Department of Economics and Finance

Working Paper No. 2220

Economics and Finance Working Paper Series

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

US HOUSE PRICES BY CENSUS DIVISION: PERSISTENCE, TRENDS AND STRUCTURAL BREAKS

Guglielmo Maria Caporale
Brunel University London, UK

Luis Alberiko Gil-Alana
University of Navarra, Pamplona, Spain
Universidad Francisco de Vitoria, Madrid, Spain

December 2022

Abstract

This paper uses fractional integration methods to examine persistence, trends and structural breaks in 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 starts around May 2020.

Keywords: US house prices; fractional integration; persistence; trends; structural breaks

JEL Classification: 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 the Grant PID2020-113691RB-I00 funded by MCIN/AEI/ 10.13039/501100011033, and from an internal Project from the Universidad Francisco de Vitoria.

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 two series diverged over time, which suggested the existence of a bubble.

The second category of papers focuses instead on the univariate properties of house prices. Earlier studies carried out 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, this approach, which is based on the dichotomy between I(0) stationary and I(1) non-stationary processes, imposes rather restrictive assumptions on the behaviour of the series of interest. A more general framework, which is more informative about properties such as persistence and mean reversion, is represented by fractional integration (see Granger, 1980; Granger and Joyeux, 1980 and Hosking, 1981). In this case, the series of interest is modelled as an I(d) process, with d being allowed to take any real value, including fractional ones. This approach encompasses a wide range of stochastic behaviours. More specifically, if $d = 0$, the process is said to exhibit "short memory", with its autocorrelations (if non-zero) decaying at an exponential rate; if $d > 0$, the process is characterised by "long memory", 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.

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

2. Data Description and Modelling Framework

We analyse monthly data on the Federal Housing Finance Agency (FHFA) House Price Index for Census Divisions and the US as a whole. The Census Divisions are East North Central (ENC), East South Central (ESC), Middle Atlantic (MA), Mountain (M), New England (NE), Pacific (P), South Atlantic (SA), West North Central (WNC), West South Central (WSC), and the sample period goes from January 1991 to August 2022. The series are not seasonally adjusted and have been obtained from https://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index-Datasets.aspx

The model is specified as follows:

$$y_t = \alpha + \beta t + x_t, \qquad (1 - B)^d x_t = u_t, \qquad u_t = \rho u_{t-12} + \varepsilon_t. \qquad (2)$$

where y_t stands for the series of interest, α and β denote the constant and the coefficient on a linear time trend respectively, B is the backshift operator, i.e., $Bx_t = x_{t-1}$, and u_t is a short-memory process which is integrated of order 0.

Note that d is allowed to take any real value, including fractional ones, and thus, as already mentioned, the chosen framework encompasses a wide range specifications, such as the classical trend stationary I(0) model if d = 0, the unit root case if d = 1, anti-persistence if d < 0, and long memory if d is positive and has a fractional value; in the latter case, if 0 < d < 0.5 the series is still covariance stationary, and if d < 1 mean reversion occurs.

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 respectively, but report the results for the logged transformed data. The time trend is now significantly positive in all cases except for Pacific (P) and for the aggregate data, and the estimates of d are slightly smaller 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 is based on a spectral density function, whose log form approximates well that of AR structures. Tables 5 and 6 report the corresponding results for the original data, while Tables 7 and 8 display those for the log-transformed ones. The time trend is now insignificant in every single case, the intercept being the only deterministic term required in the model. As for d, its estimated values are again significantly higher than 1 in all cases, ranging from 1.41 (New England, logged values) to 2.03 (Pacific, original data).

TABLES 5 – 12 ABOUT HERE

Finally, given the monthly frequency of the data, we allow for a seasonal AR(1) process in the error term. These results are reported in Tables 9 – 12 and are very similar to the previous ones obtained under the assumption of white noise errors: the time trend is not required for Mountain (M), Pacific (P) and South Atlantic (SA) as well as for the aggregate data (USA), and the degrees of integration are higher than 1 in all cases, their estimated values being bigger for the original data compared to the log-transformed ones.

The high degree of persistence implied by the estimated values of d reported in Tables 1 – 12 might be the result of misspecification due to the presence of structural breaks that have not been taken into account. In fact, given the long time span covered by the data, this is most likely to have occurred. Therefore we carry out the Bai and Perron's (2003) tests for multiple breaks as well as their version proposed for the fractional integration case by Gil-Alana (2008). The break dates detected by means of these two sets of tests were identical and are displayed in Table 13.

TABLE 13 ABOUT HERE

Two breaks are found in the case of West South Central (SWC); three in the case of East North Central (ENC) and West North Central (WNC); four in the majority of cases and five in the case of Pacific (P). A lot of the series exhibit breaks around April to June 2007, namely just before the US sub-prime mortgage crisis; April to August 2011, immediately before Operation Twist, when the Fed restructured its debt portfolio by selling short-term T-bills and buying long-term debt with the aim of flattening the yield curve and boosting the mortgage market as well as other forms of credit; and May-June 2020, following the end of the shortest US recession on record (which had been caused by the Covid-19 pandemic).

Tables 14a-14j displays the estimated coefficients for each series and each subsample. In the case of the East North Central (ENC) series, the estimates of d are now much smaller than when considering the whole sample. Thus, the unit root null hypothesis cannot be rejected in the first three subsamples (with the data ending in May 2020) and only for the last subsample (June 2020 – August 2022) the estimate of d is found to be much higher than 1. The time trend is positive in the first, second and especially in the last subsample, being negative for the time period between May 2006 and October 2011.

TABLES 14a AND 14b ABOUT HERE

Concerning the East South Central (ESC) series, the values of d are now even smaller. The unit root hypothesis is rejected in favour of mean reversion ($d < 1$) during 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.

On the whole, it is clear that there is heterogeneity between housing markets in different geographical areas of the US, which might reflect differences in the number of buyers and sellers in each case as well as other local factors; these cannot be captured by the aggregate series, and thus it is important to obtain evidence for the various Census Divisions as well. In particular, the individual series are found to be less persistent than the aggregate one, and also to be subject to structural change. The detected breaks appear to correspond to well-known economic and policy developments (such as the sub-prime mortgage crisis, changes in the Fed's debt portfolio, and the rebound after the early stages of the Covid-19 pandemic). House price persistence is transmitted to other macroeconomic and financial variables, and in particular can affect inflation persistence. Therefore accurate information on persistence is crucial for policy decisions, with different policy measures having to be taken in response to shocks depending on the degree of persistence (see Himmelberg et al., 2005). The present study offers thorough evidence on this property for US house prices in different geographical areas and time periods, and thus has important policy implications, in particular for crisis management and/or prevention.

Future research could allow for non-linearities in house prices. One possible approach would be based on Chebyshev's polynomials, which do not produce abrupt changes in the series (unlike models with structural breaks), and can be easily used in the context of fractional integration (see, e..g., Cuestas and Gil-Alana, 2016). An alternative framework would include non-linear (deterministic) trends based on Fourier transforms (Gil-Alana and Yaya, 2021), or neural networks (Yaya et al., 2021).

References

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

Bai, J. and P. Perron (2003), Computation and analysis of multiple structural change models, Journal of Applied Econometrics 18, 1, 1-22.

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.

Baxter, M., and King, R. M. (1999). Measuring business cycles: Approximate band-
Bhargava, A., (1986), On the theory of testing for unit roots in observed time series, Review of Economics and Studies 53, 3,369-384.

Beran, J., R.J. Bhansali and D. Ocker (1998), On unified model selection for stationary and nonstationary short- and long-memory autoregressive processes, Biometrika 85, 4, 921-934.

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

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

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. 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.

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.

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.

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 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.

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 interpretation. Housing Studies, 14, 733-753.

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

Robinson, P. M. (1978) Statistical inference for a randomcoefficient autoregressive model. Scandinavian Journal of Statistics 5, 163–168

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.

Yaya, O., Ogbonna, A.E., Gil-Alana, L.A. and Furuoka, F. (2021). A new unit root analysis for testing hysteresis in unemployment, Oxford Bulletin of Economics and Statistics 83(4), 960-981.

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)
WEST NORTH CENTRAL	1.04 (0.99, 1.11)	1.24 (1.20, 1.29)	**1.25 (1.21, 1.29)**
WEST SOUTH CENTRAL	1.05 (1.00, 1.11)	1.30 (1.26, 1.36)	**1.32 (1.27, 1.37)**
USA	1.06 (1.01, 1.12)	**1.70 (1.61, 1.80)**	1.70 (1.62, 1.80)

Values in bold are those from the selected specification for each series; in brackets the 95% confidence bands.

Table 2: Estimated coefficients from the selected model for each original series. White noise disturbances

Series (original)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	1.34 (1.30, 1.39)	99.407 (121.43)	0.593 (2.21)
EAST SOUTH CENTRAL	1.24 (1.20, 1.28)	99.430 (78.24)	0.758 (3.07)
MIDDLE ATLANTIC	1.31 (1.26, 1.36)	99.668 (96.34)	0.563 (1.93)
MOUNTAIN	1.53 (1.46, 1.62)	100.274 (61.51)	---
NEW ENGLAND	1.26 (1.22, 1.31)	99.278 (67.62)	0.715 (2.25)
PACIFIC	1.55 (1.50, 1.62)	99.839 (82.55)	---
SOUTH ATLANTIC	1.43 (1.38, 1.49)	99.831 (87.45)	---
WEST NORTH CENTRAL	1.25 (1.21, 1.29)	99.482 (90.00)	0.734 (3.24)
WEST SOUTH CENTRAL	1.32 (1.27, 1.37)	99.620 (92.60)	0.761 (2.39)
USA	1.70 (1.61, 1.80)	99.808 (174.67)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

.

Table 3: Estimates of the differencing parameter (logged series). White noise disturbances

Series (in logs)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	0.99 (0.93, 1.07)	1.25 (1.22, 1.30)	**1.25 (1.22, 1.30)**
EAST SOUTH CENTRAL	0.99 (0.93, 1.07)	1.12 (1.09, 1.17)	**1.13 (1.09, 1.17)**
MIDDLE ATLANTIC	0.99 (0.93, 1.07)	1.24 (1.20, 1.28)	**1.24 (1.20, 1.29)**
MOUNTAIN	0.99 (0.93, 1.07)	1.36 (1.31, 1.41)	**1.36 (1.31, 1.42)**
NEW ENGLAND	0.99 (0.93, 1.07)	1.19 (1.15, 1.24)	**1.20 (1.16, 1.24)**
PACIFIC	0.99 (0.93, 1.07)	**1.45 (1.41, 1.50)**	1.46 (1.42, 1.51)
SOUTH ATLANTIC	0.99 (0.93, 1.07)	1.33 (1.29, 1.38)	**1.34 (1.30, 1.39)**
WEST NORTH CENTRAL	0.99 (0.93, 1.07)	1.16 (1.12, 1.20)	**1.16 (1.12, 1.20)**
WEST SOUTH CENTRAL	0.99 (0.92, 1.07)	1.17 (1.13, 1.21)	**1.17 (1.13, 1.22)**
USA	0.99 (0.92, 1.07)	**1.52 (1.47, 1.59)**	1.52 (1.47, 1.59)

Values in bold are those from the selected specification for each series; in brackets the 95% confidence bands.

Table 4: Estimated coefficients from the selected model for each logged series. White noise disturbances

Series (in logs)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	1.25 (1.22, 1.30)	4.601 (1022.48)	0.0033 (3.62)
EAST SOUTH CENTRAL	1.13 (1.09, 1.17)	4.601 (741.05)	0.0036 (5.48)
MIDDLE ATLANTIC	1.24 (1.20, 1.29)	4.603 (867.52)	0.0030 (2.91)
MOUNTAIN	1.36 (1.31, 1.42)	4.604 (759.88)	0.0042 (1.93)
NEW ENGLAND	1.20 (1.16, 1.24)	4.600 (626.40)	0.0034 (2.96)
PACIFIC	1.45 (1.41, 1.50)	4.603 (853.90)	---
SOUTH ATLANTIC	1.34 (1.30, 1.39)	4.601 (866.00)	0.0036 (2.20)
WEST NORTH CENTRAL	1.16 (1.12, 1.20)	4.602 (895.61)	0.0036 (5.55)
WEST SOUTH CENTRAL	1.17 (1.13, 1.22)	4.602 (968.32)	0.0037 (5.96)
USA	1.52 (1.47, 1.59)	4.603 (1685.17)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

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)	**1.59 (1.50, 1.73)**	1.61 (1.51, 1.77)
MIDDLE ATLANTIC	1.10 (1.02, 1.20)	**1.56 (1.46, 1.78)**	1.58 (1.46, 1.81)
MOUNTAIN	1.22 (1.16, 1.30)	**1.90 (1.55, 2.34)**	1.86 (1.56, 2.31)
NEW ENGLAND	1.08 (1.01, 1.18)	**1.45 (1.38, 1.55)**	1.48 (1.39, 1.58)
PACIFIC	1.18 (1.11, 1.27)	**2.03 (1.82, 2.23)**	2.06 (1.85, 2.24)
SOUTH ATLANTIC	1.18 (1.11, 1.28)	**1.65 (1.52, 1.90)**	1.68 (1.54, 1.94)
WEST NORTH CENTRAL	1.09 (1.01, 1.19)	**1.49 (1.41, 1.61)**	1.52 (1.43, 1.66)
WEST SOUTH CENTRAL	1.12 (1.04, 1.21)	**1.51 (1.41, 1.65)**	1.54 (1.44, 1.66)
USA	1.09 (1.04, 1.15)	**1.63 (1.51, 1.91)**	1.68 (1.52, 1.96)

Values in bold are those from the selected specification for each series; in brackets the 95% confidence bands.

Table 6: Estimated coefficients from the selected model for each original series. Bloomfield disturbances

Series (original)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	1.51 (1.42, 1.64)	99.641 (136.09)	---
EAST SOUTH CENTRAL	1.59 (1.50, 1.73)	99.765 (98.64)	---
MIDDLE ATLANTIC	1.56 (1.46, 1.78)	99.914 (110.86)	---
MOUNTAIN	1.90 (1.55, 2.34)	100.53 (73.15)	---
NEW ENGLAND	1.45 (1.38, 1.55)	99.385 (75.91)	---
PACIFIC	2.03 (1.82, 2.23)	99.722 (98.56)	---
SOUTH ATLANTIC	1.65 (1.52, 1.90)	99.791 (98.56)	---
WEST NORTH CENTRAL	1.49 (1.41, 1.61)	99.840 (105.16)	---
WEST SOUTH CENTRAL	1.51 (1.41, 1.65)	100.08 (105.26)	---
USA	1.63 (1.51, 1.91)	99.818 (169.44)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 7: Estimates of the differencing parameter (logged series). Bloomfield disturbances

Series (in logs)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	0.98 (0.88, 1.12)	**1.46 (1.39, 1.55)**	1.47 (1.39, 1.58)
EAST SOUTH CENTRAL	0.98 (0.89, 1.12)	**1.48 (1.40, 1.56)**	1.49 (1.41, 1.59)
MIDDLE ATLANTIC	0.98 (0.90, 1.12)	**1.53 (1.45, 1.66)**	1.55 (1.47, 1.68)
MOUNTAIN	0.98 (0.89, 1.11)	**1.67 (1.53, 1.86)**	1.66 (1.54, 1.80)
NEW ENGLAND	0.98 (0.88, 1.10)	**1.41 (1.34, 1.50)**	1.42 (1.35, 1.52)
PACIFIC	0.98 (0.89, 1.11)	**1.94 (1.80, 2.09)**	1.97 (1.83, 2.14)
SOUTH ATLANTIC	0.98 (0.90, 1.12)	**1.60 (1.51, 1.73)**	1.63 (1.52, 1.77)
WEST NORTH CENTRAL	0.97 (0.88, 1.12)	**1.45 (1.38, 1.55)**	1.46 (1.37, 1.55)
WEST SOUTH CENTRAL	0.99 (0.88, 1.12)	**1.45 (1.36, 1.57)**	1.47 (1.38, 1.58)
USA	0.99 (0.88, 1.12)	**1.62 (1.51, 1.76)**	1.64 (1.53, 1.82)

Values in bold are those from the selected specification for each series; in brackets the 95% confidence bands.

Table 8: Estimated coefficients from the selected model for each logged series. Bloomfield disturbances

Series (in logs)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	1.46 (1.39, 1.55)	4.601 (1143.88)	---
EAST SOUTH CENTRAL	1.48 (1.40, 1.56)	4.601 (865.95)	---
MIDDLE ATLANTIC	1.53 (1.45, 1.66)	4.604 (1036.23)	---
MOUNTAIN	1.67 (1.53, 1.86)	4.608 (898.17)	---
NEW ENGLAND	1.41 (1.34, 1.50)	4.599 (702.80)	---
PACIFIC	1.94 (1.80, 2.09)	4.602 (1155.50)	---
SOUTH ATLANTIC	1.60 (1.51, 1.73)	4.603 (1015.44)	---
WEST NORTH CENTRAL	1.45 (1.38, 1.55)	4.603 (1030.14)	---
WEST SOUTH CENTRAL	1.45 (1.36, 1.57)	4.605 (1118.43)	---
USA	1.62 (1.51, 1.76)	4.603 (1771.70)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

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
EAST NORTH CENTRAL	1.03 (0.97, 1.10)	1.37 (1.32, 1.42)	**1.38 (1.33, 1.43)**
EAST SOUTH CENTRAL	1.05 (0.99, 1.12)	1.23 (1.19, 1.28)	**1.24 (1.20, 1.28)**
MIDDLE ATLANTIC	1.03 (0.97, 1.10)	1.32 (1.28, 1.38)	**1.33 (1.28, 1.39)**
MOUNTAIN	1.09 (1.04, 1.15)	**1.56 (1.48, 1.65)**	1.56 (1.49, 1.65)
NEW ENGLAND	1.08 (1.03, 1.14)	1.26 (1.21, 1.31)	**1.26 (1.22, 1.32)**
PACIFIC	1.09 (1.04, 1.15)	**1.57 (1.52, 1.64)**	1.58 (1.52, 1.64)
SOUTH ATLANTIC	1.08 (1.03, 1.14)	**1.45 (1.40, 1.51)**	1.46 (1.41, 1.52)
WEST NORTH CENTRAL	1.04 (0.98, 1.11)	1.25 (1.21, 1.30)	**1.26 (1.22, 1.31)**
WEST SOUTH CENTRAL	1.05 (0.99, 1.11)	1.32 (1.28, 1.37)	**1.33 (1.29, 1.38)**
USA	1.13 (1.05, 1.23)	**1.63 (1.51, 1.91)**	1.68 (1.52, 1.97)

Values in bold are those from the selected specification for each series; in brackets the 95% confidence bands.

Table 10: Estimated coefficients from the selected model for each original series. Sesonal AR(1) disturbances

Series (original)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	1.38 (1.33, 1.43)	99.401 (121.5)	0.582 (1.79)
EAST SOUTH CENTRAL	1.24 (1.20, 1.28)	99.430 (78.24)	0.785 (3.07)
MIDDLE ATLANTIC	1.33 (1.28, 1.39)	99.676 (96.68)	0.545 (1.79)
MOUNTAIN	1.56 (1.48, 1.65)	100.298 (62.24)	---
NEW ENGLAND	1.26 (1.22, 1.32)	99.278 (67.62)	0.715 (2.25)
PACIFIC	1.57 (1.52, 1.64)		---
SOUTH ATLANTIC	1.45 (1.40, 1.51)	99.834 (83.20)	---
WEST NORTH CENTRAL	1.26 (1.22, 1.31)	99.845 (90.19)	0.731 (3.06)
WEST SOUTH CENTRAL	1.33 (1.29, 1.38)	99.630 (92.79)	0.751 (2.24)
USA	1.63 (1.51, 1.91)	99.818 (169.44)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 11: Estimates of the differencing parameter (logged series). Seasonal AR(1) disturbances

Series (in logs)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	0.99 (0.90, 1.07)	1.30 (1.25, 1.34)	**1.29 (1.25, 1.34)**
EAST SOUTH CENTRAL	1.00 (0.91, 1.08)	1.13 (1.09, 1.17)	**1.13 (1.09, 1.17)**
MIDDLE ATLANTIC	0.99 (0.90, 1.07)	1.25 (1.21, 1.30)	**1.26 (1.22, 1.31)**
MOUNTAIN	0.98 (0,.90, 1.07)	1.38 (1.33, 1.43)	**1.38 (1.33, 1.44)**
NEW ENGLAND	0.99 (0.91, 1.07)	1.19 (1.15, 1.24)	**1.20 (1.15, 1.25)**
PACIFIC	0.99 (0.91 1.07)	**1.47 (1.42, 1.52)**	1.47 (1.43, 1.52)
SOUTH ATLANTIC	0.99 (0.91, 1.07)	1.35 (1.30, 1.40)	**1.35 (1.31, 1.40)**
WEST NORTH CENTRAL	0.99 (0.91, 1.08)	1.18 (1.14, 1.22)	**1.18 (1.14, 1.22)**
WEST SOUTH CENTRAL	0.98 (0.90, 1.07)	1.19 (1.14, 1.23)	**1.19 (1.15, 1.24)**
USA	0.97 (0.86, 1.10)	**1.61 (1.51, 1.76)**	1.64 (1.53, 1.82)

Values in bold are those from the selected specification for each series; in brackets the 95% confidence bands.

Table 12: Estimated coefficients from the selected model for each logged series. Sesonal AR(1) disturbances

Series (in logs)	No terms	An intercept	An intercept and a linear time trend
EAST NORTH CENTRAL	1.29 (1.25, 1.34)	4.600 (1029.49)	0.0034 (3.01)
EAST SOUTH CENTRAL	1.13 (1.09, 1.17)	4.601 (741.00)	0.0036 (5.48)
MIDDLE ATLANTIC	1.26 (1.22, 1.31)	4.603 (870.90)	0.0029 (2.57)
MOUNTAIN	1.38 (1.33, 1.44)	4.604 (761.84)	0.0041 (1.78)
NEW ENGLAND	1.20 (1.15, 1.25)	4.600 (626.04)	0.0034 (2.96)
PACIFIC	1.47 (1.42, 1.52)	4.603 (856.91)	---
SOUTH ATLANTIC	1.35 (1.31, 1.40)	4.601 (869.11)	0.0036 (1.98)
WEST NORTH CENTRAL	1.18 (1.14, 1.22)	4.602 (898.31)	0.0035 (4.99)
WEST SOUTH CENTRAL	1.19 (1.15, 1.24)	4.602 (971.82)	0.0037 (5.32)
USA	1.61 (1.51, 1.76)	4.603 (1756.65)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4

Table 13: Structural breaks in the series

Series (original)	N. of breaks	Break dates
EAST NORTH CENTRAL	3	April 2006; October 2011; May 2020
EAST SOUTH CENTRAL	4	November 2004; April 2007; October 2011; May 2020
MIDDLE ATLANTIC	4	June 1999; May 2007; February 2012; May 2020
MOUNTAIN	4	September 2003; February 2007; June 2011; June 2020
NEW ENGLAND	4	January 1999; November 2005; January 2012; May 2020
PACIFIC	5	February 1997; October 2003; October 2006; February 2012; May 2020
SOUTH ATLANTIC	4	January 1998; April 2007; July 2011; May 2020
WEST NORTH CENTRAL	3	June 2007; April 2011; May 2020
WEST SOUTH CENTRAL	2	July 2011; June 2020
USA	4	January 1998; April 2007; August 2011; May 2020

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-statistics for the estimated coefficients in columns 3 and 4.

Table 14b: Estimates for each subsample. East South Central

EAST SOUTH CENTRAL	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991- November 2004	0.78 (0.71, 0.87)	99.326 (229.70)	0.4240 (33.70)
December 2004 – April 2007	0.85 (0.65, 1.15)	172.238 (314.87)	0.9586 (14.40)
May 2007 - October 2011	0.38 (0.16, 0.65)	200.725 (276.07)	-0.4421 (-19.25)
November 2011 - May 2020	0.77 (0.71, 0.85)	177.178 (124.47)	0.8142 (14.40)
June 2020 - August 2022	1.38 (0.96, 1.77)	261.502 (148.12)	3.6283 (3.54)

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14c: Estimates for each subsample. Middle Atlantic

MIDDLE ATLANTIC	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991- June 1999	1.02 (0.93, 1.13)	99.873 (201.33)	0.1395 (2.54)
July 1999 – May 2007	1.15 (1.07, 1.25)	113.332 (130.25)	1.0491 (6.26)
June 2007 - February 2012	0.74 (0.62, 0.96)	218.626 (217.88)	-0.5227 (-9.31)
March 2012 - May 2020	0.93 (0.86, 1.02)	190.441 (206.96)	0.6136 (8.85)
June 2020 - August 2022	1.22 (0.37, 1.86)	248.996 (128.65)	3.1348 (4.38)

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14d: Estimates for each subsample. New England

NEW ENGLAND	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991- January 1999	1.00 (0.91, 1.12)	99.846 (119.59)	0.1538 (1.81)
February 1999 - November 2005	0.98 (0.89, 1.11)	114.514 (150.64)	1.4513 (18.72)
December 2005 - January 2012	0.74 (0.62, 0.91)	232.421 (166.55)	-0.5459 (-8.51)
February 2012 - May 2020	0.80 (0.72, 0.90)	192.446 (157.61)	0.7330 (13.28)
June 2020 - August 2022	1.18 (0.59, 1.68)	264.817 (91.77)	3.8984 (4.12)

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14e: Estimates for each subsample. Mountain

MOUNTAIN	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991- September 2003	1.01 (0.94, 1.08)	99.359 (206.70)	0.6591 (16.14)
October 2003 – February 2007	1.30 (1.17, 1.47)	199.963 (169.95)	1.9967 (3.98)
March 2007 - June 2011	1.07 (0.91, 1.31)	290.647 (140.34)	-1.7430 (-4.76)
July 2011 - June 2020	0.87 (0.78, 1.00)	194.904 (127.57)	1.8698 (21.67)
July 2020 - August 2022	1.94 (1.62, 2.24)	396.863 (150.99)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14f: Estimates for each subsample. Pacific

PACIFIC	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991- February 1997	0.92 (0.81, 1.08)	100.091 (299.37)	-0.0554 (-1.93)
March 1997 – October 2003	1.25 (1.17, 1.34)	95.902 (162.03)	1.1386 (6.31)
November 2003 - October 2006	1.55 (1.43, 1.72)	188.256 (153.80)	2.5562 (2.45)
November 2006 - February 2012	1.42 (1.33, 1.55)	276.852 (161.52)	---
March 2012 - May 2020	0.99 (0.89, 1.13)	170.539 (85.82)	1.5874 (8.28)
June 2020 – August 2022	1.91 (1.65, 2.21)	171.434 (125.23)	---

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

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 – April 2007	1.38 (1.33, 1.46)	119.982 (208.65)	0.8504 (3.13)
May 2007 - July 2011	1.02 (0.90, 1.19)	241.736 (134.07)	-1.3327 (-4.91)
August 2011 - May 2020	0.86 (0.78, 0.97)	173325 (179.51)	1.0935 (20.73)
June 2020 - August 2022	1.73 (1.29, 2.13)	290.733 (151.35)	--

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14h: Estimates for each subsample. West North Central

WEST NORTH CENTRAL	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991 - June 2007	1.08 (1.03, 1.15)	99.514 (196.79)	0.5614 (10.55)
July 2007 - April 2011	0.64 (0.45, 0.89)	213.457 (166.87)	-0.5291 (-7.99)
May 2011 - May 2020	0.80 (0.74, 0.87)	186.942 (181.16)	0.8551 (19.47)
June 2020 - August 2022	1.37 (0.75, 1.83)	279.807 (158.86)	3.2519 (3.29)

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14i: Estimates for each subsample. West South Central

WEST SOUTH CENTRAL	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991 - July 2011	1.07 (1.02, 1.13)	99.704 (136.75)	0.3507 (5.28)
August 2011 - June 2020	0.64 (0.56, 0.74)	186.911 (231.33)	1.0217 (50.09)
July 2020 - August 2022	1.50 (0.99, 1.93)	300.953 (133.65)	3.4651 (1.97)

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4.

Table 14j: Estimates for each subsample. USA

USA	d (95% band)	Intercept (t-value)	Time trend (t-value)
January 1991- January 1998	1.07 (0.90, 1.32)	99.756 (498.92)	0.2508 (8.70)
February 1998 – April 2007	1.47 (1.40, 1.56)	121.311 (389.67)	0.7091 (3.54)
May 2007 - August 2011	---	---	---
September 2011 - May 2020	1.16 (1.05, 1.33)	176.826 (338.61)	0.9780 (9.68)
June 2020 - August 2022	2.22 (1.81, 2.64)	283.332 (228.63)	3.4094 (1.79)

In brackets: the 95% confidence bands in column 2, and the t-statistics for the estimated coefficients in columns 3 and 4. Convergence was not achieved in the case of the third sub-sample, probably due to the small number of observations.