press.

Hamilton, J. (1989), “A New Approach to the Economic Analysis of Nonstationary Time Series and Business Cycle,” *Econometrica*, Vol. 57, pp. 357-84.

Koop, G., M. H. Pesaran and S. Potter (1996), “Impulse Response Analysis in Nonlinear Multivariate Models,” *Journal of Econometrics*, Vol. 74, pp. 119-47.

Lamourex, C. G. and W. D. Lastraps (1990), “Persistence in Variance, Structural Change and the GARCH Model,” *Journal of Business and Economic Statistics*, Vol. 8, pp. 225-34.

Lippi, M. and L. Reichlin (1992), “On Persistence of Shocks to Economic Variables,” *Journal of Monetary Economics*, Vol. 29, pp. 87-93.

Neftçi, S. N. (1984), “Are Economic Time Series Asymmetric over the Business Cycle?” *Journal of Political Economy*, Vol. 92, pp. 307-328.

Nelson, C. R. and C. I. Plosser (1982), “Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications,” *Journal of Monetary Economics*, Vol. 10, pp. 139-62.

Nicholls, D. F. and A. L. Pope (1988), “Bias in the Estimation of Multivariate Autoregressions,” *Australian Journal of Statistics*, Vol. 30-A, pp. 296-309.

Nummelin, E. (1984), *General Irreducible Markov Chains and Non-Negative Operators*, Cambridge, Cambridge University Press.

Pesaran, M. H. and A. Deaton (1978), “Testing Non-Nested Nonlinear Regression Models,” *Econometrica*, Vol. 46, pp. 677-94.

Pesaran, M. H. and S.M. Potter (1994), “A Floor and Ceiling Model of US Output,” University of Cambridge, mimeo.

Potter, S. M. (1995), “A Nonlinear Approach to US GNP,” *Journal of Applied Econometrics*, Vol. 10, pp. 109-25.

Potter, S. M. (1991), “Nonlinear Impulse Response Functions,” Department of Economics, UCLA, mimeo.

Rothman, P. (1991), “Further Evidence on the Asymmetric Behavior of the Unemployment Rate over the Business Cycle,” *Journal of Macroeconomics*, Vol. 13, pp. 291-98.

Saint-Paul, G. (1993), “Productivity Growth and the Structure of the Business Cycle,” *European Economic Review*, Vol. 37, pp. 861-83.

Sichel, D. (1993), “Business Cycle Asymmetry: A Deeper Look,” *Economic Inquiry*, Vol. 31, pp. 224-36.

Stout, W. F. (1974), *Almost Sure Convergence*, New York, Academic Press.

Tjøstheim, D. (1990), “Non-Linear Time Series and Markov Chains,” *Advances in Applied Probability,* Vol. 22, pp. 587-611.

Tong, H. (1990), *Non Linear Time Series: A Dynamical Systems Approach*, Oxford, Oxford University Press.

Tong, H. and K. Lim (1980), “Threshold Autoregression, Limit Cycles and Cyclical Data,” *Journal of the Royal Statistical Society Series B,* Vol. 42, pp. 245-92.

Tweedie, R. L. (1975), “Sufficient Condition for Ergodicity and Recurrence of Markov Chian on a General State Space,” *Stochastic Process and Their Applications*, Vol. 3, pp. 385-403.

Watson, M. (1986), “Univariate Detrending Methods with Stochastic Trends,” *Journal of Monetary Economics*, Vol. 18, pp. 49-75.

RECENTLY PUBLISHED "TEMI" (*)

No. 314 — *Strozzature settoriali, limiti alla crescita e pressioni inflazionistiche,* by P. CIPOLLONE and D. J. MARCHETTI (August 1997).

No. 315 — *Tassi bancari attivi e struttura finanziaria in Italia: un'analisi empirica,* by C. COTTARELLI, G. FERRI and A. GENERALE (September 1997).

No. 316 — *I canali di trasmissione della politica monetaria nel modello econometrico trimestrale della Banca d'Italia,* by S. NICOLETTI ALTIMARI, R. RINALDI, S. SIVIERO and D. TERLIZZESE (September 1997).

No. 317 — *Sensitivity of VaR Measures to Different Risk Models,* by F. DRUDI, A. GENERALE and G. MAJNONI (September 1997).

No. 318 — *Previsione delle insolvenze delle imprese e qualità del credito bancario: un'analisi statistica,* by S. LAVIOLA and M. TRAPANESE (September 1997).

No. 319 — *Da che cosa dipendono i tassi di interesse sui prestiti nelle province?,* by R. DE BONIS and A. FERRANDO (September 1997).

No. 320 — *Wherein Do the European and American Models Differ?,* by P. A. SAMUELSON (November 1997).

No. 321 — *Finance, Human Capital and Infrastructure: An Empirical Investigation of Post-War Italian Growth,* by G. FERRI and F. MATTESINI (November 1997).

No. 322 — *Long-Term Interest Rate Convergence in Europe and the Probability of EMU,* by I. ANGELONI and R. VIOLI (November 1997).

No. 323 — *Education, Infrastructure, Geography and Growth: An Empirical Analysis of the Development of Italian Provinces,* by S. FABIANI and G. PELLEGRINI (November 1997).

No. 324 — *Properties of the Monetary Conditions Index,* by G. GRANDE (December 1997).

No. 325 — *Style, Fees and Performance of Italian Equity Funds,* by R. CESARI and F. PANETTA (January 1998).

No. 326 — *Adverse Selection of Investment Projects and the Business Cycle,* by P. REICHLIN and P. SICONOLFI (February 1998).

No. 327 — *International Risk Sharing and European Monetary Unification,* by B. E. SØRENSEN and O. YOSHA (February 1998).

No. 328 — *The Behaviour of the Dollar and Exchange Rates in Europe: Empirical Evidence and Possible Explanations,* by P. DEL GIOVANE and A. F. POZZOLO (February 1998).

No. 329 — *Risultati e problemi di un quinquennio di politica dei redditi: una prima valutazione quantitativa,* by S. FABIANI, A. LOCARNO, G. ONETO and P. SESTITO (March 1998).

No. 330 — *La problematica della crescente fragilità nella "ipotesi di instabilità finanziaria" da una prospettiva kaleckiana,* by G. CORBISIERO (March 1998).

No. 331 — *Research and Development, Regional Spillovers, and the Location of Economic Activities,* by A. F. POZZOLO (March 1998).

No. 332 — *Central Bank Independence, Centralization of Wage Bargaining, Inflation and Unemployment: Theory and Evidence,* by A. CUKIERMAN and F. LIPPI (April 1998).

No. 333 — *La domanda di finanziamenti bancari in Italia e nelle diverse aree del Paese (1984-1996),* by D. FOCARELLI and P. ROSSI (May 1998).

No. 334 — *La politica fiscale nei paesi dell'Unione europea negli anni novanta,* by P. CASELLI and R. RINALDI (July 1998).

No. 335 — *Signaling Fiscal Regime Sustainability,* by F. DRUDI and A. PRATI (September 1998).

No. 336 — *Fiscal Consolidations under Fixed Exchange Rates,* by P. CASELLI (October 1998).

No. 337 — *Investimenti diretti all'estero e commercio: complementi o sostituti?,* by A. MORI and V. ROLLI (October 1998).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi – Divisione Biblioteca e pubblicazioni – Via Nazionale, 91 – 00184 Rome
(fax 39 6 47922059)