Working Paper nº 03/11

US DISPOSABLE PERSONAL INCOME AND HOUSING PRICE INDEX: A FRACTIONAL INTEGRATION ANALYSIS

Guglielmo Maria Caporale
Brunel University London
London, United Kingdom

and

Luis A. Gil-Alana
University of Navarra
Pamplona, Spain

US disposable personal income and housing price index. A fractional integration analysis
Guglielmo M. Caporale and Luis A. Gil-Alana
Working Paper No.03/11
January 2011

ABSTRACT

This paper examines the relationship between US disposable personal income (DPI) and house price index (HPI) during the last twenty years applying fractional integration and long-range dependence techniques to monthly data from January 1991 to July 2010. The empirical findings indicate that the stochastic properties of the two series are such that cointegration cannot hold between them, as mean reversion occurs in the case of DPI but not of HPI. Also, recursive analysis shows that the estimated fractional parameter is relatively stable over time for DPI whilst it increases throughout the sample for HPI. Interestingly, the estimates tend to converge toward the unit root case after 2008 once the bubble had burst. The implications for explaining the recent financial crisis and choosing appropriate policy actions are discussed.

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

Guglielmo Maria Caporale
Centre for Empirical Finance
Brunel University
West London, UB8 3PH
UK
Guglielmo-Maria.Caporale@brunel.ac.uk

1.   Introduction

Real estate bubbles are a controversial topic in economics. Whether it is possible to identify them and whether policy-makers should act to prevent them is a hotly debated issue. Mainstream economists argue that central banks should only target inflation and counter-cyclical monetary and fiscal policy should be adopted to smooth the wealth effects of bubbles only once they have occurred. In particular, in their view by adopting inflation targeting and focusing on inflationary or deflationary pressures, a central bank effectively minimises the negative side effects of short-run, extremely volatile asset prices, without having to target them directly (see, e.g., Bernanke and Gertler, 2001).

However, in a more recent study, Bernanke and Kuttner (2005) argued that the stock market is an independent source of macroeconomic volatility to which policy makers might need to respond in order to reduce inflation volatility, and the same might of course apply to other types of asset prices such as house prices. Cartensen (2004) and Cecchetti et al. (2000) also took the view that policy makers should give more consideration to asset price movements to reduce the risk of economic instability resulting from boom and bust in business cycles. Cecchetti et al., (2000), for example, argued that monetary authorities should take into account asset price movements with the aim of achieving macroeconomic stability.

Concerning house prices in particular, Post-Keynesian economists emphasise that bubbles lead to higher borrowing against increasing property values, and therefore to higher levels of debt the burden of which increases when the bubble bursts and property prices collapse, which reduces aggregate consumption and demand causing a fall in economic activity (this is the so-called debt deflation theory initially developed by Fisher, 1933); they suggest therefore that in order to avoid such a scenario methods

should be developed to identify bubbles and policy action taken to prevent them or to deflate existing ones. Various housing market indicators have in fact been constructed with the aim of detecting possible bubbles. These include housing affordability measures, housing debt measures, housing ownership and rent measures, as well as housing price indices. Of the latter, one of the most commonly used in the US is the HPI produced by the Federal Housing Finance Agency.

Even before the US housing bubble which led to the financial crisis starting in 2007 there had been a lot interest in the behaviour of house prices in the OECD countries, given their sharp increase since the 1990s. Some studies had expressed the concern that the observed divergence between house prices and fundamentals driving them, in particular household income, indicated the existence of a bubble (see, e.g., Case and Shiller, 2003 and McCarthy and Peach, 2004). Other authors had previously highlighted the fact that the macroeconomic effects of bubbles can differ considerably across countries depending on their housing and financial market institutions (see Maclennan et al., 1998) or the linkages between housing and labour markets (see Meen, 2002).

The subprime mortgage crisis starting in the US in 2007 has made the issue of the relationship between disposable personal income (DPI) and the housing price index (HPI) even more crucial, since there is wide agreement that the significant discrepancy between these two variables in the US was one of the main factors triggering off a global financial crisis of unprecedented severity.

From an econometric point of view, the existence of a long-run relationship between these two variables implies that they should be cointegrated, namely that although the two individual series might be nonstationary I(1) a linear combination of the

two is stationary I(0). This paper uses US monthly data to test whether this holds empirically. We carry out long-range dependence tests which allow for fractional degrees of differentiation (clearly including the special cases of 1 and 0 degrees of integration). The main finding of our analysis is that cointegration cannot hold between these series since they exhibit different degrees of integration. Specifically, DPI is found to be I(d) with 0.5 < d < 1 (d being the fractional degree of differentiation), implying that it is nonstationary but mean reverting. On the contrary, in the case of HPI, d is strictly above 1, with a value around 1.4, implying that this series is both nonstationary and non-mean reverting. Thus, while the effects of shocks to DPI will tend to disappear in the long run, those to HPI have permanent effects, and require decisive policy measures to bring about mean reversion. This is one of the main results obtained in this paper. We also investigate whether it is a consequence of the crisis of 2007 or it has its origins earlier. For this purpose we implement recursive procedures starting in 2000 and ending at the end of the sample in 2010.

The layout of the paper is as follows. Section 2 outlines the econometric methodology. Section 3 describes the data and presents the empirical results. Section 4 offers some concluding remarks.

2.   Econometric Methodology

In this paper we characterise the nonstationarity of the series in terms of a long memory process. Two definitions of long memory can be adopted, one in the time domain and the other in the frequency domain.

Let us consider a zero-mean covariance stationary process $\{ x_t, \ t = 0, \pm 1, ... \}$ with autocovariance function $\gamma_u = E(x_t x_{t+u})$. The time domain definition of long memory

states that $\sum\limits_{u=-\infty}^{\infty}|\gamma_u| = \infty$. Now, assuming that x_t has an absolutely continuous spectral

distribution, so that it has spectral density function

$$f(\lambda) = \frac{1}{2\pi}\left(\gamma_0 + 2\sum_{u=1}^{\infty}\gamma_u\cos(\lambda u)\right), \qquad (1)$$

the frequency domain definition of long memory states that the spectral density function

is unbounded at some frequency λ in the interval $[0,\pi)$. Most of the empirical literature

has concentrated on the case where the singularity or pole in the spectrum takes place at

the 0-frequency. This is the case of the standard I(d) models of the form:

$$(1-L)^d x_t = u_t, \qquad t = 0,\pm 1,\dots,$$
$$x_t = 0, \qquad t \leq 0, \qquad (2)$$

where L is the lag-operator ($Lx_t = x_{t-1}$) and u_t is $I(0)$.[1] However, fractional integration

may also occur at other frequencies away from 0, as in the case of seasonal/cyclical

models.

Note that the polynomial $(1–L)^d$ in (2) can be expressed in terms of its Binomial

expansion, such that, for all real d,

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

and thus

$$(1-L)^d x_t = x_t - d\,x_{t-1} + \frac{d(d-1)}{2}x_{t-2} - \dots.$$

In this context, d plays a crucial role since will be an indicator of the degree of

dependence of the time series: the higher the value of d is, the higher the level of

---

[1] An I(0) process is defined as a covariance stationary process with spectral density function that is positive and finite at all frequencies.

association will be between the observations. Processes with d > 0 in (2) display the property of "*long memory*", with the autocorrelations decaying hyperbolically slowly and the spectral density function being unbounded at the origin. If d = 0 in (2), $x_t = u_t$, and the series is I(0). If d belongs to the interval (0, 0.5) the series is still covariance stationary but the autocorrelations take a longer time to disappear than in the I(0) case. If d is in the interval [0.5, 1) the series is no longer stationary, however, it is still mean-reverting in the sense that shocks affecting the series disappear in the long run. Finally, if d ≥ 1 the series is nonstationary and non-mean-reverting. Thus, by letting d take real values, one allows for a richer degree of flexibility in the dynamic specification of the series, not achieved when using the classical I(0) and I(1) representations. These processes (with non-integer d) were introduced by Granger (1980, 1981), Granger and Joyeux (1980) and Hosking (1981) and since then have been widely employed to describe the behaviour of many economic and financial time series data (see, e.g., Diebold and Rudebusch, 1989; Sowell, 1992; Gil-Alana and Robinson, 1997).[2]

In this paper, we estimate the fractional differencing parameter d using the Whittle function in the frequency domain (Dahlhaus, 1989) along with a testing procedure developed by Robinson (1994) that allows to test the null hypothesis H_o: d = d_o in equation (2) for any real value d_o, where x_t can be the errors in a regression model of the form:

$$y_t = \beta^T z_t + x_t, \qquad t = 1, 2, ..., \qquad (3)$$

---

[2] See Gil-Alana and Hualde (2009) for an up-to-date review of fractional integration and cointegration in macroeconomic time series.

where y_t is the time series we observe, β is a (kx1) vector of unknown coefficients and z_t is a set of deterministic terms that might include an intercept (i.e., $z_t = 1$), an intercept with a linear time trend ($z_t = (1, t)^T$), or any other type of deterministic processes.

In the final part of the analysis we also apply a method based on a semiparametric local Whittle estimator (see Robinson, 1995). The estimator is implicitly defined by:

$$\hat{d} = \arg\min_{d} \left(\log \overline{C(d)} - 2d \frac{1}{m} \sum_{s=1}^{m} \log \lambda_s \right), \qquad (4)$$

$$\overline{C(d)} = \frac{1}{m} \sum_{s=1}^{m} I(\lambda_s) \lambda_s^{2d}, \qquad \lambda_s = \frac{2\pi s}{T}, \qquad \frac{m}{T} \to 0,$$

where $I(\lambda_s)$ is the periodogram of the raw time series, x_t, given by:

$$I(\lambda_s) = \frac{1}{2\pi T} \left| \sum_{t=1}^{T} x_t \, e^{i\lambda_s t} \right|^2,$$

and $d \in (-0.5, 0.5)$. Under finiteness of the fourth moment and other mild conditions, Robinson (1995) proved that:

$$\sqrt{m} \, (\hat{d} - d_o) \to_d N(0, 1/4) \quad as \ T \to \infty,$$

where d_o is the true value of d.

Although there are further refinements of this procedure (e.g., Velasco, 1999, Phillips and Shimotsu, 2004, 2005; etc.), these methods require additional user-chosen parameters, and therefore the estimation results for d can be very sensitive to the choice of these parameters. In this respect, the "local" Whittle method of Robinson (1995), which is computationally simpler, appears to be preferable.

3.   Data and empirical results

The series used for the analysis are US Disposable Personal Income (DPI), monthly, seasonally adjusted, obtained from the St. Louis Federal Reserve Bank database, and the US House Price Index (HPI), constructed by the Federal Housing Finance Agency (http://www.fhfa.gov). For both series the sample starts in 1991m1 and ends in 2010m6.

[Insert Figures 1 and 2 about here]

Figure 1 displays the time series plots of the log-transformed data and of their first differences (their growth rates). It can be seen that both log-prices behave in a very similar way till April 2007 where HPI starts falling. Instead the corresponding growth rates (in the bottom panels of the Figure) exhibit a very different pattern: that of DPI is relatively stable over time, whilst that of HPI is quite stable until 1995, then increases till mid-2005 when it falls, followed by another sharper fall at the beginning of 2007 coinciding with the start of the crisis. Figure 2 displays the correlograms of the four series. The slow decrease in the sample autocorrelation values of the log-prices series clearly suggests that they are nonstationary. However, the correlograms of the growth rates indicate again a very different pattern in the two series: whilst in the case of the growth rate of DPI most of the values are within the 95% confidence interval, in case of HPI the decay is very slow, which may be consistent with an I(d) model and d > 0.[3]

Table 1 reports the estimates of d in the following model,

$$y_t = \alpha + \beta t + x_t; \qquad (1-L)^d x_t = u_t, \qquad t = 1, 2, ..., \qquad (5)$$

assuming that the error term u_t is white noise and autocorrelated in turn. In the latter case, we consider first an AR(1) process, and then the exponential spectral model of

---

[3]   Also, the significant negative first sample autocorrelation value in the correlogram of the DPI growth rate indicates that this series might be now overdifferenced.

Bloomfield (1973), this being a non-parametric approach that produces autocorrelations decaying exponentially as in the AR(MA) case. Finally, we also allow for a seasonal (monthly) AR(1) process. Along with the (Whittle) estimates of d we also display the 95% confidence band of the non-rejection values of d using Robinson's (1994) parametric approach.

Starting with the log of the DPI (in the top panel of Table 1), it can be seen that if no regressors are included (i.e., $\alpha = \beta = 0$ in (5)), the unit root null hypothesis (i.e., $d = 1$) cannot be rejected for the cases of white noise, Bloomfield and seasonal AR disturbances. Moreover, if u_t is AR(1) the non-rejection values of d are strictly above 1. On the other hand, when assuming more realistically that the process contains an intercept and/or a linear trend, the unit root null is rejected in all cases in favour of smaller degrees of integration. The estimates of d are between 0.821 and 0.873 in the case of an intercept, and between 0.752 and 0.840 with a linear time trend, depending on the type of specification adopted for the disturbances.

[Insert Table 1 about here]

The lower panel in Table 1 presents the estimates of d for the log of HPI. When deterministic terms are not included the results are very similar to those for the DPI series and the I(1) hypothesis cannot be rejected in most of the cases. However, with an intercept and/or a linear time trend, the unit root is now decisively rejected for all types of disturbances, the estimated values of d ranging between 1.427 and 1.492 for the case of an intercept, and between 1.424 and 1.483 with a linear trend.

The above results indicate very different stochastic properties of the two series: whilst DPI exhibits an order of integration strictly below 1 implying mean reversion, the order of integration of HPI is strictly above 1, indicating lack of mean reversion and

implying that the growth rate still displays long memory. These features of the data invalidate cointegration analysis between the two series, noting that in the bivariate case a necessary condition for the existence of a long-run equilibrium relationship is that the two series display the same degree of integration.

To corroborate the result that the two series exhibit different orders of integration we also apply the semiparametric Whittle method of Robinson (1995). Figure 3 displays the corresponding results The bandwidth parameter is reported on the horizontal axis, the estimated value of d is shown on the vertical one.[4] We also display in the figure the 95% confidence intervals corresponding to the I(1) case. The results are consistent with those based on the parametric models. Specifically, for the log of DPI (displayed in the top panel) the estimates of d are within or below the I(1) interval, depending on the choice of the bandwidth parameter m: when this is small most of the estimates are within the interval; however, when increasing its value the estimates are strictly below the interval.[5]

[Insert Figure 3 about here]

Focusing now on the log of HPI (in the lower panel of Figure 3), it is evident that all the estimates are above the I(1) interval for all the bandwith parameters.

The final issue examined is parameter stability, in particular whether or not the degree of persistence of each series has changed over time. This is important to establish whether there were any signs of the impending crisis.[6] For this purpose, we estimate again the value of d for each series with a sample ending initially at 1999m12. Then, we

---

[4] In the case of the Whittle semiparametric estimator of Robinson (1995), the use of optimal values for the bandwidth parameter has not been theoretically justified. Some authors, such as Lobato and Savin (1998), use an interval of values for m. We have chosen instead to report the results for the whole range of values of m.
[5] This clearly shows the trade-off between bias and variance with respect to the choice of the bandwidth parameter.
[6] For instance, it might be useful to determine if cointegration holds for shorter periods of time before the crisis.

re-estimate d adding one observation each time in a recursive manner. First, we conduct the analysis based on white noise disturbances (see Figure 4), then for the case of seasonal AR disturbances (see Figure 5).[7]

[Insert Figures 4 and 5 about here]

Starting with the log of DPI with white noise errors, it is found that the estimated value of d remains relatively stable till mid-2008 when it starts increasing. As for the log of HPI, the value of d keeps increasing from the beginning of 2000 till the end of 2008, with a particularly sharp increase in 2007 and 2008; they it starts decreasing in 2009. Very similar results are found in the case of seasonal AR(1) errors (see Figure 5). Note also that after mid-2008 the estimated value of d increases for the log of DPI whilst it sharply decreases for the log of HPI, which may be a seen as a correction of the disequilibrium between the two series once the bubble had burst. In fact, if the estimates of d for the two series tend to 1, the unit root null will not be rejected and both variables might be cointegrated in the future.

4. Conclusions

The US subprime mortgage crisis and the following worldwide financial crisis starting in 2007 have led to an even greater interest in housing bubbles, methods to detect them, and appropriate policy actions to prevent them or smooth their effects once they have occurred. This paper has analysed the relationship between disposable personal income (DPI) and the housing price index (HPI) in the US applying fractional integration and long-range dependence techniques to monthly data. The empirical findings obtained for the time period 1991m1 – 2010m6 indicate that the stochastic properties of the two series

---

[7] It would have also been interesting to carry out the analysis for the sub-period 1990-1995. However, the

are such that cointegration cannot hold between them, as mean reversion occurs in the case of DPI but not of HPI. This provides useful information for explaining the recent financial crisis and choosing appropriate policy actions: the divergence between the two series might be related with the bubble in the housing sector, which might have its origin in the mid or late 90s. Note that the lack of available data before the 90s precludes us from examining the possibility of cointegration over a longer time span. In any case, what seems to be clear from these results is that, after the bubble had burst, the estimates of d tend to one for both series.

The implications of the analysis for crisis management and/or prevention are not obvious, although visual inspection of the growth rates of the two series (see the bottom panels in Figure 1) clearly shows a divergence from 1995. Therefore, it might be argued that perhaps it would have been desirable to implement active policies to burst the bubble already at that time.[8]

One interesting extension of this paper would be to examine if similar features can also be found in other developed countries. In particular, the degrees of integration of the log-DPI and log-HPI series could be analysed to establish whether mean reversion occurs or instead the unit root null cannot be rejected. Work along these lines is currently in progress.

---

number of observations would then be too limited to apply fractional integration techniques.
[8] Note, however, that this is a rather strong conclusion to draw on the basis of univariate tests.

References

Bernanke, B. and M. Gertler (2001). "Should Central Banks Respond to Movements in Asset Prices?", *American Economic Review,* 91 (May), pp. 253-57.

Bernanke, B. and K.N. Kuttner, (2005). "What Explains the Stock Market's Reaction to Federal Reserve Policy?," Journal of Finance, American Finance Association, vol. 60(3), pages 1221-1257, 06.

Bloomfield, P., 1973, An exponential model in the spectrum of a scalar time series, Biometrika 60, 217-226.

Carstensen, K., (2004), "Stock Market Downswing and the Stability of European Monetary Union Money Demand", Journal of Business and Economic Statistics, Vol 24, October, pp 395-402

Case, K.E. and R. Shiller (2003), "Is there a bubble in the housing market?", Brookings Papers on Economic Activity, 2, 299-342.

Cecchetti, S., H. Genberg, J. Lipsky and S. Wadhwani (2000), "Asset Prices and Central Bank Policy", Geneva Reports on the World Economy, 2, International Centre for Monetary and Banking Studies and Centre for Economic Policy Research.

Dahlhaus, R. (1989) Efficient parameter estimation for self-similar process. Annals of Statistics 17, 1749-1766.

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

Fisher, I. (1933), "The debt deflation theory of Great Depressions", Econometrica, 1, 3, 337-357.

Gil-Alana, L.A. and J. Hualde (2009) Fractional integration and cointegration. An overview with an empirical application. The Palgrave Handbook of Applied

Econometrics, Volume 2. Edited by Terence C. Mills and Kerry Patterson, MacMillan Publishers, pp. 434-472.

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

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

Granger, C.W.J., 1981, Some properties of time series data and their use in econometric model specification, Journal of Econometrics 16, 121-130.

Granger, C.W.J. and R. Joyeux, 1980, An introduction to long memory time series and fractionally differencing, Journal of Time Series Analysis 1, 15-29.

Hosking, J.R.M., 1981, Fractional differencing, Biometrika 68, 168-176.

Lobato, I.N. and N.E. Savin, 1998. Real and spurious long memory properties of stock market data. *Journal of Business and Economic* Statistics 16, 261-283.

McCarthy, J. and R.W. Peach (2004), "Are home prices the next "bubble"?", Economic Policy Review, 10, Federal Reserve Board of New York.

Maclennan, D., Mullbauer, J. and M. Stephens (1998), " Asymmetries in housing and financial markets institutions and EMU", Oxford Review of Economic Policy, 14, 54-80.

Meen, G. (2002), " The time-series behaviour of house prices: a transatlantic divide?" Journal of Housing Economics, 11, 1-23.

Phillips, P.C.B. and K. Shimotsu, 2004. Local Whittle estimation in nonstationary and unit root cases. *Annals of Statistics* 32, 656-692.

Phillips, P.C.B. and K. Shimotsu, 2005. Exact local Whittle estimation of fractional integration. *Annals of Statistics* 33, 4, 1890-1933.

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

Robinson, P.M., 1995. Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* 23, 1630-1661.

Sowell, F., 1992, Modelling long run behaviour with the fractional ARIMA model, Journal of Monetary Economics 29, 277-302.

Velasco, C., 1999. Gaussian semiparametric estimation of nonstationary time series, *Journal of Time Series Analysis* 20, 87-127.
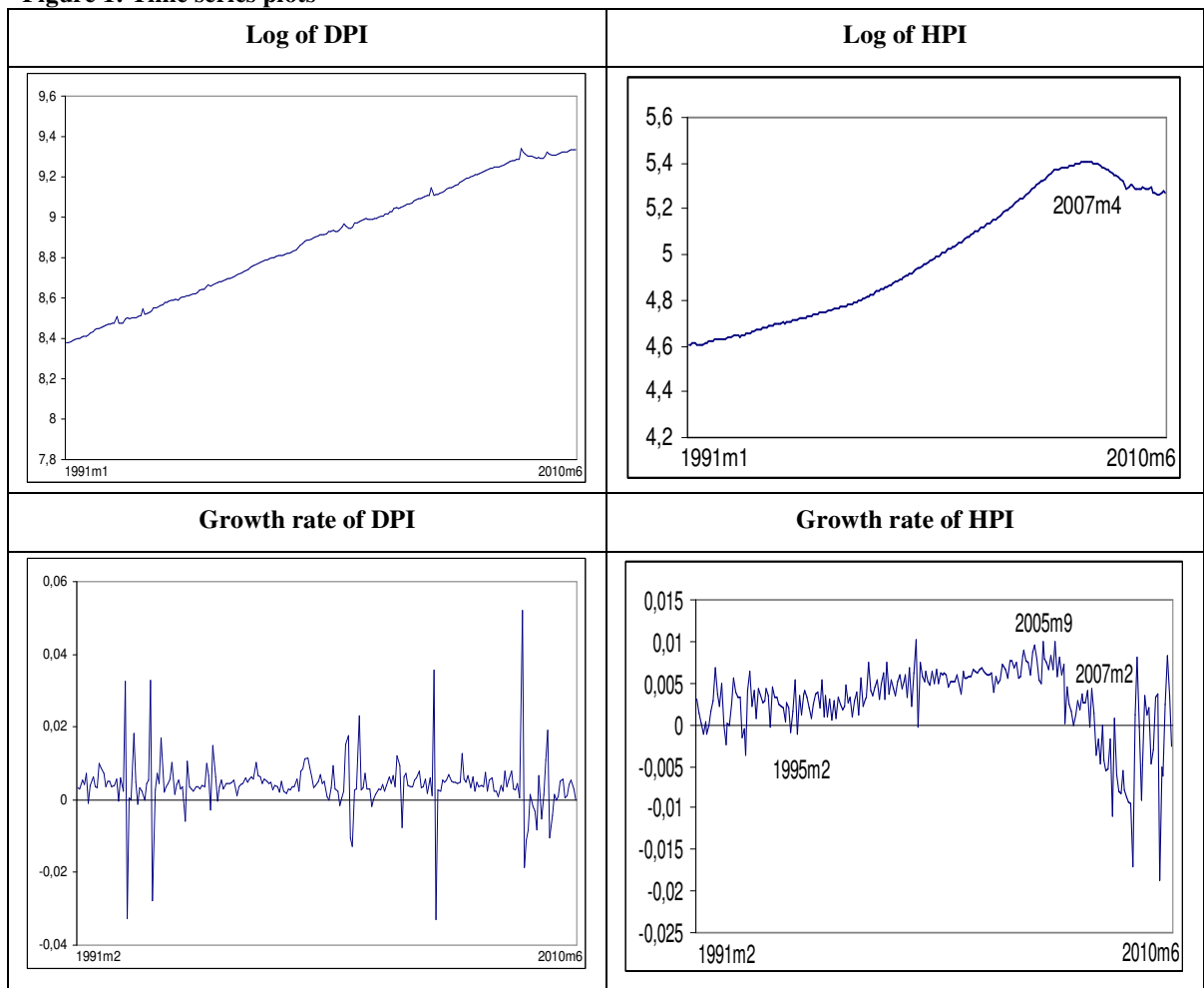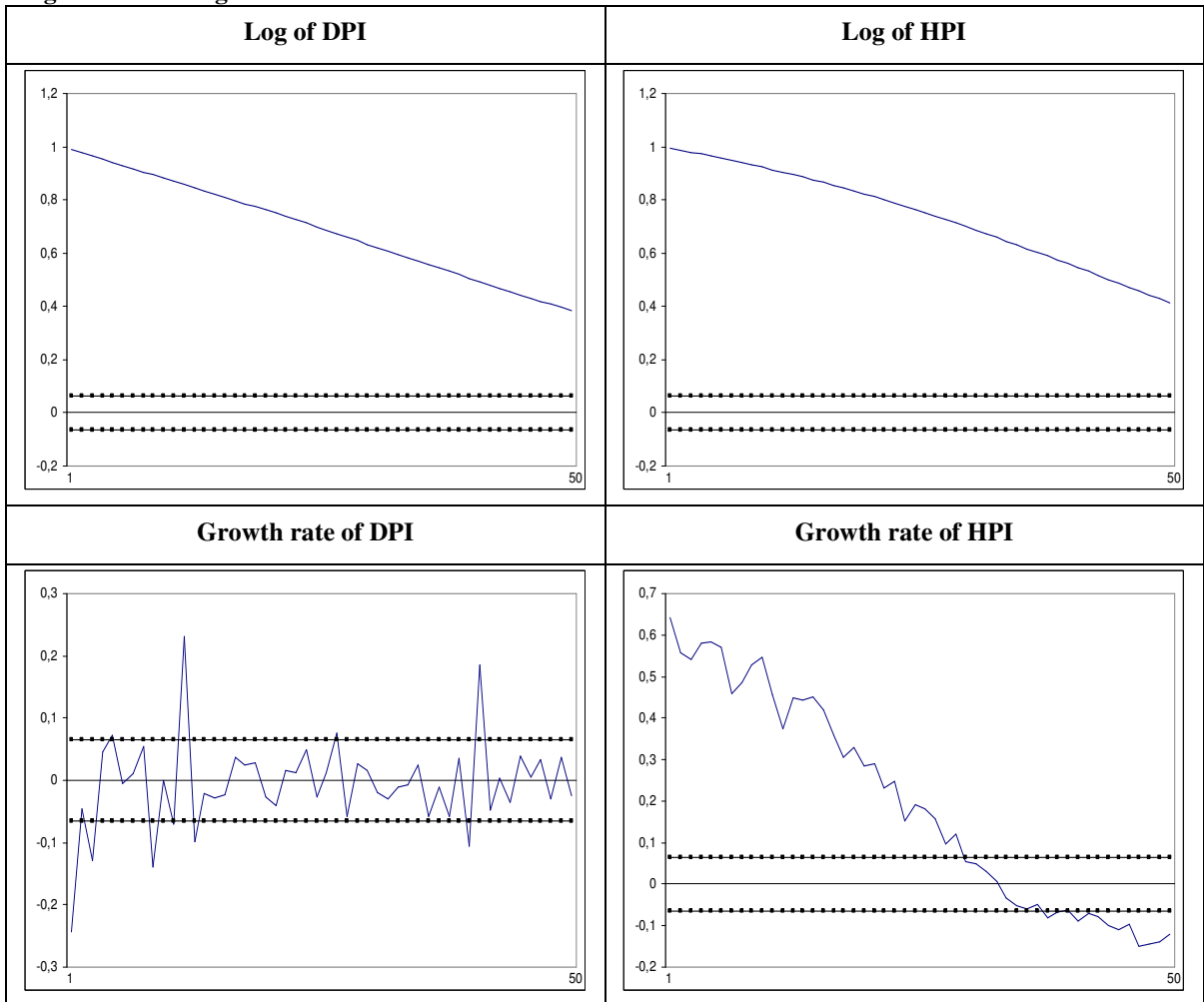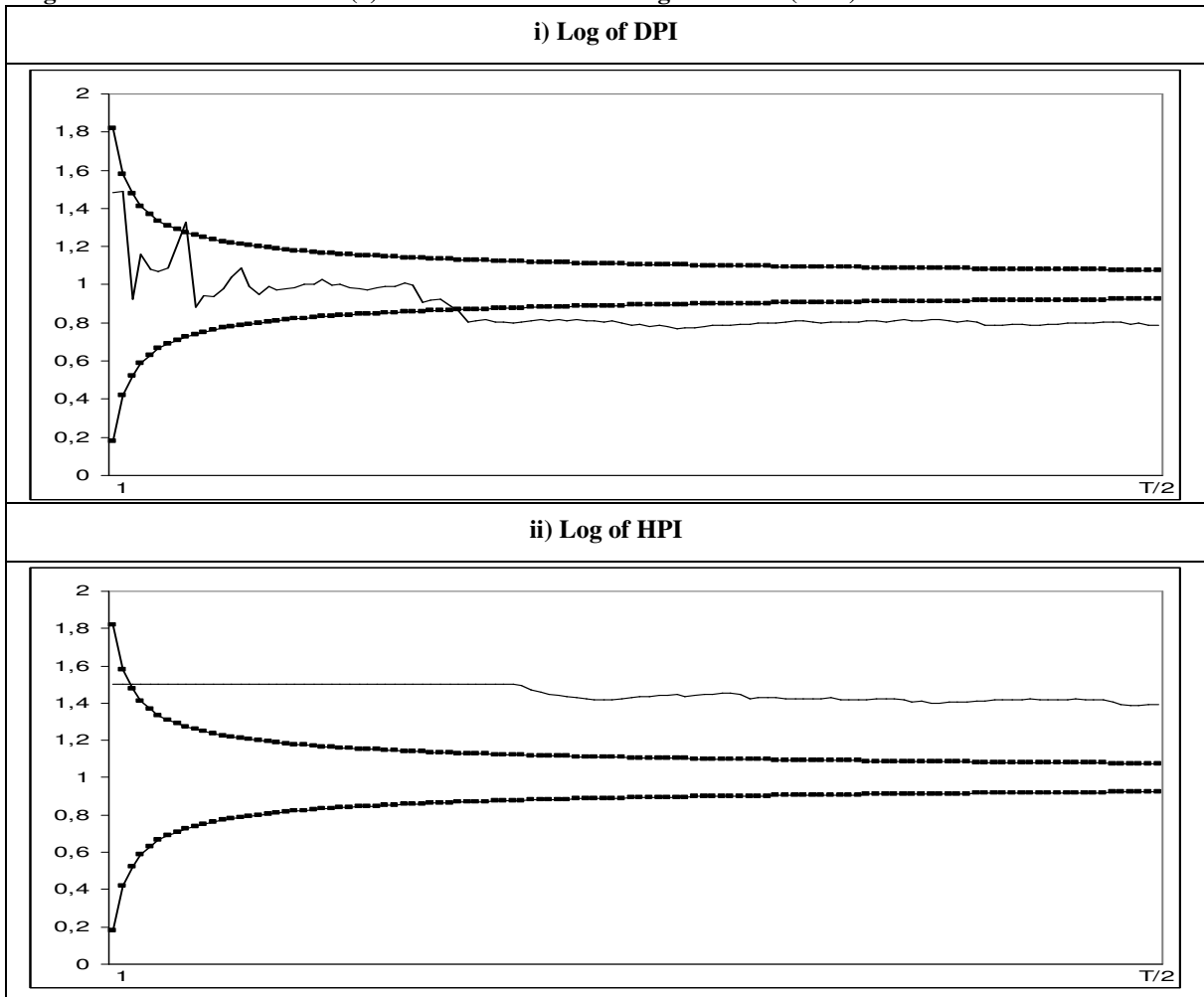
Figure 1: Time series plots

Log of DPI	Log of HPI

| Growth rate of DPI	Growth rate of HPI

Figure 2: Correlograms of the time series

The thick lines correspond to the 95% confidence bands for the null hypothesis of no autocorrelation.

Figure 3: Estimates of d and I(1) confidence intervals using Robinson (1995)

i) Log of DPI

ii) Log of HPI

The horizontal axis reports the bandwidth parameter and the vertical one the estimated value of d. The thick lines correspond to the 95% confidence bands for the I(1) null hypothesis.

Figure 4: Recursive estimates of d using white noise disturbances
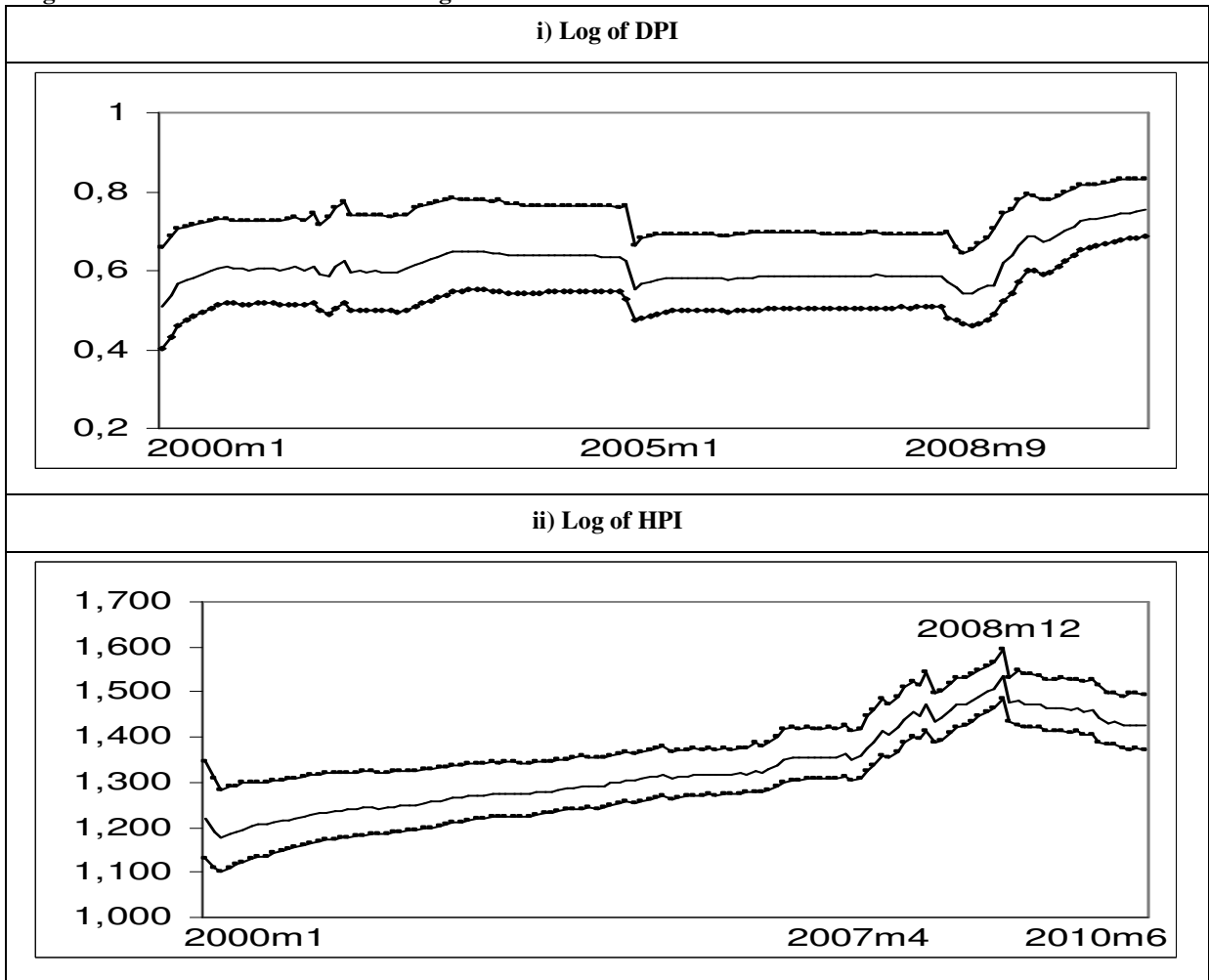
The thick lines refer to the 95% confidence intervals.

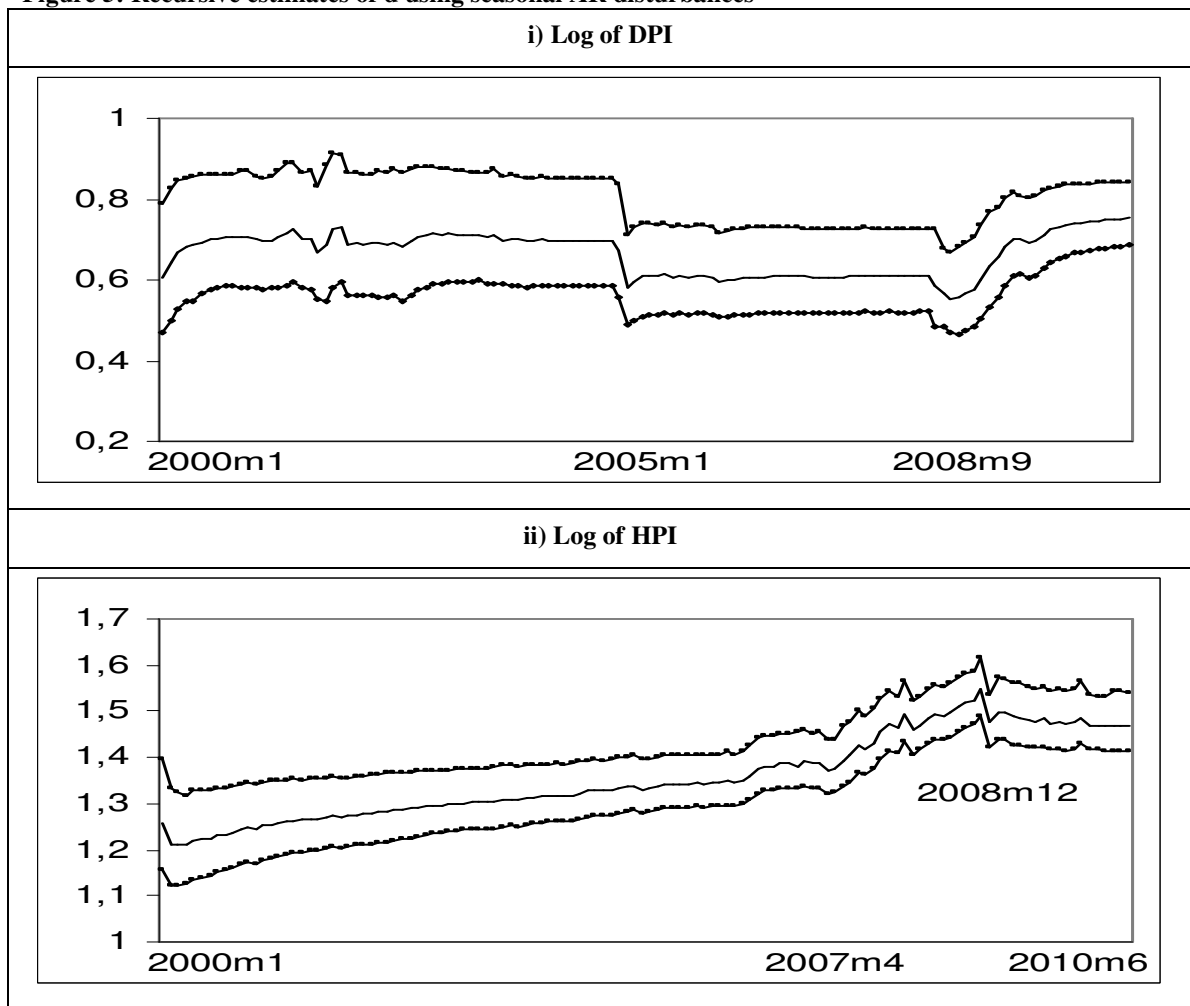**Figure 5: Recursive estimates of d using seasonal AR disturbances**

The thick lines correspond to the 95% confidence bands.

Table 1: Estimates of d and 95% confidence bands for each series

a) Log of DPI			
	No regressors	An intercept	A linear trend
White noise	0.983 (0.906, 1.083)	0.831 (0.804, 0.871)	0.752 (0.689, 0.833)
AR (1)	1.382 (1.247, 1.565)	0.873 (0.834, 0.942)	0.808 (0.709, 0.929)
Bloomfield (1)	0.963 (0.831, 1.128)	0.873 (0.830, 0.965)	0.840 (0.735, 0.972)
Seasonal AR (1)	0.978 (0.882, 1.084)	0.821 (0.791, 0.868)	0.754 (0.686, 0.842)
a) Log of HPI			
	No regressors	An intercept	A linear trend
White noise	0.980 (0.903, 1.082)	1.427 (1.375, 1.496)	1.424 (1.372, 1.493)
AR (1)	1.382 (1.247, 1.566)	1.475 (1.410, 1.559)	1.471 (1.406, 1.556)
Bloomfield (1)	0.961 (0.829, 1.135)	1.492 (1.414, 1.591)	1.483 (1.405, 1.581)
Seasonal AR (1)	0.978 (0.882, 1.082)	1.472 (1.415, 1.543)	1.469 (1.413, 1.540)

The estimates of d are based on the Whittle function in the frequency domain. The values in parentheses are the non-rejection values of d at the 95% level using Robinson's (1994) approach.