Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

# Working Paper nº 18/06

## Non-linearities and fractional integration in the US unemployment rate

Luis A. Gil-Alana
Facultad de Ciencias Económicas y Empresariales
Universidad de Navarra

and

Guglielmo M. Caporale
Faculty of Economics and Business Administration
Brunel University

Nonlinearities and fractional integration in the US unemployment rate
Luis A. Gil-Alana and Guglielmo M. Caporae
Working Paper No.18/06
December 2006

ABSTRACT

This paper proposes a model of the US unemployment rate which accounts for both its asymmetry and its long memory. Our approach introduces fractional integration and nonlinearities simultaneously into the same framework, using a Lagrange Multiplier procedure with a standard null limit distribution. The empirical results suggest that the US unemployment rate can be specified in terms of a fractionally integrated process, which interacts with some non-linear functions of labour demand variables such as real oil prices and real interest rates. We also find evidence of a long-memory component. Our results are consistent with a hysteresis model with path dependency rather than a NAIRU model with an underlying unemployment equilibrium rate, thereby giving support to more activist stabilisation policies. However, any suitable model should also include business cycle asymmetries, with implications for both forecasting and policy-making.

Luis A. Gil-Alana
University of Navarra
Faculty of Economics
Edificio Biblioteca, E. Este
E-31080 Pamplona, Spain
alana@unav.es

Guglielmo M. Caporale
Centre for Empirical Finance
Brunel University
Uxbridge, Middlesex UB8 3PH
ENGLAND
Guglielmo-Maria.Caporale@brunel.ac.uk

## 1. Introduction

Two well-known facts about the unemployment rate are (i) the high persistence of shocks, or hysteresis (see Blanchard and Summers, 1987), which is a feature, among others, of "insider" models (see Lindbeck and Snower, 1988), or of models in which fixed and sunk costs make current unemployment a function of past labour demand (see Cross, 1994, 1995), and (ii) its asymmetric behaviour, namely the fact that unemployment appears to rise faster in recessions than it falls during recoveries. Both are well documented, especially in the case of the US (see, e.g., Rothman, 1991). The former can be modelled using a fractional integration framework, where the number of differences required to achieve I(0) stationary series is a real value. As for the latter, one possible explanation is the presence of asymmetric adjustment costs of labour, such as hiring and firing costs, which have been shown to account well for movements in the unemployment rate in Europe after 1973 (see Bentolilla and Bertola, 1990), even though, as pointed out by Hamermesh and Pfann (1996), asymmetry at firm level does not necessarily imply asymmetry at macro level. Other suggested explanations include asymmetry in job destruction (i.e., the fact that jobs disappear at a higher rate during recessions than expansions – see Caballero and Hammour, 1994), and/or in capital destruction (see Bean, 1989).

Any satisfactory model of the unemployment rate has to be able to account for these two properties, i.e. long memory and non-linearity. In particular, overlooking non-linearities can result in misleading in-sample diagnostics (see Potter, 1995). Further, non-linear specifications might lead to an improvement over conventional linear forecasts (see, e.g., Parker and Rothman, 1997, Montgomery et al, 1998, and Rothman, 1998).

Moreover, the fact that most standard models for the US unemployment rate assume either a unit root (I(1)) or a stationary I(0) process with the autoregressive (AR) root close to 1, restricts the analysis to the case of integer orders of differentiation (0 or 1). Fractional integration allows for a much wider variety of model specifications that include the above cases as particular cases.

Various non-linear models have already been estimated in the literature, starting with the seminal paper by Neftci (1984) (see the extensive survey by Pfann, 1993, and also Potter, 1995). In a number of cases models with a single or infrequent shifts in the mean of the unemployment rate have been adopted. Prominent examples are Bianchi and Zoega (1998), whose Markov-switching model only allowed for a switch in the intercept in order to analyse the issue of multiple equilibria, and Papell et al. (2000), who tested for multiple structural changes. Several studies are based on smooth transition mechanisms. These include Rothman (1998), who estimated AR, (S)TAR (smooth transition autoregressive) and bilinear models for predicting US unemployment, and Hansen (1997), who fitted a TAR (threshold autoregressive) model to US unemployment. Other contributions using different approaches are Parker and Rothman (1998), who applied Beaudry and Koop's (1993) current depth of recession approach; and Franses and Paap (1998), who developed AR models with censored latent effect parameters. More recently, Coakley et al. (2001) have tried to complement the regime shift literature with business cycle asymmetries. Specifically, they combine a single regime shift in the equilibrium level with asymmetries in the speed of adjustment, which are modelled using a momentum threshold autoregression (M-TAR) model characterised by fast-up, slow-down dynamics over the business cycle. In a more theoretical paper, Caner and Hansen (2001) examine a two-regime TAR(k) model with an autoregressive unit root. They develop an appropriate asymptotic theory, and show that the joint application of two tests – for a threshold and for a unit root – enables one to distinguish between nonlinear and nonstationary processes. They find that the US unemployment rate is a stationary nonlinear threshold autoregression.

An interesting study is due to Skalin and Teräsvirta (2002), who argue that the observed asymmetry can be captured by a simple model based on the standard logistic smooth transition autoregressive (LSTAR) model for the first difference of unemployment, but also including a lagged level term. Such a specification allows for asymmetry by introducing "local" nonstationarity in a globally stable model. They stress that their analysis has implications for policy-makers, who should take into account the fact that asymmetric forecast densities mean that the probability of erring is also asymmetric. Further, there are implications for multivariate modelling: if the unemployment rate is in fact a stationary nonlinear process, linear VARs based on the assumption that it is a I(1) variable and including cointegrating relationships with other I(1) regressors will be mis-specified. Therefore, some papers analyse the joint dynamics of US output and unemployment in the context of nonlinear VARs. For instance, Altissimo and Violante (2001) estimate a threshold VAR model of output and unemployment in the US, in which nonlinearity arises from including a feedback variable measuring the depth of the current recession, and the threshold growth rate separating the two regimes (expansions and recessions) is endogenously determined.

Further evidence on nonlinearities in unemployment has been obtained by estimating linear models, and then carrying out the time domain test of time reversibility (TR) on the residuals introduced by Ramsey and Rothman (1996). For instance, Rothman (1999) finds that ARMA models of US unemployment display TR, indicating that the true DGP is not linear, a result which appears to be robust to differencing and linear detrending when the model allows for conditional mean nonlinearity; however, it is not robust to allowing for GARCH effects.

All these studies typically assume that the disturbances follow an I(0) stationary process, and adopt an AR, MA or ARMA specification for the error term. One of the few exceptions is the study by van Dijk et al. (2002), where a fractional integration smooth transition autoregression time series [FISTAR] model is estimated and shown to outperform rival specifications. In this paper we also model unemployment as a non-linear process, and allow for the disturbances to be fractionally integrated. However, unlike van Dijk et al. (2002), who employ a sequential procedure, we introduce both fractional integration and nonlinearities simultaneously into the same framework, which has the obvious advantage of requiring a single procedure for testing the order of integration of the series. Moreover, the suggested test is a Lagrange Multiplier (LM) one, and, therefore, it has a standard null limit distribution. A limitation of our approach lies in the specification of the non-linear (in the variables) process, which is such that it becomes linear in the parameters to avoid the interaction with the fractional differencing parameter. Specifically, we use non-linear transformations of the variables, which are regressed in a linear model and do not involve non-linear estimation. Thus, the parameters to be estimated and tested are those corresponding to the short-run components of the series and the order of integration respectively. However, despite this limitation, our specification does enable us to account not only for asymmetry (as other nonlinear models do), but also for the high persistence of shocks and the long memory of the unemployment process. Candelon and Gil-Alana (2003) showed that fractional integration can be used to reproduce business cycle characteristics in the US and other countries. The present study goes further in the sense that we also incorporate non-linearities to take into account the asymmetries typical of business cycles.

**(Insert Figure 1 about here)**

As previously mentioned, most time series models taking into account non-linearities (e.g. Markov-switching, threshold autoregressive or smooth transition autoregressive models) assume the presence of two or more regimes within the sample, with the series being modelled either in levels or in first differences. Our fractional integration framework enables us to examine the dynamic structure of the series in a much more flexible way. As a simple illustration, we consider the following time series model:

$$y_t \;=\; f(\theta) \;+\; x_t; \quad (1-L)^d\, x_t \;=\; u_t, \qquad t \;=\; 1, 2, \ldots, \tag{1}$$

where $f(\theta) = a\, I(y_{t-1} > y_{t-2})$, $I(\cdot)$ is the indicator function, and $u_t$ is assumed to be white noise. Figure 1 in the paper displays simple realisations of the model in (1), with T = 100, assuming (in the left-hand side plots) that a = 0 (i.e. without non-linearities), and a = 2 (right-hand side plots). We set d equal to 0, 0.25, 0.75 and 1, and find that the higher d is, the higher is the dependence between the observations. When allowing for non-linearities, we note that the cyclical structure changes along with dependence between the observations.

The outline of the paper is as follows: Section 2 presents the model and the procedure for testing the degree of integration of the series. In Section 3, the procedure is applied to the US unemployment rate. Section 4 discusses model selection, whilst Section 5 focuses on the forecasting properties of the selected model. Section 6 concludes.

## 2. Testing of I(d) hypotheses in non-linear models

Let us suppose that $\{y_t, t = 1, 2, ...T\}$ is the time series we observe (in our case, unemployment) and that it is related to some exogenous components from both the demand and the supply side, $z_t$, through the relationship:

$$y_t \;=\; f(z_t;\theta) \;+\; x_t, \qquad t \;=\; 1, 2, ..., \tag{2}$$

where θ represents the unknown coefficients and $x_t$ is driven by:

$$(1 - L)^d \; x_t \;=\; u_t, \qquad t \;=\; 1, 2, ..., \tag{3}$$

with $x_t = 0$ for $t \leq 0$, where d may be a real value and $u_t$ is I(0).[1] Note that the fractional polynomial can be expressed in terms of its Binomial expansion, such that:

$$(1 \;-\; L)^d \;=\; \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^j \;=\; 1 \;-\; d\,L \;+\; \frac{d\,(d-1)}{2} L^2 \;-\; ....$$

for all real d. Clearly, if d = 0 in (3) $x_t = u_t$, and a 'weakly autocorrelated' $x_t$ is allowed for. However, if d > 0, $x_t$ is said to be a long memory process, also called 'strongly autocorrelated', because of the strong association between observations widely separated in time. If d is an integer value, $x_t$ will be a function of a finite number of past observations, while if d is real, $x_t$ depends strongly upon values of the time series far away in the past (see, e.g. Granger and Ding, 1996; Dueker and Asea, 1998).

The time series literature has usually focused on the cases of d = 0 (weak dependence) or d = 1 (a unit root). However, to correctly determine d is crucial from a statistical viewpoint. If $d \in (0, 0.5)$ in (3), $x_t$ is covariance-stationary and mean-reverting, having auto-covariances which decay at a much slower rate than those of an ARMA

---

[1] For the purpose of the present paper, we define an I(0) process as a covariance stationary process with spectral density function that is positive and finite at the zero frequency.

process - in fact, so slowly as to be non-summable. If d $\in$ [0.5, 1), the series is no longer covariance-stationary, but it is still mean-reverting, with the effects of shocks disappearing in the long run. Finally, d $\geq$ 1 implies non-stationarity and non-mean-reversion. Therefore, the fractional differencing parameter d plays a crucial role in describing the persistence behaviour of the series: the higher d is, the higher will be the association between the observations.

Robinson (1994) proposed a Lagrange Multiplier (LM) test of the null hypothesis:

$$H_o : \; d \; = \; d_o . \qquad (4)$$

for any real given value $d_o$ in a model given by (3), where $x_t$ may be the errors from the regression (linear) model:

$$y_t \; = \; \beta' z_t \; + \; x_t , \qquad t \; = \; 1, \, 2, \, .... \qquad (5)$$

The test is based on the null differenced model in (3) – (5):

$$(1 - L)^{d_o} y_t \; = \; \beta' (1 - L)^{d_o} z_t \; + \; u_t , \qquad t \; = \; 1, \, 2, \, ..., \qquad (6)$$

and its functional form can be found in various empirical applications (e.g., Gil-Alana and Robinson, 1997; Gil-Alana, 2000, 2001).

In this paper, we extend Robinson's (1994) procedure to the case of non-linear regression models, i.e., testing $H_o$ (4) in a model given by (3) and (5). Note that under the null hypothesis given by (4): d = $d_o$, (2) and (3) become:

$$(1 \, - \, L)^{d_o} y_t \; = \; (1 \, - \, L)^{d_o} \, f(z_t \, ; \theta) \; + \; u_t , \qquad t \; = \; 1, \, 2, \, ... \; . \qquad (7)$$

The main problem with this equation lies in the interaction between the fractional polynomial $(1 - L)^{d_o}$ and the possibly non-linear function f, and the estimation of the parameters involved in such a relationship. For the purpose of the present study, let us assume that f($z_t$; θ) = θ g($z_t$), where g is of a non-linear nature. In such a case, (7) becomes:

$$(1 - L)^{d_o} y_t \;=\; \theta' w_t \;+\; u_t, \qquad t \;=\; 1, 2, \ldots, \tag{8}$$

where $w_t$ = $(1 - L)^{d_o} g(z_t)$, and hence, the "non-linearity" is not in terms of the parameters, but in terms of certain nonlinear function of the variables $z_t$. We can obtain the OLS estimate of θ and residuals:

$$\hat{u}_t \;=\; (1 - L)^{d_o} y_t \;-\; \hat{\theta}' w_t\,, \qquad \hat{\theta} = \left(\sum_{t=1}^{T} w_t\, w_t'\right)^{-1} \sum_{t=1}^{T} w_t (1-L)^{d_o} y_t\,,$$

and the same type of analysis as in Robinson (1994) can be conducted here. Denoting the periodogram of $u_t$,

$$I(\lambda_j) \;=\; \frac{1}{2\pi T}\left|\sum_{t=1}^{T} \hat{u}_t\, e^{i\lambda_j t}\right|^2,$$

the test statistic takes the form:

$$\hat{R} \;=\; \hat{r}^2; \quad \hat{r} \;=\; \left(\frac{T}{\hat{A}}\right)^{1/2} \frac{\hat{a}}{\hat{\sigma}^2}, \tag{9}$$

where T is the sample size and

$$\hat{a} \;=\; \frac{-2\pi}{T}\sum_{j=1}^{T-1}\psi(\lambda_j)\, g(\lambda_j;\hat{\tau})^{-1}\, I(\lambda_j); \qquad \hat{\sigma}^2 \;=\; \sigma^2(\hat{\tau}) \;=\; \frac{2\pi}{T}\sum_{j=1}^{T-1} g(\lambda_j;\hat{\tau})^{-1}\, I(\lambda_j);$$

$$\hat{A} \;=\; \frac{2}{T}\left(\sum_{j=1}^{T-1}\psi(\lambda_j)^2 \;-\; \sum_{j=1}^{T-1}\psi(\lambda_j)\,\hat{\varepsilon}(\lambda_j)' \times \left(\sum_{j=1}^{T-1}\hat{\varepsilon}(\lambda_j)\,\hat{\varepsilon}(\lambda_j)'\right)^{-1} \times \sum_{j=1}^{T-1}\hat{\varepsilon}(\lambda_j)\psi(\lambda_j)\right)$$

$$\psi(\lambda_j) = \log\left|2\sin\frac{\lambda_j}{2}\right|; \quad \hat{\varepsilon}(\lambda_j) \;=\; \frac{\partial}{\partial\tau}\log g(\lambda_j;\hat{\tau}); \quad \lambda_j \;=\; \frac{2\pi\, j}{T}; \quad \hat{\tau} \;=\; \arg\min_{\tau\in H}\sigma^2(\tau),$$

where H is a compact subset of the $R^q$ Euclidean space, and the function g above is a known function coming from the spectral density function of $u_t$,

$$f(\lambda;\sigma^2;\tau) = \frac{\sigma^2}{2\pi} g(\lambda;\tau), \qquad -\pi < \lambda \le \pi.$$

Note that these tests are purely parametric, and, therefore, they require specific modelling assumptions about the short-memory specification of $u_t$. Thus, if $u_t$ is white noise, then g $\equiv$ 1, and if $u_t$ is an AR process of form $\phi(L)u_t = \varepsilon_t$, $g = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\varepsilon_t)$, so that the AR coefficients are a function of $\tau$.

It is clear then that $\hat{\theta}$ is a consistent estimate of $\theta$, $\hat{u}_t$ satisfying the same properties as in Robinson (1994), and thus, under certain regularity conditions:[2]

$$\hat{R} \rightarrow_d \chi_1^2, \qquad as \quad T \rightarrow \infty. \tag{10}$$

Consequently, unlike in other procedures, we are in a classical large-sample testing situation for the reasons explained by Robinson (1994), who also showed that the tests are efficient in the Pitman sense against local departures from the null. Because $\hat{R}$ involves a ratio of quadratic forms, its exact null distribution could have been calculated under Gaussianity via Imhof's algorithm. However, a simple test is approximately valid under much wider distributional assumptions: an approximate one-sided 100$\alpha$% level test of $H_o$ (4) against the alternative: $H_a$: $d > d_o$ ($d < d_o$) will be given by the rule: "Reject $H_o$ if $\hat{r} > z_\alpha$ ($\hat{r} < - z_\alpha$)", where the probability that a standard normal variate exceeds $z_\alpha$ is $\alpha$.

To capture nonlinear features in a time series, one can choose from a wide variety of nonlinear models (see Franses and Van Dijk, 2000, for a recent survey). A model which enjoys a fair amount of popularity, mainly due to its empirical tractability, is the smooth transititon autoregressive (STAR) model, that is,

[2] These conditions are very mild and concern technical assumptions to be satisfied by $\psi(\lambda)$.

$$y_t = (\theta_{10} + \theta_{11} y_{t-1} + ... + \theta_{1p} y_{t-p})(1 - G(z_t; \gamma; c)) + (\theta_{20} + \theta_{21} y_{t-1} + ... + \theta_{2p} y_{t-p}) G(z_t; \gamma; c) + \varepsilon_t,$$

where $\varepsilon_t$ is a white noise process and the transition function $G(z_t; \gamma; c)$ usually is assumed to be the logistic function:

$$G(z_t; \gamma; c) = (1 + \exp\{-\gamma (z_t - c) / \sigma_{zt})^{-1} \qquad (11)$$

with $\gamma > 0$, and where $z_t$ is the transition variable (possibly a set of exogenous regressors), $\sigma_{zt}$ is the standard deviation of $z_t$, $\gamma$ is a slope parameter and c is a location parameter. The parameter restriction $\gamma > 0$ is an identifying restriction. The value of the logistic function (11), which is bounded between 0 and 1, depends on the transition variable $z_t$ as follows: $G(z_t; \gamma; c) \to 0$ as $z_t \to -\infty$, $G(z_t; \gamma; c) = 0.5$ for $z_t = c$, and $G(z_t; \gamma; c) \to 1$ as $z_t \to +\infty$.[4]

In our application we do not consider the parameters affecting (11) because of the interaction with the fractional integration polynomial, and thus we assume that $\gamma = 1$ and $c = 0$.[5] This is a further restriction in the model but is done in order to obtain a more tractable approach of the nonlinear fractionally integrated model. Moreover, in this way we do not have to take into account the lag structure of the dependent variable $y_t$, since this will be contained in the (possible) weak autocorrelation structure of $u_t$ in (3). Thus, a simple smooth transition model is:

$$g(z_t) = \theta_{10} [1 - G(z_t)] + \theta_{20} G(z_t); \qquad G(z_t) = \frac{1}{1 + \exp\left\{\frac{-z_t}{s_{z_t}}\right\}}; \qquad s_{z_t} = \frac{1}{T} \sum_{t=1}^{T} z_t^2,$$

[4] Applications of the STAR model and the closely related TAR model to unemployment rates can be found in Montgomery et al. (1998); Koop and Potter (1999); Caner and Hansen (2001) and Skalin and Teräsvirta (2000) among others.

[5] In the empirical application carried out in the following section, we work with demeaned series to avoid the influence of the location parameter.

where $z_t$ represents each of the variables affecting unemployment. Clearly, $G(z_t)$ does not involve the estimation of any parameters, and thus the model under the null becomes:

$$(1 - L)^{d_o} y_t = \sum_{i=1}^{k} \left( \theta_{10} S_{1t}^{i} + \theta_{20} S_{2t}^{i} \right) + u_t, \qquad t = 1, 2, \ldots,$$

where $S_{1t}^{i} = (1 - L)^{d_o} [1 - G(z_{it})]$ and $S_{2t}^{i} = (1 - L)^{d_o} G(z_{it})$. Under $H_o$ (4), the disturbances $u_t$ are assumed to be I(0), and therefore standard techniques can be applied.

Finally, in this section, we examine the implications of testing the order of integration when non-linearities of the form given in (11) are present but are not taken into account, and also the reverse case, i.e. assuming a non-linear structure (with fractional integration) when that is not present in the data. In both cases we use the parametric procedure described above, reporting the results in Table 1.

We assume that the true model is given by

$$y_t = 0.5 S_{1t} + 0.8 S_{2t} + x_t; \qquad (1 - L)^{0.5} x_t = u_t,$$

where $S_{1t}$ = (1 – 1/exp(-$z_t$/$s_t$)); $S_{2t}$ = 1 - $S_{1t}$; and $u_t$ and $z_t$ are white noise independent processes.

In Table 1 we compute the rejection frequencies of the test statistic given by (9) in the model given by (5) and (3), with $z_t$ = ($S_{1t}$, $S_{2t}$)’ and $d_o$ in (4) equal to 0, 0.10, …, 1. We use sample sizes T = 200 and 400, and Gaussian series were generated by the routines GASDEV and RAN3 of Press, Flannery, Teukolsky and Vetterling (1986).

**(Insert Table 1 about here)**

Case a) in Table 1 is the case where we truly identified the non-linear and the fractionally integrated structures. Thus, testing $H_o$ (4) with $d_o$ = 0.50 gives us the empirical size of the test. We see that the values are slightly upward biased (6.4% with T = 200 and 5.7% with T = 400, for a significance level of 5%). However, as we depart from the null, the rejection frequencies substantially increase, and they are close to 1 for $d \leq 0.10$ or $d \geq 0.70$ (with T = 400). Case b) refers to a situation where we test for fractional integration ignoring the existence of a non-linear structure. In other words, we test $H_o$ (4) in (5) and (3) assuming that $z_t = 0$. In such a case we note that the lowest rejection frequencies do not occur when d = 0.5 but rather for a slightly smaller value, d = 0.4, (10.8% with T = 200 and 11.6% with T = 400), implying that there is a bias in favour of smaller orders of integration. Finally, if we test for fractional integration and the non-linearities in a model without a non-linear structure (Case c)), we see that the procedure correctly identifies the order of integration, an obvious result if we note that the coefficients in (5) are then correctly estimated around 0.

## 3. The US case

In this section the testing procedure described in Section 2 is used to identify the dynamics of the US unemployment rate. The main relevance of the analysis from an economic viewpoint is whether it can shed any light on the adequacy of hysteresis models with path dependency (see, e.g., Blanchard and Summers, 1987) versus NAIRU models (see, e.g., Friedman, 1968), as discussed more in detail in the conclusions. The unemployment series used is the logistic transformation of the unemployment rate in the US[6], and we also consider real oil prices and real interest rates, quarterly, for the time period 1960q1 to 2002q3. Specifically, we use an oil price index (the industrial price index for refined petroleum and coal products, which is the available series with the

[6] We use a logistic transformation on the dependent variable to avoid the problem of boundedness of the unemployment rate. Note that, in the context of fractional integration, bounded variables may be in theoretical contradiction with the explosive behaviour of I(d) process for some values of d (see Wallis, 1987 for a justification based on the logistic transformation being defined between $\pm\infty$ so that standard

longest time span), and the 5-year benchmark government bond yield (end of the month). The real oil price and real interest rates series have been constructed using the GDP deflator. All series are seasonally unadjusted, and are taken from Datastream.

The variables employed are the same as in Carruth et al. (1998). In that paper, the authors examine the relationship between these three variables by means of classical cointegration techniques. We use the term "classical" in the sense that it is assumed that all individual series are nonstationary I(1), while the equilibrium long-run relationship is stationary I(0). Carruth et al. (1998) assume that causality in the model is uni-directional: only prices matter, while real interest rates are also included as another relevant variable operating at the world level, and hence causality links may also be bi-directional. If one wanted to rationalise it in terms of general equilibrium, one would say that the US is an economy with a stable set of supply-side policies implying a high degree of wage flexibility in the labour market. The main variables that have shifted the long-run labour demand up the ("wage-curve" or efficiency wage) labour supply would be changes at world level in input prices and in the cost of capital. (Note that real interest rates are implicitly assumed to have no or at most a weak effect on the labour supply via intertemporal substitution).

In this paper, we depart from the Carruth et al. (1998) model from an econometric viewpoint: rather than assuming a linear relationship, we introduce non-linearities. Moreover, instead of using integer orders of integration, we allow for the possibility of fractional values. This is motivated by earlier work reported in Caporale and Gil-Alana (2002), who found cointegration between the same set of variables for Canada in the presence of autocorrelated disturbances, suggesting that their relationship also has a

---

distributions apply). In any case, when using the original data (i.e., the US unemployment rate), the dynamic component. Furthermore, they reported evidence of fractional (as opposed to classical) cointegration, which implies long memory and slow reversion to equilibrium.

Denoting the logistic transformation of the US unemployment rate by UNE, real oil prices by ROP, and real interest rates by RIR, we employ the model:

$$UNE_t \;=\; \theta_{10}^{1}\left[1 - G(ROP_t)\right] + \theta_{20}^{1}\, G(ROP_t) + \theta_{10}^{2}\left[1 - G(RIR_t\right] + \;\; x_t\,, \qquad (12)$$

and (3), testing $H_o$ (4) for values $d_o$ ranging from 0 to 2 with 0.2 increments, using white noise and autocorrelated disturbances.[7]

Table 2 reports the values of the one-sided statistic $\hat{r}$ in (9). We observe that if we assume that $u_t$ is white noise, the only value of $d_o$ for which $H_o$ cannot be rejected is 0.80, implying long memory and mean-reverting behaviour. However, if we allow for autoregressive (AR) behaviour in $u_t$, the unit root null cannot be rejected. We also report the results based on the Bloomfield's (1973) exponential model for the I(0) disturbances $u_t$. This is a non-parametric approach to modelling $u_t$, with the spectral density function given by:

$$f(\lambda;\tau) \;=\; \frac{\sigma^2}{2\pi}\,\exp\left(2\sum_{r=1}^{p}\tau_r\cos(\lambda r)\right),$$

where p is now a parameter describing the short run dynamics of the series. Like the stationary AR(p) model, the Bloomfield (1973) model has exponentially decaying autocorrelations, and thus can be used to model $u_t$ in (2). The formulae for Newton-type iterations for estimating $\tau_l$ are very simple (involving no matrix inversion), and so are the updating formulae when p is increased; $\hat{A}$ in the Appendix can be replaced by the population quantity:

conclusions were practically the same as those reported in the paper.

$$\sum_{l=p+1}^{\infty} l^{-2} = \frac{\pi^2}{6} - \sum_{l=1}^{p} l^{-2},$$

which indeed is constant with respect to the $\tau_j$ (unlike the AR case). Similarly to the AR case, if $u_t$ follows the Bloomfield's (1973) exponential spectral model, the unit root (i.e., $d_o = 1$) is the only non-rejected value. Finally, in view of the quarterly structure of the series, we also tried seasonal autoregressions of the form:

$$u_t = \sum_{r=4}^{4p} \phi_r \, u_{t-r}, \quad t = 1, 2, ..., \tag{13}$$

with p = 1 and 2. In this case, we find that the null is rejected for all values of d smaller than or equal to 1. If p = 1, the non-rejection values occur for $d_o$ = 1.20, 1.40 and 1.60, and if p = 2, $d_o$ = 1.20 is the only non-rejection value. Thus, the results appear to be very sensitive to the specification of the I(0) disturbances, values of d smaller than, equal to, or higher than 1 being obtained depending on whether the disturbances are white noise, non-seasonally and seasonally autocorrelated.

**(Insert Tables 2 and 3 around here)**

Table 3 displays, for each type of disturbances, the 95%-confidence intervals of those values of $d_o$ for which $H_o$ cannot be rejected. These intervals were constructed as follows: first, we choose a value of d from a grid. Then, we form the test statistic testing the null for this value. If the null is rejected at the 5% level, we discard this value of d. Otherwise, we keep it. An interval is then obtained after considering all the values of d in the grid. Along with the intervals, we also report in the table the value corresponding to the lowest statistic in absolute value, ($d_o^*$), which will be an approximation to the

---

[7] Note that we do not include in the regression model G(RIR) to avoid the problem of exact multicollinearity.

maximum likelihood estimator.[8] We see that if $u_t$ is white noise, all values are below unity. If $u_t$ follows an AR process, the intervals include the unit root and the same happens with the Bloomfield model, while d is higher than 1 for seasonal autoregressions.

The large differences observed in the values of d when seasonal autoregressions are taken into account suggest that seasonality should also be considered. Seasonal dummy variables were first included in the regression model (12), but the coefficients corresponding to the dummies were found to be insignificantly different from zero. Note that the tests of Robinson (1994) are based on the null differenced model, which exhibits short memory, and thus standard t-tests apply. On the other hand, the large values of d observed in Table 2 when $u_t$ is a seasonal AR process may suggest that seasonality is of a nonstationary nature.[9] Therefore, we decided also to use another version of Robinson's (1994) tests, which is based on the model:

$$(1 - L^4)^d \; x_t \; = \; u_t, \qquad t \; = \; 1, 2, \ldots \; . \tag{14}$$

In such a case, $\hat{r}$ takes a similar form to (9), but $\hat{u}_t$ is now defined as:

$$\hat{u}_t \; = \; (1 - L^4)^{d_o} y_t \; - \; \hat{\theta}' w_t \, ,$$

and

$$\psi(\lambda_j) \; = \; \log\left| sin\frac{\lambda_j}{2} \right| + \log\left( 2\cos\frac{\lambda_j}{2} \right) + \log\left| 2\cos\lambda_j \right|; \qquad \hat{\varepsilon}(\lambda_j) \; = \; \frac{\partial}{\partial \tau} \log g(\lambda_j; \hat{\tau}),$$

and the test statistic still has the same standard null limit distribution. Ooms (1995) also proposed tests based on seasonal fractional models. They are Wald tests, requiring

[8] Note that the LM procedure employed in this paper is based on the Whittle function, which is an approximation to the likelihood function.

[9] Several studies conducted by Montanari, Rosso and Taqqu (1995, 1996, 1997) in a hydrological context showed that the presence of periodicities might influence the reliability of the estimators of the fractional differencing parameter at the zero frequency.

efficient estimates of the fractional differencing parameter. He used a modified periodogram regression estimation procedure due to Hassler (1994). In addition, Hosoya (1997) established the limit theory for long memory processes with the singularities not restricted at the zero frequency, and proposed a set of quasi log-likelihood statistics to be applied to raw time series. Unlike these methods, the tests of Robinson (1994) do not require estimation of the long-memory parameter, since the differenced series have short memory under the null.[10]

**(Insert Tables 4 and 5 about here)**

Table 4 reports the results for the same values of $d_o$ and the same type of disturbances as in Table 2, but using (12) along with the new model (14). We see that if $u_t$ is white noise, the unit root null hypothesis is rejected in favour of higher orders of integration, and $H_o$ (4) cannot be rejected when $d_o$ = 1.20, 1.40, 1.60 and 1.80. If $u_t$ is AR(1), the non-rejection values are $d_o$ = 0.80 and 1.00, and if it is AR(2) the values are slightly higher: 1, 1.20 and 1.40. Using the Bloomfield exponential spectral model, the results are the same with one or two parameters, and $H_o$ cannot be rejected at $d_o$ = 0.80, 1, 1.20 and 1.40. Finally, including seasonal AR processes of the form given by (13), the values coincide with those using white noise disturbances, i.e., 1.20, 1.40, 1.60 and 1.80. Table 5 is the counterpart to Table 3 with seasonal fractional integration, reporting the confidence intervals and the values of $d_o^*$ for each type of disturbances. If $u_t$ is white noise or seasonal AR, the values are higher than 1. For the remaining four cases (AR and Bloomfield $u_t$) the values are around 1. In the following section, we try to select the best model specification from all these potential rival specifications.

[10] Empirical applications based on this version of Robinson's (1994) tests can be found, among others, in

## 4. Model selection

First, we focus on the models presented in Tables 2 and 3 and choose, for each type of disturbances, the model with the lowest statistic in absolute value. The selected models are described in the upper part of Table 6 (denoted by NS#).[11] Simple visual inspection of the residuals for the models NS1-NS3 suggests that these are not adequate specifications, in view of the seasonal structure still apparent in the residuals (the charts are not included in the paper for reasons of space). Thus, we only compare the models NS4 and NS5 on the basis of their diagnostics.

**(Insert Table 6 about here)**

The lower part of Table 6 describes the selected models in Tables 4 and 5 based on seasonal fractional integration. They are now denoted by S#. Here, we observe that S4 and S5 (the models based on seasonal autoregressions) produce results very similar to S1 (based on a white noise $u_t$) in terms of the estimated coefficients of the non-linear variables. Moreover, the coefficients of the seasonal AR parameters are in both cases close to zero, suggesting that seasonal autoregressions are not required in the context of seasonal fractional integration. Therefore, we have five potential models to describe the series of interest: NS4, NS5, S1, S2 and S3. We test for no serial autocorrelation by means of a slight modification of the test proposed by Eitrheim and Teräsvirta (1996) for the standard STAR model. In particular, the null hypothesis of no autocorrelation in the residuals $\varepsilon_t$ can be tested against the alternative of serial dependence up to order q, that is, under the alternative $\varepsilon_t$ satisfies:

---

Gil-Alana and Robinson (2001) and Gil-Alana (2002).

[11] Note that the models based on Bloomfield (1973) disturbances are not considered since they do not have a parametric formula for the weak dependence structure.

$$\varepsilon_t = \alpha_1 \varepsilon_{t-1} + \ldots + \alpha_q \varepsilon_{t-q} + e_t,$$

where $e_t$ ~ i.i.d. $(0, \sigma^2)$. The null hypothesis is given by $H_o$: $\alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$ which, following Eitrheim and Teräsvirta (1996), is tested by means of an LM test. Here the only difference compared to that test is that one needs to include the gradient of $e_t$ with respect to the fractional differencing parameter d, evaluated under $H_o$. Under the null $\varepsilon_t = e_t$, so that

$$\frac{\partial \varepsilon_t}{\partial d} = \frac{\partial e_t}{\partial d} - \sum_{j=1}^{t-1} \frac{\varepsilon_{t-j}}{j}.$$

Performing the tests on the five selected models, the results reject the null hypothesis of no serial correlation in all models except S3, suggesting that a plausible model might be:

$$\underset{}{UNE_t} = \underset{(0.418)}{-2.212}\left[1 - G(ROP_t)\right] + \underset{(0.155)}{3.254}\, G(ROP_t) + \underset{(0.111)}{0.567}\left[1 - G(RIR_t)\right] + x_t;$$

$$(1 - L^4)^{1.15} x_t = u_t; \quad u_t = 0.714 u_{t-1} + 0.101 u_{t-2} + \varepsilon_t, \qquad (15)$$

(standard errors in parentheses), with the implication that unemployment is nonstationary and non-mean-reverting.[12] These findings allow us to discriminate between rival unemployment theories. Specifically, a natural rate model would require the process to obey mean reversion, the effects of shocks dying away and the unemployment rate reverting to its underlying equilibrium level. By contrast, in a hysteresis model the short-run equilibrium level depends on actual past levels, as shocks are not mean reverting, at least in a finite time horizon. The evidence presented here clearly gives support to the latter type of model, and to arguments in favour of more active stabilisation policies. [13]

---

[12] Note that even though d is higher than 1 in this model, the unit root null (d = 1) cannot be rejected at the 5% level (see Table 5).

[13] Harding and Pagan (2002) assess the usefulness of non-linear models (specifically, a simple Markov-chain process, and one exhibiting duration dependence) for replicating the business cycle features of US GDP, and find little evidence that non-linear effects are important to the nature of the cycle. However, theirs is a univariate approach, and as such it is not directly comparable to our multivariate model.

A limitation of the procedure we follow is that it imposes the same order of integration at zero and the seasonal frequencies. Note that the polynomial ($1-L^4$) can be decomposed into ($1-L$)($1+L$)($1+L^2$), where each of these polynomials correspond to the zero, the annual ($\pi$) and the bi-annual ($\pi/2$ and $3\pi/2$) frequencies. Thus, the large coefficient of the fractional differencing parameter may be partly due to the joint effect of the trend and the seasonal components. The tests of Robinson (1994) described in Section 2 also allow us to consider the case of different orders of integration at each of these frequencies (see, e.g. Gil-Alana, 2003), but this is not within the scope of the present paper.

## 5. Forecasting properties

In this section we compare the model selected in the previous section with another model with a linear structure. In particular, we consider the same class of models as in Table 5 but replacing the non-linear specification by a linear one, namely:

$$UNE_t \;=\; \beta_0 \, RIR_t \;+\; \beta_1 \, ROP_t \;+\; x_t . \qquad (16)$$

Note that, although only actual values of the input variables are explicitly presented in the regression model (16), the lagged structure is included through the fractional polynomials ($(1-L)^d$ and $(1-L^4)^d$) and the autoregressive terms.

**(Insert Table 7 about here)**

The selected models are described in Table 7, the selection criteria being the same as before. It can be seen that, when using non-seasonal specifications (NS#, i.e., $(1-L^4)^d$), the orders of integration are very similar to those of Table 6. They are smaller than 1 if $u_t$ is white noise or AR(2); exactly 1 for AR(1) disturbances; and higher than 1 for seasonal autoregressions. When using the seasonal fractional polynomial (S#), the orders of integration vary substantially depending on how we specify the I(0) term: d is equal to 1.51 for a white noise $u_t$; it is close to 0 (d = 0.13) with AR(1) disturbances, and higher than 1 in the remaining cases. Performing the same tests as in Section 4, we reach the conclusion that the best model is the seasonal fractional one with AR(2) disturbances, i.e.,

$$\underset{}{UNE_t} = -\underset{(0.037)}{0.033}RIR_t + \underset{(0.308)}{3.348}ROP_t + x_t; \ (1-L^4)^{1.18} x_t = u_t; \ u_t = 0.688 u_{t-1} + 0.124 u_{t-2} + \varepsilon_t \quad (17)$$

Next, we compare the two models (i.e. the non-linear and the linear one), on the basis of their forecast accuracy. We use data from 2002q3 to 2005q1 for the out-of-sample forecasting exercise. We could also have employed other non-linear and linear models. However, in another recent application, Candelon and Gil-Alana (2003) showed that simple fractional models could better characterise macroeconomic series than other more complex models.

The accuracy of different forecasting methods is a topic of continuing interest and research (see, e.g., Makridakis et al., 1998 and Makridakis and Hibon, 2000, for a review of the forecasting accuracy of competing forecasting models). Note, however, the criticism of Clements (2002), who emphasises that the forecast performance of dynamic models including some exogenous variables may not be a good guide to their adequacy.

Since the two specifications (models (15) and (17)) are based on dynamic models, we use predictions of the actual values of the dependent variables. Note that the two models impose a seasonally fractionally integrated structure on these variables, and, therefore, predictions can be easily obtained through Binomial expansions.

**(Insert Table 8 about here)**

Table 8 displays the k-period ahead forecast errors of the two models. It can be seen that the non-linear one (model (15)) produces better results in practically all cases. Also, the RMSE is lower in the non-linear case. Of course, this measure of forecast accuracy is a purely descriptive device. There exist several statistical tests for comparing different forecasting models. One of these tests, widely employed in the time series literature, is the asymptotic test for a zero expected loss differential due to Diebold and Mariano (1995). On the basis of this test, we can reject the null hypothesis that the forecast performance of models (15) and (17) is equal in favour of the one-sided alternative that model (15) outperforms its rival at the 5% significance level.

## 6. Conclusions

This paper has proposed a model of the US unemployment rate which can account for both its asymmetry and its long memory. Our approach, which is based on the tests of Robinson (1994), introduces fractional integration and nonlinearities simultaneously into the same framework, unlike earlier studies employing a sequential procedure (see van Dijk et al, 2002). Conveniently, ours is instead a single-step procedure based on the Lagrange Multiplier, therefore following a standard null limit distribution. The empirical results indicate that the US unemployment rate can be specified in terms of a fractionally integrated process, which interacts with some non-linear functions of the labour demand variables (real oil prices and real interest rates). We find that the order of integration of the series is higher than 1, implying that, even when taking first differences, they still possess a component of long memory behaviour, with the autocorrelations decaying slowly (hyperbolically) to zero. Although $d = 1.15$, the unit root hypothesis cannot be rejected. Also, given the fact that the logistic transformation we are considering is unbounded, its observed nonstationary behaviour does not raise any difficulties in terms of economic interpretation. Moreover, it is consistent with other studies that model unemployment in terms of a cointegrating relationship.[14]

On the whole, our findings suggest that a hysteresis model with path dependency (see, e.g., Blanchard and Summers, 1987) is suitable for the US unemployment rate. This implies that there exists no constant long-run equilibrium rate, with the effects of exogenous shocks not dying away within a finite time horizon, and unemployment being nonstationary. Evidence of nonstationarity was also reported, within a standard unit root framework, by Mitchell and Wu (1995), Carruth et al. (1998), and Strazicich et al. (2001) inter alia, whilst Wilkins (2003) found an order of integration higher than 1 at the seasonal frequencies. By contrast, in a NAIRU (Non Accelerating Inflation Rates of Unemployment) model, in which shocks are not long-lived, the unemployment rate reverts back to its underlying equilibrium level (see, e.g., Friedman, 1968). The implications for policy-makers are of great importance, as, on the basis of our results, activist policies to combat unemployment can be pursued. In particular, monetary policy can be effectively used without immediate inflationary consequences, since it can affect the microeconomic foundations of the labour market equilibrium. However, our analysis also confirms that any adequate model should include business cycle asymmetries, which might arise for a variety of micro- or macro-economic reasons (see, e.g., Bentolilla and Bertola, 1990, and Caballero and Hammour, 1994). The existence of such nonlinearities should be an essential feature of empirical models of the unemployment rate, and represents important information for both forecasters and policy-makers. For instance, it

[14] Note that the I(d) structure observed in the process might be a consequence of the non-linear transformations that are being applied to the original unemployment series (see Dittman and Granger, 2002).

implies that the probability of erring in forecasting is asymmetric, and so are the costs in terms of foregone output and higher output variability for a given objective function. This should be clearly taken into account when formulating stabilisation policies.

Other approaches, such as the semiparametric techniques developed by Beran, Geng and Ocker (1999) and Beran and Ocker (2001), or even the nonlinear cointegration technique of Granger and Hallman (1991), could also be used. It should be stressed, however, that the approach employed in this paper is not concerned with the estimation of the fractional differencing parameter involved in the nonlinear relationship of interest, but simply computes diagnostics for departures from any real value d. Thus, it is not surprising that, when fractional hypotheses are considered, many non-rejection values are found. It may also be worthwhile to obtain point estimates of the parameters of interest by means of maximum likelihood or Whittle approximations, though our expectation is that the results would be in line with those reported here. Furthermore, the tests for the order of integration are dependent on the particular type of nonlinearity assumed (i.e. STAR), which is not tested against a linear alternative but simply assumed. However, the coefficients corresponding to the selected model are all significant, suggesting the validity of such a model.

## References


Altissimo, F. and G.L. Violante (2001), "The nonlinear dynamics of output and unemployment in the US", Journal of Applied Econometrics, 16, 4, 461-486.

Bean, C.R. (1989), "Capital shortage", Economic Policy, 8, 11-53.

Beaudry, P. and G. Koop (1993), "Do recessions permanently change output?", Journal of Monetary Economics, 31, 149-163.

Bentolilla, S. and G. Bertola (1990), "Firing costs and labour demand: how bad is eurosclerosis?", Review of Economic Studies, 57, 381-402.

Beran, J., Y. Feng and D. Ocker (1999), "SEMIFAR models", Technical Report 3/1999, SFB 475, University of Dortmund.

Beran, J. and D. Ocker (2001) "Volatility of stock market indices_ an analysis based on SEMIFAR models", Journal of Business and Economics Statistics, 19, 103-116.

Bianchi, M., and G. Zoega (1998), "Unemployment persistence: does the size of the shock matter?", Journal of Applied Econometrics, 13, 283-304.

Blanchard, O.J. and L.H. Summers (1987), "Hysteresis in unemployment", European Economic Review, 31, 288-295.

Bloomfield, P. (1973), "An exponential model in the spectrum of a scalar time series", Biometrika, 60, 27-226.

Caballero, R.J. and M.L. Hammour (1994), "The cleansing effect of recessions", American Economic Review, 84, 1350-1368.

Candelon, B. and L.A. Gil-Alana (2003) "Fractional integration and business cycle characteristics", Empirical Economics, 29, 343-359.

Caner, M. and B.E. Hansen (2001), "Threshold autoregression with a unit root", Econometrica, 69, 1555-1596.

Caporale, G.M. and L.A. Gil-Alana (2002), "Unemployment and input prices: a fractional cointegration approach", Applied Economics Letters, 9, 347-351.

Carruth, A., A.J. Oswald, and M. Hooker (1998), “Unemployment equilibria and input prices”, Review of Economics and Statistics, 80, 621-628.

Clements, M.P., 2002, “Why forecast performance does not help us choose a model”, Working Paper, Department of Economics, University of Warwick.

Coakley, J., Fuertes, A.-M. and G. Zoega (2001), “Evaluating the persistence and structuralist theories of unemployment from a nonlinear perspective”, Studies in Non-linear Dynamics and Econometrics, 5, 1078-1091.

Cross, R.B. (1994), “The macroeconomic consequences of discontinuous adjustment: selected memory of non-dominated extrema”, Scottish Journal of Political Economy, 41, 212-221.

Cross, R.B. (1995), The Natural Rate of Unemployment: Reflections on 25 Years of the Hypothesis, Cambridge, UK, Cambridge University Press.

Diebold, F.X. and R.S. Mariano (1995) Comparing predictive accuracy, Journal of Business, Economics and Statistics 13, 253-263.

Dittman, I. and C.W.J. Granger (2002), "Properties of nonlinear transformations of fractionally integrated processes", Journal of Econometrics, 110, 113-133.

Dueker, M.J. and P.K. Asea (1998). “Non-monotonic LM dynamics in Black-market premia”, Typescript. The Federal Reserve Bank of St. Louis.

Eitrheim O. and T. Teräsvirta (1996), “Testing the adequacy of smooth transition autoregressive models”, Journal of Econometrics, 74, 59-76.

Franses, P.H. and R. Paap (1998), “Modelling asymmetric persistence over the business cycle”, Erasmus University Rotterdam, Econometric Institute Report 9852/A.

Franses, P.H. and D. van Dijk, (2002), “A nonlinear long memory model for US unemployment”, Preprint.

Friedman, M. (1968), “The role of monetary policy”, American Economic Review, 58, 1-17.

Gil-Alana, L.A. (2000), “Mean reversion in the real exchange rates”, Economics Letters, 69, 285-288.

Gil-Alana, L.A. (2001), “A fractionally integrated exponentially spectral model for the UK unemployment”, Journal of Forecasting, 20, 329-340.

Gil-Alana, L.A. (2002), “Seasonal long memory in the aggregate output”, Economics Letters, 74, 333-337.

Gil-Alana, L.A. (2003), “Seasonal misspecification in the context of fractionally integrated univariate time series”, Computational Economics, 22, 65-74.

Gil-Alana, L.A. and P.M. Robinson (1997), “Testing of unit root and other nonstationary hypotheses in macroeconomic time series”, Journal of Econometrics, 80, 241-268.

Gil-Alana, L.A. and P.M. Robinson (2001), “Testing seasonal fractional integration in the UK and Japanese consumption and income”, Journal of Applied Econometrics, 16, 95-114.

Granger, C.W.J. and Z. Ding (1996), “Varieties of long memory models”, Journal of Econometrics,73, 61-77.

Granger, C.W.J. and J. Hallmann (1991), "Long memory series with attractors", Oxford Bulletin of Economics and Statistics, 53, 11-26.

Hamermesh, D.S. and G.A. Pfann (1996), “Adjustment costs in factor demand”, Journal of Economic Literature, 34, 1264-1292.

Hansen, B.E. (1997), “Inference in TAR models”, Studies in Nonlinear Dynamics and Econometrics, 2, 1-14.

Harding, D. and A. Pagan (2002), “Dissecting the cycle: a methodological investigation”, Journal of Monetary Economics, 49, 365-381.

Hassler, U. (1994) “Misspecification of long memory seasonal time series”, Journal of Time Series Analysis, 15, 19-30.

Hosoya, Y. (1997), “A limit theory for long range dependence and statistical inference on related models”, Annals of Statistics, 25, 105-137.

Koop, G. and S.M. Potter (1999), “Dynamic asymmetries in the US unemployment, Journal of Business and Economic and Statistics, 17, 298-313.

Lindbeck, A. and D.J. Snower (1988), The Insider-Outsider Theory of Employment and Unemployment, Cambridge, Mass., MIT Press.

Makridakis, S., S. Wheelwright and R. Hyndman (1998), “Forecasting methods and applications”, 3rd Edition, John Wiley & Sons.

Makridakis, S. and M. Hibon (2000) “The M-3 competition: results, conclusions and implications”, International Journal of Forecasting, 16, 451-476.

Mitchell, W. and P. Wu (1995), “Evidence of the unit root hypothesis in quarterly unemployment rates”, mimeo, Business and Law Faculty, University of Newcastle, NSW, Australia.

Montanari, A., R. Rosso and M.S. Taqqu (1995), "A seasonal fractional differenced ARIMA model: An application to the River Nile monthly flows at Aswan", Preprint.

Montanari, A., R. Rosso and M.S. Taqqu (1996), "Some long-run properties of rainfall records in Italy", Journal of Geoophysical Research Atmospheres, 101, 431-438.

Montanari, A., R. Rosso and M.S. Taqqu (1997), "Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation and simulation", Water Resources Research, 33l, 1035-1044.

Montgomery, A.L., Zarnowitz, V., Tsay, R.S. and G.C. Tiao (1998), “Forecasting the US unemployment rate”, Journal of the American Statistical Association, 93, 478-493.

Neftci, S. (1984), “Are economic time series asymmetric over the business cycle?”, Journal of Political Economy, 92, 307-328.

Ooms, M. (1995), “Flexible seasonal long memory and economic time series”, Discussion Paper 9951/A. Presented at the World Congress Econometric Society 1995, Tokyo, Econometric Institute, Erasmus University Rotterdam.

Papell, D.H., C.J. Murray and H. Ghiblawi (2000), “The structure of unemployment”, Review of Economics and Statistics, 82, 309-315.

Parker, R.E. and P. Rothman (1998), “The current depth-of-recession and unemployment-rate forecasts”, Studies in Nonlinear Dynamics and Econometrics, 2, 151-158.

Pfann, G.A. (1993), “The thin end of the wedge: non-linearity in stochastic labor market dynamics”, in J.C. van Ours, G.A. Pfann, and G. Ridder (eds.), Labor Demand and Equilibrium Wage Formation, 1-26, Amsterdam, North-Holland.

Potter, S.M. (1995), “A nonlinear approach to US GNP”, Journal of Applied Econometrics, 10, 109-125.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1986), “Numerical recipes: The Art of Scientific Computing”, Cambridge University Press, Cambridge.

Ramsey, J.B. and P. Rothman (1996), “Time irreversibility and business cycle asymmetry”, Journal of Money, Credit and Banking, 28, 1-21.

Robinson, P.M. (1994), “Efficient tests of nonstationary hypotheses”, Journal of the American Statistical Association, 89, 1420-1437.

Rothman, P. (1991), “Further evidence on the asymmetric behaviour of unemployemnt rates over the business cycle”, Journal of Macroeconomics, 13, 292-298.

Rothman, P. (1998), “Forecasting asymmetric unemployment rates”, Review of Economics and Statistics, 80, 164-168.

Rothman, P. (1999), “Time irreversible unemployment rates”, Working Papaer 9903, Department of Economics, East Carolina University.

Skalin, J. and T. Teräsvirta (2002), “Modelling asymmetries and moving equilibria in unemployment rates”, Macroeconomic Dynamics 6, 202-241.

Strazicich, M.C., M. Tieslau and J. Lee (2001), “Hysteresis in unemployment? Evidence from panel unit root tests with structural change”, mimeo, Department of Economics, University of North Texas, Denton, Texas.

Van Dijk, D.J.C., Franses, P.H. and R. Paap, (2002), “A nonlinear long memory model for US unemployment”, Macroeconomic Dynamics, 6, 202-241.

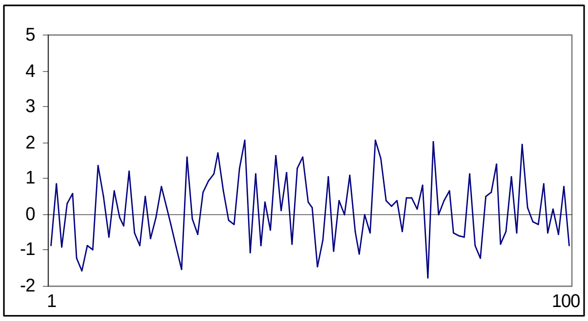
Wallis, K., (1987), “Time series analysis of bounded economic variables”, Journal of Time Series Analysis, 8, 115-123.

Wilkins, N. (2003), “Fractional integration at a seasonal frequency with an application to quarterly unemployment rates”, mimeo, School of Finance and Economics, UTS, Sydney.

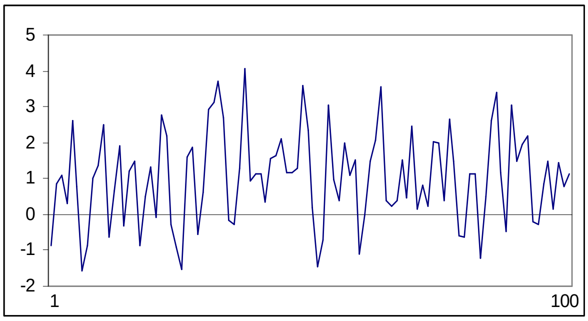
**FIGURE 1**

Fractional integration with non-linear models

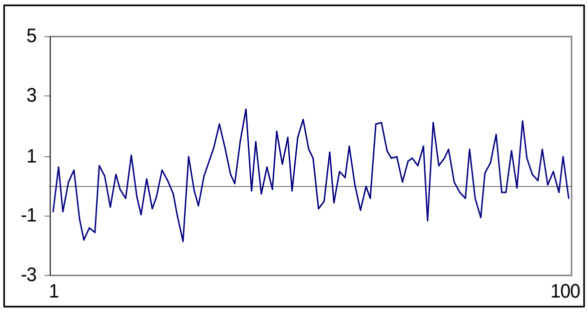
d = 0 and a = 0



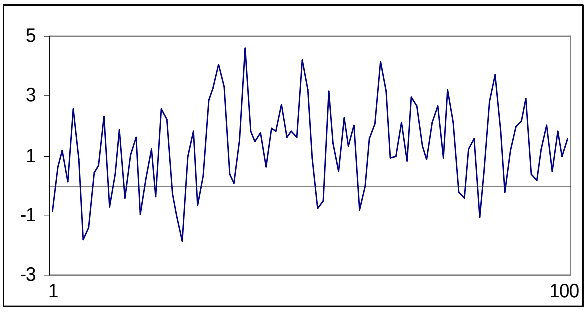

d = 0 and a = 2



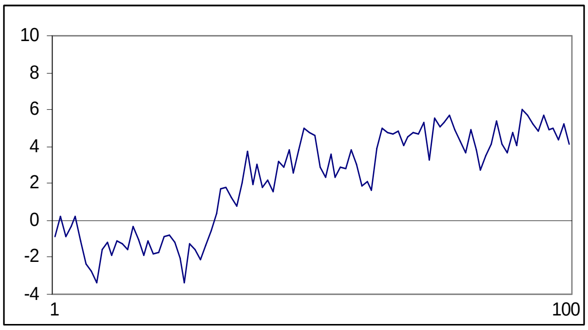

d = 0.25 and a = 0



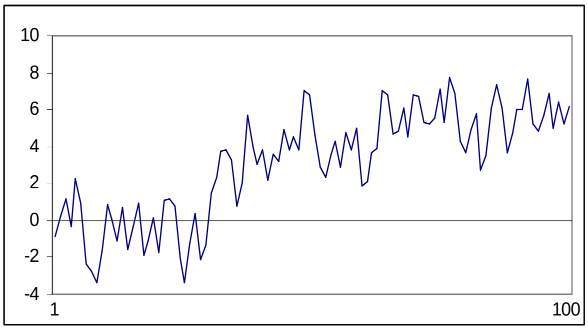

d = 0.25 and a = 2



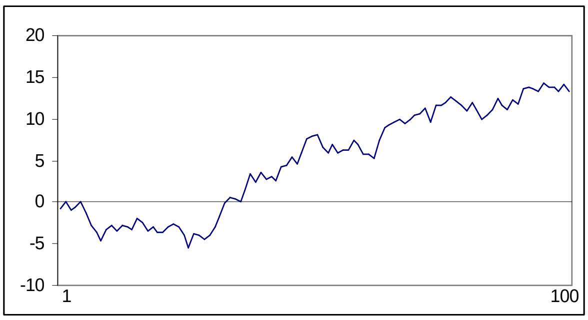

d = 0.75 and a = 0



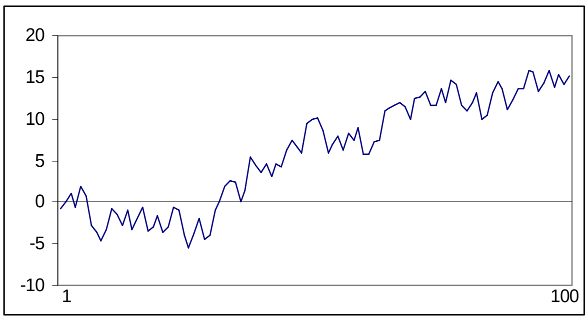

d = 0.75 and a = 2




d = 1.00 and a = 0




d = 1.00 and a = 2



| TABLE 1 | | | | | | |
|---|---|---|---|---|---|---|
| Rejection frequencies of the procedure in Section 2: Fractional integration and non-linearities | | | | | | |
| | **T = 200** | | | **T = 400** | | |
| | **Case a)** | **Case b)** | **Case c)** | **Case a)** | **Case b)** | **Case c)** |
| **0.00** | **0.996** | **0.967** | **0.995** | **1.000** | **0.998** | **1.000** |
| **0.10** | **0.972** | **0.862** | **0.972** | **0.999** | **0.988** | **1.000** |
| **0.20** | **0.871** | **0.602** | **0.872** | **0.994** | **0.891** | **0.993** |
| **0.30** | **0.567** | **0.264** | **0.567** | **0.880** | **0.477** | **0.881** |
| **0.40** | **0.181** | **0.108** | **0.180** | **0.361** | **0.116** | **0.361** |
| **0.50** | **0.064** | **0.261** | **0.065** | **0.057** | **0.403** | **0.057** |
| **0.60** | **0.234** | **0.609** | **0.234** | **0.396** | **0.882** | **0.395** |
| **0.70** | **0.599** | **0.888** | **0.599** | **0.904** | **0.996** | **0.905** |
| **0.80** | **0.884** | **0.982** | **0.885** | **0.996** | **1.000** | **0.997** |
| **0.90** | **0.982** | **0.998** | **0.983** | **1.000** | **1.000** | **1.000** |
| **1.00** | **0.998** | **0.999** | **0.999** | **1.000** | **1.000** | **1.000** |

Case a) refers to the case where we truly identified the non-linear and the fractionally integrated structures. Case b) refers to the situation where we test for fractional integration ignoring the existence of the non-linear structure. In Case c) we test for fractional integration and non-linearities in a model without a non-linear structure.

| TABLE 2 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Testing the order of integration with the tests of Robinson (1994) in a fractional model | | | | | | | | | | | |
| | **0.00** | **0.20** | **0.40** | **0.60** | **0.80** | **1.00** | **1.20** | **1.40** | **1.60** | **1.80** | **2.00** |
| **White noise** | **15.55** | **11.98** | **8.22** | **3.98** | -0.05 | **-2.83** | **-4.40** | **-5.27** | **-5.80** | **-6.15** | **-6.41** |
| **AR (1)** | **5.01** | **4.08** | **3.42** | **2.80** | **1.79** | -0.09 | **-1.69** | **-2.75** | **-3.39** | **-3.77** | **-4.00** |
| **AR (2)** | **3.51** | **2.96** | **2.48** | **1.90** | 0.08 | -1.22 | **-2.75** | **-4.00** | **-4.90** | **-5.50** | **-5.92** |
| **Bloomfield (1)** | **5.08** | **3.15** | **2.95** | **2.49** | **1.97** | 0.10 | **-2.10** | **-3.95** | **-5.35** | **-6.38** | **-7.17** |
| **Bloomfield (2)** | **7.01** | **4.96** | **3.68** | **3.44** | **2.72** | 0.13 | **-2.89** | **-5.45** | **-7.37** | **-8.80** | **-9.89** |
| **Seasonal** | **11.36** | **9.22** | **7.07** | **5.23** | **3.70** | **2.41** | 1.26 | 0.19 | -0.82 | **-1.79** | **-2.69** |
| **Seasonal** | **16.10** | **15.22** | **13.58** | **10.40** | **6.16** | **2.85** | 0.20 | **-1.85** | **-3.36** | **-4.41** | **-5.15** |

In bold, the non-rejection values of the null hypothesis at the 5% level.

| TABLE 3 | | |
|---|---|---|
| Confidence Intervals of the non-rejection values of $d_o$ at the 95% significance level | | |
| **Disturbances** | **Confidence Intervals** | **d** |
| **White noise** | **[0.72 - 0.90]** | **0.80** |
| **AR (1)** | **[0.82 - 1.19]** | **0.99** |
| **AR (2)** | **[0.69 - 1.05]** | **0.82** |
| **Bloomfield (1)** | **[0.85 - 1.15]** | **1.01** |
| **Bloomfield (2)** | **[0.89 - 1.11]** | **1.01** |
| **Seasonal AR (1)** | **[1.14 - 1.76]** | **1.44** |
| **Seasonal AR(2)** | **[1.09 - 1.37]** | **1.22** |

| TABLE 4 | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Testing the order of integration with the tests of Robinson (1994) in a seasonal fractional model | | | | | | | | | | | |
| | **0.00** | **0.20** | **0.40** | **0.60** | **0.80** | **1.00** | **1.20** | **1.40** | **1.60** | **1.80** | **2.00** |
| **White noise** | **6.64** | **6.09** | **5.53** | **4.80** | **3.73** | **2.55** | 1.54 | 0.74 | -0.08 | -0.47 | **-1.96** |
| **AR (1)** | **3.14** | **2.89** | **2.44** | **2.20** | -0.86 | -1.44 | **-1.92** | **-2.32** | **-2.65** | **-2.94** | **-3.18** |
| **AR (2)** | **6.61** | **6.24** | **5.19** | **3.85** | **2.33** | 0.92 | -0.26 | -1.23 | **-1.98** | **-2.55** | **-2.97** |
| **Bloomfield (1)** | **2.34** | **2.19** | **1.88** | **1.73** | 0.89 | 0.05 | -0.73 | -1.38 | **-1.94** | **-2.42** | **-2.84** |
| **Bloomfield (2)** | **2.13** | **2.01** | **1.92** | **1.76** | 0.92 | 0.05 | -0.75 | -1.42 | **-1.99** | **-2.49** | **2.92** |
| **Seasonal** | **4.73** | **3.98** | **3.16** | **3.16** | **2.95** | **2.59** | 1.50 | 1.39 | 1.17 | 0.26 | **-1.92** |
| **Seasonal** | **6.17** | **4.09** | **2.91** | **2.15** | **1.95** | **1.69** | 1.20 | 1.00 | 0.17 | -1.34 | **-2.33** |

**In bold, the non-rejection values of the null hypothesis at the 5% level.**

| TABLE 5 | | |
|---|---|---|
| Confidence Intervals of the non-rejection values of $d_o$ at the 95% significance level | | |
| **Disturbances** | **Confidence Intervals** | **d** |
| **White noise** | **[1.22 - 1.91]** | **1.59** |
| **AR (1)** | **[0.70 - 1.13]** | **0.94** |
| **AR (2)** | **[0.90 - 1.50]** | **1.15** |
| **Bloomfield (1)** | **[0.66 - 1.50]** | **1.06** |
| **Bloomfield (2)** | **[0.69 - 1.48]** | **1.05** |
| **Seasonal AR (1)** | **[1.17 - 1.90]** | **1.70** |
| **Seasonal AR(2)** | **[1.08 - 1.84]** | **1.62** |

| TABLE 6 | |
|---|---|
| Selected models from Tables 1 and 2 | |
| **NS1** | $UNE_t = -2.136V_1 - 0.960V_2 + 2.624V_3 + x_t;\ (1-L)^{0.80}x_t = \varepsilon_t$ <br> (1.001) (1.110) (1.050) |
| **NS2** | $UNE_t = -2.073V_1 - 1.462V_2 + 2.870V_3 + x_t;\ (1-L)^{0.99}x_t = u_t;\ u_t = -0.268u_{t-1} + \varepsilon_t$ <br> (1.017) (1.223) (1.097) |
| **NS3** | $UNE_t = -2.132V_1 - 1.033V_2 + 2.665V_3 + x_t;\ (1-L)^{0.82}x_t = u_t;\ u_t = -0.090u_{t-1} + 0.165u_{t-2} + \varepsilon_t$ <br> (0.987) (1.106) (1.038) |
| **NS4** | $UNE_t = -1.755V_1 - 1.494V_2 + 2.824V_3 + x_t;\ (1-L)^{1.44}x_t = u_t;\ u_t = 0.839u_{t-4} + \varepsilon_t$ <br> (0.678) (0.895) (0.766) |
| **NS5** | $UNE_t = -1.926V_1 - 1.606V_2 + 2.908V_3 + x_t;\ (1-L)^{1.22}x_t = u_t;\ u_t = -0.091u_{t-4} + 0.781u_{t-8} + \varepsilon_t$ <br> (0.757) (0.781) (0.841) |
| Selected models from Tables 3 and 4 | |
| **S1** | $UNE_t = -4.268V_1 - 2.815V_2 + 4.075V_3 + x_t;\ (1-L^4)^{1.59}x_t = \varepsilon_t$ <br> (0.775) (0.712) (0.707) |
| **S2** | $UNE_t = -2.725V_1 + 2.770V_2 + 0.141V_3 + x_t;\ (1-L^4)^{0.14}x_t = u_t;\ u_t = 0.791u_{t-1} + \varepsilon_t$ <br> (0.427) (0.158) (0.440) |
| **S3** | $UNE_t = -2.212V_1 + 3.254V_2 + 0.567V_3 + x_t;\ (1-L^4)^{1.15}x_t = u_t;\ u_t = 0.714u_{t-1} + 0.110u_{t-2} + \varepsilon_t$ <br> (0.418) (0.155) (0.111) |
| **S4** | $UNE_t = -4.270V_1 - 312004V_2 + 3.882V_3 + x_t;\ (1-L^4)^{1.69}x_t = u_t;\ u_t = -0.014u_{t-4} + \varepsilon_t$ <br> (0.650) (0.627) (0.615) |
| **S5** | $UNE_t = -4.268V_1 - 2.805V_2 + 4.078V_3 + x_t;\ (1-L^4)^{1.61}x_t = u_t;\ u_t = -0.011u_{t-4} + 0.08u_{t-8} + \varepsilon_t$ <br> (0.617) (0.818) (0.729) |

Standard errors in parentheses.

| **TABLE 7** | |
|---|---|
| Selected models based on fractional non-seasonal models | |
| **NS1** | $UNE_t = -0.184\,RIR_t + 0.677\,ROP_t + x_t; \quad (1-L)^{0.80} x_t = \varepsilon_t$ <br> (0.076) (0.485) |
| **NS2** | $UNE_t = -0.194\,RIR_t + 0.311\,ROP_t + x_t; \quad (1-L)^{1.00} x_t = u_t; \quad u_t = -0.279\,u_{t-1} + \varepsilon_t$ <br> (0.078) (0.466) |
| **NS3** | $UNE_t = -0.186\,RIR_t + 0.628\,ROP_t + x_t; \ (1-L)^{0.82} x_t = u_t; \ u_t = -0.092\,u_{t-1} + 0.169\,u_{t-2} + \varepsilon_t$ <br> (0.075) (0.474) |
| **NS4** | $UNE_t = -0.176\,RIR_t + 0.062\,ROP_t + x_t; \quad (1-L)^{1.43} x_t = u_t; \quad u_t = 0.842\,u_{t-4} + \varepsilon_t$ <br> (0.052) (0.283) |
| **NS5** | $UNE_t = -0.188\,RIR_t + 0.125\,ROP_t + x_t; \ (1-L)^{1.23} x_t = u_t; \ u_t = -0.094\,u_{t-4} + 0.779\,u_{t-8} + \varepsilon_t$ <br> (0.059) (0.779) |
| Selected models based on seasonal fractional models | |
| **S1** | $UNE_t = -0.319\,RIR_t + 1.261\,ROP_t + x_t; \quad (1-L^4)^{1.51} x_t = \varepsilon_t$ <br> (0.052) (0.400) |
| **S2** | $UNE_t = -0.020\,RIR_t + 3.523\,ROP_t + x_t; \quad (1-L^4)^{0.13} x_t = u_t; \quad u_t = 0.768\,u_{t-1} + \varepsilon_t$ <br> (0.040) (0.334) |
| **S3** | $UNE_t = -0.033\,RIR_t + 3.348\,ROP_t + x_t; \ (1-L^4)^{1.18} x_t = u_t; \ u_t = 0.688\,u_{t-1} + 0.124\,u_{t-2} + \varepsilon_t$ <br> (0.037) (0.308) |
| **S4** | $UNE_t = -0.295\,RIR_t + 0.699\,ROP_t + x_t; \quad (1-L^4)^{1.65} x_t = u_t; \quad u_t = -0.566\,u_{t-4} + \varepsilon_t$ <br> (0.039) (0.301) |
| **S5** | $UNE_t = -0.288\,RIR_t + 0.665\,ROP_t + x_t; \ (1-L^4)^{1.65} x_t = u_t; \ u_t = -0.049\,u_{t-4} + 0.03\,u_{t-8} + \varepsilon_t$ <br> (0.034) (0.278) |

| TABLE 8 | | |
|---|---|---|
| Forecast prediction errors of the selected models | | |
| **Time period** | **Model 15 (non-linear)** | **Model 17 (linear)** |
| **2002q4** | **0.3202** | **0.3263** |
| **2003q1** | **0.0664** | **0.0658** |
| **2003q2** | **-0.0817** | **-0.0782** |
| **2003q3** | **-0.2351** | **-0.2281** |
| **2003q4** | **-0.1879** | **-0.1903** |
| **2004q1** | **-0.3561** | **-0.3562** |
| **2004q2** | **-0.0715** | **-0.0665** |
| **2004q3** | **0.0461** | **0.0485** |
| **2004q4** | **0.0872** | **0.0902** |
| **2005q1** | **0.3517** | **0.3583** |
| RMSE | 0.906 | 0.958 |