Department of Economics and Finance

Economics and Finance Working Paper Series

Working Paper No. 2220

Guglielmo Maria Caporale and Luis Alberiko Gil-Alana

# US House Prices by Census Division: Persistence, Trends and Structural Breaks

December 2022

http://www.brunel.ac.uk/economics

## 1. Introduction

House prices are an important factor affecting the real economy as well as financial markets. Their key importance was shown very clearly by the 2007 sub-prime mortgage crisis in the US, which was mainly caused by a housing bubble that had started in the previous decade (see Shiller, 2007). The empirical literature aiming to shed light on their behaviour comprises two main strands. The first type of studies analyses their relationship with economic fundamentals. For instance, Capozza and Helsely (1989, 1990) provided evidence on the long-run equilibrium relationship between real house prices and real income. Caporale and Gil-Alana (2015) used long-range dependence techniques to examine the long-run linkages between the Housing Price Index (HPI) and Disposable Personal Income (PDI) in the US and showed that these and its autocorrelations decay at a rather slower hyperbolic rate; if $0 < d < 0.5$, the process is covariance stationary, and as long as $d < 1$ mean reversion will occur, even if the fractional parameter is in the non-stationary range; finally, $d = 1$ corresponds to the unit root case, and $d > 1$ to explosive behaviour. Papers modelling house prices using this method include Barros et al. (2012, 2015), Gil-Alana et al. (2013, 2014), and Gupta et al. (2014). However, all these studies focus on long-run persistence only and do not allow for possible breaks. More recently, Canarella et al. (2021) have instead used a fractional integration model including both a long-run and a cyclical component to analyse persistence in both US and UK house prices over a long time span, and have also tested for breaks. They find that long-run persistence plays a greater role, and that breaks occurred at different times in the two countries being examined (earlier in the US), which implies that national factors were their main drivers of house prices.

The present study belongs to the second strand of the literature on house prices, which carries out univariate analysis, and it also follows a fractional integration approach as in the more recent contributions mentioned above. However, unlike them, it provides evidence on US house price behaviour by geographical area. More specifically, it examines data for various Census Divisions. This is an important addition to the existing body of empirical literature, since there can obviously be significant differences between the housing markets of different areas of a country (the US, in our case) which are not captured by the aggregate price series, and thus different policy prescriptions might be appropriate in each case. The other issue addressed by our analysis is the possible presence of breaks in the series under examination, which is also of key importance to understand changes in the housing market which might have occurred as a result of a variety of factors (fundamentals or others), again with implications for the design of effective stabilisation policies.

## 3. Empirical Results

Table 1 displays the estimated values of d from the model given by equation (1) under the assumption that the error term, $u_t$, is a white noise process. Following the standard literature on unit roots (see, e,g., Bharghava, 1986; Schmidt and Phillips, 1992; etc.), we consider three specifications including respectively: (i) no deterministic terms, i.e. $\alpha = \beta = 0$ (see column 2 for the corresponding results); (ii) a constant only, i.e. $\beta = 0$ (see column 3); (iii) both a constant and a linear time trend, i.e. $\alpha \neq 0$ and $\beta \neq 0$ (see column 4). In all cases we display the estimates of d along with the 95% confidence bands; those in bold are from the preferred specification selected on the basis of the statistical significance of the regressors.

It can be seen that the coefficient on the time trend is significant for six out of the nine Census Divisions examined (i.e., in all cases except Mountain (M), Pacific (P) and South Atlantic (SA)), whilst it is insignificant for the US aggregate data. Table 2 displays the estimated regression coefficients for the selected specification in each case. The biggest ones on the time trend are found for West South Central (WSC, 0.761) and East South Central (ESC, 0.758); the estimated values of d are significantly higher than 1 in all cases, ranging from 1.24 (East South Central, ESC) and 1.25 (West North Central) to 1.53 (Mountain) and 1.55 (Pacific). For the US aggregate data, the time trend is insignificant and the order of integration is 1.70, much higher than for the individual Census Divisions, which is probably due to the aggregation effect on the degree of integration of the series (Robinson, 1978, Granger, 1980).

**TABLES 1 – 4 ABOUT HERE**

Tables 3 and 4 are similar to Tables 1 and 2 respecac than before (between 1.13 for ESC and 1.45 for P), and again higher for the aggregate series (d = 1.52), the unit root null hypothesis being rejected in all cases in favour of d > 1 – in other words, mean reversion does not occur in any single case, and thus shocks have permanent effects.

The results considered so far might be biased owing to the strong assumption that the residuals are a white noise process. Thus, in what follows, we allow for autocorrelation; in particular, rather than imposing a parametric ARMA model that would require specifying the correct AR and MA orders (which is not straightforward in the context of fractional integration, see Beran et al., 1998) we apply the non-parametric modelling approach of Bloomfield (1973), which

the first, third and fourth subsamples, while it cannot be rejected during the second and the last subsamples. All the time trend coefficients are significant, being positive in all subsamples except the third one going from May 2007 to October 2011. The highest coefficient on the time trend again corresponds to the last subsample.

As for the Middle Atlantic (MA) series, there are also four breaks and thus five subsamples. The estimates of d are between 0.74 (June 2007 – February 2012) and 1.22 (June 2020 – August 2022) and, as in other cases, the time trends are all significantly positive, except the third one for the period starting in June 2007. Once again the estimated time trend coefficient is significant and particularly high in the last subsample.

**TABLES 14c AND 14d ABOUT HERE**

Very similar results are obtained for New England, though now mean reversion (i.e., significant evidence of d smaller than 1) is found for the third and four subsamples (December 2005 – January 2012, February 2012 – May 2020) and a negative trend for the third subsample (December 2005 – January 2012). The positive trend coefficients are equal to 0.1538 for the first subsample; 1.4513 for the second one; 0.7330 for the fourth subsample, and 3.8984 for the final one starting in June 2020.

**TABLES 14e AND 14f ABOUT HERE**

In the case of the Mountain (M) series the results are slightly different: mean reversion is not found in any single case, and d is statistically higher than 1 in the second and last subsamples, in the latter case being insignificant. Five breaks are detected in the case of the Pacific (P) series; mean reversion does not occur in any subsample, and d is estimated to be much higher than 1, especially in the last subsample. The time trend is negative in the first subsample, positive in the second, third and fifth, and insignificant in the fourth and sixth.

**TABLES 14g AND 14h ABOUT HERE**

Regarding the South Atlantic (SA) series, breaks are detected in January 1998, April 2007, July 2011, and May 2020. Mean reversion occurs in the fourth subsample (from August 2011 to May 2022) and the time trend is insignificant in the last subsample. In the case of the West North Central (WNC) series, mean reversion takes place in the second (July 2007 – April 2011) and third (May 2011 – May 2020) subsamples, with a significant negative trend in the former. There are only two breaks (July 2011 and June 2020) in the West South Central (WSC) series; mean reversion occurs in the second subsample, and the time trend is significantly positive in all three subsamples.

**TABLES 14i AND 14j ABOUT HERE**

Finally, there are four breaks in the US aggregate series (January 1998, April 2007, August 2011 and May 2020), and no mean reversion in any single case. The time trend coefficients are all positive, although convergence cannot be achieved for the third subsample (May 2007 - August 2011), probably as a result of the small number of observations. In the other cases the time trend coefficient is significantly positive, again being particularly high in the last subsample.

## 4. Conclusions

This paper uses fractional integration methods to analyse the behaviour of US house prices, more specifically 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. The full sample estimates imply that the order of integration of the series is above 1 in all cases, and is particularly high for the aggregate series. However, when the possibility of structural breaks is taken into account, segmented trends are detected; the subsample estimates of the fractional differencing parameter tend to be lower, with mean reversion occurring in a number of cases, and the time trend coefficient being at its highest in the last subsample, which in most cases start around May 2020.

Cuestas J.C. and L.A. Gil-Alana (2016). “A Non-Linear Approach with Long Range Dependence Based on Chebyshev Polynomials”, Studies in Nonlinear Dynamics and Econometrics, 23, 445–468.

Gil-Alana, L.A. (2008) Fractional integration and structural breaks at unknown periods of time, Journal of Time Series Analysis 29, 1, 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.

Granger, C.W.J. (1980), Long memory reln/Subships 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.

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

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.

Himmelberg, C., Mayer, C. and T. Sinai (2005), Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions, Journal of Economic Perspectives 19(4): 67-92.

Holmes, M., and Grimes, A., (2008). Is There Long-run Convergence among Regional Meen, G. (1999), Regional house prices and the ripple effect: A new interpretn/Sub. Housing Studies, 14, 733-753.

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

Robibson, P. M. (1978)

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.

**Table 1: Estimates of the differencing parameter (original series). White noise disturbances**

| Series (original) | No terms | An intercept | An intercept and a linear time trend |
|---|---|---|---|
| EAST NORTH CENTRAL | 1.04 (0.99, 1.10) | 1.33 (1.29, 1.38) | **1.34 (1.30, 1.39)** |
| EAST SOUTH CENTRAL | 1.05 (0.99, 1.10) | 1.23 (1.20, 1.27) | **1.24 (1.20, 1.28)** |
| MIDDLE ATLANTIC | 1.04 (0.99, 1.10) | 1.30 (1.26, 1.35) | **1.31 (1.26, 1.36)** |
| MOUNTAIN | 1.10 (0.99, 1.10) | **1.53 (1.46, 1.62)** | 1.54 (1.47, 1.62) |
| NEW ENGLAND | 1.07 (0.99, 1.10) | 1.26 (1.21, 1.30) | **1.26 (1.22, 1.31)** |
| PACIFIC | 1.10 (0.99, 1.10) | **1.55 (1.50, 1.62)** | 1.56 (1.50, 1.62) |
| SOUTH ATLANTIC | 1.09 (1.04, 1.15) | **1.43 (1.38, 1.49)** | 1.43 (1.38, 1.49) |

**Table 5: Estimates of the differencing parameter (original series). Bloomfield disturbances**

| Series (original) | No terms | An intercept | An intercept and a linear time trend |
|---|---|---|---|
| EAST NORTH CENTRAL | 1.09 (1.01, 1.19) | **1.51 (1.42, 1.64)** | 1.54 (1.45, 1.68) |
| EAST SOUTH CENTRAL | 1.14 (1.06, 1.24) | | |

**Table 9: Estimates of the differencing parameter (original series). Seasonal AR(1) disturbances**

| Series (original) | No terms | An intercept | An intercept and a linear time trend |
|---|---|---|---|

**Table 13: Structural breaks in the series**

| Series (original) | N. of breaks | Break dates |
|---|---|---|

**Table 14a: Estimates for each subsample. East North Central**

| EAST NORTH CENTRAL | d (95% band) | Intercept (t-value) | Time trend (t-value) |
|---|---|---|---|
| January 1991 - April 2006 | 0.95 (0.90, 1.01) | 99.463 (276.89) | 0.5119 (24.52) |
| May 2006 - October 2011 | 0.90 (0.77, 1.10) | 193.605 (193.66) | -0.5820 (-5.94) |
| November 2011 - May 2020 | 0.99 (0.90, 1.12) | 157.294 (224.40) | 0.7586 (11.45) |
| June 2020 - August 2022 | 1.58 (1.00, 1.96) | 233.831 (181.58) | 3.0793 (2.63) |

In brackets: the 95% confidence bands in column 2, and the t-

**Table 14d: Estimates for each subsample. New England**

| NEW ENGLAND | d | (95% band) | Intercept |
| --- | --- | --- | --- |

**Table 14g: Estimates for each subsample. South Atlantic**

| SOUTH ATLANTIC | d (95% band) | Intercept (t-value) | Time trend (t-value) |
|---|---|---|---|
| January 1991- January 1998 | 0.83 (0.71, 1.01) | 99.704 (314.63) | 0.2359 (12.74) |
| February 1998 – | | | |

**Table 14**