Department of Economics and Finance

Economics and Finance Working Paper Series

Working Paper No. 2319

Guglielmo Maria Caporale and Luis Alberiko Gil-Alana

Long-run Trends and Cycles in US House Prices

October 2023

LONG-RUN TRENDS AND CYCLES IN US HOUSE PRICES

Guglielmo Maria Caporale

Brunel University London, UK

Luis Alberiko Gil-Alana

University of Navarra, NCID, DATAI, Pamplona, Spain

and Universidad Francisco de Vitoria, Madrid, Spain

October 2023

Abstract

This paper analyses US nominal house prices at an annual frequency over the period from 1927 to 2022 by means of a very general time series model. This includes both a (linear and non-linear) deterministic and a stochastic component, with the latter allowing for fractional orders of integration at both the long-run and the cyclical frequencies. The results are heterogeneous depending on the model specification and on whether or not the series have been logged. Specifically, a linear model appears to be more appropriate for the logged data whilst a non-linear one appears to be a better fit for the original ones. Further, the order of integration at the zero or long-run frequency is much higher than at the cyclical one. The former is in fact around 1 in all specified models, which implies a high degree of persistence of this component. Finally, the order of integration of the cyclical structure implies that cycles have a periodicity of about 8 years, but it is almost insignificant in all cases.

Keywords: US house prices; trends; cycles; persistence; long memory; fractional integration

JEL Classification: C15; C22; E30

Corresponding author: Professor Guglielmo Maria Caporale, Department of Economics and Finance, Brunel University London, Uxbridge, UB8 3PH, UK. Email: Guglielmo-Maria.Caporale@brunel.ac.uk; https://orcid.org/0000-0002-0144-4135

* Luis A. Gil-Alana gratefully acknowledges financial support from an Internal Project of the Universidad Francisco de Vitoria and the MINEIC-AEI-FEDER PID2020-113691RB-I00 project from ‘Ministerio de Economía, Industria y Competitividad’ (MINEIC), `Agencia Estatal de Investigación' (AEI) Spain and `Fondo Europeo de Desarrollo Regional' (FEDER).

1. Introduction

House prices are a key variable whose fluctuations can have a significant impact on both the real and the financial sectors of the economy, as documented, among others, by Case et al. (2005), Davis and Heathcote (2005), Leamer (2007), Attanasio et al. (2011), Carroll et al. (2011), Funke and Paetz (2013), Chen et al. (2018). Their crucial importance became even more apparent as a result of the global financial crisis (GFC) of 2007-08. This originated from the US housing market, where the issuance of sub-prime mortgages had become widespread and led to a housing bubble and serious financial turmoil when it eventually burst (see, e.g., Shiller, 2007). Consequently, numerous empirical studies have been carried out to understand the behaviour of house prices. Broadly speaking, two main approaches have been followed in the literature for this purpose, the first focusing on their drivers, the second on their stochastic properties. Among studies belonging to the first category, Capozza and Helsely (1989, 1990) analysed the impact of real income on real house prices, whilst Caporale and Gil-Alana (2015) used fractional integration methods to show that the US Housing Price Index (HPI) and Disposable Personal Income (PDI) do not converge over time, presumably owing to the existence of a bubble.

The second type of studies carry out univariate analysis of the house price series. The early literature used unit root tests (see, e.g. Meen, 1999, for UK regional prices, and Cook and Vougas, 2009 for aggregate prices in the presence of structural breaks; Clark and Coggin, 2011, and Zhang et al., 2017, for the US; Arestis and Gonzales, 2014, for 18 OECD countries; etc.). However, it is well known that this type of tests has very low power against specific alternatives such as structural breaks (Campbell and Perron, 1991); trend-stationary models (DeJong et al., 1992), regime-switching (Nelson et al., 2001), or fractional integration (Diebold and Rudebusch, 1991; Hassler and Wolters, 1994; Lee and Schmidt, 1996; etc.). By contrast, a fractional integration framework (see Granger, 1980;

Granger and Joyeux, 1980 and Hosking, 1981) is much more general, since it is not based on the dichotomy between I(0) stationary and I(1) non-stationary series, which is very restrictive. Instead the differencing parameter d is allowed to take any real value, including fractional ones. This approach encompasses a wide range of stochastic behaviours, including the unit root case, and provides evidence on whether or not the series of interest is mean-reverting (and thus on whether exogenous shocks have permanent or transitory effects) and on its degree of persistence. It has been used in some studies on house prices, such as Barros et al. (2012, 2015), Gil-Alana et al. (2013, 2014), and Gupta et al. (2014) to analyse long-run persistence.

An important issue in this context is the possible presence of structural breaks. Caporale and Gil-Alana (2023) allow for them within a fractional integration framework to model the monthly Federal Housing Finance Agency (FHFA) House Price Index for Census Divisions, and the US as a whole, over the period from January 1991 to August 2022. Their analysis detects segmented trends, with the subsample estimates of the fractional differencing parameter being lower and mean reversion occurring in several cases.

Other recent papers argue that it is also essential to allow for both a long-run and a cyclical component in house prices. Such a modelling approach is followed by Canarella et al. (2021) to examine persistence in both US and UK house prices over a long time span. Their conclusion is that the long-run component dominates, and also that there are breaks in the series corresponding to different domestic factors. Compared to that study, the present one adopts an even more general specification, since it includes not only a stochastic component allowing for fractional integration at both the long-run and cyclical frequencies, but also a deterministic one which can be either linear or non-linear, the two being jointly modelled (see the specification in Section 2); moreover, autocorrelation of a general form as in Bloomfield (1973) is allowed in the error term. This framework is applied to analyse US nominal house prices at an annual frequency over the period from 1927 to 2022.

The layout of the paper is the following. Section 2 outlines the modelling framework. Section 3 describes the data and presents the empirical results. Section 4 offers some concluding remarks.

2. The Econometric Model

The model estimated in this study is more general than those used in the previous literature on house prices. Specifically, it includes both a deterministic and a stochastic component, with the latter allowing for fractional degrees of integration at both the long-run and cyclical frequencies.

The deterministic part of the model is specified as follows:

$$y(t) \;=\; f(z(t);\, \psi) \;+\; x(t), \qquad t \;=\; 1, 2, \dots, \tag{1}$$

where y(t) stands for house prices (either the original or the logged series), and f is a function that can be linear, for instance including an intercept and a linear time trend, (Bhargava, 1986, Schmidt and Phillips, 1992) as in the following equation:

$$f(z(t);\, \psi) \;=\; \alpha \;+\; b\, t \tag{2}$$

or non-linear, for example including Chebyshev polynomials in time of the following form:

$$f(z(t);\, \psi) \;=\; \sum_{i=0}^{m} \theta_i P_{iT}\,(t), \tag{3}$$

where *m* denotes the number of coefficients of the Chebyshev polynomial in time $P_{i,T}(t)$ defined as:

$$P_{0,T}(t) = 1, \text{ and } P_{i,T}(t) \;=\; \sqrt{2}\, cos(i\pi(t-0.5)/T),$$

Hamming (1973) and Smyth (1998) provided a detailed description of these polynomials, whilst Bierens (1997) and Tomasevic and Stanivuk (2009) argued that it is possible to approximate highly non-linear trends with rather low degree polynomials. If $m = 0$ the model contains an intercept, and if $m \geq 1$, it becomes non-linear - the higher m is, the less linear the approximated deterministic component becomes.

Concerning the stochastic terms, x(t) in (1) is assumed to be a process characterised by two orders of integration, one for the long-run or zero frequency, which captures possible stochastic trends, and the other for the cyclical structure of the data. More precisely, x(t) is defined as follows:

$$(1-L)^{d_1}\,(1-2\,\cos w(r)\,L+L^2)^{d_2}\,x(t) \;=\; u(t)_t, \qquad t=1,2,\ldots, \tag{4}$$

where L is the lag operator (i.e., Lx(t) = x(t-1); d_1 and d_2 are real parameters, w(r) = 2πr/T, r = T/j, with j indicating the number of periods per cycle, and u(t) being a short memory or I(0) process, defined as a covariance stationary one with a spectral density function that is positive and finite at all frequencies. Thus, u(t) may be a white noise process with zero mean and constant variance, but it may also include some type of weak autocorrelation as in the case of the stationary and invertible AutoRegressive Moving Average (ARMA)-type of models. Here, we impose autocorrelation by applying the non-parametric approach of Bloomfield (1973), which involves using a spectral density function of the following form:

$$f(\lambda;\,\tau) \;=\; \left[\frac{\sigma 2}{2\pi}\right]\exp[\,2\textstyle\sum_{i=0}^{n}\tau_i\cos(\lambda\, i)\,]\;, \tag{5}$$

where σ^2 is the variance of the error term and n denotes the number of short-run dynamic terms. Bloomfield (1973) showed that, given a stationary and invertible ARMA (p, q) process of the following form:

$$u(t) \;=\; \sum_{r=1}^{p}\varphi_r\,u(t-r) \;+\; \varepsilon_t \;+\; \sum_{s=1}^{q}\theta_s\varepsilon(t-s)\,,$$

where ε_t is a white noise process, its spectral density function is given by:

$$f(\lambda;\tau) \;=\; \frac{\sigma^2}{2\pi}\;\left|\frac{1\;+\;\sum_{s=1}^{q}\theta_s e^{i\lambda s}}{1\;-\;\sum_{r=1}^{p}\varphi_r e^{i\lambda r}}\right|^2 .$$

According to Bloomfield (1973), the log of the above expression can be well approximated by Eq. (5) when p and q are small values, and thus it does not require the estimation of as many parameters as in the case of ARMA models. In addition, Bloomfield's (1973) model has the advantage of being stationary for all its values (see Gil-Alana, 2004).

Let us now consider further Eq. (4). Note that the first polynomial can be expanded for any real value d_1 as

$$\sum_{j=0}^{\infty}\binom{d_1}{j}(-1)^j L^j \;=\; 1 - d_1 L + \frac{d_1(d_1-1)}{2}L^2 - \cdots$$

In this context, d_1 indicates the degree of persistence of the series in relation to the long-run or zero frequency. Thus, if $d_2 = 0$ in Eq. (4), x(t) can be expressed as

$$x(t) \;=\; d_1 x(t-1) \;-\; \frac{d_1(d_1-1)}{2}\,x(t-2) \;+\; \ldots \;+\; u(t)$$

and the higher the value of d_1 is, the higher is the degree of dependence between the observations. Moreover, if d_1 is positive, x(t) displays the property of long memory since in that case its spectral density function becomes

$$f(\lambda;\;\tau) = \frac{\sigma^2}{2\pi}\left|\frac{1}{1-e^{i\lambda}}\right|^{d_1},$$

which tends to infinity as $\lambda \to 0^+$.

This specification allows us to consider a wide range of cases including, among others, the following ones:

i) anti-persistence, if $d_1 < 0$,

ii) short memory, if $d_1 = 0$

iii) long memory, though covariance stationary processes, if $0 < d_1 < 0.5$,

iv) 1/f noise, if $d_1 = 0.5$,

v) nonstationary mean reverting processes, if $0.5 \leq d_1 < 1$,

vi) unit roots, if $d_1 = 1$,

vii) explosive processes, if $d_1 > 1$.

Next we focus on the cyclical structure of x(t) which is captured by the second polynomial in (4). Gray et al. (1989) showed that, by denoting $\mu = w(r)$, this polynomial can be expressed in terms of the orthogonal Gegenbauer terms $C_{i,d_2}(\mu)$, such that for all real d2 $\neq 0$,

$$(1 - 2\mu L + L^2)^{-d_2} = \sum_{i=0}^{\infty} C_{i,d_2}(\mu) L^i,$$

where $C_{i,d_2}(\mu)$ can be defined recursively as:

$$C_{0,d_2}(\mu_r^j) = 1, \quad C_{1,d_2}(\mu_r^j) = 2\,\mu\, d,$$

and

$$C_{i,d_2} = 2\,\mu\left(\frac{d_2 - 1}{j} + 1\right) C_{i-1,d_2}(\mu) - \left(2\frac{d_2 - 1}{i} + 1\right) C_{i-2d_2}(\mu)\,.$$

This type of process was introduced by Andel (1986), and authors such as Gray et al. (1989, 1994), Giraitis and Leipus (1995), Chung (1996a, 1996b), Gil-Alana (2001), Dalla and Hidalgo (2005), Caporale and Gil-Alana (2013) and others subsequently used it to analyse time series data.

3. Data Description and Empirical Results

We analyse nominal house prices for the US, at an annual frequency, from 1927 to 2022, which have been obtained from the Federal Reserve Bank of St. Louis database and compiled by Robert Shiller in http://www.econ.yale.edu//~shiller/data.htm.

INSERT FIGURES 1 – 3 ABOUT HERE

Figure 1 displays time series plots of the original series, its logged transformation, and the first differences of both. It can be seen that the series in levels, whether logged or not, exhbit an upward trend throughout the sample period under examination, except for a sharp drop corresponding to the global financial crisis (GFC) of 2007-08, after which prices recovered and returned to their growth path. The first differenced series are much more volatile (especially the logged one), but again one can observe a fall coinciding with the GFC, which is followed by a swift recovery.

Figure 2 shows the correlograms of all four series. It can be seen that the values for the original series and their log transformations decay very slowly, which may indicate the presence of unit roots, whilst the values for the first differenced data suggest the presence of a cyclical pattern. The first 20 values of the periodograms are reported in Figure 3. Similarly to the correlograms, these are large and positive at the long-run or zero frequency in the case of the series in levels (see the upper panel), which might indicate the presence of long memory (i.e., $d_1 > 0$); however, the plots for the differenced series (see the lower panel also suggest possible cyclical patterns.

The first estimated model focuses only on the long-run or zero frequency and is specified as follows:

$$y(t) = \alpha + \beta t + x(t), \qquad (1-L)^d x(t) = u(t), \qquad (6)$$

where d is the fractional differencing parameter, α and β are jointly estimated with d, t stands for a linear time trend, and u(t) follows the exponential spectral model of Bloomfield (1973) implicitly defined by equation (5).

Table 1 displays the estimates of d along with the corresponding 95% confidence intervals, under the assumption of i) no deterministic terms (α = β = 0 in (6)); ii) an intercept only (β = 0 a priori) and iii) an intercept and a linear time trend. The preferred specification is chosen on the basis of the statistical significance of the estimated coefficients. We report the results for both the original and log-transformed data in levels.

TABLES 1 AND 2 ABOUT HERE

Table 2 displays the estimated parameters from the selected model for each of the two series. The time trend is statistically significant in both cases with a positive coefficient, and the estimates of d are 0.85 for the original data and 0.97 for the log-tansformed ones. However, the confidence intervals imply that the unit root null hypothesis (i.e., d = 1) cannot be rejected for either series.

Next, we consider a non-linear specification with Chebyshev polynomials in time. Specifically, the estimated model is now the following:

$$y(t) \;=\; \sum_{i=0}^{m} \theta_i P_{iT}\,(t) \;+\; x(t), \qquad (1-L)^d\, x(t) \;=\; u(t), \qquad (7)$$

where P_{iT} are the Chebyshev polynomials defined above and u(t) again follows the exponential spectral model of Bloomfield (1973). The results, for m = 3, are displayed in Table 3.

TABLE 3 ABOUT HERE

It can be seen that they now differ depending on the series analysed. More precisely, for the original data the estimate of d is about 0.52, though the confidence interval is extremely large and it includes both the I(0) and the I(1) hypotheses. In addition, the coefficients for the non-linear trends are statistically significant. However, for the logged values, these coefficients (θ_2 and θ_3) are insignificant and the estimate of d is 0.84 (0.34, 1.56), such that the unit root null hypothesis cannot be rejected.

Next we allow for a cyclical component. First we consider the linear case and thus estimate the following model:

$$y(t) \;=\; \alpha \;+\; \beta\, t \;\;+\; x(t),$$

$$(1-L)^{d_1}\,(1-2\;\cos w(r)\,L + L^2)^{d_2}\, x(t) \;\;=\;\; u(t)\,, \qquad t = 1, 2, \ldots, \qquad (8)$$

where d_1 refers to the long-run or zero frequency and d_2 to the order of integration of the cyclical component. On the basis of the plots of the periodograms displayed in Figure 3, we assume that r in equation (8) is constrained between 4 and 20, which corresponds to cycles between 5 and 24 years.

Table 4 reports the results based once more on the assumption of u(t) following the exponential spectral model of Bloomfield (1973), again in the case of i) no deterministic terms, ii) a constant, and iii) a constant and a linear time trend. The time trend is again statistically significant for both the original and the logged data, with the estimated values of r being 11 and 12 in those two cases, which corresponds to cycles of approximately 8 years (T = 96/11 = 8.72, and 96/12 = 8). As for the differencing parameters, their values from the selected specifications (marked in bold in the table) are 0.94 and 0.91 for d_1 in the case of the original data and the logged ones respectively, with the confidence intervals including the unit root case; the corresponding ones for d_2 are 0.07 and 0.09 respectively, with the confidence intervals being very wide, and thus the null of $d_2 = 0$ not being rejected. This implies that there is no significant cyclical component in the series under examination.

TABLES 4 AND 5 ABOUT HERE

Table 5 displays the corresponding results when allowing for non-linearities in the form of Chebyshev polynomials in time. As in the previous case, the estimates of r are equal to 11 and 12 for the original data and the logged ones respectively, which implies the presence of cycles of about 8 years. The estimates of d_1 and d_2 are now slightly higher than in the linear case. Specifically, they are equal to 1.10 and 1.06 for d_1 in the case of the original and logged data respectively, which implies once more that the unit root null hypothesis, i.e., $d_1 = 1$, cannot be rejected in either case. As for the estimated value of d_2, this is now positive but only slightly significant in the case of the original data ($d_2 = 0.29$)

while insignificant (d_2 = 0.20) for the logged series. Finally, in line with the results reported in Table 3 for a model with a single order of integration, the non-linear coefficients are statistically significant in the case of the original data.

4. Conclusions

This paper proposes a very general time series model not previously used in the literature on house prices to analyse their behaviour in the US from 1927 to 2022. Specifically, the adopted fractional integration framework includes both a deterministic and a stochastic component (the latter modelling both long-run and cyclical behaviour), and also allows for non-linearities. It has the advantage of encompassing a wide range of stochastic processes (including the standard unit root case) and provides useful information on properties such as mean reversion and persistence.

The results are heterogeneous depending on the model specification and on whether or not the series have been logged. Specifically, a linear model appears to be more appropriate for the logged data whilst a non-linear one seems to be a better fit for the original ones. Further, the order of integration at the zero or long-run frequency is much higher than at the cyclical one. The former is in fact around 1 in all specified models, which implies a high degree of persistence of this component. Finally, the order of integration of the cyclical structure implies that cycles have a periodicity of about 8 years, but it is almost insignificant in all cases.

These results are broadly consistent with those of Canarella et al. (2021), who had analysed a longer sample period from 1830 to 2016 in the case of the US (from 1845 to 2016 in the case of the UK) and found evidence of significant cyclical persistence only in the first sub-sample, the dominant break in their sample corresponding to some important post-WWII developments in US housing policy, such the National Housing

Act of 1949 with the following 1955 Amendment, and the Housing and Urban Development Act of 1965. Therefore the more general framework we employ appears to confirm that cyclicality has more recently become a less crucial issue for US house prices, and that it might not be necessary to account explicitly for it when bulding forecasting models. In addition, our findings are important for policy makers, since they imply that their focus should be on long-run persistence rather than cyclical one in the case of house prices.

Our analysis could be extended in several ways. In particular, tests for structural breaks could be carried out using the approach of Bai and Perron (2003), or the one developed by Gil-Alana (2008) specifically in the context of fractional integration, and then sub-sample estimates could be obtained. Non-linearities could also be modelled using other methods such as Fourier transform functions (Gil-Alana and Yaya, 2021; Caporale et al., 2023) or neural networks (Yaya et al., 2021), all of them in the context of fractional integration. Finally, the robustness of the results could be checked using de-seasonalised data as well as estimating the models at different frequencies. Future work will address these issues.

References

Andel, J. (1986), Long memory time series models. Kybernetika 22, 105-123.

Arestis, P. and A.R. González, A.R. (2014), Modelling the housing market in OECD countries, International Review of Applied Economics 28, 131-153.

Attanasio, O.P, Leicester, A. and M. Wakefield (2011), Do house prices drive consumption .growth? The coincident cycles of house prices and consumption in the UK. Journal of the European Economic Association 9, 399–435.

Bai, J. and P. Perron P (2003), Computation and analysis of multiple structural change models, Journal of Applied Econometrices, 18, 1, 1-22. https://doi.org/10.1002/jae.65

Barros, C. P., Gil-Alana, L. A., and J.E. Payne (2012), Comovements among US state housing prices: Evidence from fractional cointegration, Economic Modelling 29: 936 942.

Barros, C. P., Gil-Alana, L. A., and J.E. Payne (2015), Modeling the Long Memory Behavior in US Housing Price Volatility, Journal of Housing Research 24, 87-106.

Bhargava, A. (1986), On the Theory of Testing for Unit Roots in Observed Time Series, Review of Economic Studies 53, 3, 369-384.

Bierens, H. J. (1997), Testing the unit root with drift hypothesis against nonlinear trend stationarity with an application to the US price level and interest rate. Journal of Econometrics 81, 29-64.

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

Campbell, J.Y. and P. Perron (1991), Pitfalls and opportunities: What macroeconomists should know about unit roots, NBER Macroeconomic Annual 1141-201.

Canarella, G., Gil-Alana. L.A., Gupta, R. and S.M. Miller (2021), Persistence and cyclical dynamics of US and UK house prices: Evidence from over 150 years of data, Urban Studies, 58, 1, 53-72.

Caporale, G.M., J. Cunado and L.A. Gil-Alana, L.A. (2013), Modelling long trends and cycles in financial time series data, Journal of Time Series Analysis 34, 3, 405-421.

Caporale, G.M. and L.A. Gil-Alana (2015), US disposable personal income and housing price index: a fractional integration analysis, Journal of Housing Research, 24, 1, 73-86.

Caporale, G.M. and L.A. Gil-Alana (2023), "US house prices by Census Division: persistence, trends and structural breaks", International Advances in Economic Research, 29, 79-90, https://doi.org/10.1007/s11294-023-09868-9

Caporale, G.M., Gil-Alana, L. and O.S. Yaya (2023), "Modelling persistence and non-linearities in the US Treasury 10-year bond yields", forthcoming, Economics Bulletin.

Capozza, D., and R. Helsley (1989), The fundamentals of land prices and urban growth, Journal of Urban Economics 26, 295-306.

Capozza, D., and R. Helsley (1990), The stochastic city, Journal of Urban Economics, 28, 187-203.

Carroll, D.C., Otsuka, N. and J. Slacalek (2011), How large are housing and financial wealth effects? A new approach. Journal of Money, Credit and Banking 43, 55–79.

Case, K.E., Quigley, J. and R.J. Shiller (2005), Comparing wealth effects: The stock market versus the housing market. Advances in Macroeconomics 5: 1235–1235.

Chen, H., Michaux, M. and N. Roussanov (2018), Houses as ATMs? Mortgage refinancing and macroeconomic uncertainty. NBER Working Paper No.19 421.

Chung, C.-F. (1996a), A generalized fractionally integrated autoregressive moving-average process. Journal of Time Series Analysis 17, 111-140.

Chung, C.-F. (1996b), Estimating a generalized long memory process. Journal of Econometrics 73, 237-259.

Clark, S.P., and T.D. Coggin (2011), Was there a US house price bubble? An econometric analysis using national and regional panel data. Quarterly Review of Economics and Finance 51, 189-200.

Cook, S., and D. Vougas (2009), Unit root testing against an ST-MTAR alternative: finite sample properties and an application to the UK housing market, Applied Economics 41, 1397 404.

Dalla, V., and Hidalgo, J. (2005), A parametric bootstrap test for cycles. Journal of Econometrics 129, 219-261.

Davis, M.A. and J. Heathcote (2005), Housing and the business cycle. International Economic Review 46: 751–784.

De Jong, D., J. Nankervis, N.E. Savin and C.H. Whiteman (1992), Integration versus trend stationarity in time series, Econometrica 60, 423-433.

Diebold, F.X. and G.D. Rudebusch (1991) On the power of Dickey-Fuller test against fractional alternatives, Economics Letters 35, 155-160.

Funke, M. and M. Paetz (2013), Housing prices and the business cycle: An empirical application to Hong Kong. Journal of Housing Economics 22, 62–71.

Gil-Alana, L.A. (2001), Testing stochastic cycles in macroeconomic time series. Journal of Time Series Analysis 22, 411-430.

Gil-Alana, L.A. (2004), The use of the Bloomfield model as an approximation to ARMA processes in the context of fractional integration. *Mathematical and Computer Modelling*. 39: 429-436. doi: 10.1016/S0895-7177(04)90515-8.

Gil-Alana, L.A. (2008), Fractional integration and structural breaks at unknown periods of time. Journal of Time Series Analysis, 29, 163-185.

Gil-Alana, L. A., Aye, G., and R. Gupta (2013), Testing for persistence in South African house prices, Journal of Real Estate Literature 21, 293-314.

Gil-Alana, L. A., Barros, C. P., and N. Peypoch (2014), Long memory and fractional integration in the housing price series of London and Paris, Applied Economics 46, 3377-3388.

Gil-Alana, L.A. and O. Yaya (2021), Testing fractional unit roots with non-linear smooth break approximations using Fourier functions, Journal of Applied Statistics 48 (13-15), 2542-2559.

Giraitis, L., and Leipus, R. (1995), A generalized fractionally differencing approach in long memory modeling. Lithuanian Mathematical Journal 35, 65-81.

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 R. Joyeux (1980), An Introduction to Long Memory Time Series Models and Fractional Differencing, Journal of Time Series Analysis 1, 1, 15-29.

Gray, H.L., Zhang, N.-F. & Woodward, W.A. (1989), On generalized fractional processes. Journal of Time Series Analysis, 10, 233–257.

Gray, H.L., N.-F. Zhang, W. A. Woodward (1994), On generalized fractional processes-a correction, Journal of Time Series Analysis 15 (5), 561–562.

Gupta, R., André, C., and L.A. Gil-Alana (2014), Comovements in Euro area housing prices. A fractional cointegration approach, Urban Studies 52, 3123-3143.

Hamming, R. W. (1973). Numerical Methods for Scientists and Engineers. Dover.

Hassler, U. and J. Wolters (1994) On the power of unit root tests against fractional alternatives, Economics Letters 45, 1-5.

Hosking, J.R.M. (1981), Fractional differencing.Biometrika, 68, 168–176.

Leamer, E.E. (2007), Housing is the business cycle. Proceedings of the Federal Reserve Bank of Kansas City. NBER Working Paper No. 13428, 149–233.

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.

Meen, G. (1999), Regional House Prices and the Ripple Effect: A New Interpretation, Housing Studies 14, 6, 733-753.

Nelson, C.R., J. Piger and E. Zivot (2001), Markov regime-switching and unit root tests, Journal of Business Economics and Statistics 19, 404-415.

Schmidt, P. and P.C.B. Phillips (1992), LM tests for a unit root in the presence of deterministic trends, Oxford Bulletin of Economics and Statistics 54, 257-287.

Shiller, R. J. (2007), The Subprime Solution, Princeton University Press, Princeton.
Smyth, G.K. (1998). Polynomial Approximation. John Wiley & Sons, Ltd, Chichester.

Smyth, G.K. (1998). Polynomial Approximation. In P. Armitage & T. Colton (Eds.), *Encyclopedia of Biostatistics*. Chichester: John Wiley & Sons, Ltd.

Tomasevic, N. M. and Stanivuk, T. (2009), Regression Analysis and approximation by means of Chebyshev Polynomial. Informatologia 42, 3, 166-172.

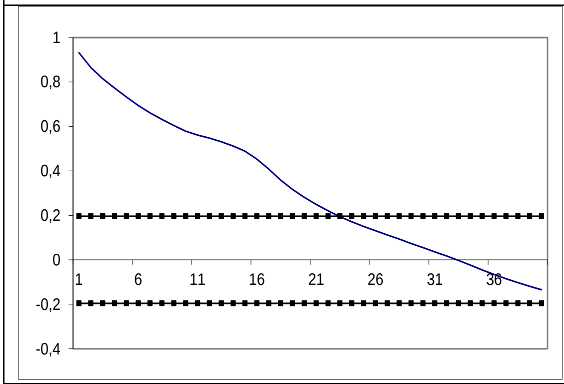
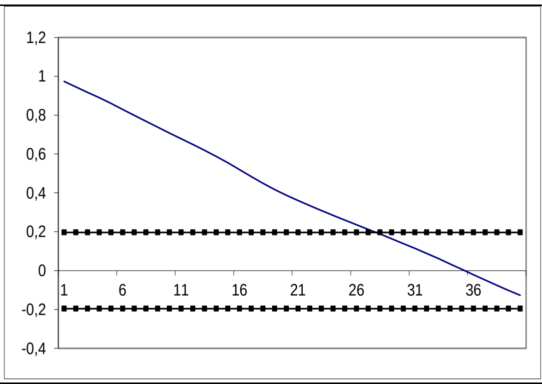
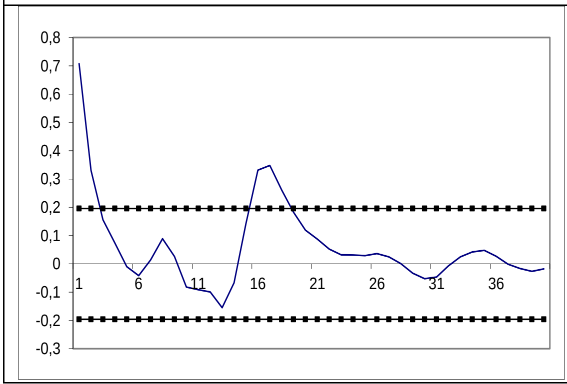
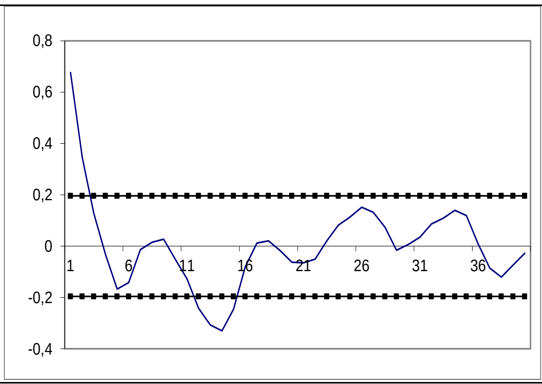
Zhang, H., Hudson, R., Metcalf, H., and V. Manahov (2017), Investigation of institutional changes in the UK housing market using structural break tests and time-varying parameter models, Empirical Economics 53, 617-640.

Yaya, O.S., Ogbonna, A.E., Furuoka, F., & Gil-Alana, L.A. (2021). A new unit root test for unemployment hysteresis based on the autoregressive neural network. Oxford Bulletin of Economics and Statistics, 83(4), 960-981.

Figure 1: Time series plots

Note: The displayed series are annual US nominal house prices from 1927 to 2022.

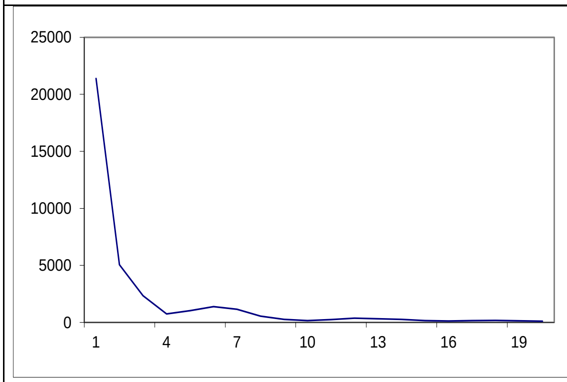
Figure 2: Correlograms of the series

House prices	House prices (in logs)
	
(1 – L) House prices	(1 – L) House prices (in logs)
	

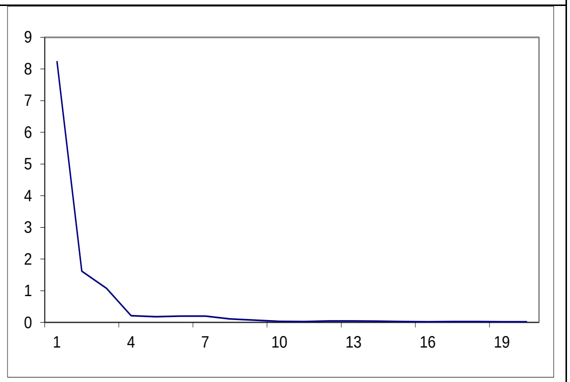
Note: The values in black are the 95% confidence bands for the autocorrelations.

Figure 3: Periodograms of the series

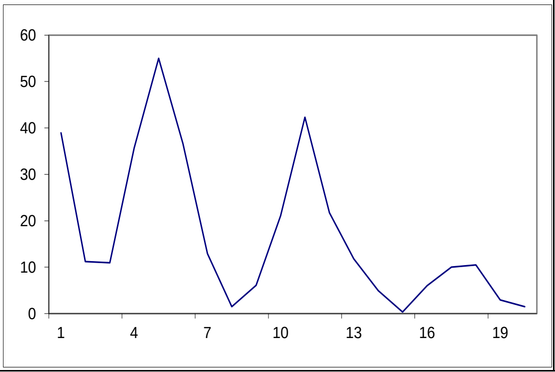
House prices



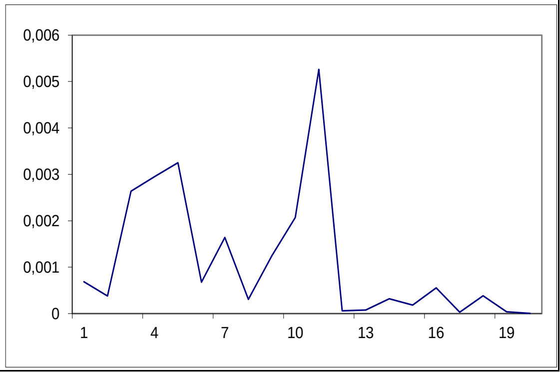
House prices (in logs)



(1 – L) House prices

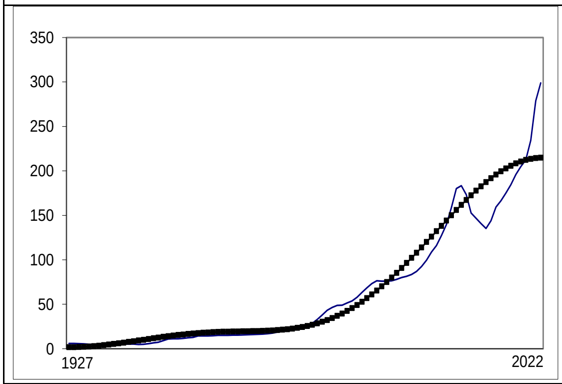
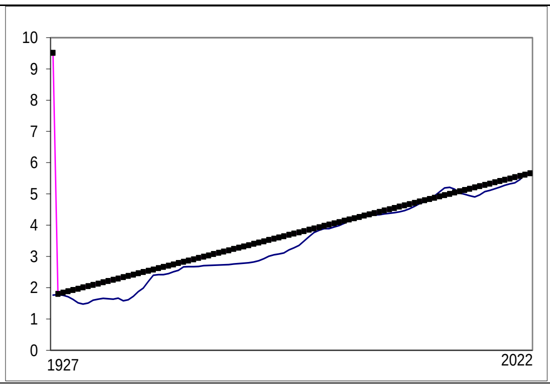


(1 – L) House prices (in logs)



Note: The displayed values are the discrete frequencies $\lambda_j = 2\pi j/T$, for j = 1, … 20

Figure 4: Data and estimated time trend

House prices (non-linear)	House prices (in logs) (linear)
	

Note: Estimated non-linear (left) and linear (right) trends for the original and logged values respectively.

Table 1: Estimates of d at the long-run frequency with a linear trend model

Series	No deterministic terms	With an intercept	With an intercept and a linear trend
Original data	0.83 (0.62, 1.66)	0.78 (0.65, 1.47)	**0.85 (0.57, 1.45)**
Logged data	0.78 (0.55, 1.28)	0.97 (0.80, 1.41)	**0.97 (0.67, 1.41)**

Note: The reported values are the estimates of the differencing parameter d in the model given by Equation (8). The values in parenthesis are the 95% confidence bands, and those in bold are the ones corresponding to the selected models.

Table 2: Estimated coefficients from the selected models in Table 1

Series	d (95% conf. band)	Intercept (tv)	Time trend (tv)
Original data	0.85 (0.57, 1.45)	0.263 (2.05)	2.783 (9.91)
Logged data	0.97 (0.67, 1.41)	1.717 (39.37)	0.041 (10.56)

Note: The values in column 2 are the estimates of d in the model given by Equation (8). In parenthesis, the 95% confidence intervals. The values in columns 3 and 4 are the estimated α and β in the same equation with their associated t-statistics in parenthesis.

Table 3: Estimates of d at the long-run frequency with a non-linear trend model

Series	d	θ_0	θ_1	θ_2	θ_3
Original data	0.52 (-1.58, 1.44)	66.486 (14.80)	-62.863 (-19.70)	29.452 (10.41)	-12.545 (-4.89)
Logged data	0.84 (0.34, 1.56)	3.586 (19.04)	-1.271 (-11.77)	0.039 (0.61)	-0.068 (-1.47)

Note: The values in column 2 are the estimates of d in the model given by Equation (9). In parenthesis, the 95% confidence interval. Those in columns 3 – 6 are the Chebychev coefficients with their associated t-statistics.

Table 4: Estimates in a model with two orders of integration. Linear case

Linear case	No terms		An intercept		**An intercept and a linear time trend**	
Original data	j = 11		j = 12		**j = 12**	
	d_1 = 0.95 (0.61, 1.38)	d_2 = -0.26 (-0.47, 0.31)	d_1 = 1.02 (0.58, 1.41)	d_2 = -0.20 (-0.55, 0.41)	**d_1 = 0.94 (0.43, 1.39)**	**d_2 = 0.07 (-0.31, 0.66)**
	Intercept	Time trend	Intercept	Time trend	**Intercept**	**Time trend**
	---	---	5.76 (3.81)	---	**0.255 (17.89)**	**2.556 (4.35)**
Logged data	j = 12		j = 11		**j = 11**	
	d_1 = 0.91 (0.62, 1.33)	d_2 = -0.24 (-0.61, 0.59)	d_1 = 0.91 (0.70, 1.33)	d_2 = -0.05 (-0.31, 0.40)	**d_1 = 0.91 (0.56, 1.42)**	**d_2 = 0.09 (-0.47, 0.68)**
	Intercept	Time trend	Intercept	Time trend	**Intercept**	**Time trend**
	---	---	1.77 (2.88)	---	**1.689 (15.43)**	**0.035 (2.34)**

Note: d_1 and d_2 are the orders of integration at the long-run and cyclical frequencies respectively as described in Equation (10). In parenthesis the 95% confidence bands. J refers to the frequency with a singularity in the spectrum, such that T/j indicates the number of periods (years) per cycle.

Table 5: Estimates in a model with two orders of integration. Non-linear case

Non-Linear case	j = 11		θ_0	θ_1	θ_2	θ_3
Original data	d_1 = 1.10 (0.42, 1.77)	d_2 = 0.29 (0.00, 0.67)	61.332 (11.54)	-44.182 (-22.31)	34.231 (20.08)	-15.415 (-2.00)
Non-Linear case	j = 11		θ_0	θ_1	θ_2	θ_3
Logged data	d_1 = 1.06 (0.49, 1.82)	d_2 = 0.20 (-0.19, 0.51)	3.549 (19.33)	-1.199 (-9.89)	0.051 (0.13)	-0.061 (-1.03)

Note: d_1 and d_2 are the orders of integration at the long-run and cyclical frequencies respectively as described in Equation (10). In parenthesis the 95% confidence bands. The other values are the Chebychev coefficients with their associated t-statistics.