

Further evidence on the statistical properties of Real GNP

Laura Mayoral*

Universitat Pompeu Fabra

First version: May 2005

This version: February 2006

Abstract

The well-known lack of power of unit root tests has often been attributed to the short length of macroeconomic variables and also to DGP's that depart from the $I(1)$ - $I(0)$ alternatives. This paper shows that by using long spans of annual real GNP and GNP per capita (133 years) high power can be achieved, leading to the rejection of both the unit root and the trend-stationary hypothesis. This suggests that possibly neither model provides a good characterization of these data. Next, more flexible representations are considered, namely, processes containing structural breaks (SB) and fractional orders of integration (FI). Economic justification for the presence of these features in GNP is provided. It is shown that the latter models (FI and SB) are in general preferred to the ARIMA ($I(1)$ or $I(0)$) ones. As a novelty in this literature, new techniques are applied to discriminate between FI and SB models. It turns out that the FI specification is preferred, implying that GNP and GNP per capita are non-stationary, highly persistent but mean-reverting series. Finally, it is shown that the results are robust when breaks in the deterministic component are allowed for in the FI model. Some macroeconomic implications of these findings are also discussed.

KEY WORDS: GNP, unit roots, fractional integration, structural change, long memory, exogenous growth models.

*The author wants to thank two anonymous referees and the editor for their valuable comments. She also acknowledges financial support from the Spanish Ministry of Education through grants SEC2003-04429 and SEC2003-04476 and also from the Barcelona Economics Program of CREA.

1. Introduction

Questions about the nature of the trend component and the persistence of shocks in macroeconomic series and, in particular, in GNP have occupied a very important place in economics and have given rise to a vast literature on the subject. In spite of this fact, important issues remain unclear. Until the early 80's, movements in output were traditionally viewed as representing temporary fluctuations about a stable linear time trend (the so-called trend-stationary models, T-ST henceforth). According to this view, innovations have no impact on long-run forecasts of GNP. Nevertheless after the influential article of Nelson and Plosser (1982) many economists argued that GNP was better characterized as a differenced-stationary model (D-ST) given by the sum of a stochastic trend that shifts every period (a unit root component) and a transitory term. Under this characterization, there is not reversion to a deterministic trend path because any stochastic shock has a permanent effect on the process and therefore, they will affect output forecasts into the indefinite future.

The widespread acceptance of the unit root hypothesis that followed Nelson and Plosser's article was mainly due to the fact that unit root tests could not reject the latter hypothesis. However, various explanations not related with the existence of a unit root in the process were soon put forward to account for such phenomenon. Firstly, some authors claimed that the main difference between these models is basically of an asymptotic character and, given the relative short length of the available macroeconomic data, unit root tests lack power to distinguish between T-ST and D-ST models (see Christiano and Eichenbaum (1990), Rudebusch (1993), etc.). Moreover, this is not only true when the models are close (local alternatives) but also when the hypotheses are distant. In particular, Rudebusch (1993) used quarterly real GNP per capita to show that the best-fitting T-ST and D-ST models implied very different medium and long term dynamics. Then, he computed the small sample distributions of the Dickey-Fuller (DF henceforth) test corresponding to both best-fitting models and showed that they were very similar despite the very different dynamics of these processes. This finding contributed to the "we don't know" literature initiated by Christiano and Eichenbaum (1990), since it implied that the DF test has very low power

to distinguish between the relevant hypotheses.

Secondly, other group of authors considered the T-ST and D-ST models as two extreme possibilities that in some circumstances were not suitable. The former implies that all shocks are transitory and that the trend component is always the same (deterministic) while the latter has the opposite predictions: all shocks are permanent and the trend component shifts every period. Thus, models different from the D-ST or T-ST were postulated as DGPs and it was shown that the classical unit root tests have very low power under these new formulations. Among the processes departing from the traditional models those containing structural breaks and fractionally orders of integration (as opposed to integer ones) have attracted a great deal of attention (see Banerjee and Urga (2005) for a recent survey on both topics).

Perron (1989) initiated the literature on structural breaks by arguing that the movement in the trend component could be well-explained by a few permanent shocks related to very significant events (such as wars, deep economic crisis, etc.), the so-called structural breaks (SB hereafter), in an otherwise stable linear trend. All remaining shocks had a transitory character. He showed that when the DGP is of this type, standard unit root tests cannot reject the D-ST model. Using quarterly data for the post-war GNP, he presented evidence against a unit root in favor of a trend-shift alternative where the shift was associated to the 1973 oil crisis. This idea had a great impact and has stimulated a very active research in this area (see Perron (2005) for a recent survey).

On the other hand, fractionally integrated (FI) models were introduced by Granger and Joyeux (1980) and Hosking (1981) who showed that, by allowing for fractional orders of integration (and not only integer ones as in the ARIMA methodology), a richer description of the persistence of shocks could be achieved. These models are able to fill the gap between the short-lasting and the permanent effect of shocks in the T-ST and D-ST models, respectively, by allowing for intermediate behaviors (such as long memory, non-stationary mean-reversion, etc.). Interestingly, it has been shown analytically that fractional integration can arise in aggregate output just by allowing for some degree of heterogeneity at the micro level. Furthermore, empirical support for the existence of these features in GNP and

other variables (not only in macroeconomics but also in finance and other social sciences), is well-documented (see Henry and Zaffaroni (2002) for surveys regarding these contributions).

The goal of this paper is to shed further light on the controversy about the statistical properties of real GNP and real GNP per capita taking into account simultaneously the two criticisms described above. Firstly, in order to have good power properties a data set that covers a very long time span (133 years) will be analyzed. Diebold and Senhadji (1996) used first these data to show that Rudebusch's (1993) conclusions could be reversed. They replicated his analysis using annual U.S. GNP and they showed that the unit root could be clearly rejected in favor of trend stationarity (T-ST) and that the DF test had power in that case to distinguish between these two models. As it is well-known, rejecting a hypothesis in favor of an alternative one does not imply that the latter is a good description of the data. This paper starts by noticing that, if the same procedure and the same data set as in the paper above is used for testing T-ST against D-ST, evidence against the former hypothesis also arises (see Section 2). This suggests that possibly neither the $I(1)$ nor the $I(0) + trend$ models are suitable for these data.

Next, we enlarge the set of DGP's considered by exploring if there exists evidence in favor of FI and SB models in these data set. It turns out that when tested against the D-ST and T-ST, the former models are preferred. A well-known statistical problem arises here: despite the fact that FI and SB imply very different medium and long term dynamics, it is difficult to identify them. This is so because short memory models containing some type of trends and/or breaks display spurious persistence that may look very similar to that generated by $FI(d)$ models. The opposite is also true, that is, conventional procedures for detecting and dating structural changes tend to find spurious breaks, usually in the middle of the sample, when in fact there is only fractional integration in the data. Surprisingly, although there exist quite a few papers that look for FI behavior or for the existence of structural breaks in output, to the best of our knowledge this is the first contribution that directly test these hypotheses.

The main results can be summarized as follows. It is first shown that high power can be achieved if long spans of data are used. In particular both the T-ST and D-ST are clearly

rejected in this data set in favor to other alternatives such as FI and SB models. When the latter models are directly tested, the former is preferred. These results are also robust to the different econometric methods employed and to the inclusion of trends and/or breaks in the FI model. It turns out that real GNP and real GNP per capita are well characterized as non-stationary but mean-reverting processes. The implications of these conclusions are discussed along the paper.

The outline of the paper is as follows. Section 2 introduces the data and considers some classical unit roots tests of $I(1)$ vs. $T - ST$ and vice versa. Section 3 analyzes the existence of structural breaks in the data for the case where the innovations are weakly dependent. Section 4 describes the main characteristics and the economic mechanisms that are able to generate FI in economic data. Several procedures for estimating and testing the FI hypothesis against both the $I(1)$ and the $I(0)$ ones are also considered. Section 5, in turn, tests the hypotheses of FI versus a breaking-trend model. Some robustness checks the consider the possibility of breaks/trends and FI are also included. Finally, Section 6 put forward some macroeconomic implications of the findings of this article and concludes.

2. The data and preliminary tests

We consider the same data set employed in Diebold and Senhadji (1996) (DS henceforth): the annual real GNP series reported in table 1.10 of the National Income and Product Accounts of the United States, measured in billions of 1987 dollars, ranging from 1929 to 2001 (8 new observations have been added with respect to DS analysis). As in DS, these series has been spliced to the 1869-1929 real GNP series of Balke and Gordon (1989) or Romer (1989) given rise to two different series of Real GNP, each containing 133 annual observations. Per capita GNP has also been considered and in order to construct the series, total population residing in the United States (in thousands of people) has been taken from table A-7 of *Historical Statistics of the United States* for years ranging from 1869 to 1970 and for the years 1971-2001, data has been taken from the *Census Bureau's Current Population Reports*, Series P-25. All series are in natural logs.

We keep the same notation as in DS and we define:

GNP-BG (“GNP-Balke-Gordon”). Gross national product. Pre-1929 values from Balke-Gordon.

GNP-R (“GNP-Romer”). Gross national product. Pre-1929 values from Romer.

GNP-BGPC (“GNP-Balke-Gordon, per capita”). Gross national product per capita. Pre-1929 values from Balke-Gordon.

GNP-RPC (“GNP-Romer, per capita”). Gross national product per capita. Pre-1929 values from Romer.

Figure 1 contains the plots of these data. The post-1929 series values of both the GNP-R and GNP-BG are identical but pre-1929 values differ slightly due to the differing assumptions underlying their construction. The same is true for the per capita series GNP-RPC and GNP-BGPC.

[Figure 1 about here]

Some preliminary unit root tests

Before beginning the analysis, it is illustrative to look at some basic properties of D-ST and T-ST models. A trended process, as those in Figure I, may be represented by

$$y_t = \mu + \beta t + u_t. \quad (1)$$

If u_t is a linear weakly-dependent process, that is,

$$u_t = C(L)\varepsilon_t \quad (2)$$

with $\varepsilon_t \sim i.i.d(0, \sigma^2)$ and $C(L) = \sum_{j=0}^{\infty} c_j L^j$ such that $\sum_{j=0}^{\infty} j |c_j| < \infty$, then y_t is a T-ST process. In this case, the trend component $\tau_t^{T-ST} = \mu + \beta t$ is completely deterministic and stable over time, while the cycle $c_t = u_t$ has a short effect on the trend because the correlation function of u_t decays to zero at an exponential rate.

On the other hand, a unit root in y_t can be modelled as,

$$(1 - L)u_t = C(L)\varepsilon_t. \quad (3)$$

In this case, y_t is a unit root process with a drift equal to β and an initial condition given

by μ . Noticing that $C(L)$ can always be written as $C(L) = C(1) + (1-L)C^*(L)$,¹ where $C^*(L) = \sum_{i=0}^{\infty} c_i^* L^i$ and $c_i^* = -\sum_{j=i+1}^{\infty} c_j$, it follows that y_t is given by,

$$y_t = \mu + \beta t + C(1) \sum_{j=0}^{t-1} \varepsilon_{t-j} + a_t, \quad (4)$$

where $a_t = C^*(L)\varepsilon_t$ is a weakly-dependent stationary processes. Expression (4) corresponds to the trend-cycle decomposition (Beveridge and Nelson (1981)). The trend component is given by $\tau_t^{D-ST} = \mu + \beta t + C(1) \sum_{j=0}^{t-1} \varepsilon_{t-j}$ and, as opposed to τ_t^{T-ST} it has a stochastic character since it shifts every period in an unpredictable way.

Since the major difference between T-ST and D-ST models is the duration of shocks, it is clear that data covering long data spans are needed in order to have a reasonable power. This is the approach followed by DS (1996) who contested Rudebusch (1993) conclusions by using the above-described long data set. They followed Rudebusch' approach and computed the best fitting T-ST and D-ST models for each of the four series. Then, for each series they computed the exact finite sample distribution of the corresponding t-statistics from an augmented Dickey-Fuller (ADF) regression under both best-fitting models. They showed that the p-value associated to the former statistic was very small under the D-ST model but quite large under the T-ST one and therefore it was possible to reject the null hypothesis of a unit root with reasonable power.

Nevertheless, as noticed by the authors rejecting the null hypothesis does not mean that the alternative is a good characterization of this data set. Table I below show that, when the hypotheses are reversed, the T-ST hypothesis is rejected too for the same data set. This table also presents the results of other classical unit root tests. The first three columns have the unit root as null hypothesis. In addition to the Augmented Dickey-Fuller test employed by DS, also the Phillips-Perron (P-P) (1988) and the efficient DF-GLS method proposed by Elliott, Rothenberg and Stock (1996) are considered. Not surprisingly, the $I(1)$ hypothesis is rejected for the four extended series employed in this article, confirming DS findings. The last column reports the result of the KPSS test for the null hypothesis of trend stationarity

¹See Johansen (1995), Lemma 4.1.

versus a unit root. Interestingly, the $I(0)$ +trend hypothesis is also rejected in all the four series.²

[Table I about here]

Figure 2 is similar to DS' Figure 2 but contains in this case the exact finite-sample distribution of the KPSS statistic under the best-fitting T-ST and D-ST models for real GNP per capita (Pre-1929 values are from Romer, (1989)).³ This plot confirms DS conclusions in the sense that the KPSS value is very unlikely under the best-fitting D-ST formulation: only 4% of the values coming from the D-ST distribution are smaller than that one. But remarkably, it shows that the sample KPSS is also unlikely under the T-ST model since only 19% of the values are bigger than that quantity.

It is interesting to notice that the chronic problem of lack of power does not seem to apply here since both hypotheses can be rejected. It is natural to ask now why this is happening. A plausible explanation stems from considering DGP's that differ from the T-ST and D-ST ones. It is well-known that the tests above do have power against some alternative formulations. This is the case of Dickey-Fuller tests against FI processes (see Diebold and Rudebusch, 1991) and also against some types of breaks (cf. Perron, 1989). Similar properties have been found to hold for the KPSS test (Lee and Schmidt (1996)).

In the following sections we will explore the plausibility of processes that can be considered to be mid-way between the T-ST and the D-ST formulations: those containing structural breaks and fractional orders of integration.

3. Models containing Structural Breaks

It follows from the discussion below equation (1) that the unit root versus T-ST problem can be viewed as addressing the question “do the data support the view that the trend

²For the DF and the DF-GLS tests, the number of lags was chosen according to the BIC criterion and in all cases a number equal to 2 was set. For the remaining tests, also two lags were considered to compute the Newey-West variance-covariance estimator. Different values were tried and the results remained qualitatively identical.

³The best-fitting T-ST and D-ST models are reproduced in Table XI in the Appendix.

never changes or is it changing every period?”. Clearly, these alternatives are not mutually exclusive and can be seen as too extreme in some contexts. A more interesting question to ask would be “how frequent permanent shocks are?”. Perron (1989) initiated this line of research postulating that only a few shocks, those related with very significant events, could have a permanent character while the remaining ones would only last for a short time. The specific number of permanent shocks becomes then case-specific. Under these assumptions, the process y_t can be represented by the sum of two components: a deterministic trend whose parameters are allowed to change at each break date and a cycle, similar to that described in (2). More specifically, using quarterly data for the postwar U.S. GNP, Perron (1989) favored the model $y_t = \tau_t + u_t$ where the trending component is given by,

$$\tau_t = \alpha_1 + \beta_1 t + (\beta_2 - \beta_1) (t - T^*) 1_{(t > T^*)},$$

where T^* coincides with the 1973 oil crisis, $1_{(t > T^*)}$ is an indicator function and β_2 captures the slow down in the growth rate after the crisis.

We now explore if there exists evidence in favor of a breaking trend in the data set considered in this article. Given the long time span considered here, it seems reasonable to allow for several permanent shocks in the data occurring at unknown break dates. We employ the method proposed by Bai and Perron (1998, 2003a, b), henceforth BP,⁴ for multiple structural breaks. BP propose three types of tests. The $\text{supF}_T(k)$ test considers the null hypothesis of no breaks against the alternative of k breaks. The $\text{supF}_T(l+1/l)$ test, takes the existence of l breaks, with $l = 0, 1, \dots$, as H_0 against the alternative of $l+1$ changes. Finally, the so-called “double maximum” tests, UDmax and WDmax, test the null of absence of structural breaks versus the existence of an unknown number of breaks. Bai and Perron

⁴Some of the procedures developed in Bai and Perron (1998) are not valid when trending regressors are allowed for, as it is the case of these data. Nevertheless, the consistency, the rate of convergence and the confidence intervals of the estimated breaks points still hold. The case with trending data is discussed in Bai (1999) and yields different asymptotic distributions for the tests of no break versus a fixed number of breaks. But, as discussed in Bai and Perron (2003), the asymptotic distributions in the two cases are fairly similar, especially in the tail (where critical values are obtained), and simulations confirm that the size distortions are minor.

(2003b) suggest beginning with the $\text{supF}_T(k)$ test. If no break is detected, they recommend checking this result with the UDmax and WDmax tests to see if at least one break exists. When this is the case, they recommend continuing with a sequential application of the $\text{supF}_T(l+1/l)$ test, with $l = 1, \dots$. To compute the tests, a Gauss code (available from Perron's webpage at <http://econ.bu.edu/perron/code.html>) has been employed. To test the changes in the trend of the series, the following representation has been considered,

$$y_t = \tau_t + u_t,$$

where τ_t is a linear time trend and u_t is a general weakly dependent linear process. The autocorrelation in u_t has been corrected by introducing lagged terms of the dependent variable. A number of breaks equal or less than 5 is considered and both the mean and the trend are allowed to change, while the remaining parameters were assumed to remain fixed.

Table II gathers the main results.⁵ They can be summarized as follows. Real GNP shows little evidence of structural breaks. The null of no-break is only rejected at the 10% level versus the alternative of two breaks, occurring, for both GNP-R and GNP-BG around the 1929 crisis and the end of World War II. On the other hand, one break is found in per-capita series located around 1939, coinciding with the beginning of World War II. Since there is no evidence of a change in real GNP around that date, the latter finding suggests that it may be due to a demographic shock. Tests of no-break versus an unknown number of breaks and sequential tests were performed and their conclusions were similar as the above-described ones.

[Table II about here]

4. Fractionally Integrated models

General description

As described above, memory is infinite in D-ST models (past shocks are perfectly remembered) while it decays exponentially fast in T-ST ones. Fractionally integrated (FI)

⁵For the sake of brevity, we only report the results of the tests with three breaks or less, but the hypotheses of 4 and 5 breaks were also considered and rejected in favor of the no-break hypothesis in all cases.

processes constitute an interesting alternative to this dichotomy since they are able to bridge the gap between these two possibilities by introducing a continuum of degrees of memory. To achieve that, the same model as in (1) can be entertained, that is

$$y_t = \mu + \beta t + u_t, \quad (5)$$

where in this case u_t is defined as

$$(1 - L)^d u_t = C(L) \varepsilon_t. \quad (6)$$

The parameter d determines the integration order of u_t and governs the memory of the process. The higher the value of d , the more persistent shocks are. Notice that if $d = \{0, 1\}$, the T-ST and D-ST models are found respectively. However, if fractional values of d are considered, richer memory properties can be obtained. This provides for parsimonious, yet flexible, modeling of the low frequency variation. The process u_t is stationary if $d < 1/2$ and invertible whenever $d > -1$ (see Hosking (1981) and Odaki (1993)). *Long memory* occurs whenever d belongs to the $(0, 0.5)$ interval, since in this case the autocorrelation function, $\rho(k)$, verifies that,

$$\rho(k) \sim ck^{2d-1} \text{ for large } k, d \in (0, 0.5), \quad (7)$$

where c is a constant. In other words, long memory is implied by a hyperbolic decay of correlations, as opposed to the case $d = 0$ where correlations decay exponentially fast. For values of $d \in [0.5, 1)$, the process is non-stationary since the variance is unbounded. Shocks are very persistent but do not have a permanent effect implying that the process is mean reverting despite its non-stationary character. Values of d greater or equal than 1 imply a permanent behavior of shocks.

Operationally, a binomial expansion of the operator $(1 - L)^d$ is used in order to differentiate fractionally a process:

$$(1 - L)^d = \sum_{i=0}^{\infty} \pi_i(d) L^i,$$

where,

$$\pi_i(d) = \Gamma(i-d) / \Gamma(-d) \Gamma(j+1), \quad (8)$$

and $\Gamma(\cdot)$ denotes the gamma function. When $d = 1$, this is just the usual first-differencing filter. For non-integer d , the operator $(1 - L)^d$ provides an infinite-order lag-operator polynomial with coefficients that decay very slowly.⁶

Fractional integration in GNP

Why output could display fractional integration? Some authors have pointed out that one plausible explanation for the existence of FI behavior in GNP data is that production shocks themselves also display this behavior. The argument is based on the fact that strongly dependent shocks (possibly inherited from underlying geophysical processes) in a real business cycle model of the economy can account for the presence of such behavior in aggregate income series (see Kydland and Prescott (1982)). The existence of that type of shocks in geophysical variables, such as rainfall, riverflow or climatic series is well documented (see Mandelbrot and Wallis (1969), Lawrence and Kottegoda (1977) and Hipel and McLeod (1978) among others).

However, a more satisfactory explanation is provided by economic models that produce FI in output despite white noise innovations. All these models make use of results obtained by Granger (1980) and Robinson (1978) who showed how FI arises as a consequence of aggregation over heterogeneous entities (see also Zaffaroni (2004)).

Michelacci and Zaffaroni (2000) showed that FI in GNP per capita arises in a Solow-Swan growth model just by allowing for cross sectional heterogeneity in the speed with which different units in the same countries adjust. Abadir and Talmain (2002) consider a monopolistically-competitive Real Business Cycle model and, by allowing for heterogeneity at the firm level, show that GDP turns out to be very persistent although mean-reverting. Haubrich and Lo (2001) discuss a multiple-sector real business cycle model along the lines

⁶Since the expansion is infinite, in practice a truncation is needed to differentiate fractionally a series (see Dolado, Gonzalo and Mayoral (2002) for details on the consequences of the truncation).

of Long and Plosser (1983) and show analytically that GNP behaves as a FI process.

From an empirical point of view, several papers have also tested the existence of FI in output with somehow mixed conclusions. Both Diebold and Rudebusch (1989) and Sowell (1992a) analyzed quarterly post-war data and obtain estimates of d below unity. Nevertheless, their results are in line with Rudebusch's (1993) and Christiano and Eichenbaum (1990) conclusions in the sense that the confidence interval of the estimated value of d includes the unit root and, in the Sowell's case, also the T-ST model.

In the next subsections, we contribute to the previous literature by analyzing the empirical plausibility of the FI hypothesis in real GNP and real GNP per capita. The use of data that covers longer time spans and a wider set of econometric techniques will allow us to obtain more robust conclusions.

Estimation of $FI(d)$ models

There is a broad literature on the parametric and semiparametric estimation of FI models. In the following we consider some of the most representative techniques in both fields. ARFIMA processes have been chosen to specify the parametric model. This amounts to consider that the polynomial $C(L)$ in expression (6) admits an ARMA representation. Exact Maximum Likelihood (ML) (Sowell, 1992b), Minimum Distance (MD) (Mayoral, 2004), and the Whittle estimator with tapered data (WT) (Velasco and Robinson, 2000) are employed. The semiparametric techniques proposed by Geweke and Porter Hudak (GPH) (1983) and by Teverovsky and Taqqu (TT) (1997) have also been applied. Table III gathers the main results.⁷

Two main conclusions can be drawn from the inspection of the table above. Firstly, the memory parameter is a fractional number below unity for all series across all techniques

⁷All parametric models have chosen according to the BIC criterion. The exact ML, the GPH and the TT estimator have been computed in first differences and unity has been added to the estimated value of d . Following Velasco and Robinson (2000), tapering has been employed to compute the Whittle estimator since non-stationary was suspected. Finally, the number of frequencies used in the calculus of the GPH estimator was set equal to $T^{0.5}$.

employed. Secondly, although the values differ slightly across techniques, values of d in the interval $(0.5,1)$ are always found. More specifically, all parametric methods show values of d around 0.6-0.7 whereas the semiparametric GPH estimator delivers higher values around 0.9. This fact is not surprising since it is well-known that short-run autocorrelation may bias upwards this estimator (Sowell (1992b)). On the other hand, if the GPH is computed on the first differences of the series, the estimates change radically (they are, after adding unity, around 0.2-0.3). This difference may be explained by noticing that deterministic trends can substantially bias the estimator (cf. Sibbertsen, (2003)). Summarizing, the finding of fractional integration seems to be very robust in all the four series, with an integration order around 0.7. This implies that the series are non-stationary with very persistent although mean-reverting shocks.

[Table III about here]

Figure 3 reproduces Figure 2 above but the exact distribution of the best-fitting $FI(d)$ model (according to the exact ML procedure) has been included. It can be seen that the sample KPSS occupies a central position in the distribution under the FI hypothesis, implying that this is a very likely value under this distribution. More specifically, the probability mass on the left side of this value is equal to 0.44.

[Figure 3 about here]

Although Table III provides some evidence in favor of the hypothesis of non-stationary mean-reversion ($0.5 \leq d < 1$), tests based on confidence intervals around the estimated values are known to have very low power to reject alternatives with an integer order of integration. Then, the hypothesis of $FI(d)$ would be directly tested against the T-ST and the D-ST alternatives in the following section.

Testing $I(1)$ versus $FI(d)$

Two popular tests have been employed: Lagrange Multiplier (LM) and Wald type tests. The former were introduced by Robinson (1994) and Tanaka (1999) in the frequency and

time domain respectively.⁸ In contrast to classical unit root tests, the statistics have standard asymptotic distributions and under gaussianity are locally optimal. A drawback, however, is that they are fully parametric and the results may rely heavily in the parametric specification of the model. A different approach was introduced by Dolado, Gonzalo and Mayoral (2002, 2004) who generalized the traditional Dickey-Fuller test of $I(1)$ against $I(0)$ to the more general framework of $I(1)$ versus $FI(d)$. The so-called Fractional Dickey-Fuller (FDF) test is based upon the t-ratio associated to the coefficient of $(1-L)^d y_{t-1}$ in a regression of $(1-L)y_t$ on $(1-L)^d y_{t-1}$, and possibly, some lags of $(1-L)y_t$, to account for the short run autocorrelation of the process, and/or some deterministic components if the series displays a trending behavior or initial conditions different from zero. Besides its simplicity, two additional features stand out in this approach: as in the Dickey-Fuller approach (see Said and Dickey, (1984)), the test allows for a semiparametric specification of the short term structure and secondly, although not locally optimal under gaussianity, it presents a higher power in finite samples that, in general, outperforms the above-mentioned methods.

Table IV presents the outcome of the tests. With respect to the FDF test, the invariant regression described in Dolado et al. (2004) has been performed against several non-stationary fractional hypotheses.⁹ For all the values of d considered under the alternative (from 0.6 to 0.9), the null hypothesis of a unit root was rejected. Similar results were obtained by applying the time domain LM test.¹⁰ Then, both tests support the hypothesis that these

⁸See also Breitung and Hassler (2002) and Demetrescu et al. (2005) for asymptotically equivalent versions of these tests.

⁹The testing regression that has is $\Delta y_t = \alpha_1 + \alpha_2 \tau_{t-1}(d) + \alpha_3 \tau_{t-1}(d-1) + \phi \Delta^d y_{t-1} + \sum_{j=1}^k \Delta y_{t-j} + a_t$ and a number of lags of Δy_t equal to two was chosen according to the BIC criterion. Coefficients α_1, α_2 and α_3 are associated to different deterministic components that result from introducing a constant and a time trend in the DGP (see Dolado et al., 2004). The deterministic trends are given by $\tau_t(d) = \sum_{i=0}^{t-1} \pi_i(d)$ and $\tau_t(d-1) = \sum_{i=0}^{t-1} \pi_i(d-1)$, where the coefficients $\pi_i(\delta)$ come from the expansion of $(1-L)^\delta$ as defined in equation (8).

¹⁰Tanaka's (1999) time domain version has been computed instead of Robinson's (1994) original frequency domain test since Monte Carlo simulations show that the former slightly outperforms its frequency domain counterpart in finite samples (cf. Tanaka, (1999)). The test statistic is given by $\sqrt{T} \sum_{k=1}^{T-1} \frac{1}{k} \hat{\rho}_k$, where $\hat{\rho}_k$ is the autocorrelation function of the residuals of a $FI(d)$ parametric model, and it is asymptotically normally

series do not contain a unit root.

[Table IV about here]

Testing FI versus Trend stationarity

Once the unit root has been rejected, we consider the problem of directly testing for *FI* versus T-ST. As mentioned above, the KPSS test is consistent against some types of long memory processes (Lee and Schmidt, (1996)), therefore the rejection of the null of $I(0)$ reported in Table I is consistent with the finding of fractional integration in the data reported in Tables III and IV. To formally test these hypotheses, LM tests have again been considered, setting in this case the T-ST as null hypothesis versus the alternative of a higher integration order. Table V presents the outcome of the LM test for testing the hypothesis of $I(0)$ vs. a bigger integration order. Not surprisingly, the null hypothesis of T-ST is rejected.

[Table V about here]

Bhattacharya et al. (1983) first noticed that processes of the form: $y_t = f(t) + u_t$ where u_t is a weakly dependent process and $f(t)$ is a deterministic trend of the form $k(m+t)^{d-1/2}$ could produced spurious long memory with an integration order equal to $d \in (0, 0.5)$.¹¹ The same test was computed again, this time considering trends of the type $\alpha + \beta t^{d-1/2}$ for different values of $d \in (0, 0.5)$. In all cases, the null hypothesis of $I(0)$ innovations could be rejected.¹²

From Kwiatkowski et al. (1992) it is known the importance of reversing the null and the alternative hypothesis. To do so, we next employ a test recently introduced in Mayoral (2005). It is a likelihood-ratio test that allows for testing non-stationary FI vs. T-ST. The test rejects the null hypothesis of non-stationary *FI* for large values of $L(d, \sigma)|_{H_1} - L(d, \sigma)|_{H_0}$, where L represents the likelihood function under the alternative and null hypothesis, respectively. The distribution of the test is non-standard and simulated critical values can be distributed.

¹¹See also Künsch (1986) and Giraitis et al. (2001).

¹²For the sake of brevity, the corresponding figures are omitted but they are available upon request.

found in Mayoral (2005). Finally, a nice feature of this technique is that it is not needed to specify a parametric model, since a non-parametric correction can be introduced to account for the short term correlation.¹³

Table VI gathers the results of applying this test on GNP data. The first column reports the values of the test for the null hypothesis $H_0 : d \geq 0.5$ vs. the alternative of T-ST while the remaining ones consider the simple hypothesis problems $H_0 : d = d_0$ vs. T-ST for $d_0 = \{0.6, 0.7, 0.8, 0.9, 1\}$, respectively. To compute the test in the first column, the likelihood function needs to be evaluated at an estimated value of d under H_0 which has been taken from Table III (second row).

The following conclusions may be drawn. In all cases, the null hypothesis of non-stationary FI cannot be rejected against T-ST (first column). To interpret the remaining columns, it is important to bear in mind that if the order of integration under H_0 , d_0 , exceeds the true integration order, d^* , the test rejects H_0 , at least asymptotically. Thus, it turns out that this tests confirms the output from the estimation methods reported in Table III that reported estimated values of d around 0.6-0.7. The null hypothesis of FI cannot be rejected for values of d_0 around these values. Nevertheless, for higher values of d_0 , H_0 is rejected, as expected.

[Table VI about here]

5. Fractional Integration or Structural Breaks

So far, two important conclusions can be drawn. First, the rejection of the D-ST and the T-ST hypotheses is very robust across the various alternatives considered, the different econometric techniques employed and the type of pre-1929 data (Balke-Gordon or Romer). This suggests that the problem of lack of power is not to be very relevant here. Second, empirical evidence supporting the hypothesis of non-stationary fractional integration has been

¹³More specifically, the test statistic should be multiplied by an estimate of the quantity $(\gamma_0/\lambda^2)^{-1}$, where $\lambda = \sigma C(1)$ and $\gamma_0 = \sigma^2 \sum_{i=0}^{\infty} c_j^2$. In particular, γ_0 has been estimated as $\sum (\Delta^{d_0} y_t)^2 / T$ whereas to estimate λ^2 the Newey-West estimator has been employed. The lag truncation, q has been chosen according to Andrews (1991).

found and similar evidence in favor of a breaking trend in an otherwise weakly dependent data has also been reported.

In spite of the very different medium and long term dynamics that these two formulations imply, the latter finding is not surprising since it is well-known that these models share some statistical properties that make their identification a difficult task. Some authors have shown analytically that the existence of trends and/or changes in some parameter values that have not been explicitly accounted for can produce spurious persistence properties similar to those of FI processes (see Bhattacharya et al. (1983), Künsch (1986), Teverovsky and Taqqu (1997), Giraitis et al. (2001), Mikosch and Starica, (2004)). The opposite effect is also well-documented, that is, standard methods for detecting and locating structural breaks tend to find spurious breaks when the DGP is a FI process (see Hsu (2001), Nunes et al. (1995) and Krämer and Sibbertsen (2002)).

A very intuitive way to see why this happens is to notice that both FI and SB models share the property of being able to accommodate a few number of “permanent” shocks while the remaining ones are transitory. In this context, “permanent” should be understood in a broad sense as in Perron (2005), that is, a shock is permanent if, given a sample of data, it lasts for a long time and in particular is still in effect at the end of that sample. To see this, consider the decomposition of a FI process obtained by using the property $C(L) = C(1) + (1-L)C^*(L)$. The process y_t , defined as in (5) and (6) can be rewritten as,

$$y_t = \mu + \beta t + C(1) \sum_{j=0}^{t-1} \pi_j(-d) \varepsilon_{t-j} + a_t^*, \quad (9)$$

where $a_t^* = (1-L)^{1-d} C^*(L) \varepsilon_t$ is stationary and invertible for values of $d > 0$. As a particular case, notice that if $d = 1$, then $\pi_j(-1) = 1$ for all j implying that all shocks are permanent and expression (9) is identical to the trend-cycle decomposition in (4). Otherwise, $\pi_j(-d) \approx \Gamma(d)^{-1} j^{d-1}$ for large j . This means that for values of $d < 1$, shocks tend to vanish but at a very slow rate, specially for high values of d .

The process $x_t = C(1) \sum_{j=0}^{t-1} \pi_j(-d) \varepsilon_{t-j}$ is very persistent while a_t^* is not. This implies that a shock happening at time T^* has an effect on y_{T^*+h} given approximately by

$C(1)\pi_h(-d)\varepsilon_{T^*} \approx C(1)\Gamma(d)^{-1}h^{d-1}\varepsilon_{T^*}$ for large h . This quantity can be large even at very distant h if ε_{T^*} is also large. For instance, if $d = 0.7$, the effect of ε_{T^*} on y_{T^*+h} is approximately given by $0.32C(1)\varepsilon_{T^*}$, $0.24C(1)\varepsilon_{T^*}$ and $0.20C(1)\varepsilon_{T^*}$ for $h = 20, 50$ and 100 , respectively.

This property allows FI processes (with values of $d < 1$) to mimic the behavior of SB models. After a sufficiently large number of periods, the effect of most shocks will be small and not significantly different from zero. However, large shocks can retain a considerable impact on the process even at very distant horizons. In particular, given the relative short length of most macroeconomic variables, large shocks can still be in effect at end of the sample. Parke (1999) reached similar conclusions using a different approach. He introduced the so-called Error Duration (*ED*) model where the process y_t is given by the cumulation of shocks that switch to 0 after a random delay following a power law distribution. He showed that under certain conditions y_t presents a behavior similar to a $FI(d)$ process with $d \in (0, 1]$.¹⁴ In this framework, it is easy to compute the number of shocks that, on average, will remain alive at the end of the sample. For instance, he showed that in a sample of 100 observations and a value of $d = 0.6$, on average, 4.5 shocks would still be in effect at the end of the sample while the rest would have already switched to zero.

In spite of sharing the above-described property, SB and FI processes present very different dynamics because the “transitory” shocks behave very differently. For instance, forecasts from the SB model revert to the deterministic trend quickly in sharp contrast to those from $FI(d)$ processes, whose shocks are much more persistent, specially for non-stationary values of d .

The problem of distinguishing between SB and FI models has been largely overlooked, despite the fact that an erroneous identification of the process has the same undesirable consequences as in the $I(1)$ case where, in contrast, a dense literature is available. To the best of our knowledge, this paper is the first contribution that directly tests the hypothesis of *FI* versus *SB* in GNP data. Similar to most of the $I(1)$ vs. SB literature, we consider

¹⁴More specifically, if the probability that a shock survives for k periods, p_k , decreases with k at the rate $p_k = k^{2d-2}$ for $d \in (0, 1]$, Parke shows that the ED model generates long memory for $0 < d \leq 1$.

this testing problem without allowing for structural breaks under the *FI* hypothesis. Then, we check whether the main conclusions hold when this assumption is relaxed.

Testing for FI versus SB

Two procedures will be implemented to test for FI vs. SB. The first one is a likelihood ratio test that extends the technique used in Section 4 (see Mayoral (2005)). It is designed for testing null hypotheses of the type $H_0 : d \geq 0.5$ or $H_0 : d = d_0$ for some $d_0 \geq 0.5$ vs. the alternative of $H_1 : d = 0$ allowing for a single break in (some or all of the components of) the deterministic trend. The break date is unknown and is estimated as the value that maximizes the likelihood under H_1 . The asymptotic distributions are non-standard and critical values can be found in Mayoral (2005). A non-parametric correction has been introduced to account for the short-term correlation as it was explained in Section 4.

Table VII present the results. The main conclusion is that, for all the series considered, the null hypothesis of non-stationary fractional integration cannot be rejected (first column, $H_0 : d_0 \geq 0.5$). Columns 2 to 5 also report some interesting information. They contain the output from the tests where the null hypothesis is a single point $d = d_0$, for various values of d_0 . It should be remembered that if the process is *FI* (d^*), whenever the value used as null hypothesis, d_0 , is bigger than the true integration order, the test rejects the hypothesis of d_0 at least asymptotically. It follows that the results of the tests are in good agreement with the estimated values of d reported in Table III. For all series, the test is not able to reject FI in the range of values of d around 0.7 or smaller. For higher values of d , the results differ. The null hypothesis $d_0 = 0.8$ is rejected in per-capita series, although it is not in real GNP. Finally, for a null hypothesis closer to one, $d = 0.9$, the test rejects fractional integration in favor of a structural break in all the four series.

[Table VII about here]

Next, we apply a different technique that permits to reverse the hypothesis. We use an asymptotically equivalent version of the LM test of Robinson (1994) and Tanaka (1999) recently introduced by Demetrescu et al. (2005). This technique has the advantage of being

very simple to implement since is completely regression-based and the autocorrelation of the short memory component can be captured by introducing an increasing number of lagged terms of the dependent variable, in the spirit of Said and Dickey (1984). As a drawback, breaks in the trend component can only be introduced at dates that should be pre-specified and this could be an important limitation of this procedure (see Banerjee et al. (1992) and Zivot and Andrews (1992) for a discussion). We consider the three dates that have usually been postulated as candidates for break dates, namely, the 1929 crash, World War II and the oil crisis of 1973. Breaks in the level, in the constant or both have been considered. For the sake of brevity, only values corresponding to GNP-R and GNP-RPC are reported (those related to GNP-BG and GNP-BGPC were very similar). For comparison, values of the test obtained without allowing the possibility of breaks in the trend are also reported (bottom row). Critical values are taken from a standard normal distribution and large values of the statistic favor an order of integration greater than 0.

The results are presented in Table VIII. As it was found in Section 3, the null hypothesis of T-ST is rejected for both GNP and GNP per-capita (bottom row). This is also the case for real GNP across the different types of breaks and dates considered. Values of the statistic are large and positive which implies that the test favors higher values of d . The result is more ambiguous in the case of real GNP per capita. In this case, the test only rejects a T-ST model with a break in the slope occurring in the first half of the sample at a 10% signification level. In spite of this result, it can be concluded that in general FI models are preferred to SB ones.

[Table VIII about here]

Some robustness checks

In line with most of the $I(1)$ vs. SB testing literature, in the section above FI and SB models have been treated as mutually exclusive alternatives. Considering a pure FI model is economically meaningful since there exists economic underpinnings for the existence of such a behavior in output (Section 4). It is also justified from a statistic point of view

since, as explained in Parke (1999), FI models are able to represent processes that possess a few number of “permanent” shocks while the rest are transitory. Since this is the main characteristic of the type of SB models considered here, it makes sense to consider which model is preferred for these data.

Nevertheless, it could be the case that both FI and SB coexist in the data. If this was the case and breaks in deterministic component were not taken into account, estimates of the order of integration, d , are likely to be biased upwards. This would lead to incorrect forecasts and biased measures of persistence.

In the following we check whether there is evidence of overestimation of d stemming from unaccounted shifts in the deterministic components. We carry out two types of analysis. Firstly, we reestimate the processes according to the exact ML and MD procedures described in Section 4, but this time dummy variables have been introduced in the trend component to capture possible breaks in the trend parameters. Changes in the level and/or the rate of growth have been allowed at three pre-specified dates: the 1929 crash, the beginning of World War II and the 1973 oil crisis. Figures in Table IX reproduce the results of estimating GNP-R and GNP-RPC series for the case where both the level and the slope of the trend component are allowed to change. The results when only some of the components were allowed to change or those associated to GNP-BG are very similar and for the sake of brevity are omitted. It can be seen that the introduction of dummy variables do not change significantly the outcome of the estimation procedures.

[Table IX about here]

The results above depend on the break date that has been selected in a rather arbitrary way. More robust methods are also available. Iacone (2005) has recently shown that the local Whittle estimator can still yield consistent estimates of d in the presence of some trending components and shifts in the mean.¹⁵ We assume that the DGP is given as in (5) and (6), where u_t can be non-stationary. Since stationary of the stochastic component

¹⁵More specifically, the estimator is consistent if trends of the form $\kappa t^{\rho_0 - 1/2}$ for some $\rho_0 < d_0 + 1/2$ are included and also in the presence of mean shifts.

is required, first differences have been taken before applying this technique.¹⁶ Table X presents the estimated values of d (unity has been added to the original estimation). It can be seen that the estimated values do not change significantly from those that have been presented along the article.¹⁷

Secondly, the LM test introduced by Demetrescu et al. (2005) has been applied again, this time setting as null hypothesis different values of d and the same breaking trend pattern as that described in Table IX (that is, breaks are allowed at three possible dates and both in the slope and the level). Again, the results do not vary significantly: in general, for values of the null hypothesis around 0.6-0.7, this hypothesis cannot be rejected whereas it can if higher values of d are used.

The previous results suggest that the estimated orders of integration provided in Table III, were no breaks were allowed for, are not overestimated. Clearly, this does not rule out the possibility of breaks in the trend component. There is an incipient literature dealing with the problem of testing for structural breaks in the presence of (stationary) long memory innovations (see Hidalgo and Robinson, (1996) and Lazarova (2005)). Unfortunately, these techniques are not well-suited here since non-stationary FI processes are suspected and adapting them for this problem is well beyond the scope of the paper. Nevertheless, this will an interesting avenue for future research.

Concluding remarks

This paper has tried to shed further light on the controversy about the statistical properties of real GNP. Taking as starting point the conclusions in Diebold and Senhadji (1996), we have complemented their analysis by first, considering a wider range of models, both under

¹⁶Estimation of d_0 is carried out by maximixing the Whittle log-likelihood in a neighborhood of the zero frequency. Following Iacone (2005), the number of frecuencies included in the criterion function was $m = 16$ whereas the parameter in charge of the trimming of the lowest frequencies, l , was set equal to 3.

¹⁷The modified version of the variance-type estimator introduced in Teverovsky and Taqqu (1997) was also tried. Nevertheless, given the short length of the data, the resulting plots were too scattered to obtain any conclusion.

the null and under the alternative hypothesis. In agreement to their results, the unit root hypothesis was robustly rejected for all the alternatives and across the different techniques. But, interestingly, when the hypotheses were reversed also trend-stationarity was rejected. This led us to analyze in depth some generalizations of the considered models: FI and processes containing breaks. These models differ in many respects and possess very different implications in terms of duration of shocks, long-run implications, etc. Nevertheless, they may be difficult to identify since they shared some statistical properties such as a similar correlation structure. Applying a wide set of recent techniques, various tests of FI vs. structural break have been implemented. The final conclusion is that the finding of fractional integration is robust. Therefore, this may suggest that the evidence that has been found in other studies supporting the existence of breaks in GNP could be the result of the existence of fractional integration in the data, and therefore, these breaks could be spurious.

From an economic point of view, the implications of these findings are important. As we have seen, long memory can appear in macroeconomic series after aggregating heterogeneous individual entities. This suggests that moving from the representative agent assumption to a multiple-sector real business cycle model introduces not unmanageable complexity, but qualitatively new behavior that should be taken into account. On the other hand, calibrations aimed at matching only a few first and second order moments can similarly hide major differences between models and the data, missing long range dependence properties (which is basically characterized by the slow rate of decay of covariances). Finally, the lack of structural breaks in the data together with the finding of integration orders of around 0.7 for per capita series imply that the growth rate of GDP per capita (first differences of the logarithm of real GNP per capita) is well characterized as a process with little persistence and a constant mean. As Jones (1995) first suggested, this evidence is inconsistent with endogenous growth theories for which permanent changes in certain policy variables have permanent effects on the rate of economic growth.

References

- Abadir, K. and G. Talmain (2002), "Aggregation, Persistence and Volatility in a Macro model", *Review of Economic Studies* 69, 749-779.
- Andrews, D. W. K. (1991), "Heteroskedasticity and autocorrelation consistent covariance matrix estimation", *Econometrica* 59, 817-854.
- Bai, J. (1999), "Likelihood Ratio tests for multiple structural changes", *Journal of Econometrics*, Vol. 91, N^o 2, 299-323.
- Bai J. and P. Perron (1998), "Testing for and estimation of multiple structural changes", *Econometrica*, 66, 47-79.
- Bai J. and P. Perron (2003a), "Critical values of Multiple Structural Change Tests", *Econometrics Journal*, 6, 72-78.
- Bai J. and P. Perron (2003b), "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics* 18, 1-22.
- Baillie, R.T. (1996), "Long memory processes and fractional integration in economics and finance," *Journal of Econometrics*, 73, 15-131.
- Balke, N. and R. J. Gordon (1989), "The estimation of prewar gross national product: methodology and new evidence", *Journal of Political Economy*, 97 (1), 38-92.
- Banerjee, A., Lumsdaine, R.L., and Stock, J.H. (1992), "Recursive and sequential tests of the unit root and trendbreak hypothesis," *Journal of Business and Economic Statistics*, 10, 271-288.
- Banerjee A. and G. Urga, "Modelling structural breaks, long memory and stock market volatility: an overview," *Journal of Econometrics*, 129, 1-34.
- Bhattacharya, R.N., Gupta, V.K., E. Waymire (1983), "The Hurst effect under trends," *Journal of Applied Probability* 20, 649-662.
- Beveridge S. and C. Nelson C. (1981), "A New Approach to Decomposition of Economic Time Series," *Journal of Monetary Economics* 7, 151-174.
- Breitung, J., and U. Hassler (2002), "Inference on the cointegration rank in Fractionally Integrated processes", *Journal of Econometrics*, 110, 167-185.

Christiano, L. J. and M. Eichenbaum (1990), "Unit root in real GNP: do we know and do we care?", *Carnegie-Rochester Conference Series on Public Policy*, 32, 7-61.

Demetrescu, M., V. Kuzin and U. Hassler (2005), "Long Memory Testing in the Time Domain," Preprint.

Diebold, F.X, and G. Rudebusch, (1989), "Long memory and persistence in aggregate output", *Journal of Monetary Economics*, 24, 189-209.

Diebold, F.X, and G. Rudebusch (1991), "On the power of the Dickey-Fuller tests against fractional alternatives", *Economics Letters*, 35, 55-160.

Diebold, F. X and A.S. Senhadji (1996), "The uncertain unit root in Real GNP: Comment", *The American Economic Review*, Vol.86, N^o 5, 291-1298.

Dolado J.J., J. Gonzalo and L. Mayoral (2002), "A Fractional Dickey-Fuller test for unit roots", *Econometrica*, Vol 70, 5.

Dolado J.J., J. Gonzalo and L. Mayoral (2004), "Testing I(1) against I(d) alternatives in the presence of deterministic components", Mimeo.

Elliott, G., T.J. Rothenberg and J.H. Stock (1996) "Efficient tests for an autoregressive unit root," *Econometrica* 64, 813-836.

Geweke, J., and S. Porter-Hudak (1983), "The estimation and application of long memory time series models" *Journal of Time Series Analysis*, 4, 221-238.

Giraitis, L., P. Kokoszka, R. Leipus (2001), "Testing for long memory in the presence of a general trend," *Journal of Applied Probability* 38, 1033 - 1054.

Giraitis, L., Leipus, R. and Phillipe, A. (2003) "A test for stationarity versus trends and unit roots for a wide class of dependent errors," Preprint.

Granger, C.W.J. (1980), "Long memory relationships and the aggregation of dynamic models", *Journal of Econometrics* 14, 227-238.

Granger, C.W.J and K. Joyeux, (1980), "An introduction to long memory series", *Journal of Time series Analysis*, 1, 15-30.

Haubrich, J. G and A. W. Lo, (2001), "The sources and nature of long-term memory in aggregate output", *Economic Review (Q II)* 15-30, Federal Reserve Bank of Cleveland.

Henry, M. and P. Zaffaroni (2002), "The Long Range Dependence paradigm for Macro-

conomics and Finance”, in *Long range dependence: Theory and applications*, P. Doukhan, G. Oppenheim and M. Taqqu (ed). Birkhäuser, Boston.

Hidalgo, J. and P. Robinson, (1996), “Testing for Structural change in a long-memory environment”, *Journal of Econometrics*, 70, (1), 159-174.

Hipel, K.W., and I. McLeod, (1978), “Preservation of the rescaled adjusted range”, *Water Resources Research*, 14, 509-518.

Hosking, J.R.M. (1981), “Fractional differencing”. *Biometrika*, 68, 165-176.

Hsu, C. C. (2001), “Change point estimation in regressions with $I(d)$ variables,” *Economics Letters* 70, 147-155.

Iacone, F. (2005) “Local Whittle estimation of the memory parameter in presence of deterministic components,” Preprint.

Jones, C. I. (1995), “Time series of endogenous growth models”, *The Quarterly Journal of Economics*, 110, 2, 495-525.

Krämer W. and P. Sibbertsen (2002), “Testing for structural changes in the presence of long memory”, *International Journal of Business and Economics*, Vol.1, N^o 3, 235-242.

Künsch, H. (1986), “Discrimination between deterministic trends and long range dependence”, *Journal of Applied Probability* 23, 1025-1030.

Kydland, F. E. and E.C. Prescott (1982), “Time to build an aggregate fluctuation”, *Econometrica*, vol. 50 n^o 6, p. 1279-313.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin, (1992), “Testing the null hypothesis of stationarity against the alternative of a unit root”, *Journal of Econometrics*, 54, 159-178.

Lawrence, A. J. and N. T. Kotegoda (1977), “Stochastic modelling of riverflow time series”, *Journal of the Royal Statistical Association* 140, 1-47.

Lazarova, S. (2005), “Testing for structural change in regression with long memory processes”, *Journal of Econometrics*, 129.

Lee, D. and P. Schmidt, (1996), “On the power of the KPSS test of stationarity against fractionally-integrated alternatives”, *Journal of Econometrics* 73, 285-302.

Long, J. and C.I. Plosser, (1983), “Real Business Cycles”, *Journal of Political Economy*,

vol 91. n^o 1, p. 39-69.

Mandelbrot, B. and J.Wallis (1969), "Robustness of the rescaled range R/S in the measurement of non-cyclic long run statistical dependence", *Water Resources Research*, 5, 867-988.

Mayoral, L. (2004), "A new minimum distance estimator for ARFIMA processes", Working paper, Universitat Pompeu Fabra..

Mayoral, L.(2005), "Is the observed persistence spurious or real? A test for fractional integration versus structural breaks", Working Paper, Universitat Pompeu Fabra.

Michelacci, C. and P. Zaffaroni (2000), "Fractional Beta-convergence", *Journal of Monetary Economics*, 45,129-153.

Mikosch, T. and C. Starica (2004), "Non-stationarities in financial time series, the long range dependence and the IGARCH effects, *Review of Economics and Statistics*, 86 (1), 378-390.

Nelson C. R. and C. I. Plosser (1982), "Trends and Random walks in macroeconomic time series: Some evidence and implications", *Journal of Monetary Economics*, 10, 139-162.

Nunes , L. C., C. M. Kuan and P. Newbold (1995), "Spurious Breaks", *Econometric Theory* 11, 736-749.

Odaki, M (1993) "On the invertibility of fractionally differenced ARIMA processes," *Biometrika* V. 80, 3, 703-709.

Parke, W.R. (1999), "What is fractional Integration," *The Review of Economics and Statistics*, 81 (4) , p. 632-638.

Perron, P. (1989), "The Great Crash, the Oil Price shock and the unit root hypothesis", *Econometrica* 58, 1361-1401.

Perron, P. (2005), "Dealing with Structural Breaks," Forthcoming in the *Palgrave Handbook of Econometrics, Volume 1: Econometric Theory*.

Phillips, P.C.B and P. Perron (1988), "Testing for unit roots in time series regression", *Biometrika*, núm. 75, pp. 335- 346.

Phillips P.C.B. and Z. Xiao, (1998), "A primer on unit root testing", *Journal of Economic Surveys*, Vol. 12, N^o 5, p.423-469.

Robinson, P.M. (1978), "Statistical inference for a random coefficient autoregressive Model", *Scandinavian Journal of Statistics*, 5, 163-168.

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

Romer, C. D. (1989), "The prewar business cycle reconsidered: New estimates of Gross National Product, 1869-1908", *Journal of Political Economy* 97(1), p.1-37.

Rudebusch, G. D., (1993) "The uncertain unit root in real GNP", *American Economic Review*, 83, 1, p. 264-272.

Said, S.E., and D. A. Dickey (1984): "Testing for unit roots in autoregressive moving average models of unknown order," *Biometrika* , 71, 599-608.

Sibbertsen, P. (2003) "Long-Memory versus structural breaks: an overview (2003), Technical Report 28/2001, SFB 475, Universität Dortmund (erscheint in Statistical Papers).

Sowell, F. (1992a), "Modeling long-run behavior with the fractional ARIMA model", *Journal of Monetary Economics*, 29, 227-302.

Sowell, F. (1992b), "Maximum likelihood estimation of stationary univariate fractionally integrated time series", *Journal of Econometrics*, 53.

Tanaka, K. (1999) "The nonstationary fractional unit root," *Econometric Theory*, 15, 549-582.

Teverovsky V. and M. Taqqu (1997), "Testing for long-range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimator", *Journal of Time Series Analysis*, Vol. 18, N^o 3, 279-304

Velasco, C., and P. M. Robinson (2000), "Whittle pseudo-Maximum likelihood estimation for nonstationary time series", *Journal of the American Association*, 95, 1229-1243.

Zaffaroni, P. (2004), "Contemporaneous aggregation of linear dynamic models in large economies", *Journal of Econometrics*, 120, 75-102.

Zivot and Andrews (1992), "Further evidence on the Great Crash, the oil-price shock and the unit root hypothesis", *Journal of Business and Economic Statistics*, Vol. 10,3, 251-270.

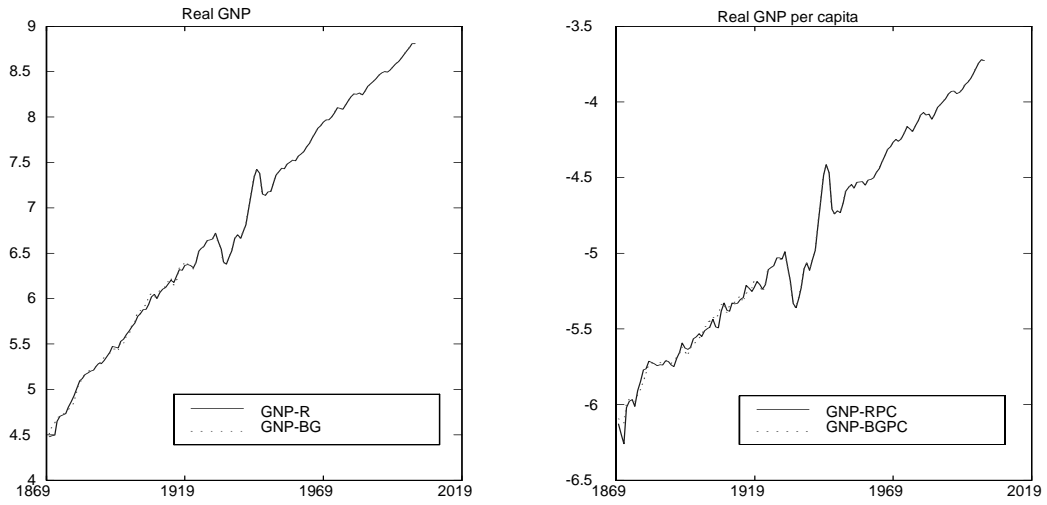


FIG. 1. Annual US real GNP and GNP per-capita 1869-2001

TABLE I

UNIT ROOT TESTS

<i>Data</i> Test	H_0 : Unit Root			H_0 : T-ST
	<i>ADF</i>	P-P	DF-GLS	KPSS
GNP-R	-4.70**	-3.53*	-3.34*	0.190*
GNP-RPC	-4.63**	-4.27*	-4.61**	0.168*
GNP-BG	-4.16**	-3.65*	-3.34*	0.181*
GNP-BGPC	-4.79**	-3.65*	-4.76**	0.182*

*Rejection at the 5% level;**Rejection at the 1% level;

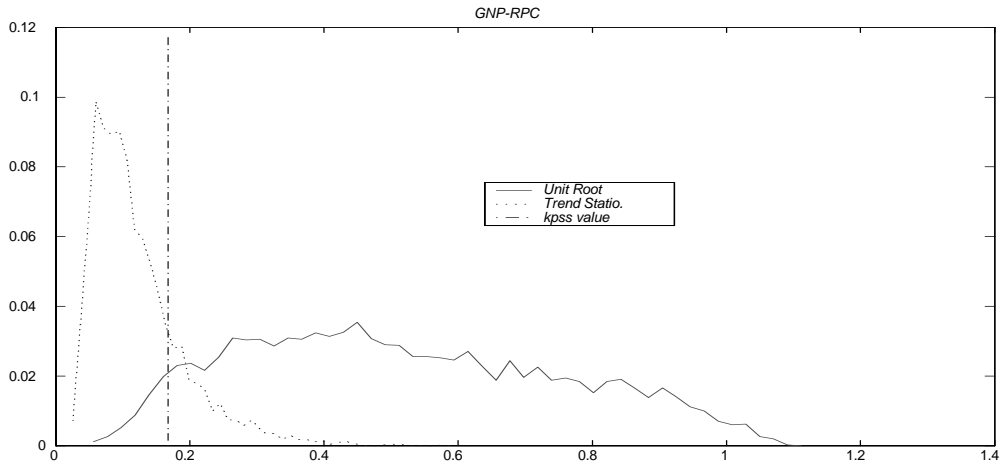


FIG. 2. Exact Distribution of the KPSS statistic computed under the best-fitting T-ST and D-ST of real GNP per capita (pre-1929 values from Romer).

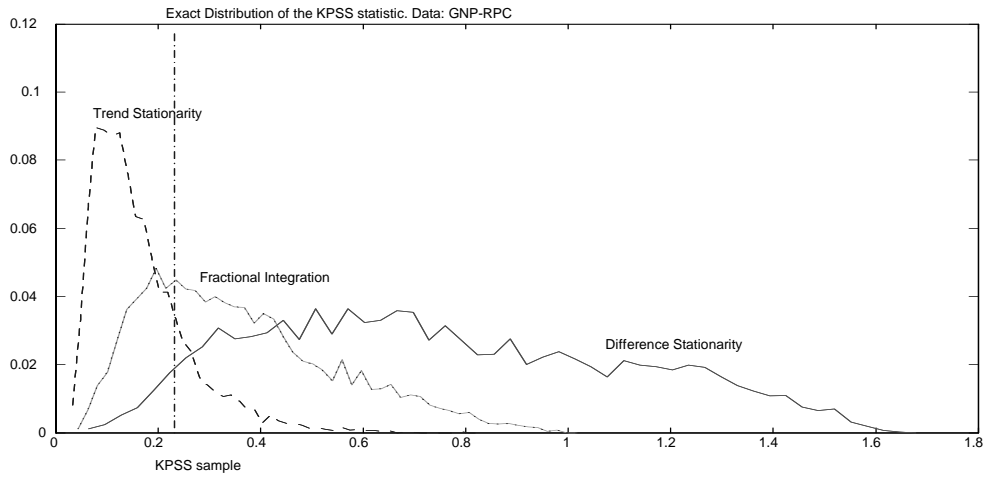


FIG. 3. Exact Distribution of the KPSS statistic computed under the best-fitting T-ST, D-ST and FI(d) models of real GNP per capita (pre-1929 values from Romer).

TABLE II

SUP F TESTS FOR A FIXED NUMBER OF BREAKS.

 H_0 : no break

	GNP-R	GNP-RPC	GNP-BG	GNP-BGPC
H_1 : 1 break	6.53	15.84***	5.53	16.26***
H_1 : 2 breaks	9.44*	8.06	9.43*	7.39
H_1 : 3 breaks	7.35	6.13	7.07	5.49
Break date (BD)	—	1941	-	1939
Conf. Int. of	(1927 – 1929)	(1938-1956)	(1928 – 1930)	(1936-1958)
Break Dates	(1945 – 1956)		(1945 – 1956)	

*Rejection at the 10% level; **Rejection at the 5% level; ***Rejection at the 1% level.

TABLE IIIESTIMATION OF $FI(d)$ MODELS

Method-Data	GNP-R	GNP-BG	GNP-RPC	GNP-BGPC
MD	$d = 0.59$ (0.21)	$d = 0.711$ (0.15)	$d = 0.570$ (0.20)	$d = 0.581$ (0.21)
	$\phi_1 = 0.84$ (0.13)	$\phi_1 = 0.554$ (0.12)	$\phi_1 = 0.797$ (0.14)	$\phi_1 = 0.757$ (0.13)
	$\phi_2 = -0.18$ (0.12)	—	$\phi_2 = -0.198$ (0.11)	$\phi_2 = -0.208$ (0.11)
ML	$d = 0.68$ (0.16)	$d = 0.65$ (0.16)	$d = 0.65$ (0.15)	$d = 0.63$ (0.15)
	$\phi_1 = 0.60$ (0.11)	$\phi_1 = 0.63$ (0.12)	$\phi_1 = 0.61$ (0.12)	$\phi_1 = 0.62$ (0.12)
Whittle	$d = 0.731$ (0.22)	$d = 0.627$ (0.24)	$d = 0.731$ (0.24)	$d = 0.731$ (0.26)
	$\phi_1 = 0.625$ (0.21)	$\phi_1 = 0.63$ (0.21)	$\phi_1 = 0.62$ (0.21)	$\phi_1 = 0.621$ (0.21)
GPH	$d = 0.93$ (0.27)	$d = 0.91$ (0.27)	$d = 0.89$ (0.263)	$d = 0.88$ (0.271)
TT	$d = 0.66$	$d = 0.66$	$d = 0.83$	$d = 0.83$

TABLE IVTEST OF I(1) VERSUS FI(d). S.L: 5%

$H_1 :$	FD-F Test				LM Test (Tanaka,1999)
	$d = 0.6$	$d = 0.7$	$d = 0.8$	$d = 0.9$	$d < 1$
GNP-R	-4.06*	-3.92*	-3.75*	-3.58*	-1.71*
GNP-RPC	-3.85*	-3.67*	-3.49*	-3.32*	-1.91*
GNP-BG	-2.89*	-2.78*	-2.66*	-3.51*	-1.91*
GNP-BGPC	-3.99*	-3.81*	-3.63*	-3.46*	-1.89*
Crit. Values (5% S.L.)	-2.45	-2.23	-2.11	-1.97	-1.64

*Rejection at the 5% level;

TABLE VTEST OF T-ST VERSUS FI(d)

LM test, (Tanaka (1999))

$H_0 : d = 0$ vs. $H_1 : d > 0$	
GNP-R	1.95*
GNP-RPC	1.82*
GNP-BG	1.99*
GNP-BGPC	1.88*
Crit. Values (5% S.L.)	1.64

*Rejection at the 5% level;

TABLE VILR TESTS FI(d) VS I(0). SL:5%

$H_0 :$	$d_0 > 0.5$	$d_0 = 0.6$	$d_0 = 0.7$	$d_0 = 0.8$	$d_0 = 0.9$	$d_0 = 1.0$
GNP-R	0.361	0.578	0.261	0.118	0.053	0.022*
GNP-RPC	0.395	0.5102	0.2309	0.103	0.045*	0.019*
GNP-BG	0.402	0.57803	0.2570	0.115	0.051	0.022*
GNP-BGPC	0.462	0.5102	0.2304	0.102	0.045*	0.019*
Critical Values (5%)	(♣)	(0.469)	(0.209)	(0.096)	(0.047)	(0.024)

*Rejection at the 5% level.

(♣) The crit. val. used are 0.295 and 0.327 for GNP-R and GNP-RPC and 0.327 and 0.351 for GNP-BG and GNP-BGPC.

Truncation lag to compute the Newey-West estimator was equal to 3.

TABLE VIILR TESTS FI(d) VS I(0) WITH ONE BREAK. S.L.:5%

$Data/H_0 : d = d_0$	$d_0 \geq 0.5$	$d_0 = 0.6$	$d_0 = 0.7$	$d_0 = 0.8$	$d_0 = 0.9$
GNP-R	0.623	1.4925	0.5050	0.1677	0.0551*
GNP-RPC	0.651	1.1236	0.3787	0.1262	0.0416*
GNP-BG	0.734	1.4085	0.4784	0.1607	0.0534*
GNP-BGPC	0.763	1.0638	0.3610	0.1209	0.0409*
Critical Values (5%)	(♣)	0.6135	0.2994	0.1515	0.0821

*Rejection at the 5% level;

(♣) Crit.val. are 0.353, 0.489 for GNP-R and GNP-RPC and 0.489 and 0.521 for GNP-BG and GNP-BGPC.

TABLE VIIILM TESTS $H_0 : I(0) + \textit{Breaking}$ trend.

<i>Break date \ Type of break</i>	GNP-R			GNP-RPC		
	<i>Level</i>	<i>Slope</i>	<i>Both</i>	<i>Level</i>	<i>Trend</i>	<i>Both</i>
1929	2.63***	2.30**	2.25**	1.76**	1.75**	1.73**
1939	2.97***	2.20**	2.38***	1.68**	1.61*	1.58*
1973	2.25**	2.30**	2.28**	1.79**	1.78**	1.76**
T-ST (no break)	2.53***			2.01**		

*Rejection at the 10% level; **Rejection at the 5% level; ***Rejection at the 5% level

TABLE IXRE-ESTIMATION OF $FI(d)$ MODELS ALLOWING FOR BREAKS

Method-Data	GNP -R		GNP-RPC	
	MD	ML	MD	ML
No break	$d = 0.59$ (0.21)	$d = 0.68$ (0.16)	$d = 0.57$ (0.21)	$d = 0.65$ (0.15)
1929	$d = 0.57$ (0.20)	$d = 0.55$ (0.16)	$d = 0.52$ (0.21)	$d = 0.66$ (0.15)
1939	$d = 0.57$ (0.21)	$d = 0.59$ (0.16)	$d = 0.51$ (0.20)	$d = 0.66$ (0.15)
1973	$d = 0.62$ (0.21)	$d = 0.54$ (0.16)	$d = 0.56$ (0.21)	$d = 0.64$ (0.15)

TABLE XWHITTLE LOCAL ESTIMATION OF d

GNP-R	GNP-BG	GNP-RPC	GNP-BGPC
0.78	0.79	0.69	0.67

Appendix

TABLE XI

ESTIMATED BEST FITTING TREND AND DIFF. STAT. MODELS

Data series	Regressor				
	c	t	y_{t-1}	y_{t-2}	Δy_{t-1}
Trend stationary (dependent variable y_t)					
GNP-R	0.797 (0.187)	0.0052 (0.0013)	1.306 (0.082)	-0.489 (0.07)	—
GNP-RPC	-1.21 (0.25)	0.0032 (0.007)	1.306 (0.07)	-0.492 (0.077)	—
GNP-BG	0.805 (0.187)	0.0052 (0.0013)	1.207 (0.042)	-0.377 (0.187)	—
GNP-BGPC	-1.12 (0.023)	0.0032 (0.007)	1.306 (0.077)	-0.481 (0.077)	—
Difference-stationary (dependent variable Δy_t)					
GNP-R	0.019 (0.0053)	—	—	—	0.401 (0.083)
GNP-RPC	0.001 (0.080)	—	—	—	0.315 (0.080)
GNP-BG	0.022 (0.0053)	—	—	—	0.30 (0.083)
GNP-BGPC	0.011 (0.080)	—	—	—	0.39 (0.080)