# Appendix of: The Nature of Occupational Unemployment Rates in the United States: Hysteresis or Structural?*

B. Candelon†, A. Dupuy‡ and L. Gil-Alana§

June 26, 2008

## 1 A Simulation Study

In this section, we present some Monte Carlo experiments describing the performance of the procedure of Section 3. Initially, we consider that the true data generating process is given by Equation ??, with $D_t = I(t > T_b)$; $T_b = T/2$, $\beta = 10$, and $d_1 = d_2 = 1$. In other words, the true model is:

$$u_t = 10D_t + x_t, \qquad \rho(L;d)x_t = w_t, \qquad \rho(L;d) = (1 - L^{12})(1 - L) \tag{1}$$

with white noise $w_t$, and $T = 120, 240, 360, 480, 600$ and 720. For this purpose, we generate Gaussian series using the routines GASDEV and RAN3 of Press, Flannery, Teukolsky and Wetterling (1986),? and 1,000 replications are used in each case.

*We thank participants at a seminar at Maastricht University. Luis Gil-Alana thanks the financial support from the Ministerio de Ciencia y Tecnologia (SEJ2005-07657/ECON, Spain).
†Department of Economics, Maastricht University, Maastricht, the Netherlands.
‡Corresponding author. Research Centre for Education and the Labour Market, (ROA) and Department of Economics, Maastricht University, P.O. Box 616, MD 6200 MD, Maastricht, the Netherlands. E-mail: a.dupuy@roa.unimaas.nl. Tel: +31 43 3883735.
§Department of Economics, University of Navarre, Pamplona, Spain.

Across Table 1 we report the probabilities of correctly determining the time of the break and the fractional differencing parameters, using a grid of values for the time break $T^* = T/10, T/10 - 1, ..., (1), ..., 9T/10 - 1$ and $9T/10$, and $(d_{10}, d_{20})$-values from 0 to 2 with 0.2 increments. The most noticeable feature observed in this table is that the procedure accurately determines the break date in all cases, and we find zero-probabilities for all values of $d_{10}$ and $d_{20}$ if $T^*$ is different from the true time of the break. Thus, we only report across the table the values of $d_{10}$ and $d_{20}$ where we observe a non-zero probability. We see that if $T = 120$, higher probabilities are obtained at other values than the true ones (e.g. $d_{10} = d_{20} = 0.8$), however, if $T > 120$, the highest probabilities are obtained in all cases at $d_{10} = d_{20} = 1$, and if $T = 600$, the probabilities corresponding to the true parameters are higher than 0.9. Note that these probabilities are based on the grid employed for the orders of integration and hence the probabilities becomes lower as the range of values for the increments in the $d$'s is reduced. On the other hand, larger increments produce larger probabilities of detecting the true values.

**(Tables 1 - 3 about here)**

We also performed the experiment with other values for the time break and the fractional differencing parameters. Table 2 displays the results for $T_b = T/4, d_{10} = 0.4$ and $d_{20} = 0.8$ (i.e. stationarity in the first subsample but nonstationarity after the shock), while Table 3 refers to the case of $T_b = 3T/4$, $d_{10} = 0.8$ and $d_{20} = 0.4$ (i.e. stationarity only after the break). The results are similar in both cases, and they are completely in line with those given in Table 1. Thus, if $T = 600$, the probability of correctly determining the true parameters exceeds 0.9 in all cases.

The accuracy in the estimation of the break date in the results presented so far might be a consequence of the coefficient used for the break dummy in equation (7). Thus, in Table 4, we examine the probability of correctly determining the break for different coefficients for the dummy variable. We now assume that the break date takes place at $T/2$, with $d_1 = d_2 = 1$, and look at the probability of detecting the true break date for a grid of values $(T/10, T/10+1, ..., 9T/10-1, 9T/10)$, using as coefficients for the deterministic break, $\beta = 10, 5, 3$ and 1.

**(Table 4 about here)**

We observe in this table that if $\beta = 10$ or 5, the procedure correctly determines the break date in the 100% of the cases even for a sample size of $T = 120$. However, reducing the magnitude of the coefficient, the probabilities are very small in some cases, especially if the sample size is small.

## 2 Robinson (1994) Score Test for Seasonal and Long-run Fractional Integration

The set-up in Robinson (1994) is the model given by equations (3) − (5), i.e, It is supposed that $w_t$ has spectral density given by:

$$u_t = \beta' z_t + x_t, \quad ,(1-L^{12})^{d_1}(1-L)^{d_2} x_t = w_t, \tag{2}$$

and suppose that the $w_t$ above has a spectral density given by:

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

where the scalar $\sigma^2$ is known and $g$ is a function of known form, which depends on frequency $\lambda$ and the unknown $(q \times 1)$ parameter vector $\tau$.

Unless $g$ is a completely known function (e.g., $g \equiv 1$, as when $w_t$ is white noise), we have to estimate the nuisance parameter $\tau$, for example by $\hat{\tau} = argmin_{\tau \in T^*} \sigma^2(\tau)$ , where $T^*$ is a suitable subset of $R^q$ Euclidean space, and:

$$\sigma^2(\tau) = \frac{2\pi}{T}\sum_{s=1}^{T-1} g(\lambda_s;\tau)^{-1} I_{\widehat{w}}(\lambda_s)$$

where $I_{\hat{w}}$ is the periodogram of $\hat{w}_t = (1-L^{12})^{d_1}(1-L)^{d_2}\hat{u}_t - \hat{\beta}' s_t$,

$$\hat{\beta} = (\sum_{t=1}^{T} s_t^2)^{-1} \sum_{t=1}^{T} s_t (1-L^{12})^{d_{10}}(1-L)^{d_{20}} y_t,$$

$$s_t = (1-L^{12})^{d_{10}}(1-L)^{d_{20}} z_t,$$

evaluated at the discrete frequencies:

$$\lambda_s = \frac{2\pi s}{T},$$

which is given by

$$I_{\hat{w}}(\lambda_s) = |(2\pi T)^{-1/2} \sum_{t=1}^{T} \widehat{w}_t e^{i\lambda_s t}|^2.$$

Note that the tests are purely parametric, requiring specific modelling assumptions regarding the short memory specification of $w_t$. Thus, for example, if $w_t$ is an AR process of form: $\phi(L)w_t = \epsilon_t$, then $g = |\phi(e^{i\lambda})|^{-2}$, with $\sigma^2 = V(\epsilon_t)$, so that the AR coefficients are a function of $\tau$.

The test statistic, which is derived via Lagrange Multiplier (LM) principle, adopts the form:

$$\hat{R} = \frac{T}{\hat{\sigma}^4}\hat{a}'\hat{A}^{-1}\hat{a}, \tag{3}$$

where $T$ is the sample size, and

$$
\begin{aligned}
\hat{\sigma^2} &= \sigma^2(\hat{\tau}) = \frac{2\pi}{T}\sum_{s=1}^{T-1}\psi(\lambda_s)g(\lambda_s;\hat{\tau})^{-1}I(\lambda_s); \\
\hat{a} &= \frac{-2\pi}{T}\sum_{s}^{*}\psi(\lambda_s)g(\lambda_s;\hat{\tau})^{-1}I(\lambda_s); \\
\hat{A} &= \frac{2}{T}(\sum_{j}^{*}\psi(\lambda_s)\psi(\lambda_s)' \\
&\quad -\sum_{j}^{*}\psi(\lambda_s)\hat{\epsilon}(\lambda_s)'(\sum_{j}^{*}\hat{\epsilon}(\lambda_s)\hat{\epsilon}(\lambda_s)'))^{-1}\sum_{j}^{*}\hat{\epsilon}(\lambda_s)\psi(\lambda_s')) \\
\psi(\lambda_s)' &= [\psi_1(\lambda_s),\psi_2(\lambda_s)]; \\
\psi_1(\lambda_s) &= log|2sin\frac{\lambda_s}{2}|; \\
\psi_2(\lambda_s) &= log|2(sin\lambda_{s/2}| + log(2cos(\lambda_{s/2})) \\
&\quad +log|2cos\lambda_s| + log|2(cos\lambda_s - cos(\pi/3)| \\
&\quad +log|2(cos\lambda_s - cos(2.\pi/3)| \\
&\quad +log|2(cos\lambda_s - cos(\pi/6)| + log|2(cos\lambda_s - cos(5.\pi/6)|; \\
\hat{\epsilon}(\lambda_s) &= \frac{\delta}{\delta\tau}logg(\lambda_s;\hat{\tau})
\end{aligned}
$$

and the summation on * in the above expressions is over $\lambda \in M$ where $M = \lambda : -\pi < \lambda < \pi,\ \lambda \notin (\rho_k - \lambda_1, \rho_k + \lambda_1)$, $k = 1, 2, ..., s$ such that $\rho_k, k = 1, 2, ..., s$ are the distinct poles of $\psi(\lambda)$ on $(-\pi, \pi]$.

# 3 Asymptotic Theory of the Test of Robinson (1994) in the Presence of a Structural Break

It is straightforward from Robinson (1994) to show that the test statistic has a standard null limit behavior. For simplicity we just concentrate here on the case of white noise $w_t$. For this purpose we need to rely on the definitions 1, 2 and 3 referring respectively to the classes F, G and H in Appendix 2 in Robinson (1994). The class F imposes a martingale difference assumption on the disturbances $w_t$, that is substantially weaker than the Gaussianity assumptionused in motivating the test. The class G imposes a mild lack of multicollinearity on the differenced series for $z_t$, which is satisfied by the dummy variables employed in the paper. The class H refers to some technical restrictions required to approximate integrals by sums. Note that under the null hypothesis of $d = d_0$, the model under analysis in (3), (4) and (5) becomes:

$$\rho(L; d)u_t = \beta\rho(L, d_0)D_t + w_t, \qquad t = 1, ..., T$$

where $w_t$ is assumed to be I(0) and thus, standard theory applies.

We call $W_t = \rho(L, d_0)D_t$ and $D = \sum_{t=1}^{T} W_t W_t'$. Then, it can be easily seen that $E\|D^{1/2}(\hat{\beta} - \beta)\|^2 = o(1)$ as $T \to \infty$, where $\hat{\beta}$ is the OLS estimation of $\beta$. The only requirement is that $D$, define as above must be a positive definite matrix for sufficiently large $T$, and this condition is satisfied by the dummy variables employed in the paper. With respect to the test statistic, we first decompose $\hat{a}$ as described in Appendix 1 into $(\hat{a} - a) + (a - a^*) + a^*$, where $a^* = \frac{-2\pi}{T}\sum_s^* \psi(\lambda_s)I_{\widehat{w}}(\lambda_s)$ and $a = -\sum_{k=1}^{t-1}\psi_k C_w(k)$, $C_w(k) = \frac{1}{t-k}\sum_{t=1}^{T-k} w_t w_{t+k}$. It follows from Theorem 1 in Robinson (1994) that $(\hat{a} - a) = o_p(T^{-1/2})$, $(a - a^*) = o_p(T^{-1/2})$ and $a^* \to_d N(0, \sigma^4\Psi)$, where $\Psi = \frac{1}{2\pi}\int_{-\pi}^{\pi}\psi(\lambda)\psi(\lambda)'d\lambda$. On the other hand, noting that $w_t^2 - \sigma^2$ are stationary martingale differences, and that $C_{\widehat{w}} - C_w \to_p 0$, then $C_w(0) \to_p \sigma^2$ and thus, it follows that $\hat{\sigma}^2 \to_p \sigma^2$. Finally, $\hat{A} \to A$ by Lemma 3 in Robinson (1994). Similar arguments can be developed with respect to the local efficiency power property of the tests (see Theorem 2 in Robinson, 1994) and with the extension to the weak autocorrelation for the I(0) disturbances $w_t$ (Theorem 3).

Table 1: Break at T/2 with d1=1 and d2=1.

| $d_1$ (Seasonality) | $d_2$ (Long run) | T=120 | T=240 | T=360 | T=480 | T=600 | T=720 |
|---|---|---|---|---|---|---|---|
| 0.6 | 0.0 | 0.010 | — | — | — | — | — |
| 0.6 | 0.8 | 0.010 | — | — | — | — | — |
| 0.8 | 0.8 | 0.234 | 0.043 | 0.010 | — | — | — |
| 1.0 | 0.8 | 0.085 | 0.141 | 0.087 | 0.021 | 0.014 | 0.007 |
| 1.2 | 0.8 | 0.010 | — | — | — | — | — |
| 1.0 | 0.2 | 0.021 | 0.010 | — | — | — | — |
| 0.6 | 1.0 | 0.010 | — | — | — | — | — |
| 0.8 | 1.0 | 0.106 | 0.032 | 0.032 | 0.043 | 0.023 | 0.006 |
| 1.0 | 1.0 | 0.127 | 0.293 | 0.500 | 0.641 | 0.905 | 0.972 |
| 1.2 | 1.0 | 0.010 | 0.054 | — | 0.010 | — | — |
| 1.4 | 1.0 | 0.010 | — | — | — | — | — |
| 0.8 | 1.2 | 0.053 | 0.065 | 0.021 | 0.011 | — | — |
| 1.0 | 1.2 | 0.074 | 0.217 | 0.250 | 0.271 | 0.058 | 0.015 |
| 1.2 | 1.2 | 0.010 | 0.021 | 0.010 | — | — | — |
| 0.6 | 1.4 | 0.021 | — | — | — | — | — |
| 0.8 | 1.4 | 0.148 | 0.065 | 0.043 | — | — | — |
| 1.0 | 1.4 | 0.063 | 0.065 | 0.011 | — | — | — |
| 1.2 | 1.4 | 0.010 | — | — | — | — | — |

Note: _ means 0−probability

Table 2: Break at T/4 with d1=0.8 and d2=0.4.

| $d_1$ (Seasonality) | $d_2$ (Long run) | T=120 | T=240 | T=360 | T=480 | T=600 | T=720 |
|---|---|---|---|---|---|---|---|
| 0.6 | 0.2 | 0.106 | 0.174 | 0.043 | 0.010 | — | — |
| 0.8 | 0.2 | 0.159 | 0.281 | 0.173 | 0.065 | 0.043 | 0.020 |
| 1.0 | 0.2 | 0.021 | 0.010 | — | — | — | — |
| 0.4 | 0.4 | 0.063 | — | — | — | — | — |
| 0.6 | 0.4 | 0.393 | 0.141 | 0.109 | 0.066 | 0.054 | 0.018 |
| 0.8 | 0.4 | 0.256 | 0.359 | 0.663 | 0.847 | 0.902 | 0.962 |
| 1.0 | 0.4 | 0.011 | 0.032 | 0.011 | 0.011 | — | — |

Note: _ means 0−probability

Table 3: Break at 3T/4 with d1=0.4 and d2=0.8.

| $d_1$ (Seasonality) | $d_2$ (Long run) | T=120 | T=240 | T=360 | T=480 | T=600 | T=720 |
|---|---|---|---|---|---|---|---|
| 0.0 | 0.4 | 0.085 | 0.010 | — | — | — | — |
| 0.2 | 0.4 | 0.022 | — | — | — | — | — |
| 0.0 | 0.6 | 0.095 | 0.043 | — | — | — | — |
| 0.2 | 0.6 | 0.297 | 0.163 | 0.087 | 0.019 | — | — |
| 0.4 | 0.6 | 0.127 | 0.271 | 0.163 | 0.043 | 0.044 | 0.021 |
| 0.6 | 0.6 | 0.023 | 0.010 | — | — | — | — |
| 0.8 | 0.6 | — | 0.011 | — | — | — | — |
| 0.2 | 0.8 | 0.138 | 0.032 | 0.076 | 0.054 | 0.055 | 0.011 |
| 0.4 | 0.8 | 0.128 | 0.380 | 0.597 | 0.837 | 0.891 | 0.960 |
| 0.6 | 0.8 | 0.009 | 0.009 | 0.010 | 0.010 | — | — |
| 0.2 | 1.0 | — | 0.009 | — | — | — | — |
| 0.4 | 1.0 | 0.020 | 0.021 | — | 0.010 | — | — |
| 0.6 | 1.0 | — | 0.010 | 0.043 | 0.021 | 0.010 | 0.008 |
| 0.2 | 1.2 | 0.053 | 0.020 | 0.021 | — | — | — |

Note: _ means 0−probability

Table 4: Probability of detecting the true break fraction for different break coefficients, beta.

| $\beta$ | T=120 | T=240 | T=360 | T=480 | T=600 | T=720 |
|---|---|---|---|---|---|---|
| 10 | 100 | 100 | 100 | 100 | 100 | 100 |
| 5 | 100 | 100 | 100 | 100 | 100 | 100 |
| 3 | 83.33 | 92.30 | 100 | 100 | 100 | 100 |
| 1 | 10.71 | 13.09 | 17.98 | 20.5 | 33.39 | 33.57 |