

A discusión

UNEMPLOYMENT AND HYSTERESIS: A NONLINEAR UNOBSERVED COMPONENTS APPROACH*

Silvestro Di Sanzo and Alicia Pérez-Alonso**

WP-AD 2005-34

Corresponding author: A. Pérez-Alonso. Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, Campus de Sant Vicent del Raspeig, Apartado de Correos 99, 03080 Alicante, Spain. E-mail: alicia@merlin.fae.ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Diciembre 2005
Depósito Legal: V-99-2006

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

* We would like to thank Máximo Camacho, Juan Mora and Gabriel Pérez-Quirós for their valuable comments, which resulted in a great improvement of an earlier version of this paper. Useful comments were also received from participants at the QED Jamboree in Copenhagen, the 20th Annual Congress of the European Economic Association (EEA) and seminar participants at the University of Alicante. All errors are our own responsibility.

** S. Di Sanzo and A. Pérez-Alonso: Departamento de Fundamentos del Análisis Económico, Universidad de Alicante.

**UNEMPLOYMENT AND HYSTERESIS:
A NONLINEAR UNOBSERVED COMPONENTS APPROACH**

Silvestro Di Sanzo and Alicia Pérez-Alonso

ABSTRACT

The aim of this paper is to find a possible hysteresis effect on unemployment rate series from Italy, France and the United States. We propose a definition of hysteresis taken from Physics which allows for nonlinearities. To test for the presence of hysteresis we use a nonlinear unobserved components model for unemployment series. The estimation methodology used can be assimilated into a threshold autoregressive representation in the framework of a Kalman filter. To derive an appropriate p-value for a test for hysteresis we propose two alternative bootstrap procedures: the first is valid under homoskedastic errors and the second allows for general heteroskedasticity. We investigate the performance of both bootstrap procedures using Monte Carlo simulation.

JEL Classification: C12; C13; C15; C32; E24.

Keywords: Hysteresis; Unobserved Components Model; Threshold Autoregressive Models; Nuisance parameters; Bootstrap.

1 Introduction

The business press make continual references to the high unemployment that characterizes European countries. For example, in France and Italy unemployment since the mid 1970s has steadily increased without any significant decrease or evident tendency to revert to a stable underlying unemployment rate. It remains very high at around 10% (see Figure 1). Many theories have emerged to provide an economic explanation which could account for this observed unemployment persistence. Most of the work in the relevant literature assumes that it can be attributed to changes in the natural rate of unemployment and/or changes in the cyclical rate of unemployment. Based on this framework, two main approaches are the natural rate theory and the unemployment hysteresis theory.

The first approach supposes that output fluctuations generate cyclical movements in the unemployment rate, which in the long run, will tend to revert to its equilibrium. The crux of the natural rate hypothesis is that cyclical unemployment and natural unemployment evolve independently. Hence, the tendency of the natural rate to remain at a high level is the result of permanent shocks on the structure of the labour market such as increased unemployment benefits, strong trade unions, legislative restrictions on dismissal and minimum wage laws (see Friedman [6]).

The second approach supposes that cyclical unemployment rate and natural rate do not evolve independently. The basic idea of the hysteresis hypothesis is that a change in the cyclical component of the unemployment rate may be permanently propagated to the natural rate. Based on this idea, an increase in the cyclical unemployment rate may lead to an increase, over time, in the level of the natural rate (see Phelps [22]). A direct corollary of hysteresis is that short-run adjustment of the economy can take place over a very long period. Then, aggregate demand policy, traditionally considered as ineffective in changing the natural rate of unemployment, can have a permanent effect on it.

In this paper, we focus on this second approach. The word hysteresis derives from the Greek $\nu\sigma\tau\epsilon\rho\epsilon\omega$, which means *to come later*. The physicist James Alfred Ewing was the first to introduce this term into scientific literature to explain the behaviour of electromagnetic fields in ferric metals. As pointed out in Amable *et al.* [1], a mathematical modelling of hysteresis requires us to consider a system subject to an external action, that is an input-output system. Hysteresis is defined as a particular type of response of the system when one modifies the value of the input: *the system is said to exhibit some remanence when there is a permanent effect on output after the value of the input has been modified and brought back to its initial position*. This formal definition implies that a hysteretic process is characterized by the following properties:

1. It is necessary to know the history of the system in order to assess its position. Hence, the history of the system matters. This implies the presence of a unit root in the process.

2. There is a remanence effect. If one transitory shock is followed by a second of the same intensity in the opposite direction the system does not revert to its former equilibrium. Hence, a transitory shock has a permanent effect on the system's equilibrium, since the system retains traces of past shocks on it even after those influences have ceased to apply. It must be noted that

this property is only present in non linear systems.

To sum up, hysteresis occurs in nonlinear systems that exhibit multiplicity of equilibria and the remanence property.

The first attempt to introduce a measure of hysteresis into unemployment theory was made by Blanchard and Summers [3], to explain the insider-outsider dynamics in wage formation. They argue that unemployment exhibits hysteresis when current unemployment depends on past values with the sum of their coefficients equal to or very close to unity. That is, hysteresis in unemployment arises when unemployment series has a unit root. The presence of a unit root in the process means it is path dependent. That is, any shock is entirely incorporated into the series level. Therefore, hysteresis is assimilated into the concept of "full persistence". Based on this framework, a great number of studies have investigated whether unemployment series exhibits a unit root, measuring persistence by the sum of coefficients in an autoregressive process with a constant mean value parameter (see, for example, León-Ledesma and McAdams [18] and Papell *et al.* [21]) or by means the coefficients of lagged unemployment in an ARMA(p,q) process (see Layard *et al.* [16], Sachs [25] and Summers [26]). Therefore, the dominant approach in the empirical literature to determine whether hysteresis exists focuses on testing for the existence of a unit root in a linear process.

Two problems arise with this kind of model. The first is that natural and cyclical shocks are summarized in the innovation with no distinction. As pointed out above, hysteresis in unemployment arises when a change in cyclical unemployment induces a permanent change in the natural rate. Having said that, the presence of a unit root in the unemployment rate is a necessary condition for the existence of hysteresis but not a sufficient one since the unit root could be generated by accumulation of natural shocks and be completely independent of whether there is hysteresis. Hence, separating the respective effects of transitory and permanent shocks on the natural rate of unemployment is the only way to assess if changes in it are due to cyclical (this is the case of hysteresis) or natural shocks or both. The second problem refers to the practice of checking for the presence of hysteresis using a linear model. A linear model lacks the property of remanence. Under linearity, a shock on the system followed by a second one of the same intensity in the opposite direction will bring the system back to its initial position. In this context, it is incorrect to use the term hysteresis, and we should refer rather to persistence. There is a major difference between persistence and hysteresis. In a system exhibiting persistence the response to impulses is a linear function, which is not the case for a system with hysteresis.

A number of papers have studied methods for checking for presence of hysteresis in a nonlinear framework. They employ a battery of unit root tests that control for the possible existence of nonlinear behaviour in unemployment series. Papell *et al.* [21] test for unit roots in autoregressive models with structural changes and León-Ledesma [17] implement a unit root test in a threshold autoregressive model. Though these models incorporate nonlinearities to model the behaviour of unemployment rate series they have the same weak point as the linear models described above: it is not possible to tell whether a change in the natural rate is due to transitory or permanent shocks.

So if our goal is to check for the presence of a hysteresis effect on the unemployment rate we

need a nonlinear econometric model that discriminates between natural and cyclical sources of influence on the unemployment rate.

Jaeger and Parkinson (henceforth JP, see [13]) put this idea into perspective and adopt an unobserved components (UC) model¹ to test the validity of the hysteresis hypothesis. They generate a pure statistical decomposition of the actual unemployment rate into a natural rate component and a cyclical component, which are both treated as latent variables. They also assume a particular structure to describe the variation over time of these latent variables. The hysteresis effect is introduced by allowing cyclical unemployment to have a lagged effect on the natural rate, which is assumed to contain a unit root. They only consider symmetric responses of the natural rate as regards cyclical unemployment fluctuations. In this way, they implicitly assume hysteresis is a linear phenomenon. This approach is insufficient to correctly identify hysteresis. Though it allows the different sources of shocks (i.e., cyclical or permanent) to be identified, it does not take into account the existence of nonlinear dynamics in unemployment series: this is necessary to capture the remanence property of a hysteretic process. Under JP's model it is only possible to establish whether delayed cyclical unemployment has a significant impact on the natural rate. This describes persistence but it does not correspond to hysteresis.

In order to take into account this nonlinear feature of a hysteretic process, we propose an extended version of JP's model. In particular, we allow past cyclical unemployment to have a different effect on the natural rate, which depends on the regime of the economy. It is thus possible to capture the stylized fact that natural rate does not decrease in cyclical expansion periods as much as it increases in cyclical recession periods. This provides a plausible explanation for the tendency of the natural rate to remain at a high level. The parameters of the model are estimated by maximum likelihood using a modified Kalman filter that incorporates the methodology implemented for the estimation of the threshold autoregressive (TAR) models (Tong [27]) in order to split the sample into two groups, which we may call regimes.

Under this new framework, the problem of testing for hysteresis becomes a problem of testing for linearity. The relevant null hypothesis is a one-regime model (i.e. the non presence of hysteresis) against the alternative of two regimes (i.e. the presence of hysteresis). The absence of a body of finite sample theory for nonlinear models means that empirical research must rely either on asymptotic theory or bootstrapping for inference. Testing for the econometric hypothesis of interest in the context of nonlinear models raises a particular problem known in statistics literature as hypothesis testing when a nuisance parameter is not identified under the null hypothesis (see, among others, Andrews and Ploberger [2], Chan [4] and Hansen [11]). In particular, the threshold parameter and the delay lag of the threshold variable are not identified under the null of linearity. If the model is not identified, the asymptotic distributions of standard tests are unknown, which means that tabulation of critical values is not possible. There is no shortcut solution for this problem. A trick to avoid the presence of unidentified nuisance parameters under the null hypothesis of interest is to let the variance of the natural shock be regime-dependent. Under this tenuous

¹See Harvey [12] for a detailed description of the Unobserved Component models.

assumption the model is identified and the chi-squared distributional approximation for standard tests statistics is valid. This argument appears in the existing literature to identify the transition probabilities of the Markov-Switching model (Hansen [9] and [10]; Karame [15]) under the null. It sounds convincing but confuses the relevant testing problem. That is, a linear model versus a nonlinear one. Under this formulation of the null hypothesis, even if we accept the null we are not accepting a linear model, since the variance is still a source of nonlinearity. Hence, to really test for the presence of nonlinearity in our model the variance must be kept constant while the mean parameter shifts across regimes.

Hansen [11] derives an asymptotic distribution free of nuisance parameters that is useful for testing and inference in TAR models. He shows that critical values are easily approximated via Monte Carlo simulation. As far as we know, no distributional theory is available to implement a linearity test in the framework of a UC model with nonlinear dynamics described by the TAR methodology. With this in mind, we propose a simulation method for calculating a bootstrap p -value for the relevant test of linearity. Then we use this method to check for the presence of hysteresis in Italy, France and the United States.

The paper proceeds as follows. Section 2 briefly describes JP's model and introduces an extended version that accounts for nonlinearity. Section 3 proposes two alternative bootstrap procedures to compute the p -value for a linearity test under the framework of interest. It also discusses the design of the Monte Carlo experiments that are used to investigate the small sample performance of the bootstrap version of the test and presents the results of the experiments. Empirical results for Italy, France and the United States are presented in Section 4. Section 5 concludes. Estimation methods are relegated to Appendix A. Appendix B contains all the tables and figures.

2 An extension of Jaeger and Parkinson's model

JP propose a pure statistical decomposition of the unemployment rate to evaluate the data for evidence of hysteresis effects. They assume the actual unemployment rate to be the sum of two unobservable components: a non-stationary natural rate component, U^N , and a stationary cyclical component, U^C ,

$$U_t = U_t^N + U_t^C. \quad (1)$$

In order to contemplate the necessary condition for hysteresis, i.e. the presence of a unit root in the process, they first test for the presence of a unit root in the unemployment series and then impose it in the model. Having said that, the natural rate component is defined as a random walk plus a term capturing possible hysteresis effects,

$$U_t^N = U_{t-1}^N + \alpha U_{t-1}^C + \epsilon_t^N. \quad (2)$$

Coefficient α measures, in percentage points, how much the natural rate increases if the economy experiences a cyclical unemployment rate of 1.0 percent. The size of this coefficient is their measure of hysteresis.

The cyclical² component of the unemployment rate is defined as a stationary second-order autoregressive process,

$$U_t^C = \phi_1 U_{t-1}^C + \phi_2 U_{t-2}^C + \epsilon_t^C. \quad (3)$$

The system is completed by augmenting the model with a version of Okun's law, which relates cyclical unemployment and output growth,

$$D_t = \sum_{k=1}^p \beta_k D_{t-k} + \delta U_t^C + v_t, \quad (4)$$

where D_t stands for the output growth rate at date t . Equation (4) defines the output growth rate as an autoregressive process of order p plus a term capturing the influence of the cyclical rate of unemployment. Since the cyclical component is assumed to be stationary, we consider U_t^C instead of ΔU_t^C as in JP's model in order to avoid a problem of over-differentiation. Although JP suggested setting $p = 1$, in our empirical work we choose the value of p that fits the data well. In particular, we find $p = 2$.

The disturbances ϵ_t^N , ϵ_t^C and v_t are assumed to be mutually uncorrelated shocks which are normally distributed with variances σ_N^2 , σ_C^2 and σ_v^2 , respectively.

In order to test the hysteresis hypothesis, i.e. past cyclical movements on unemployment have a permanent impact on the natural rate, JP perform a significance test on parameter α ,

$$H_0 : \alpha = 0.$$

If parameter α is significantly different from zero, they argue there exists a hysteresis effect on the unemployment rate. It is important to note that JP's model is linear, since it implies that past cyclical unemployment changes have the same impact, in absolute terms, on the natural unemployment rate. For example, a variation in the cyclical component of (1%) or (-1%) causes a variation in the natural rate of ($\alpha\%$) or ($-\alpha\%$) respectively. Again, we remark that this linear context lacks the property of remanence, so it is not possible to observe hysteresis. We would do better to refer to persistence rather than hysteresis.

At this point, we want to relax the assumption of linearity and we introduce nonlinearities into JP's model. This extension allows us to detect whether hysteresis is present in unemployment series. Nonlinearities are introduced by allowing past cyclical unemployment to have a different impact on the natural rate, which depends on the regime of the economy. To that end, equation (2) becomes

$$U_t^N = \begin{cases} U_{t-1}^N + \alpha_1 U_{t-1}^C + \epsilon_t^N & \text{if } q_{t-1} \geq \gamma \\ U_{t-1}^N + \alpha_2 U_{t-1}^C + \epsilon_t^N & \text{if } q_{t-1} < \gamma \end{cases}, \quad (2')$$

where q_t is the threshold variable and γ stands for the threshold parameter. Equations (1), (3) and (4) remain the same together with assumptions about shocks.

²As in JP, we find that AR(2) processes for the cyclical component fit the data well for all the countries under study.

This kind of model is estimated via maximum likelihood in the framework of the Kalman filter³. We employ a modified Kalman filter in order to incorporate a deterministic cut-off of the sample that corresponds to a raw indicator for favorable and unfavorable periods, which is based on the methodology implemented for the estimation of TAR models. We choose the long difference $U_{t-1} - U_{t-d}$, with $d \in \{2, 3\}$, as our threshold variable. This variable is an indicator of the state of the economy to identify the regimes. The integer d is called the threshold delay lag. Whether this variable is lower or higher than the threshold parameter γ determines whether an observation belongs to one regime or the other. We consider an economy with two regimes, one related to high long differences (regime 1), i.e. an *unfavorable* regime, and the other with low long differences (regime 2), i.e. a *favorable* regime. Parameters d and γ are unknown so they must be estimated along with the other parameters. The maximization is best solved through a grid search over two-dimensional space (γ, d) . To execute a grid search we need to fix a region over which to search. It is important to restrict the set of threshold candidates *a priori* so that each regime contains a minimal number of observations. We restrict the search to values of γ lying on a grid between τ th and $(1 - \tau)$ th quantiles of q_{t-1} for each value of d . In our applications we choose $\tau = 0.30$. Then we estimate the model for each pair (γ, d) belonging to this grid and retain the one that provides the highest log-likelihood value.

As mentioned in the previous section, in this context a test for hysteresis becomes a test for linearity, i.e. a test for a single regime against the alternative of two regimes. The null hypothesis we are interested in is

$$H_0 : \alpha_1 = \alpha_2.$$

At this point, a remark is needed. If the unemployment rate displays a nonlinear behaviour, JP's model is misspecified and any inference based on the parameters of this model may lead us to wrong conclusions. This reflection suggests that the following testing strategy should be implemented.

The starting point is the extended JP model, where we test the null hypothesis $\alpha_1 = \alpha_2$. If we reject it, we are accepting the presence of hysteresis in unemployment series. If it is not rejected, hysteresis is not present in the unemployment rate but there is still a place for the presence of persistence. Once this point is reached, the next step is to estimate the linear model proposed by JP and test for persistence following the strategy they propose.

As pointed out in the previous section, when we perform the test of linearity a problem of unidentified nuisance parameters under the null hypothesis arises. That is, there exists a set of parameters that are not restricted under the null hypothesis. In particular, the null hypothesis $\alpha_1 = \alpha_2$ does not restrict the threshold parameter γ and the delay d . As a result, conventional statistics do not have an asymptotic standard distribution. In order to circumvent this problem, we employ a bootstrap technique to compute the p -value associated with the test of interest.

³See Appendix A for a detailed description of this estimation methodology.

3 Experimental design for computing the bootstrap p -value for the linearity hypothesis test

Our aim in this section is to approximate the distribution of the test statistic of interest by a consistent bootstrap procedure. In particular, we implement a Wald test statistic. The difficulty is that there is no well-accepted bootstrap method that is appropriate in the present framework. We propose two bootstrap procedures: the first is valid if the errors in our model are homoskedastic and the second allows for the presence of general heteroskedasticity.

3.1 Homoskedastic bootstrap

STEP 1: We estimate the *supW* test. To compute this test we need only to estimate the model under the alternative hypothesis of nonlinearity. The parameters of interest are $\{\theta_1 = (\sigma_N, \sigma_C, \sigma_v, \alpha_1, \alpha_2, \phi_1, \phi_2, \delta, \beta_1, \beta_2, \dots, \beta_p)', \gamma, d\}$. For each given value of (γ, d) belonging to the grid Δ described in the previous section, we obtain the maximum likelihood estimates $\widehat{\theta}_1(\gamma, d) = (\widehat{\sigma}_N, \widehat{\sigma}_C, \widehat{\sigma}_v, \widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{\phi}_1, \widehat{\phi}_2, \widehat{\delta}, \widehat{\beta}_1, \widehat{\beta}_2, \dots, \widehat{\beta}_p)'$ and compute the pointwise Wald test statistic,

$$W(\gamma, d) = R\widehat{\theta}_1(\gamma, d)(R\widehat{Var}(\widehat{\theta}_1(\gamma, d))R')^{-1}(R\widehat{\theta}_1(\gamma, d))',$$

where $R = (0 \ 0 \ 0 \ 1 \ -1 \ 0 \ 0 \ 0 \ 0 \dots 0)$ is an $(1 \times (8 + p))$ vector, and $\widehat{Var}(\widehat{\theta}_1)$ is the heteroskedasticity-robust maximum likelihood estimator of the variance-covariance matrix. Then, arguing from the union-intersection principle, Davies [5] proposes the statistic

$$supW = \sup_{(\gamma, d) \in \Delta} W(\gamma, d).$$

STEP 2: We compute the residuals under the null hypothesis of linearity: $\alpha_1 = \alpha_2$. Under the null, the model is reduced to JP's specification. We construct estimates of the vector of unknown parameters $\theta_0 = (\sigma_N \ \sigma_C \ \sigma_v \ \alpha \ \phi_1 \ \phi_2 \ \delta \ \beta_1 \ \beta_2 \dots \beta_p)'$ and the vector of unobserved variables $(U_t^N \ U_t^C)'$ using the Kalman filter methodology. Let $\widetilde{\theta}_0 = (\widetilde{\sigma}_N \ \widetilde{\sigma}_C \ \widetilde{\sigma}_v \ \widetilde{\alpha} \ \widetilde{\phi}_1 \ \widetilde{\phi}_2 \ \widetilde{\delta} \ \widetilde{\beta}_1 \ \widetilde{\beta}_2 \dots \widetilde{\beta}_p)'$ and $(\widetilde{U}_t^N \ \widetilde{U}_t^C)'$ denote the maximum likelihood estimates of coefficients and unobserved components, respectively. We compute the residuals

$$\begin{aligned} \widetilde{\epsilon}_t^C &= \widetilde{U}_t^C - \widetilde{\phi}_1 \widetilde{U}_{t-1}^C - \widetilde{\phi}_2 \widetilde{U}_{t-2}^C; \\ \widetilde{\epsilon}_t^N &= \widetilde{U}_t^N - \widetilde{U}_{t-1}^N - \widetilde{\alpha} \widetilde{U}_{t-1}^C; \\ \widetilde{v}_t &= D_t - \widetilde{\beta}_1 D_{t-1} - \widetilde{\beta}_2 D_{t-2} - \dots \widetilde{\beta}_p D_{t-p} - \widetilde{\delta} \widetilde{U}_t^C; \end{aligned}$$

where $t = \max\{2, p\} + 1, \dots, T$. Let T^* denote the sample size for the bootstrap.

STEP 3: We generate B bootstrap samples $Z_{T^*b}^* = \{Z_{tb}^* = (U_t^*, D_t^*)' : t = 1, \dots, T^*\}$, $b = 1, \dots, B$, by first generating unobserved bootstrap components

$$\begin{aligned}
U_t^{*N} &= U_{t-1}^{*N} + \tilde{\alpha}U_{t-1}^{*C} + \epsilon_t^{*N}; \\
U_t^{*C} &= \tilde{\phi}_1U_{t-1}^{*C} + \tilde{\phi}_2U_{t-2}^{*C} + \epsilon_t^{*C};
\end{aligned}$$

and next computing the bootstrap version of the series

$$\begin{aligned}
U_t^* &= U_t^{*N} + U_t^{*C}; \\
D_t^* &= \tilde{\beta}_1D_{t-1}^* + \tilde{\beta}_2D_{t-2}^* + \dots + \tilde{\beta}_pD_{t-p}^* + \tilde{\delta}U_t^{*C} + v_t^*;
\end{aligned}$$

where the bootstrap errors $\{\epsilon_t^{*C}, t = 1, \dots, T^*\}$, $\{\epsilon_t^{*N}, t = 1, \dots, T^*\}$ and $\{v_t^*, t = 1, \dots, T^*\}$ are independent values obtained by resampling, with replacement, from the set of residuals under H_0 . The parameters used to construct the data are the parameter values estimated in Step 2. We need to establish a set of initial values for $(U_0^{*N}, U_0^{*C}, U_{-1}^{*C}, D_0^*, D_{-1}^*, \dots, D_{-(p-1)}^*)$. We take the simple approach of conditioning, where applicable, on the observed values. Hence, the underlying natural rate at time 0 is assumed to be equal to the observed value of the unemployment series at time 0, whereas the cyclical elements are assumed to be zero. That is, $(U_0, 0, 0, D_0, D_{-1}, \dots, D_{-(p-1)})$. These initial conditions are kept fixed throughout the bootstrap replications.

STEP 4: Each bootstrap sample $\{Z_{Tb}^* : b = 1, \dots, B\}$ is then used to re-estimate the parameters under H_1 . The algorithm employed to estimate the bootstrap threshold parameter γ^* and the delay lag d^* proceed as in Step 1, where the threshold variable is given by $U_{t-1}^* - U_{t-d}^*$. Let $\hat{\theta}_1^*(\gamma^*, d^*)$ denote the estimator of θ_1 when using the bootstrap sample. We then compute the pointwise Wald test statistic associated with the bootstrap sample as

$$W^*(\gamma^*, d^*) = R\hat{\theta}_1^*(\gamma^*, d^*)(R\widehat{Var}(\hat{\theta}_1^*(\gamma^*, d^*)))R')^{-1}(R\hat{\theta}_1^*(\gamma^*, d^*))',$$

and the $supW^*$ test as

$$supW^* = \sup_{(\gamma^*, d^*) \in \Delta^*} W^*(\gamma^*, d^*).$$

STEP 5: Repeating this for $b = 1, \dots, B$ gives a sample $\{supW^* : b = 1, \dots, B\}$ of $supW$ values. This sample mimics a random sample of draws of $supW$ under the null hypothesis. We compute the bootstrap p -value as $p_B = card(supW^* \geq supW)/B$, that is the fraction of $supW^*$ values that are greater than the observed value $supW$.

We carry out $B = 1000$ bootstrap replications.

3.2 Heteroskedastic bootstrap

Our aim here is to calculate a bootstrap distribution of the Wald test allowing for the possibility of general heteroskedasticity. The algorithm is similar to the one described above, but replacing

the resampling scheme in Step 3. In particular, the resampling we propose is based on the idea of the wild bootstrap, which was studied for the first time by Wu [28] in the context of variance estimation in heteroskedastic linear models. In our context, it looks like this:

Step 3': We generate B bootstrap samples $Z_{T^*b}^* = \{Z_{tb}^* = (U_t^*, D_t^*)' : t = 1, \dots, T^*\}$, $b = 1, \dots, B$.

To do this, we propose the following algorithm:

I. Generate η_t^N independent and identically distributed variables from a fixed distribution⁴ such that $E(\eta_t^N) = 0$ and $E[(\eta_t^N)^2] = E[(\eta_t^N)^3] = 1$. Define $\epsilon_t^{N*} = \tilde{\epsilon}_t^N \eta_t^N$, where $\tilde{\epsilon}_t^N$ is the t th residual calculated in Step 2. The bootstrap error ϵ_t^{N*} satisfies $E^*(\epsilon_t^{N*}) = 0$ and $E^*[(\epsilon_t^{N*})^2] = (\tilde{\epsilon}_t^N)^2$. Similarly, generate η_t^C and η_t^v and construct $\epsilon_t^{C*} = \tilde{\epsilon}_t^C \eta_t^C$ and $v_t^* = \tilde{v}_t \eta_t^v$.

II. We set the initial conditions $(U_0^{*N}, U_0^{*C}, U_{-1}^{*C}) = (U_0, 0, 0)$ and, for $t = 1, 2, \dots, T^*$, set $(U_t^{*N}, U_t^{*C}) = (\tilde{U}_t^N, \tilde{U}_t^C)$, that is, unobserved bootstrap components are generated with conditionally set design on the estimated unobserved components in Step 2,

$$\begin{aligned} U_t^{*N} &= \tilde{U}_{t-1}^N + \tilde{\alpha} \tilde{U}_{t-1}^C + \epsilon_t^{N*}; \\ U_t^{*C} &= \tilde{\phi}_1 \tilde{U}_{t-1}^C + \tilde{\phi}_2 \tilde{U}_{t-2}^C + \epsilon_t^{C*}. \end{aligned}$$

III. To define the bootstrap observations, we set initial conditions $(D_0^*, D_{-1}^*, \dots, D_{-(p-1)}^*) = (D_0, D_{-1}, \dots, D_{-(p-1)})$ and use a conditional resampling on (D_1, \dots, D_{T^*}) ,

$$\begin{aligned} U_t^* &= U_t^{*N} + U_t^{*C}; \\ D_t^* &= \tilde{\beta}_1 D_{t-1} + \tilde{\beta}_2 D_{t-2} + \dots + \tilde{\beta}_p D_{t-p} + \tilde{\delta} U_t^{*C} + v_t^*. \end{aligned}$$

The parameters used to construct the data are the parameter values estimated in Step 2.

3.3 Monte Carlo evidence

In this section we report on a Monte Carlo simulation study designed to evaluate the small sample performance of both homoskedastic and heteroskedastic bootstrap procedures in the problem of testing for linearity. We start with a brief description of the design of the experiment, then we proceed with the discussion of the results.

3.3.1 Design of the Experiment

In each Monte Carlo experiment, we generate observations from the linear state-space model

$$\begin{aligned} \begin{bmatrix} U_t \\ D_t \end{bmatrix} &= \begin{bmatrix} 1 & 1 & 0 \\ 0 & \delta & 0 \end{bmatrix} \begin{bmatrix} U_t^N \\ U_t^C \\ U_{t-1}^C \end{bmatrix} + \begin{bmatrix} 0 \\ \beta \end{bmatrix} D_{t-1} + \begin{bmatrix} 0 \\ v_t \end{bmatrix} \\ \begin{bmatrix} U_t^N \\ U_t^C \\ U_{t-1}^C \end{bmatrix} &= \begin{bmatrix} 1 & \alpha & 0 \\ 0 & \phi_1 & \phi_2 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} U_{t-1}^N \\ U_{t-1}^C \\ U_{t-2}^C \end{bmatrix} + \begin{bmatrix} \epsilon_t^N \\ \epsilon_t^C \\ 0 \end{bmatrix}, \end{aligned} \tag{5}$$

⁴In particular, the variable η_t^N was sample from Mammen's ([19], p.257) two-point distribution attaching masses $(5 + \sqrt{5})/10$ and $(5 - \sqrt{5})/10$ at the points $-(\sqrt{5} - 1)/2$ and $(\sqrt{5} + 1)/2$, respectively.

with ϵ_t^N i.i.d. $N(0, \sigma_N^2)$, ϵ_t^C i.i.d. $N(0, \sigma_C^2)$ and v_t i.i.d. $N(0, \sigma_v^2)$. In order to illustrate the performance of the bootstrap test under the alternative of nonlinearity, we also consider an extension of model (5), allowing coefficient α to differ if $U_{t-1} - U_{t-d} \geq \gamma$ or $U_{t-1} - U_{t-d} < \gamma$ as in (2'). Let M_0 and M_1 denote the class of linear and nonlinear state-space models, respectively.

We test the null hypothesis of linearity. As discussed at the end of Section 2, the null hypothesis is true if and only if $\alpha_1 = \alpha_2$. Hence, M_0 is nested in M_1 . We use the statistic $supW$ based on an estimated M_1 setting $d = 1$, and compute the p -value using both the homoskedastic and the heteroskedastic bootstrap procedures with $B = 99$ bootstrap replications. The estimation of the rejection probabilities is calculated from $R = 500$ simulation runs. The processing time becomes excessive when greater values of B or R are used.

We first explore the size of the supW statistic under the null hypothesis of a single regime. This involves generating data from the linear model M_0 . The parameter values for the M_0 data-generating process (DGP) are $\sigma_N = 0.35$, $\sigma_C = 0.04$, $\sigma_v = 0.68$, $\delta = 10.23$, $\beta = 0.33$, $\phi_1 = -0.007$, $\phi_2 = -1.13 \times 10^{-5}$ and $\alpha = 0.36$. We next explore the power of the test against the two-regime alternative. Hence, we generate data from M_1 . In the case of the M_1 DGP, the parameter values are $\sigma_N = 0.5$, $\sigma_C = 0.05$, $\sigma_v = 0.7$, $\delta = 2.56$, $\beta = 0.24$, $\phi_1 = 0.14$, $\phi_2 = 0.8$, $\alpha_2 = 0.8$ and $\gamma = 0.07$. We vary α_1 between (1, 1.2, 1.4, 1.6, 1.8). In both cases, the selected parameters are chosen according to the corresponding estimated model for Italy.

The experiments proceed by generating artificial series of length $T + 50$ according to M_0 or M_1 with $T = 150$, and initial values set to zero. We then discard the first 50 pseudo-data points in order to minimize the effect of the initial conditions and the remaining T points are used to compute the test statistic. We simulate the proportion of rejections of the test at the 5%, 10% and 20% significance levels.

3.3.2 Simulation results

In Table 1 we present simulation evidence concerning the empirical size and power of the test. We observe a reasonable approximation of the level at all significance levels considered. It is interesting to note that the heteroskedastic bootstrap design yields a slightly better approximation of the nominal level than the homoskedastic design. Deviations from the null hypothesis are detected with high probability across the various parameterizations. We observe that in all cases under consideration the test based on the homoskedastic bootstrap approach yields substantially lower rejection probabilities than the heteroskedastic bootstrap test. It should be emphasized that this happens even though the model generated is homoskedastic. As expected, the performance of both bootstrap procedures improves as the difference between the values of parameters in the two regimes increases.

4 Empirical results

Our study concerns Italy, France and the United States. The economic series employed are the quarterly unemployment rate (U) and real gross domestic product (GDP). Data for Italy (running from 1970:1 through 2002:1), France (1978:1-2002:1) and U.S. (1965:1-2002:1) come from OECD *Main Economic Indicators*. All data are obtained as seasonally adjusted and all the variables except the unemployment rate are in natural logs.

We have decomposed the unemployment rate assuming that the natural rate contains a unit root. This assumption must be tested. To do this, we employ the Phillips-Perron test for unit roots. We obtain that unemployment rates display non-stationary behavior for all countries. We also perform the unit root test for the GDP series, which also displays non-stationary behaviour for all countries. Results are presented in Table 2.

Tests for hysteresis are reported in Table 3. The p -values presented in Table 3 are calculated following the bootstrap technique described in Section 3. Diagnosis checking of the residuals of the linear model⁵ leads us to implement a heteroskedastic bootstrap for the cases of Italy and the U.S. and a homoskedastic bootstrap for France. According to bootstrap p -values, the hysteresis effect is significant at the 5% level for Italy and France.

Results concerning the estimated models for Italy and France are available in Tables 4 and 5. At this point it should be emphasized that in this paper we describe an algorithm for implementing the maximum likelihood estimation of the parameters of the nonlinear model, but this is not a theory of inference. We do not provide a proof of consistency of the estimators, or a distribution theory. Hence, parameters and standard errors should be interpreted somewhat cautiously. For the case of Italy, the maximum likelihood estimate of the threshold parameter is $\hat{\gamma} = 0.07$, with a 90% bootstrap confidence interval $[-0.008, 0.143]$. Our estimate of the delay parameter is $\hat{d} = 2$. Hence, the threshold model splits the regression into two regimes depending on whether or not the threshold variable is higher than this threshold parameter. That is, we consider we are in regime 1 when $U_{t-1} - U_{t-2} \geq 0.07$ and in regime 2 when $U_{t-1} - U_{t-2} < 0.07$ (see Figure 2). For Italy, there are less observations in regime 1 (44%) than in regime 2 (56%), which means that this country spent more periods of time in the *favorable* regime. This is also the case for France. Analyzing the estimated hysteresis parameter, we observe a point of great interest. Both parameters are positive, and the one associated with Regime 1 is greater than that of Regime 2. This points to asymmetric responses of the natural rate as regards cyclical unemployment movements in the following direction: the natural rate does not decrease in *favorable* cyclical periods as much as it increases in *unfavorable* cyclical periods. This agrees with the economic intuition of the hysteresis mechanism. For example, for the case of Italy, the natural rate does not decrease as much when cyclical unemployment decreases (0.797%) as it rises when cyclical unemployment rises (1.366%).

It is worth analyzing the U.S. separately. The information concerning the model estimated

⁵The assumptions underlying the errors of the linear model are tested via appropriate autocorrelation, heteroskedasticity and normality test statistics, which are available from the authors upon request. We find evidence in favour of non-autocorrelation in all countries. Evidence against homoskedasticity is only found in Italy and the U.S..

is provided in Table 6. According to the hysteresis test, we cannot reject the null of linearity. However, as we mentioned in Sections 1 and 2, there is still a place for persistence. In fact, we find evidence in favour of it, given that parameter α is significant at 5%. Hence, though there is no hysteresis, cyclical shocks have a significant impact on the natural rate. In particular, if the economy experiences a cyclical unemployment rate of 1% the natural rate increases by 1.337%.

5 Implications for the theory of unemployment

At this point, it is important to emphasize that in this paper we focus on the properties of unemployment series from a purely statistical point of view. We neither derive an economic explanation which could account for any of the stylized facts of unemployment nor estimate any particular model drawn from a well defined economic theory. Having said that, it is interesting to address whether our empirical results agree with the recent theoretical economic literature about hysteresis.

Many mechanisms are listed in economic literature as giving rise to hysteresis. For instance, after a negative shock, firms may reduce capital stock along with employment. The latter will cause unemployment to persist because the firm may not re-open its plant once the shock is removed. Second, long periods of unemployment may cause workers to lose skill, which could lead to long-term unemployed workers losing the possibility of returning to the labour market. Moreover, long term unemployment may have a demoralising effect on search behaviour, contributing to a less efficient matching process. Third, after a negative shock, the insider (currently employed worker) has the power to push up wages due to the cost to the firm of labour turnover and this increase in wages may permanently raise the unemployment rate (insider-outsider theory, see Blanchard and Summers [3]). Therefore, a cyclical shock that reduces the number of insiders leads to a permanent change in the natural rate.

Recently, many authors have provided theoretical models to account for these mechanisms. These models are based on the idea that hysteresis is a nonlinear phenomenon associated with multiple stable equilibria (see Røed [24] and Mortensen [20]). In models of this kind, shocks can move unemployment from one equilibrium to another. For instance, Pissarides [23] shows how search activity coupled with some kind of insider-wage determination causes hysteresis in a bargaining model. If insider-workers set wages, we expect wages to be higher the lower the average level of search activity among the unemployed is. A search decrease among those currently unemployed implies a high probability of employment for those currently in work if they lose their jobs. Since each period of unemployment contributes to the demoralisation of the work-force, the present optimal wage is positively related to past rates of unemployment. Under certain conditions, large shocks may move the economy from a low-unemployment equilibrium to high-unemployment equilibrium. If these theoretical models are correct, we should expect high levels of persistence and frequent unemployment equilibrium changes in the face of shocks that have affected European countries.

In this paper, we find hysteresis in all the European countries included in our analysis. This

suggests that the dynamic of the unemployment rate is characterized by nonlinear behaviour with frequent shifts from one equilibrium to the other. We conclude that theoretical models with multiple natural equilibria appear consistent with the data.

6 Conclusions

The aim of this paper is twofold. First, to take into account the nonlinear feature of a hysteretic process we propose a definition of hysteresis taken from Physics. To provide an operational statistical framework for our concept of hysteresis we use the unobserved components approach, which decomposes unemployment rate into a non-stationary natural component and a stationary cyclical component. We extend the model of Jaeger and Parkinson [13] by introducing nonlinearities into the specification of the natural rate component. We do this by allowing past cyclical unemployment to have a different effect on the current natural rate depending on the regime of the economy. To estimate the model we use a modified Kalman filter that incorporates a sample partition that corresponds to two different regimes. The procedure for identifying these regimes is related to the TAR methodology. Under this framework, a test for hysteresis becomes a test for linearity. Second, when we implement a test for linearity a problem of unidentified nuisance parameters arises since the threshold parameter and the delay lag of the threshold variable are only identified under the alternative hypothesis of hysteresis. As a result, the standard asymptotic distributions of the classical tests are unknown under the null. Our objective is to implement a correct test for the relevant null hypothesis of a one-regime model. We rely on bootstrapping techniques to calculate an appropriate p-value for the decision rule. We propose two bootstrap procedures: the first is valid if the errors in our model are homoskedastic and the second one allows for general forms of heteroskedasticity. In a Monte Carlo simulation study, both bootstrap approximations of the linearity test are investigated in greater detail, and we find that they work quite well. Our study concerns Italy, France and the United States. For European countries, we reject the null of linearity. This is related to the presence of hysteresis. On the other hand, for the United States we reject the hysteresis hypothesis. We find symmetric responses of the natural rate as regards cyclical fluctuations in unemployment.

References

- [1] Amable, B., Henry, J., Lordon, F. & Topol, R. (1995), *Hysteresis revisited: a Methodological Approach*. In Cross, R. (ed): The natural rate of Unemployment. Cambridge University Press.
- [2] Andrews, D.W.K. and Ploberger, W. (1994), *Optimal tests when a nuisance parameter is present only under the alternative*. *Econometrica* 62(6), 1383-1414.
- [3] Blanchard, O.J. & Summers, L.H. (1986), *Hysteresis and the European Unemployment Problem*. NBER Macroeconomics Annual, Vol. I, MIT Press, Cambridge MA, 15-78.
- [4] Chan, K.S. (1990), *Testing for Threshold Autoregression*. *The Annals of Statistics* 18(4), 1886-1894
- [5] Davies, R.B. (1987), *Hypothesis testing when a nuisance parameter is present only under the alternative*. *Biometrika* 74(1), 33-43.
- [6] Friedman, M. (1968), *The Role of Monetary Policy*. *The American Economic Review*, 1-17.
- [7] Hall, P. (1992), *The Bootstrap and Edgeworth expansion*. Springer Series in Statistics.
- [8] Hamilton, J.D. (1994), *Time Series Analysis*. Princeton.
- [9] Hansen, B.E. (1992), *The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP*. *Journal of Applied Econometrics*, 7, S61-S82.
- [10] Hansen, B.E. (1996), *Erratum: The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP*. *Journal of Applied Econometrics*, 11(2), 195-98.
- [11] Hansen, B.E. (1996), *Inference when a nuisance parameter is not identified under the null hypothesis*. *Econometrica* 64(2), 413-30.
- [12] Harvey, A.C. (1989), *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
- [13] Jaeger, A. and Parkinson, M. (1994), *Some evidence on Hysteresis in Unemployment Rates*. *European Economic Review* 38, 329-42.
- [14] Kalman, R.E. (1960), *A New Approach to Linear Filtering and Prediction Problems*. *Journal of Basic Engineering, Transactions of the ASME Series D* 82, 35-45.
- [15] Karamé, F. (1999), *Unemployment Persistence: the Hysteresis Assumption Revisited*. Mimeo EUREQua.
- [16] Layard, R., Nickell, S. & Jackman, R. (1991), *Unemployment, Macroeconomic Performance and the Labour Market*. Oxford University Press, Oxford.

- [17] León-Ledesma, M. (2002), *Nonlinearities and unit roots in OECD unemployment*. Revista Brasileira de Economia de Empresa (Brazilian Review of Economics and Business), 2(3), 23-30.
- [18] León-Ledesma, M. & McAdam, P. (2004), *Unemployment, Hysteresis and Transition*. Scottish Journal of Political Economy, 51(3), 277-401.
- [19] Mammen, E. (1993), *Bootstrap and Wild Bootstrap for High-Dimensional Linear Models*. The Annals of Statistics, 21, 255-85.
- [20] Mortensen, D.T. (1989), *The persistence and indeterminacy of unemployment in search equilibrium*. Scandinavian Journal of Economics, 91, 347-70.
- [21] Papell, D.H., Murray, C.J. & Ghiblawi, H. (2000), *The structure of unemployment*. The Review of Economics and Statistics, 82(2), 309-315.
- [22] Phelps, E.S. (1972), *Inflation Policy and Unemployment theory: The Cost-Benefit Approach to Monetary Planning*. W.W. Norton, New-York.
- [23] Pissarides, C.A. (1986), *Unemployment and vacancies in Britain*. Economic Policy 3, 499-540.
- [24] Røed, K. (1997), *Hysteresis in Unemployment*. Journal of Economic Surveys 11, 389-418.
- [25] Sachs, J. (1987), *High unemployment in Europe Diagnosis and policy implications*. H. Claes (ed.), Unemployment in Europe, Timbro, Stockholm.
- [26] Summers, L.H. (1990), *Understanding Unemployment*. MIT Press, Cambridge, MA.
- [27] Tong, H. (1990), *Nonlinear Time Series: a Dynamical System Approach*. Oxford University Press.
- [28] Wu, C.F.J. (1986), *Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis*. The Annals of Statistics, 14, 1261-95.

Appendix A: Estimation procedures

In this appendix we present different filters which have been proposed in the relevant literature for estimating the sort of model described in Section 2. Firstly, we examine the Kalman filter, which allows us to estimate JP's model. Secondly, we introduce the threshold Kalman filter, which is a Kalman filter modified to include a threshold state equation.

The Kalman filter

In 1960, R.E. Kalman [14] published a famous paper describing a recursive solution to the discrete data linear filtering problem. Since that time, due largely to advances in digital computing, the Kalman filter has been the subject of extensive research and applications, particularly in the area of autonomous or assisted navigation.

The Kalman filter⁶ is a set of mathematical equations that provides an efficient recursive computational procedure for estimating the state of a process, in a way that minimizes the mean squared error (MSE). The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the system modelled is unknown.

To start with, consider an $(n \times 1)$ vector of observed variables at date t , y_t . These observable variables are related to a possibly unobserved $(r \times 1)$ vector h_t , known as the *state vector*, via a *measurement equation*,

$$y_t = H'h_t + A'x_t + w_t, \quad (6)$$

where H' and A' are matrices of parameters of dimension $(n \times r)$ and $(n \times k)$, respectively, x_t is an $(k \times 1)$ vector containing exogenous or lagged dependent variables, and the $(n \times 1)$ vector w_t is a white noise disturbance vector with covariance matrix given by:

$$E(w_t w'_\tau) = \begin{cases} R & \text{for } t = \tau \\ 0 & \text{otherwise} \end{cases},$$

where R is an $(n \times n)$ matrix.

Despite the fact that the variables of h_t are, in general, not observable, they are known to be generated by a first-order Markov process,

$$h_t = Fh_{t-1} + \Pi'x_t + v_t, \quad (7)$$

where F and Π' are matrices of parameters of dimension $(r \times r)$ and $(r \times k)$, respectively. The $(r \times 1)$ vector v_t is a white noise disturbance vector:

$$E(v_t v'_\tau) = \begin{cases} Q & \text{for } t = \tau \\ 0 & \text{otherwise} \end{cases},$$

⁶See Hamilton ([8], Chapter 13) and Harvey ([12], Chapter 3) for a more detailed description of the Kalman filter.

where Q is an $(r \times r)$ matrix. Equation (7) is the *transition equation*.

The disturbances v_t and w_t are assumed to be uncorrelated at all lags:

$$E(v_t w'_\tau) = 0 \text{ for all } t \text{ and } \tau.$$

Further assumptions on measurement and transition disturbances are as follows: i) they are uncorrelated with the exogenous variables; ii) they are assumed to be normally distributed in order to calculate the likelihood function.

The state space form (SSF) that represents the dynamics of the univariate time series y_t is composed of equations (6) and (7). The information set at time $t - 1$ is given by matrix $\Psi_{t-1} \equiv (y'_{t-1}, y'_{t-2}, \dots, y'_1, x'_{t-1}, x'_{t-2}, \dots, x'_1)'$. Note that there are two set of unknowns: the parameters of the model in H' , A' , R , F , Π' and Q (these matrices will be referred as the system matrices) and the elements of the state vector h_t . The goal of the Kalman filter procedure is to form a forecast of the unobserved state vector at time t based on the information at date $t - 1$. For now, we will assume that the particular numerical values of the system matrices are known. Let $\hat{h}_{t|t-1}$ denote the linear forecast of the state vector h_t based on (x_t, Ψ_{t-1}) , and $P_{t|t-1}$ denote the MSE matrix associated with this forecast.

Because the filter is a recursion, it is started assuming initial values for the mean and variance of the state variables, $\hat{h}_{1|0}$ and $P_{1|0}$, respectively. We can therefore conduct the Kalman filter in four major steps. First, we calculate the one-period-ahead forecast of the unobserved state vector and the associated MSE matrix :

$$\begin{aligned} \hat{h}_{t|t-1} &= E[h_t | x_t, \Psi_{t-1}] = F\hat{h}_{t-1|t-1} + \Pi'x_t, \\ P_{t|t-1} &= E[(h_t - \hat{h}_{t|t-1})(h_t - \hat{h}_{t|t-1})' | \Psi_{t-1}] = FP_{t-1|t-1}F' + Q. \end{aligned}$$

The next step is to calculate the one-step forecast of the measurement variable y_t at date $t - 1$ knowing information up to and including $t - 1$,

$$\hat{y}_{t|t-1} = E(y_t | x_t, \Psi_{t-1}) = H'\hat{h}_{t|t-1} + A'x_t. \quad (8)$$

Once the new observation y_t becomes available at date t , we can calculate the forecast error on the observed variable and its MSE:

$$\begin{aligned} \lambda_t &= y_t - \hat{y}_{t|t-1}, \\ \Lambda_t &= E[(y_t - \hat{y}_{t|t-1})(y_t - \hat{y}_{t|t-1})' | \Psi_t] = H'P_{t|t-1}H + R, \end{aligned} \quad (9)$$

and it is possible to update the estimated state vector and its MSE matrix:

$$\begin{aligned} \hat{h}_{t|t} &= E[h_t | \Psi_t] = \hat{h}_{t|t-1} + \Phi_t \lambda_t, \\ P_{t|t} &= (I - \Phi_t H')P_{t|t-1}, \end{aligned}$$

where $\Phi_t = P_{t|t-1}H(\Lambda_t)^{-1}$ is known as the *filter gain* since a certain fraction of the difference between the observable and predicted states is added to the previous prediction. These last two terms that are generated using the updating equations are the inputs of the next filter iteration.

Hence, if the system matrices are known the Kalman filter will yield as outcome the sequences $\{\widehat{h}_{t|t-1}\}_{t=1}^T$ and $\{P_{t|t-1}\}_{t=1}^T$. We can view the Kalman filter as a sequential updating procedure that consists of forming a prior guess about the state of nature and then adding a correction to that guess, this correction being determined by how well the guess has performed in predicting the next observation. However, the state space model is not entirely estimated since we do not usually know the parameters of the system matrices. Considering that $\{v_t, w_t\}_{t=1}^T$ are Gaussian, then the distribution of y_t conditional on (x_t, Ψ_{t-1}) is Gaussian with mean given by equation (8) and variance given by equation (9). We use the *prediction error decomposition* to construct the logarithm of the distribution function, so that for a Gaussian model it has the form:

$$\ln f(y_t|x_t, \Psi_{t-1}) = -\frac{N}{2} \ln 2\pi - \frac{1}{2} \ln |\Lambda_t| - \frac{1}{2} \lambda_t' \Lambda_t^{-1} \lambda_t.$$

Finally, to estimate the parameters of the system matrices we maximize the log-likelihood function

$$\ln L = \sum_{t=1}^T \ln f(y_t|x_t, \Psi_{t-1})$$

with respect to the underlying unknown parameters using nonlinear optimization techniques.

The Threshold Kalman filter

Nonlinearities can be introduced into state space models in a variety of ways. One of the most important classes of models has Gaussian disturbances but allows the system matrices to depend on past observations available at time $t-1$. This class of models is known in time series literature as *conditionally Gaussian*⁷. These models have the attractive property of still being tractable by the Kalman filter. In our model, we only introduce regime-switching in the state equation. The state space representation is the following:

$$\begin{aligned} y_t &= H'h_t + A'x_t + w_t \\ h_t &= F(q_{t-1})h_{t-1} + \Pi'x_t + v_t, \end{aligned}$$

where $q_{t-1} = U_{t-1} - U_{t-d}$ stands for the threshold variable. Despite the fact that the coefficient matrix associated with h_{t-1} depends on observations up to and including $t-1$, it may be regarded as being fixed once we are at time $t-1$. The same hypotheses about the disturbance vectors v_t and w_t are retained. Hence the derivation of the Kalman filter proceeds as in the previous section but a simple modification is introduced.

As mentioned above, the goal of the Kalman filter procedure is to derive a forecast of the unobserved state vector h_t based on the information set Ψ_{t-1} . Here the goal is to form a forecast of h_t conditional not only on Ψ_{t-1} but also on the regime of the economy. Let j be a dummy variable that refers to the regime of the economy:

⁷See Harvey ([12], Section 3.7.) for a more detailed description of this class of models.

$$j = \begin{cases} 1 & \text{if } q_{t-1} \geq \gamma \\ 2 & \text{if } q_{t-1} < \gamma \end{cases}.$$

We calculate the conditional forecast of the state variables and its conditional error covariance, or MSE, matrix as follows:

$$\begin{aligned} \hat{h}_{t|t-1}^j &= F_j \hat{h}_{t-1|t-1}^j + \Pi' x_t \\ P_{t|t-1}^j &= F_j P_{t-1|t-1}^j F_j' + Q, \end{aligned}$$

where F_j refers to the transition matrix in each regime.

The conditional forecast of observed variables is given by:

$$\hat{y}_{t|t-1}^j = H' \hat{h}_{t|t-1}^j + A' x_t.$$

Once observable variables are realized at date t , we can calculate the conditional error forecast and its conditional variance:

$$\begin{aligned} \lambda_t^j &= y_t - \hat{y}_{t|t-1}^j \\ \Lambda_t^j &= H' P_{t|t-1}^j H + R. \end{aligned}$$

Finally, we update the previous conditional forecast of unobserved variables and its conditional variance as follows:

$$\begin{aligned} \hat{h}_{t|t}^j &= \hat{h}_{t|t-1}^j + \Phi_t^j \lambda_t^j \\ P_{t|t}^j &= (I - \Phi_t^j H') P_{t|t-1}^j, \end{aligned}$$

with $\Phi_t^j = P_{t|t-1}^j H (\Lambda_t^j)^{-1}$ the *filter gain*. These last two terms correspond to the inputs of the next filter iteration.

To estimate parameters γ and d we construct a grid $\Delta = \Theta \otimes \Pi$ over the two-dimensional space (γ, d) , where Θ and Π are the grids for γ and d , respectively.

We proceed in two steps. First, we estimate the model for each candidate (γ, d) belonging to the selected grid. Then we maximize the log-likelihood function conditionally to the particular values of γ and d under consideration:

$$\ln L(\gamma, d) = \sum_{t=1}^T \ln f(y_t | \gamma, \Psi_{t-1})$$

with respect to the underlying unknown parameters, using nonlinear optimization techniques. Second, we retain the threshold parameter value and the delay lag that provide the highest log-likelihood. That is, γ and d are given by:

$$(\hat{\gamma}, \hat{d}) = \arg \max_{(\gamma, d) \in \Delta} \ln L(\gamma, d).$$

Appendix B: Tables and Figures

Table 1: Monte Carlo Results

Nominal size	5%	10%	20%
Simulated size			
<i>DGP</i> ₁			
Homoskedastic bootstrap	0.042	0.092	0.203
Heteroskedastic bootstrap	0.045	0.091	0.195
Simulated power			
<i>DGP</i> ₂ : $\alpha_1 - \alpha_2 = 0.2$			
Homoskedastic bootstrap	0.189	0.431	0.586
Heteroskedastic bootstrap	0.476	0.581	0.690
<i>DGP</i> ₃ : $\alpha_1 - \alpha_2 = 0.4$			
Homoskedastic bootstrap	0.555	0.658	0.769
Heteroskedastic bootstrap	0.664	0.747	0.813
<i>DGP</i> ₄ : $\alpha_1 - \alpha_2 = 0.6$			
Homoskedastic bootstrap	0.679	0.762	0.833
Heteroskedastic bootstrap	0.776	0.816	0.851
<i>DGP</i> ₅ : $\alpha_1 - \alpha_2 = 0.8$			
Homoskedastic bootstrap	0.779	0.829	0.869
Heteroskedastic bootstrap	0.848	0.867	0.885
<i>DGP</i> ₆ : $\alpha_1 - \alpha_2 = 1$			
Homoskedastic bootstrap	0.895	0.898	0.905
Heteroskedastic bootstrap	0.899	0.943	0.987

*DGP*₁ : $\sigma_N = 0.35$, $\sigma_C = 0.04$, $\sigma_v = 0.68$, $\delta = 10.23$, $\beta = 0.33$, $\phi_1 = -0.007$, $\phi_2 = -1.13 \times 10^{-5}$, $\alpha = 0.36$; *DGP*₂₋₆ : $\sigma_N = 0.5$, $\sigma_C = 0.05$, $\sigma_v = 0.7$, $\delta = 2.56$, $\beta = 0.24$, $\phi_1 = 0.14$, $\phi_2 = 0.8$, $\alpha_2 = 0.8$, $\gamma = 0.07$.

Table 2: Unit Root Tests

Phillips-Perron test on GDP series	
Italy	-2.743
France	0.888
U.S.	-0.365
Phillips-Perron test on Unemployment series	
Italy	-1.716
France	-2.593
U.S.	-2.089

Note 1: For the Phillips-Perron test, we use Mackinnon critical values for rejecting the hypothesis of a unit root. We do not reject the null hypothesis of a unit root at 1%, 5% or 10%.

Table 3 : Tests for the Hysteresis Assumption

$H_0 : \alpha_1 = \alpha_2$	
<i>Italy</i>	$p - value = 0.000^*$
<i>France</i>	$p - value = 0.021^*$
U.S.	$p - value = 0.665$

*Significant at 5%

Table 4: Estimation Results⁸

		ITALY	
<i>Order of the autoregressive process in the identification equation⁹</i>		$p = 2$	
		NONLINEAR MODEL	
<i>Percentage of observations</i>		$i = 1$	$i = 2$
		44%	56%
Natural Rate Equation		$i = 1$	$i = 2$
α_i		1.366 (0.571)	0.797 (0.811)
σ^N		0.472 (0.055)	
Cyclical Rate Equation			
ϕ_1		0.140 (0.186)	
ϕ_2		0.795 (0.155)	
σ^C		0.054 (0.045)	
Identification Equation			
β_1		0.490 (0.118)	
β_2		-0.010 (0.095)	
δ		2.557 (1.556)	
σ^v		0.708 (0.057)	
<i>Threshold</i>		$\gamma = 0.07$	
90% confidence ¹⁰		[-0.008, 0.143]	
<i>Delay lag</i>		$d = 2$	

⁸Standard errors are calculated from a consistent maximum likelihood estimate of the variance-covariance matrix and provided in brackets.

⁹Lagged values of the output growth rate are introduced until the residuals of the linear specification become uncorrelated over time. We need $p = 2$ for the case of Italy.

¹⁰We compute the confidence interval based on the bootstrap percentiles described by Hall [7].

Table 5 : Estimation Results¹¹

		FRANCE	
<i>Order of the autoregressive process in the identification equation¹²</i>		$p = 2$	
		NONLINEAR MODEL	
<i>Percentage of observations</i>		$i = 1$	$i = 2$
		48%	52%
Natural Rate Equation		$i = 1$	$i = 2$
α_i		2.210 (0.446)	1.704 (0.173)
σ^N		0.102 (1.715)	
Cyclical Rate Equation			
ϕ_1		-0.532 (0.254)	
ϕ_2		0.175 (0.189)	
σ^C		0.476 (0.343)	
Identification Equation			
β_1		0.339 (0.088)	
β_2		0.481 (0.093)	
δ		0.084 (0.138)	
σ^v		0.466 (0.030)	
<i>Threshold</i>		$\gamma = 0.1$	
90% confidence ¹³		[0.000, 0.323]	
<i>Delay lag</i>		$d = 3$	

¹¹Standard errors are calculated from a consistent maximum likelihood estimate of the variance-covariance matrix and provided in brackets.

¹²Lagged values of the output growth rate are introduced until the residuals of the linear specification become uncorrelated over time. We need $p = 2$ for the case of France.

¹³We compute the confidence interval based on the bootstrap percentiles described by Hall [7].

Table 6 : Estimation Results¹⁴

U.S.	
<i>Order of the autoregressive process in the identification equation¹⁵</i>	
$p = 2$	
LINEAR MODEL	
Natural Rate Equation	
$\alpha^{(a)}$	1.337 (0.244)
σ^N	0.070 (0.121)
Cyclical Rate Equation	
ϕ_1	0.498 (0.242)
ϕ_2	-0.062 (0.060)
σ^C	0.239 (0.072)
Identification Equation	
β_1	0.338 (0.078)
β_2	0.265 (0.078)
δ	-1.379 (0.517)
σ^v	0.811 (0.069)

(a) The Wald test statistic for the null hypothesis ($\alpha = 0$) is distributed chi-square with 1 degree of freedom under the null. It is significant at 5%.

¹⁴Standard errors are calculated from a consistent maximum likelihood estimate of the variance-covariance matrix and provided in brackets.

¹⁵Lagged values of the output growth rate are introduced until the residuals of the linear specification become uncorrelated over time. We need $p = 2$ for the case of U.S.

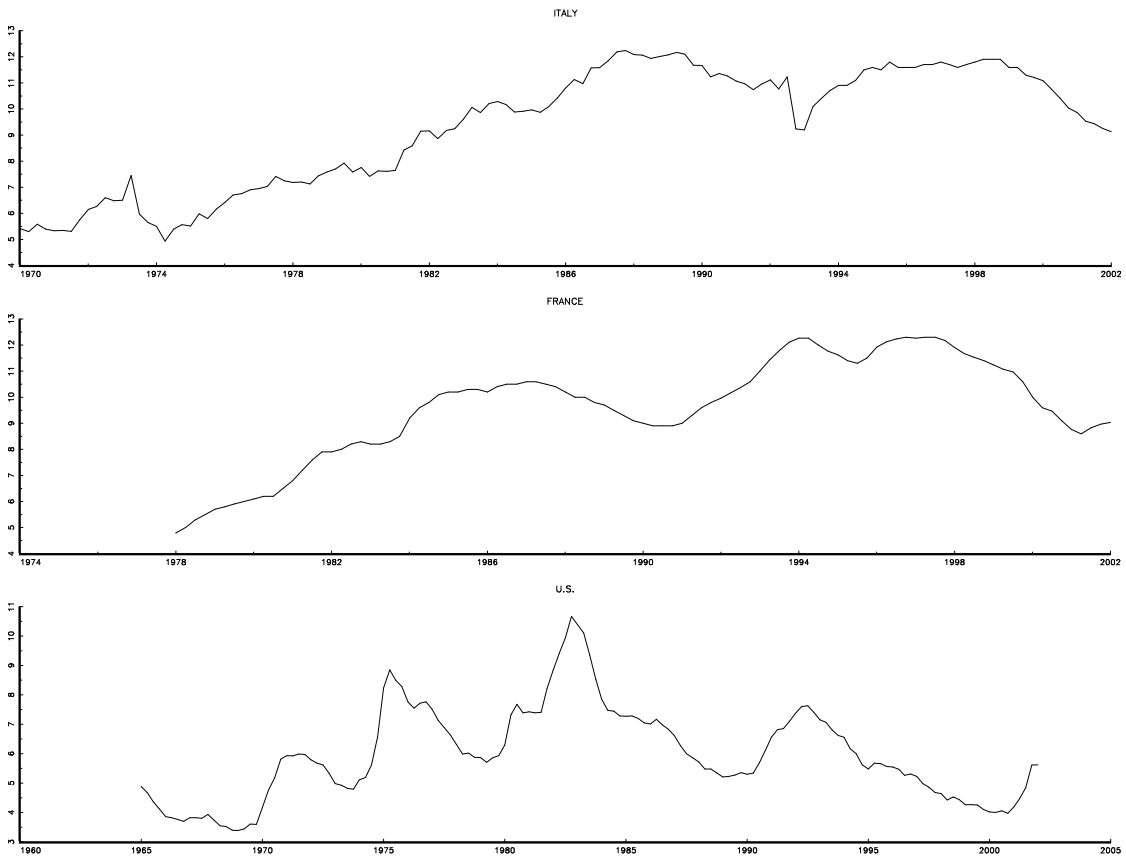


Figure 1: Unemployment Rate

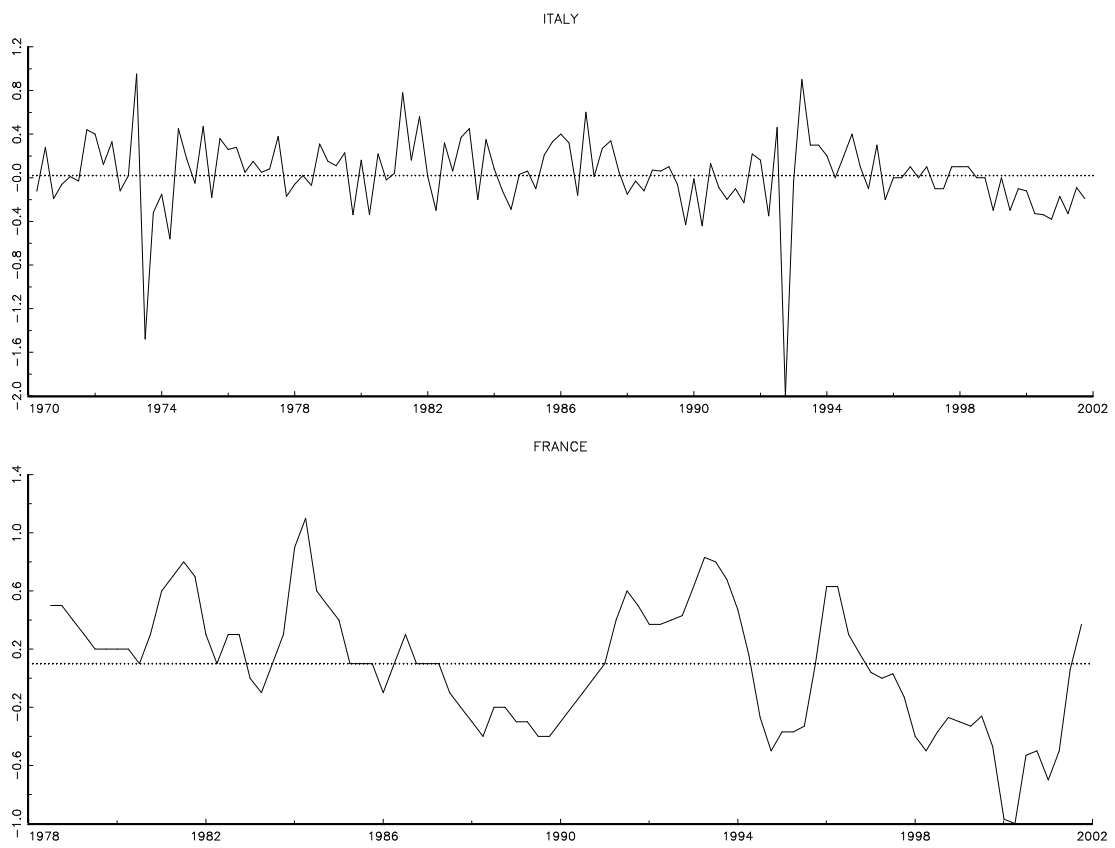


Figure 2: Threshold variable: $U_{t-1} - U_{t-\hat{d}}$ and $\hat{\gamma}$.