Journal of Economics and Finance (2023) 47:63–93
https://doi.org/10.1007/s12197-022-09598-4

Non-linear structures, chaos, and bubbles in U.S. regional housing markets

Benjamas Jirasakuldech[1] · Riza Emekter[2] · Thuy Bui[1]

Accepted: 19 July 2022 / Published online: 17 September 2022

Abstract

This study analyzes the nonlinear price pattern and its underlying source of nonlinearity for U.S. housing markets along with the plausible explanations of chaos and bubble-like characteristics during 1987 to 2019. The results from the BDS test show evidence of nonlinear dependence in overall U.S. housing markets along with home markets in twenty cities. The K-map Z-map analysis shows that nonlinear dependence in all cities is consistent with chaotic behavior. The nonlinear dependence is also substantiated with the use of Markov chain test where nonlinearity is due to the persistence of either positive or negative returns. Applying the duration dependence test on positive runs confirms that housing markets in all five regions experience some episodes of bubbles, except for home markets in Detroit and Minneapolis in Midwest region. A time reversibility test further provides supporting evidence that the mechanism generating nonlinear dependence in housing markets in all four cities in Midwest region comes from non-Gaussian innovations. Similar finding is reported in housing markets in other regions including Atlanta, Charlotte, Dallas, San Diego, and San Francisco, suggesting that a linear function with non-Gaussian error terms is appropriate for modelling these housing markets.

Keywords Nonlinear dependence · Chaos · Bubbles · Non-Gaussian innovation · And Housing markets

JEL Classification G 14 · G 15

✉ Benjamas Jirasakuldech
b.jirasakuldech@sru.edu

Riza Emekter
emekter@rmu.edu

Thuy Bui
thuy.bui@sru.edu

[1] School of Business, Slippery Rock University, 1 Morrow Way, Slippery Rock, PA 16057, USA

[2] Department of Finance, Robert Morris University, 6001 University Boulevard, Moon Township, PA 15108, USA

1 Introduction

Prior studies in housing market have shown the significant role of housing assets in the capital allocation (i.e., Mill 1989; Cannon et al 2006). Understanding the house price behavior is undoubtedly crucial for successful modeling of asset prices. Furthermore, the correct identification of true nature of data-generating process for house prices will enable policymakers to design the appropriate forecasting model for housing market which signals information about future movements in economics activity. Such appropriate identification process could lead to a better policy control. The literature in non-linear dynamics of house price is not to great extent compared to nonlinearity in security prices (Hsieh 1991; Lim and Brooks 2011; and Caraiani 2014). While existing studies on house pricing predominantly focused on using linear framework, the non-linear dynamics property in housing markets is of critical concern because it is indicative of market inefficiency and possible presence of housing bubbles. If the housing markets were indeed characterized by nonlinear property, using a linear model to forecast housing price may result in an incorrect house pricing and poor forecasting performance. Recent attempts that address the issue of nonlinearity in housing markets do not provide a clear understanding of the nature of non-linear dependence driving the home prices. In general, the nonlinearity can stem from either nonlinearity in functional form or non-Gaussian innovations. Understanding the source of asymmetry has become increasingly more important in designing a successful forecasting model. The non-Gaussian model may become a correct approach if the nonlinearity in housing markets was caused by non-Gaussian innovations as opposed to a true nonlinearity in the model.

There are many reasons that lead us to believe that U.S. housing markets may exhibit non-linear price pattern that might be consistent with bubbles. First, unlike other financial markets, housing markets possess unique characteristics such as a lack of quality information disclosure, high transaction costs, low liquidity, limited house pricing transparency, very rigid supply side and short-trading limitation, making this housing market different and more prone to exhibit non-linear price behavior and bubbles (Herring and Wachter 2002). Second, housing prices in the U.S. adjust more quickly to positive market events leading to an increase in equilibrium price rather than negative market events causing a price decline particularly during the Great Recession (Kim and Bhattacharya 2009). Third, a huge lump sum of transaction costs and the lack of short trading in the housing markets can cause a few arbitrage opportunities arising from large deviations between house prices and fundamentals. Such departure of house prices from fundamental values could create a nonlinear behavior and a bubble (Rosenthal 1999; Muellbauer and Murphy 1997). Fourth, housing market is unique because it does not clear instantly after economic shocks. It takes some time for buyers and sellers to locate each other and for the suppliers to construct new buildings to meet the demand, which can lead to prices discontinuity and non-linearity.

Motivated by the importance of U.S. housing market as an investible asset and its unique characteristics, this study comprehensively analyzes the nonlinear behavior and its underlying source of nonlinearity for aggregate U.S. house prices along with plausible explanations for chaos and bubble-like characteristics during 1987 to 2019 period. We extend our analysis of nonlinearity to housing markets in five regions including Northeast, Midwest, Southeast, Southwest, and Western using housing data across 20 different cities. Typically, house prices react more to the regional economic and demographic shocks rather than nationwide shocks. The focus on regional housing markets allows us to compare the housing dynamics across regions.

This study analyzes the following four research questions systematically. First, do the U.S. housing markets exhibit some forms of dependency? Second, if the dependency persists, is it due to linear, non-linear, or chaotic behavior? In addition, is the nonlinearity in housing markets caused by nonlinearity in the return series or by movement in macroeconomic variables? Third, is the nonlinear dependency consistent with the characteristic of a bubble? Lastly, is the nonlinearity in housing markets caused by asymmetric nonlinear behavior in the functional form or asymmetric innovations? To derive the consensus view of nonlinearity in housing markets, this paper utilizes a variety of techniques, including the BDS test, the chaos test, the Markov chain test, the duration dependence test for bubbles, and the time reversibility test for the source of nonlinearity.

The contribution of this study in discovering true underlying nature of data-generating process for house prices will improve the forecasting performance. If the house price movement indeed reflects non-linear adjustment, employing a linear model to forecast house price movement will generate an inaccurate and inefficient forecast of house prices, which in turn resulting in wrong prediction of economy as house prices typically signal the stage of real economic activities. The finding of true cause of asymmetric price pattern helps policymakers and investors to better design a forecasting model that accounts for nonlinearity possibly induced by non-Gaussian innovations as opposed to nonlinearity in a functional form. The most notable contribution of our research provides insights as to housing markets in which regional areas are more likely to exhibit asymmetry and prone to experience bubbles. The housing price cycles would help the policymakers to recognize housing imbalances in each region as the developments in housing markets have the major impact on the local economy.

Our results based on a BDS test report evidence of nonlinear dependence in the aggregate market along with housing markets in 20 cities. The K-map Z-map analysis shows that nonlinear dependency in housing markets in all cities is consistent with the chaotic behavior. Furthermore, Markov Chain test reports the persistence of either positive or negative returns in housing markets. Applying the duration dependence test on positive runs of housing returns confirms that housing markets in all cities experience some episodes of bubbles, except Detroit, MI, and Minneapolis, MN in Midwest region. A time reversibility test further reveals that the mechanism generating nonlinear dependence in the housing markets in all four cities in Midwest region comes from the non-Gaussian innovations. Similar finding is also reported for housing markets in Atlanta, GA, Charlotte, NC, Dallas, TX, and San Diego and San Francisco, CA, suggesting that a linear function with non-Gaussian error terms is appropriate for modelling these housing markets.

The remainder of the paper is organized as follows. Section 2 provides the literature review for nonlinear dependency and bubbles in housing markets. Section 3 describes our data and summarizes the descriptive statistics of the U.S. Case-Shiller housing index for national and 20 cities. Section 4 presents the descriptions and empirical results of the BDS, K-map and Z-map, Markov Chain, Duration Dependence, and Time Reversibility methodologies in details. Section 5 offers some concluding remarks.

2 Literature review

The literature in non-linear price behavior and bubbles in housing markets is relatively scant. Genesove and Mayer (2001) attributed the nonlinear behavior in housing prices in Boston to the sellers' loss aversion. Seslen (2004) showed that households behave rational and react differently in response to the upturn and downturn in housing markets causing nonlinear behavior. Similar findings are reported by Kim and Bhattacharya (2009) where nonlinear in housing market is due to the asymmetric response to different phases of expansion and contraction. Muellbauer and Murphy (1997) also suggested that a huge lump sum transaction cost can cause nonlinearity and the appreciation of house prices. Study by Enders and Siklos (2001) showed that nonlinearity in house prices is due to asymmetric adjustment in the underlying determinants of housing value such as GDP and interest rates. This is confirmed by Skalin and Teräsvirta (2002) who showed that nonlinearity in housing price is driven by the nonlinearities in underlying macroeconomic variables. Nonlinearity in housing price is also substantiated by the superior forecasting performance of nonlinear techniques (Miles 2008; Trindade et al. 2010).

There is ample research in nonlinear price behavior that supports characteristics of housing market bubbles in the U.S. and other countries (Himmelberg et al. 2005; and McCarthy and Peach 2004). Shiller (2005) showed that U.S. housing market experienced bubbles. Belke and Marcel (2005) found that housing markets in 15 states showed the largest price increase consistent with bubble characteristics. Mikhed and Zemcik (2007) revealed that U.S. house prices and rents are not cointegrated, indicating the presence of a housing bubble. Huang (2013) showed possible existence of a housing bubble during 2001 to 2004 due to violation of relationship between housing volatility and realized returns. Hott (2012) further showed that investor's herding behavior causes house prices in European and non-European OECD countries to fluctuate more than fundamentally justified, suggesting evidence of a bubble. Engsted et al. (2016) reported that 18 OECD countries except Germany and Italy experienced explosive housing market bubbles. Shi et al. (2016) found the housing bubbles in Australia before the 2008 global financial crisis. Besarria et al. (2018) implemented cointegration method between the house rental prices and selling prices and reported housing bubbles in Brazil. Asal (2019) compared long-run price with actual price of housing in Sweden and reported evidence of housing bubble in the early 2004.

In summary, the extant literature provides several pieces of evidence supporting the nonlinearity and bubbles in housing markets in U.S. and around the world. However, to the best of our knowledge, none of previous studies provides a comprehensive analysis of asymmetric price behavior that could lead to the possible chaos and price bubble in the U.S. housing market. The next section will describe our data and empirical approaches.

3 Data

This study uses monthly seasonally adjusted data of S&P Case-Shiller U.S. Home Price Index for national and 20 cities to form five regional housing markets of Midwest, Northeast, Southeast, Southwest, and Western regions. The time period studied is from January 1978 to March 2019 except the home price index for Detroit (1991:01), Atlanta (1991:01), Dallas (2000:02), Denver (2001:01), and Seattle (1990:01). The data are obtained from the Federal Reserve Bank of Saint Louis: https://fred.stlouisfed.org/. The continuously compounded monthly return of each housing index series is computed as, $\boldsymbol{R_t} = 100 * \boldsymbol{ln}(\boldsymbol{P_t}/\boldsymbol{P_{t-1}})$, where P_t represents the housing index value at the end of the month t and P_{t-1} is the prior month closing value. We also used four macroeconomic variables including economic conditions index (EC), interest rate (INT), inflation rate (INF), and U.S. population growth (PopGrowth) in the K-map and Z-map analysis. The casual relationship of these macroeconomic variables and housing markets behavior has been well documented by many studies (ie., Hepsen and Kalfa 2009; Leung and Ng 2019; Gallagher 2019; and Maynou et al. 2021).

The summary statistics for monthly U.S. aggregate housing returns and 20 cities are reported in Table 1. The housing market in Portland, OR generates the highest average returns of 0.6544%, while the housing market in Cleveland, OH has the lowest return of 0.2170%. The most volatile housing market is Las Vegas, NV which shows the highest variance of 1.2614%, while the house price changes of Charlotte, NC exhibit the least volatility of 0.4431%. All the housing markets exhibit a departure from normality as shown by significant skewness, kurtosis, and Jarque–Bera test statistics at the 1% significant level, except for New York. This is the first descriptive indication of the serial dependence in the higher moments in the housing returns in most of the cities and regions. Next section will examine this serial dependence in empirical settings.

4 Methodology and empirical results

We empirically test whether there exists some form of dependency in the return behavior of U.S. housing markets and such dependency can be explained by nonlinear, chaotic dynamics or bubble-like structures. First, the Brock-Dechert-Scheinkman (BDS) test (Brock et al. (1996)) is implemented to check for any dependency in the house prices for each city. To further uncover the nature of dependency, the housing return is first modelled as a linear autoregressive moving average (ARIMA) and then non-linear Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and the residuals are subjected to BDS test to identify whether dependence is in linear or non-linear form. The application

Table 1 Summary statistics of monthly returns of S&P Case-Shiller U.S. national and city home price index, seasonally adjusted

		Period	Obs	Mean	Standard Deviation	Minimum	Maximum	Skewness	Kurtosis	Jarque–Bera	*P*-Value
S&P Case-Shiller											
U.S. Home Price Index		1987:01–2019:03	386	0.3066	0.6492	-2.2585	2.0384	-0.6150***	1.4669***	58.9418***	(0.0000)
S&P Case-Shiller City Home Price Index											
Northeast Region											
1	Boston MA	1987:01–2019:03	386	0.2968	0.6307	-1.6479	1.8996	-0.3050**	0.1248	6.2333**	(0.0443)
2	NewYork NY	1987:01–2019:03	386	0.2620	0.6353	-1.7544	1.8281	-0.1368	0.0767	1.2998	0.7608
3	Washington DC	1987:01–2019:03	386	0.3355	0.7931	-2.1225	2.7011	-0.3418***	0.9308***	21.4504***	(0.0000)
Midwest Region											
4	Cleveland OH	1987:01–2019:03	386	0.2170	0.5761	-3.8933	2.5651	-1.2785***	8.9084***	1381.5603***	(0.0000)
5	Detroit MI	1991:01–2019:03	338	0.2325	0.9390	-3.6749	3.3829	-0.9622***	3.2994***	205.4695***	(0.0000)
6	Illinois CH	1987:01–2019:03	386	0.2569	0.7218	-3.8932	2.5650	-0.6431***	2.4294***	121.5409***	(0.0000)
7	Minneapolis MN	1989:01–2019:03	362	0.2842	0.8373	-4.7560	2.6042	-1.7277***	6.6099***	839.1129***	(0.0000)
Southeast Region											
8	Atlanta GA	1991:01–2019:03	338	0.2321	0.7346	-4.8402	2.4032	-1.8974***	9.5054***	1475.2728***	(0.0000)
9	Charlotte NC	1987:01–2019:03	386	0.2423	0.4431	-1.8726	1.4786	-1.0491***	3.3859***	255.2033***	(0.0000)
10	Miami FL	1987:01–2019:03	386	0.3333	1.0089	-4.1394	2.9592	-1.3011***	3.8129***	342.7301***	(0.0000)
11	Tampa FL	1987:01–2019:03	386	0.2717	0.8847	-3.3579	2.9182	-0.6964***	2.3996***	123.8137***	(0.0000)
Southwest Region											
12	Dallas TX	2000:02–2019:03	230	0.2783	0.5157	-1.5485	2.3075	-0.6263***	2.1199***	58.1088***	(0.0000)
13	Phoenix AZ	1989:01–2019:03	362	0.2920	1.1866	-4.5022	4.3158	-0.9779***	4.6777***	387.7446***	(0.0000)
Western Region											
14	Denver CO	1987:01–2019:03	386	0.3838	0.5127	-1.6061	1.6986	-0.5123***	0.7828***	26.7427***	(0.0000)
15	Las Vegas NV	1987:01–2019:03	386	0.2833	1.2614	-4.6463	5.4583	-0.4125***	4.4958***	336.0330***	(0.0000)
16	Los Angeles CA	1987:01–2019:03	386	0.4102	1.0760	-3.6884	3.4081	-0.5188***	1.2458***	42.2857***	(0.0000)
17	Portland OR	1987:01–2019:03	386	0.6544	0.6957	-1.9688	2.5233	-0.5613***	1.7273***	68.2578***	(0.0000)

Table 1 (continued)

		Period	Obs	Mean	Standard Deviation	Minimum	Maximum	Skewness	Kurtosis	Jarque–Bera	*P*-Value
18	San Diego CA	1987:01–2019:03	386	0.4065	1.0389	-3.3754	5.1185	-0.2798**	2.1157***	77.0344***	(0.0000)
19	San Francisco CA	1987:01–2019:03	386	0.4563	1.1679	-4.1724	3.2362	0.7880***	2.0482***	107.4248***	(0.0000)
20	Seattle WA	1987:01–2019:03	350	0.4213	1.0366	-3.6322	4.4920	-0.0239	1.8684***	50.9406***	(0.0000)

The studied period for all cities home price index is from January 1987 to March 2019 except the home price index for Detroit (1991:01), Minneapolis (1989:01), Atlanta (1991:01), Dallas (2000:01), and Phoenix (1989:01). Mean and standard deviations are expressed in percent. Asymptotic standard error of coefficient of skewness is $(6/N)^{1/2}$. Asymptotic standard error of coefficient of excess kurtosis is $(24/N)^{1/2}$. Jarque–Bera is test statistics for normality. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively of the K-map and Z-map model developed by Larrain (1991) is subsequently conducted to test whether the nonlinear stochastic dependence in housing markets is consistent with chaos form, or it is driven by underlying macroeconomic variables. In complement with the BDS test, a Markov Chain test developed by McQueen and Thorley (1991) is also implemented to test for departure from random walk and predictable patterns in housing markets. As nonlinear dependency in time series is known to be one of characteristics of bubbles, a duration dependence developed by McQueen and Thorley (1991) test is exploited to test for bubbles in housing prices. To further identify the main driver of underlying nonlinearity in housing markets, a time reversibility developed by Ramsey and Rothman (1996) is employed to uncover whether the non-linear dependence is due to the functional form or non-Gaussian error terms. Such information is useful in delivering a superior fit model for forecasting pricing of U.S. housing markets.

4.1 BDS test

The BDS test is a commonly known test in detecting dependency in many financial time series. If a time series is generated by an IID process, the probability that the distance between any pair of observation is smaller than an arbitrary number should be the same for all pairs.

$$p\left(\left|X_t - X_{t+k}\right| < \varepsilon\right) = p\left(\left|X_t - X_{t+l}\right| < \varepsilon\right) = p_1(\varepsilon) \tag{1}$$

where k and l are arbitrary numbers. Therefore, for any IID financial time series, the joint probabilities that each pair of a sample will satisfy the condition can be expressed as:

$$\prod_{d=1}^{m} p\left(\left|X_{t+d} - X_{t+k+d}\right| < \varepsilon\right) = p_m(\varepsilon) = p_1^m(\varepsilon) \tag{2}$$

The BDS test statistic is calculated as follows:

$$W_m(l) = \frac{\sqrt{T}\left[p_m(\varepsilon) - p_1^m(\varepsilon)\right]}{\sigma_m(\varepsilon)} \tag{3}$$

The results of the BDS test for dependency in aggregate and cities housing index are reported in Table 2. The null hypothesis of no dependence is consistently rejected for housing returns in all cities at the 1% significant level, suggesting that U.S. housing prices do not follow a random walk. To further investigate whether the form of dependency could be attributable to linear or non-linear structures, the residuals from the fitted linear ARIMA model and fitted nonlinear GARCH are subjected to the BDS tests. If dependency is due to the linearity or nonlinearity in the model, fitting ARIMA or GARCH models to the housing returns should remove linear or nonlinear dependency and therefore, residuals should be IID. However, the BDS test on the residuals from ARIMA model continues to reject the null hypothesis of no dependence across all cities and regions at the 1% significance level, suggesting that the use of autoregressive

Table 2 The BDS test statistics on S&P Case-Shiller U.S. national and city home price index, seasonally adjusted

		Period	Obs	BDS Test on Changes in House Price					BDS Test on Residuals from GARCH fitted Model				
				Dimension					Dimension				
				m = 2	m = 3	m = 4	m = 5	m = 6	m = 2	m = 3	m = 4	m = 5	m = 6
S&P Case-Shiller													
U.S. Home Price Index		1987:01–2019:03	386	0.1365***	0.2202***	0.2704***	0.2965***	0.3064***	0.1497***	0.2473***	0.3065***	0.3394***	0.3599***
S&P Case-Shiller City Home Price Index													
Northeast Region													
1	Boston MA	1987:01–2019:03	386	0.0872***	0.1470***	0.1781***	0.1931***	0.1976***	0.0981***	0.1653***	0.2012***	0.2202***	0.2275***
2	NewYork NY	1987:01–2019:03	386	0.1124***	0.1908***	0.2359***	0.2615***	0.2706***	0.1172***	0.2003***	0.2489***	0.2757***	0.2852***
3	Washington DC	1987:01–2019:03	386	0.1271***	0.2162***	0.2705***	0.3023***	0.3174***	0.1233***	0.2095***	0.2618***	0.2915***	0.3063***
Midwest Region													
4	Cleveland OH	1987:01–2019:03	386	0.0602***	0.1075***	0.1386***	0.1585***	0.1666***	0.0598***	0.1069***	0.1378***	0.1577***	0.1659***
5	Detroit MI	1991:01–2019:03	338	0.1011***	0.1773***	0.2253***	0.2544***	0.2726***	0.0985***	0.1700***	0.2184***	0.2473***	0.2633***
6	Illinois CH	1987:01–2019:03	386	0.0864***	0.1528***	0.1915***	0.2131***	0.2223***	0.1073***	0.1844***	0.2313***	0.2596***	0.2738***
7	Minneapolis MN	1989:01–2019:03	362	0.0983***	0.1652***	0.2018***	0.2201***	0.2254***	0.1176***	0.1940***	0.2371***	0.2593***	0.2656***
Southeast Region													
8	Atlanta GA	1991:01–2019:03	338	0.1378***	0.2420***	0.3134***	0.3611***	0.3899***	0.1509***	0.2573***	0.3301***	0.3768***	0.4035***
9	Charlotte NC	1987:01–2019:03	386	0.0434***	0.0758***	0.0963***	0.1066***	0.1080***	0.0626***	0.1041***	0.1296***	0.1414***	0.1427***
10	Miami FL	1987:01–2019:03	386	0.1430***	0.2366***	0.2945***	0.3287***	0.3480***	0.1402***	0.2355***	0.2935***	0.3275***	0.3459***
11	Tampa FL	1987:01–2019:03	386	0.1084***	0.1832***	0.2257***	0.2495***	0.2611***	0.1139***	0.1918***	0.2382***	0.2680***	0.2835***
Southwest Region													
12	Dallas TX	2000:02–2019:03	230	0.0636***	0.0997***	0.1171***	0.1263***	0.1294***	0.0778***	0.1224***	0.1459***	0.1592***	0.1639***
13	Phoenix AZ	1989:01–2019:03	362	0.1656***	0.2784***	0.3492***	0.3928***	0.4180***	0.1659***	0.2790***	0.3502***	0.3942***	0.4196***
Western Region													
14	Denver CO	1987:01–2019:03	386	0.0846***	0.1500***	0.1849***	0.2052***	0.2118***	0.0991***	0.1734***	0.2174***	0.2427***	0.2532***
15	Las Vegas NV	1987:01–2019:03	386	0.1415***	0.2376***	0.2961***	0.3306***	0.3493***	0.1426***	0.2394***	0.2993***	0.3337***	0.3524***
16	Los Angeles CA	1987:01–2019:03	386	0.1461***	0.2486***	0.3130***	0.3528***	0.3727***	0.1461***	0.2486***	0.3130***	0.3528***	0.3727***
17	Portland OR	1987:01–2019:03	386	0.0954***	0.1637***	0.2060***	0.2304***	0.2432***	0.1143***	0.1968***	0.2509***	0.2838***	0.3016***

Table 2 (continued)

		Period	Obs	BDS Test on Changes in House Price					BDS Test on Residuals from GARCH fitted Model				
				Dimension					Dimension				
				m = 2	m = 3	m = 4	m = 5	m = 6	m = 2	m = 3	m = 4	m = 5	m = 6
18	San Diego CA	1987:01–2019:03	386	0.1300***	0.2224***	0.2784***	0.3136***	0.3311***	0.1300***	0.2224***	0.2783***	0.3133***	0.3307***
19	San Francisco CA	1987:01–2019:03	386	0.1215***	0.2026***	0.2495***	0.2767***	0.2881***	0.1219***	0.2032***	0.2484***	0.2743***	0.2847***
20	Seattle WA	1987:01–2019:03	350	0.1029***	0.1684***	0.2036***	0.2154***	0.2140***	0.1097***	0.1822***	0.2218***	0.2365***	0.2367***

The studied periods for all cities home price index are from January 1987 to March 2019 except the home price index for Detroit (1991:01), Minneapolis (1989:01), Atlanta (1991:01), Dallas (2000:01), and Phoenix (1989:01). The monthly S&P Case-Shiller Home Price Index data are obtained from the Federal Reserve Bank of Saint Louis Economics Database. The numbers in each column represented the BDS test statistics on the housing return of each city and the residuals from GARCH fitted model. The embedding dimensions (m) are from 2 to 6. The bootstrapped p-value for the BDS test statistics are calculated and used because of small smaples. *** Indicates significance at the 1% level, ** Indicates significance at the 5% level, and * indicates significance at the 10% level moving model does not help to remove any potential linear dependence. To further explore the structure of dependency, the residuals from fitted nonlinear GARCH model for each city are tested for possible nonlinear dependency via the BDS test.

Table 2 reveals the estimated values of the BDS test on the GARCH residuals from each series for embedding dimensions (m) from 2 to 6. The null hypothesis of IID in residuals is still unequivocally rejected for overall U.S. housing markets and 20 cities at all dimensions at the 1% significance level.

The findings of non-IID in the residuals from nonlinear GARCH filtered series virtually suggest that the pricing behavior of overall U.S. housing markets and 20 cities could possibly be generated by a nonlinear model. Our results are in line with current literature of nonlinear dependence in housing markets (Engelhardt 2003; Seslen 2004; Kim and Bhattacharya 2009).

4.2 K-map and Z-map analysis

Although the BDS test is a popular test for detecting dependency in a time series, the pitfall of the test is that it will reject IID if the time series is chaotic and does not converge. However, Hsieh (1991) pointed out that a chaotic time series can be generated by nonlinear deterministic process which may look random, but not all non-linear dynamics exhibit a chaotic behavior. According to May (1976), a time series X_{t+1}, where $X_{t+1} = aX_t(1 - X_t)$ will be a general nonlinear process for most values of a. However, for the values of a between 3.57 and 3.8, the process will behave like a chaos. This illustrates the point that chaos is only a small subset of nonlinear process.[1]

The K-map and Z-map analysis is further conducted to determine whether the dependency in the housing markets is specifically driven by a chaotic structure. Following the study by Larrain (1991), non-linear components in the housing returns are modelled as K-map to characterize chaotic behaviors. The linear behavior components in the returns of housing are modelled as Z-map to include four macro-economic factors such as economic condition index (EC), inflation (CPI), interest rate (INT), and population growth (PopGrowth).[2] It is more logical to use more state-level heterogeneity controls in the K-map and Z-map analysis. We, therefore, include economic condition indices (EC) which measure average economic growth in each metropolitan statistical area (MSAs). These indices are first constructed by Arias et al. (2016) and include 15 variables to gauge various aspects of economic activities in the MSAs.[3] The inclusion of broader

[1] According to Devaney (1989), chaos process has three conditions. The chaos dynamics are highly dependent on the initial starting point and topologically transitive with many periodic orbits close to each other.

[2] The study period for K-map and Z-map analysis is different from other tests due to the limited data on economic condition index variable. The study period runs from February 1990 to March 2019 for all cities except for Detroit (1991:01), Atlanta (1991:01), and Dallas (2000:01).

[3] The 15 variables used in economic condition index calculation include average weekly hours worked, unemployment rate, all goods-producing employees, all private service-producing employees, all government employees, real average hourly earnings, construction permits for new private residential buildings, real average quarterly wages per employee, total real personal income per capita, industrial availability rate, office vacancy rate, return on average assets, net interest margin, loan loss reserve ratio, and gross metropolitan product.

variables in economic condition indices in the MSAs makes it a viable option as it defines Gross Metropolitan Product Growth potentially affecting home price across regions. This allows for a unified comparison across metro areas.

These four macroeconomic variables are used in this chaos test to examine whether past nonlinear housing returns and macroeconomic variables exhibit stable or chaotic effects on future behaviors of overall and each city U.S. housing prices. Studies by Peng (2016) showed that unemployment rate has a negative effect on housing markets. Similarly, interest rate also negatively affects the returns on housing markets (Peng and Tsai 2019; Stevenson 2008). Lee et al. (2017) showed that the causal relationship between inflation and housing price behavior is positive. Lastly, the increase in population growth positively affects the returns on housing markets (Stevenson and Young 2014; Otto 2007).

The K-map and Z-map of the housing returns can be modelled as follows:

$$
\begin{aligned}
R_t = c + \beta_1 R_{t-1}^1 + \beta_2 R_{t-1}^2 + \beta_3 R_{t-1}^3 + \beta_4 R_{t-1}^4 + \beta_5 (EC) \\
+ \beta_6 (Interest) + \beta_7 (Inflation) + \beta_8 (PopGrowth) + \varepsilon_t
\end{aligned}
\tag{4}
$$

If the dynamics of housing market is driven mainly by erratic and chaotic behavior, this will result in the more significant and powerful of the estimated coefficients on the K-map nonlinear return variables (c, $\beta_1, \beta_2, \beta_3$, and β_4) when compared with the estimated coefficients of the Z-map linear macroeconomic variables ($\beta_{5,} \beta_6, \beta_{7,}$ and β_8).

The first step in testing for chaos in housing markets is to determine the best degree of non-linearity in housing returns in the K-map model using a stepwise regression. The ordinary least square regression is then performed on the full model that is composed of significant non-linear K-map components and linear fundamental economic Z-map components. Table 3 showed that the nonlinear K-map coefficients are dominant for the U.S. national housing index and for all 20 cities. They are more significant than coefficients of Z-map with larger magnitude. This result indicates that home price index in all cities across five regions are driven by chaos. Surprisingly, economics condition index is the dominant and significant factor in the Z-map analysis in most of the cities. Population growth does not have any explanatory power in pricing behavior of housing in any cities. The home price index in Illinois, Dallas, San Francisco, and Seattle cannot be explained by the movement of any of four macroeconomic variables. These findings validate that nonlinear stochastic dependence in housing markets in most of the cities are predominantly driven by chaotic behavior but not the underlying macroeconomic variables.

4.3 Second-order Markov chain

The second order Markov chain test developed by McQueen and Thorley (1991) is further conducted to detect for nonlinear predictable components in the U.S. house price index. This test examines for nonlinear dependence based on price behavior occurred in the past two periods. The structure of the Markov chain test is to test the null hypothesis that house prices follow a random walk. If the housing markets follow a random walk, the probability of observing a positive or negative return in the current period should be invariant to what occurred in previous states.

A two-state second-order Markov process can be constructed by first defining I_t process as follows:

$$\boldsymbol{I_t} = \begin{cases} 1, \boldsymbol{R_t} \geq 0 \\ 0, \boldsymbol{R_t} < 0 \end{cases}$$

The I_t process is then translated into a two-order transition probability (λ_{ij}) as shown follows:

$$\lambda_{ij} = Prob\left(I_t = 0 \middle| I_{t-2} = i, I_{t-1} = j\right) \tag{5}$$

where i and j can take the value of either 1 or 0. To illustrate, λ_{00} is the probability that a negative return will continue to persist in the current period given two preceding negative returns. Consequently, $(1\text{-}\lambda_{00})$ is the probability that a sequence of two negative returns in prior states will revert to a positive return in the current period.

The random walk hypothesis postulates that the chance of either state 0 ($I_t=0$) or state 1 ($I_t=1$) occurrence in the current period should be invariant to any prior two-state sequence. Thus, a rejection of the null hypothesis is an indication of the presence of nonlinear dependence in the housing markets. Seven null hypotheses of equal transition probabilities for different states can be formed. Let λ is the set of all possible probabilities of having a negative return given the past two periods,$\lambda = \{\lambda_{00}, \lambda_{01}, \lambda_{10}, \lambda_{11}\}$. Define H as the set of null hypotheses to be tested $H = \binom{\lambda}{2}$ which is the combination of λ by two. $H_{1-6null} = \{\lambda_{00} = \lambda_{11}, \lambda_{00} = \lambda_{10}, \ldots, \lambda_{10} = \lambda_{11}\}$ versus $H_{1-6alternative} = \{\lambda_{00} \neq \lambda_{11}, \lambda_{00} \neq \lambda_{10}, \ldots, \lambda_{10} \neq \lambda_{11}\}$. The more restrictive null hypothesis to test the equality of four probabilities is $H_{7,null} : \lambda_{00} = \lambda_{01} = \lambda_{10} = \lambda_{11} versus$ $H_{7,alt} : \lambda_{00} \neq \lambda_{01} \neq \lambda_{10} \neq \lambda_{11}$.

The less restrictive null hypotheses 1 to 6 posit that the probability of observing negative returns in the current period should be the same regardless of what happened in the prior two periods. For more restrictive null hypothesis 7, the probability of a negative or positive occurrence in the current period should be same independent of pattern happened in prior two periods. To test each null hypothesis, the log likelihood ratio test is then calculated based on restricted transition probability vs. unrestricted transition probability.

As reported in Table 4, the LRT consistently rejects seven null hypotheses of randomness in the U.S. housing markets and 19 cities at the traditional significant level. This implies that U.S. housing markets exhibit a non-random walk pattern due to an inherent nonlinear dependence in housing price. Only housing market in Phoenix, AZ exhibits a random price pattern during 1987 to 2019.

Table 3 K-map and Z-map anlaysis on S&P Case-Shiller U.S. national and city home price index, seasonally adjusted

				K-Map Coefficients				Z-Map Coefficients				F Statistics	Durbin-Watson Statistics
	Period	Obs	α_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8		
$R_t = c + \beta_1 R^1_{t-1} + \beta_2 R^2_{t-1} + \beta_3 R^3_{t-1} + \beta_4 R^4_{t-1} + \beta_5(\text{EC}) + \beta_6(\text{Interest}) + \beta_7(\text{Inflation}) + \beta_8(\text{PopGrowth}) + \varepsilon_t$													
S&P Case-Shiller													
U.S. Home Price Index	1987:01–2019:03	386	0.2477***	0.9052***				0.0078	0.0145**	0.0675	-0.4156***	356.20***	0.9099
Case Shiller City Home Price Index													
Northeast Region													
1 Boston MA	1990:02–2019:03	350	-0.0921	0.8439***	0.1465***	-0.1355***		0.0438***	-0.0083	0.1267	0.2285	68.23***	2.1771
2 NewYork NY	1990:02–2019:03	350	-0.1677**	0.8073***	0.2164***	-0.0455	-0.0349	0.0425***	-0.0093	0.0262	1.5798	114.97***	2.4402
3 Washington DC	1990:02–2019:03	350	-0.0545	0.6246***	0.1926***	0.0897***	-0.0531***	0.0262**	-0.0214*	-0.0903	1.1764	155.78***	2.3499
Midwest Region													
4 Cleveland OH	1990:02–2019:03	350	0.0214	0.2955***	-0.2229***	0.0267***		0.0286**	0.0013	0.0406	1.5239	8.41***	2.1695
5 Detroit MI	1991:01–2019:03	338	0.3421**	0.5903***	-0.1954***	0.0086	0.0136**	0.0306***	0.0004	-0.2984*	-1.1609	50.06***	2.0830
6 Illinois CH	1990:02–2019:03	350	0.0831	0.6796***	-0.1595***		0.0159**	0.0193	0.0062	-0.0866	0.3588	50.77***	2.0890
7 Minneapolis MN	1990:02–2019:03	350	0.0818	0.7624***	-0.1218***	-0.0172	0.0083*	0.0574***	0.0089	0.0273	-1.2864	47.61***	2.3025
Southeast Region													
8 Atlanta GA	1991:01–2019:03	338	0.0678	0.8088***	-0.0319		0.0028*	0.0201*	0.0021	-0.0635	-0.6868	101.47***	1.8965
9 Charlotte NC	1990:02–2019:03	350	0.1356	0.5572***	-0.3677***	-0.1425**	0.1522**	0.0417***	0.0274**	-0.0274**	-0.0623	27.71***	2.1732
10 Miami FL	1990:02–2019:03	350	0.0728	0.8756***	-0.0814***	0.0095	0.0087***	0.0037	-0.0267*	-0.0152	1.0866	192.97***	2.4278
11 Tampa FL	1990:02–2019:03	350	-0.0035	0.7455***				0.0569***	-0.0067	0.1178	-0.7179	170.88***	2.3656
Southwest Region													
12 Dallas TX	2000:02–2019:03	230	0.0631	0.7824***		-0.1814***		0.0743	-0.0111	0.0904	-0.4620	28.49***	2.1114
13 Phoenix AZ	1990:02–2019:03	350	-0.0614	0.9545***	-0.0321*		0.0023**	0.0078	-0.0328***	0.1711**	1.6740	607.00***	2.0177

Table 3 (continued)

	Period	Obs	α_0	K-Map Coefficients β_1	β_2	β_3	β_4	Z-Map Coefficients β_5	β_6	β_7	β_8	F Statistics	Durbin-Watson Statistics
Western Region													
14 Denver CO	1990:02–2019:03	350	-0.0021	0.8031***		-0.0748**		0.0282***	-0.0050	0.0059	0.7092	86.19***	2.3671
15 Las Vegas NV	1990:02–2019:03	350	0.0235	0.7959***		0.0023		0.0324***	-0.0196	0.0951	-0.2768	238.81***	2.3474
16 Los Angeles CA	1990:02–2019:03	350	-0.0287	0.9532***	-0.0116	-0.0057		0.0076	-0.0247**	0.0056	1.4241	365.67***	2.3325
17 Portland OR	1990:02–2019:03	350	-0.1079	0.7657***		-0.0373*		0.0405***	0.0015	0.1203	0.8698	84.14***	2.2945
18 San Diego CA	1990:02–2019:03	350	-0.0374	0.8776***				0.0187*	-0.0222	-0.1148	1.4503	307.34***	2.4505
19 San Francisco CA	1990:02–2019:03	350	0.0578	0.9004***	-0.0613*		0.0049	0.0067	-0.0267	0.0784	0.9441	176.31***	2.2179
20 Seattle WA	1990:02–2019:03	350	0.0688***	0.6600***	0.1567***	0.0913***	-0.0487***	-0.0060	-0.0191	-0.0918	0.6853	152.42***	2.3983

The period studies for all cities home price index are from February 1990 to March 2019 except the home price index for Detroit (1991:01), Atlanta (1991:01), and Dallas (2000:01) This is due to the limited data availbility for economic condition variable for each MSAs. The monthly S&P Case-Shiller Home Price Index Data are obtained from Federal Reserve Bank of Saint Louis Economics Database. The K-map Z-map is modelled as $R_t = c + \beta_1 R^1_{t-1} + \beta_2 R^2_{t-1} + \beta_3 R^3_{t-1} + \beta_4 R^4_{t-1} + \beta_5(\text{EC}) + \beta_6(\text{Interest}) + \beta_7(\text{Inflation}) + \beta_8(\text{PopGrowth}) + \varepsilon_t$ The first four terms represent K-map non-linear components in returns, where n=1, 2, 3, and 4 and the next four terms represent Z-map variables including economic condition for each MSA, federal fund rate, inflation, and population growth. The model of the first four nonlinear terms are selected using Stepwise regression. The best model for each state housing markets is selected based on adjusted R^2. The significance of each variables is performed using ordinary least square regrressions. *** Indicates significance at the 1% level, ** Indicates significance at the 5% level, and * Indicates significance at the 10% level

The estimated values of $\widehat{\lambda}_{00}$ are larger than those of $\left(1-\widehat{\lambda}_{00}\right)$ for all 20 cities, pointing to the persistence of negative returns. For example, 83.30% $\left(\widehat{\lambda}_{00}\right)$ of the time the overall U.S. housing market exhibits three consecutive negative returns, while only 16.70% $\left(1-\widehat{\lambda}_{00}\right)$ of the time the housing market will revert to positive returns after two negative returns. In other words, the U.S. housing markets in each city tend to show continued price depreciation or persistence of negative returns. Similarly, the smaller estimated values of $\widehat{\lambda}_{11}$ than $\left(1-\widehat{\lambda}_{11}\right)$ for home markets in all cities suggest the likelihood of a house price appreciation for three consecutive periods. For overall U.S. housing markets, there is a 93.40% $\left(1-\widehat{\lambda}_{11}\right)$ chance that a positive house price change continues to occur in the current period given prior two positive prices change. This is significantly higher than the probability of observing negative price change in current period after two positive price changes of 6.60% ($\widehat{\lambda}_{11}$). Overall, these transition probabilities point to an existence of predictable components in the U.S. housing markets across five regions where negative (positive) return is more likely to occur after two consecutive sequences of negative (positive) returns. The empirical findings of a persistent pattern in positive returns indicate the possibility of a bubble in the U.S. housing markets.

4.4 Duration dependence test

A housing bubble is a phenomenon where there is a continued rise in demand driving up property price to an unsustainable level. One of the characteristics of a bubble is the tendency of a positive run up in prices causing positive returns to persist. This bubble feature can be captured by the application of a duration dependence test.

The implementation of duration dependence developed by McQueen and Thorley (1994) to detect for housing bubbles is unique to this study because it does not take into account the fundamental variables such as rental incomes as it commonly used in cointegration test for bubbles. The main premise of the cointegration test is that an existence of long-run relationship between assets value and their underlying fundamental variables would be evidence against the presence of bubbles. The problem of cointegration test is that it requires the correct identification of the fundamental variables that explain the movement of underlying asset values.[4] The duration dependence tests for a bubble by examining the relationship between positive returns and its length. This test differs from cointegration tests in that it does not require a prior correct identification of underlying fundamental factors, therefore, it provides a superior advantage of not testing the joint null hypothesis of no bubbles and no model misspecification.[5]

[4] As indicated by Evans (1991), the finding of bubble could be the result of omitting the important fundamental variables. Therefore, the empirical results from the use of cointegration test for bubbles is questionable as it is subject to testing joint null hypothesis of bubbles and model specification.

[5] Prior studies implemented duration dependence technique include Jirasakuldech et al. (2006), Lehkonen (2010), Emekter et al. (2012), Nartea and Cheema (2014), Nartea et al. (2017), and Watanapalachaikul (2021).

The duration dependence test simply examines duration of the house price increases and the likelihood that it will revert to a price decline when a bubble bursts. Putting it differently, the presence of a rational bubble would imply that the probability of obtaining a negative return (a bubble bursts) should decrease as the length of positive return increases (a bubble grows). If such rule breaks, a bubble cannot thrive in the market, which implies that the hazard rate should be negative.

The first step for a duration dependence test is to divide the housing returns into a group of positive and negative returns. A run in this case is a sequence of positive or negative returns which is defined as follows.

$$R_k \in \left\{X_i < 0 | X_{i-1} > 0, \ldots, X_{i-k} > 0, X_{i-k-1} < 0\right\} \forall i \leq \; n \; and \; k \leq i \tag{6}$$

where X_t is a time series with n abnormal returns and R_k is a positive run with length k. The numbers of positive or negative runs at a particular length k are counted. Theoretically, the hazard rate (h_i) or probability that a bubble will burst should decline as the length of the positive run increases to support the survival of a rational bubble. The hazard rate is then expressed as a function of log of the lag length:

$$h_i = \frac{1}{1 + e^{-d_i}} \tag{7}$$

$$d_i = \alpha + \beta Lni, \tag{8}$$

The relationship between the likelihood that a run will end (d_i) and lag length (i) is investigated through the estimated β parameter obtained via the logit regression. A likelihood ratio test is then carried out to test the null hypothesis of $\beta = 0$ (no bubble). An existence of a bubble would result in $\beta < 0$ or a negative hazard rate.

Table 5 shows that the LRT consistently rejects the null hypothesis of no bubble or a constant hazard rate ($\beta = 0$) in favor of $\beta < 0$ for all housing markets across four regions at the 1% level. Only housing markets in Detroit, MI and Minneapolis, MN in Midwest region do not experience the bubbles. For illustration purpose, the logit regression yields β =-0.6520 for Boston, MA, -0.6009 for New York, NY, and -0.5441 for Washington, DC, which indicate a negative relationship between the duration of the run and the probability of a run to end. As the housing price continues to appreciate in value, the probability that it will revert to a depreciation in value will diminish in support of a growing bubble. It is interesting to note that when a duration dependence test was conducted on the S&P Case-Shiller U.S. national home price index, we fail to report evidence of housing bubbles. One plausible explanation where we failed to find evidence of bubbles in U.S. national home price but reported strong evidence of bubbles in 18 cities could be attributable to

Table 4 Markov chain tests for monthy returns on S&P Case-Shiller national and city home price index, seasonally adjusted

State		S&P Case-Shiller U.S. Home Price		Northeast Region Boston-NY		Northeast Region New York-NY		Northeast Region Washington-DC		Midwest Region Cleveland-OH		Midwest Region Detroit-MI		Midwest Region Illinois-CH		Midwest Region Minneapolis-MN		Southeast Region Atlanta-GA		Southeast Region Charlotte-NC		Southeast Region Miami-FL		Southeast Region Tampa-FL	
		Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State	
Previous State		0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0	0	75	15	59	18	65	19	61	19	39	15	42	13	61	15	45	12	41	8	26	17	42	12	63	28
0	1	1	17	9	24	10	23	7	22	14	33	4	21	9	27	2	30	3	15	14	30	13	19	21	21
1	0	15	3	18	16	19	14	19	10	15	32	13	11	15	21	11	20	7	10	17	28	12	20	28	14
1	1	17	241	25	215	23	211	22	224	33	203	21	211	27	209	29	211	14	238	31	221	19	247	21	39

State	U.S. Home Price	Boston-NY	New York-NY	Washington-DC	Cleveland-OH	Detroit-MI	Illinois-CH	Minneapolis-MN	Atlanta-GA	Charlotte-NC	Miami-FL	Tampa-FL
MLE Estimates	384	384	384	384	384	336	384	360	336	384	384	235
λ_{00}	0.833	0.766	0.773	0.763	0.722	0.764	0.803	0.789	0.837	0.605	0.777	0.692
$\sigma(\lambda_{00})$	(0.039)	(0.048)	(0.045)	(0.048)	(0.061)	(0.052)	(0.045)	(0.054)	(0.053)	(0.074)	(0.056)	(0.048)
λ_{01}	0.056	0.272	0.303	0.241	0.298	0.160	0.250	0.606	0.167	0.318	0.406	0.500
$\sigma(\lambda_{01})$	(0.054)	(0.077)	(0.086)	(0.079)	(0.004)	(0.073)	(0.072)	(0.043)	(0.087)	(0.070)	(0.086)	(0.077)
λ_{10}	0.833	0.592	0.575	0.655	0.319	0.542	0.417	0.355	0.412	0.378	0.375	0.666
$\sigma(\lambda_{10})$	(0.088)	(0.085)	(0.068)	(0.088)	(0.079)	(0.107)	(0.082)	(0.086)	(0.119)	(0.072)	(0.085)	(0.073)
λ_{11}	0.066	0.104	0.098	0.089	0.140	0.091	0.114	0.120	0.056	0.123	0.071	0.350
$\sigma(\lambda_{11})$	(0.015)	(0.019)	(0.019)	(0.018)	(0.023)	(0.019)	(0.021)	(0.021)	(0.014)	(0.021)	(0.015)	(0.062)
Likelihood Ratio Hypothesis Test												
H_1: $\lambda_{00} = \lambda_{01}$ (LRT)	143.233***	80.823***	90.455***	98.535***	40.495***	71.247***	77.565***	73.642***	70.810***	30.742***	60.927***	17.588***
P-value	(0.000)	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_2: $\lambda_{00} = \lambda_{10}$ (LRT)	111.316***	70.587***	77.581***	75.571***	47.813***	56.240***	80.011***	71.210***	69.270***	26.681***	66.367***	14.415***
P-value	(0.000)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_3: $\lambda_{00} = \lambda_{11}$ (LRT)	195.566***	122.450***	12.693***	13.993***	70.242***	101.004***	127.844***	97.886***	134.498***	43.956***	117.660***	17.381***
P-value	(0.145)	(0.002)	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_4: $\lambda_{01} = \lambda_{10}$ (LRT)	102.497***	55.031***	64.053***	67.492***	29.083***	43.512***	50.9424***	37.334***	46.105***	21.131***	50.893***	13.754***
P-value	(0.000)	(0.000)	(0.001)	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_5: $\lambda_{01} = \lambda_{11}$ (LRT)	11.665***	10.382***	9.932***	10.406***	14.234***	6.113***	11.207***	11.021***	7.086***	11.834***	28.759***	3.548*
P-value	(0.000)	(0.001)	(0.002)	(0.001)	(0.079)	(0.013)	(0.000)	(0.000)	(0.007)	(0.000)	(0.000)	(0.059)
H_6: $\lambda_{10} = \lambda_{11}$ (LRT)	56.519***	37.943***	45.773***	50.741***	20.800***	28.709***	26.748***	16.868***	21.861***	25.069***	44.447***	11.441***
P-value	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
H_7: $\lambda_{00} = \lambda_{01} = \lambda_{10} = \lambda_{11}$	225.920***	131.985***	146.490***	152.829***	71.592***	112.670***	130.842***	107.278***	139.072***	53.029***	128.806***	19.954***
P-value	(0.000)	(0.006)	(0.0060)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 4 (continued)

State	Southwest Region				Western Region													
	Dallas-TX		Phoenix-AZ		Denver-CO		Las Vegas-NV		Los Angeles-CA		Portland-OR		San Diego-CA		San Francisco-CA		Seattle-WA	
	Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State		Current State	
Previous State	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0 0	19	10	53	9	31	15	57	16	101	9	32	8	75	17	82	15	49	16
0 1	3	20	8	6	12	17	9	24	7	10	8	18	13	14	7	15	14	25
1 0	10	13	9	5	15	13	16	17	9	8	8	18	17	10	15	7	16	23
1 1	20	133	6	264	16	265	25	220	11	229	17	275	14	224	15	228	25	180
MLE Estimates	228		360		384		384		384		384		384		384		348	
λ_{00}	0.655		0.855		0.674		0.781		0.918		0.800		0.815		0.845		0.753	
$\sigma(\lambda_{00})$	(0.083)		(0.096)		(0.069)		(0.048)		(0.026)		(0.045)		(0.040)		(0.036)		(0.053)	
λ_{01}	0.130		0.571		0.414		0.273		0.412		0.308		0.482		0.318		0.359	
$\sigma(\lambda_{01})$	(0.070)		(0.138)		(0.091)		(0.077)		(0.119)		(0.132)		(0.096)		(0.099)		(0.076)	
λ_{10}	0.435		0.666		0.536		0.485		0.529		0.643		0.629		0.681		0.410	
$\sigma(\lambda_{10})$	(0.103)		(0.136)		(0.094)		(0.087)		(0.121)		(0.128)		(0.093)		(0.099)		(0.078)	
λ_{11}	0.131		0.384		0.057		0.102		0.045		0.022		0.085		0.062		0.122	
$\sigma(\lambda_{11})$	(0.027)		(0.135)		(0.014)		(0.019)		(0.014)		(0.009)		(0.015)		(0.015)		(0.023)	
Likelihood Ratio Hypothesis Test																		
H_1: $\lambda_{00}=\lambda_{01}$ (LRT)	29.807***		2.6520		64.592***		77.462***		145.176***		103.460***		115.200***		127.804***		54.372***	
P-value	(0.000)		(0.103)		(0.000)		(0.003)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
H_2: $\lambda_{00}=\lambda_{10}$ (LRT)	21.219***		2.093		47.749***		71.742***		151.972***		93.192***		104.395***		112.419***		58.086***	
P-value	(0.000)		(0.148)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.001)	
H_3: $\lambda_{00}=\lambda_{11}$ (LRT)	0.841		2.263		88.376***		124.781***		287.199***		201.779***		190.205***		210.406***		92.550***	
P-value	(0.359)		(0.132)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
H_4: $\lambda_{01}=\lambda_{10}$ (LRT)	15.944***		2.267		44.769***		52.401***		123.962***		87.141***		94.662***		98.246***		41.291***	
P-value	(0.000)		(0.132)		(0.000)		(0.002)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
H_5: $\lambda_{01}=\lambda_{11}$ (LRT)	3.559*		0.856		26.901***		11.017***		21.239***		35.652***		32.182***		15.089***		16.485***	
P-value	(0.059)		(0.354)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
H_6: $\lambda_{10}=\lambda_{11}$ (LRT)	11.723***		2.334		62.024***		33.869***		46.283***		71.938***		75.147***		57.589***		30.141***	
P-value	(0.000)		(0.127)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
H_7: $\lambda_{00}=\lambda_{01}=\lambda_{10}=\lambda_{11}$	39.523***		2.831		117.545***		131.759***		290.690***		224.826***		207.871***		224.206***		96.183***	
P-value	(0.000)		(0.418)		(0.000)		(0.002)		(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	

The Markov Chain test is applied on the monthly returns of S&P Case-Shiller national and city home price index. The studied periods for each city do vary depending on the availability of the data. The period studied for all cities home price index is from January 1987 to March 2019 except the home price index for Detroit (1991:01), Minneapolis (1989:01), Atlanta (1991:01). Dallas (2000:01), and Phoenix (1989:01). The numbers in parenthesis below the Maximum Likelihood Estimates (MLE) of the transition probabilities λ_{ij} are the associated asymptotic standard errors, or $\sigma(\lambda_{ij})$. The critical values for H_1 to H_6 are χ^2 (1) = 6.63 at the 1% level, χ^2(1) = 3.84 at the 5% level, and χ^2 (1) = 2.71 at the 10% level. The critical values for H_7 is χ^2(3) = 11.3 at the 1% level, χ^2(3) = 7.82 at the 5% level, and χ^2 (3) = 6.25 at the 10% level. P-Value is the marginal significance level, which is the probability of obtaining that value of the LRT or higher under the null hypothesis. ***, **, * Indicate significance at the 1%, 5%, and 10% level, respectively the loss of information resulting in less variation when the data is aggregated.[6] Acemoglu et al. (2007) explained how the volatility from individual data series can pass to the aggregate data and lot of details could be masked.[7]

4.5 Time reversibility test

We further investigate the underlying source of nonlinearity that characterizes each housing market. Discovering the true cause of nonlinearity is crucial as it provides a blueprint for policymakers in developing an appropriate forecasting model for each housing market that could lead to a better policy control. In doing so, the non-linear time reversibility test (TR) developed by Ramsey and Rothman (1996) is conducted to enhance our understanding of mechanism driving housing dynamics. The truth about time reversibility test is that nonlinear dynamics in financial times may be stirred by either nonlinear asymmetry in the original functional form or the asymmetry in the innovation process.

Conducting time reversibility test requires an initial understanding of symmetric behavior of time series with respect to the time. The time reversibility test is built on the notion that for any times series that is symmetric, reversing the time axis will not alter the behaviors of the symmetric series. In such case, a symmetric time series is also time reversible. A rejection of time reversibility in the housing returns would indicate the presence of asymmetric patterns. Such pattern is consistent with the non-random walk behaviors.

According to Ramsey and Rothman (1996), a time series $\{X_t\}$ is time reversible if $E\left[X_t^i * X_{t-k}^j\right] = E\left[X_t^j * X_{t-k}^i\right]$ for all $i, j, k \epsilon N$, and k is the lag periods. Once the time reversible series is formed, a bicovariance function ($\gamma_{2,1}(k)$) for $i=1$ and $j=2$, is calculated as follows:

$$\gamma_{2,1}(k) = \left\{E\left[X_t^2 * X_{t-k}\right] - E\left[X_t * X_{t-k}^2\right]\right\} \tag{9}$$

Under the null hypothesis (H_0) of $\{X_t\}$ is time reversible, the expected values of sample estimate of a bicovariance function of $\widehat{\gamma}_{2,1}(k)$ are zero for all lags k. The

[6] The standard deviations of housing returns for 20 cities in our study are about 82%. However, if we average the returns of housing in 20 cities across each month, the standard deviation is reduced to 61%, which is close to standard deviation of U.S. National Home Price Index of 65%. This verifies that standard deviation decreases significantly in aggregate data, which can be attributable to the loss of information when data is aggregated. This is expected because the correlations among home price indexes from different cities are not perfectly positive. The highest correlation between home price index of Los Angeles and San Diego is 0.88, while the average correlation of home price index is 0.51 between two random cities.

[7] The duration dependence showed a negative hazard rate in negative runs across all cities except for Denver, CO and San Francisco, CA, suggesting that negative returns tend to persist. This result substantiated the findings of persistence in negative returns by Markov Chain test. To conserve space, the results are available from authors upon request.

sample estimate of this bicovariance function is then calculated based on the following formula:

$$\widehat{\gamma}_{2,1}(\mathrm{k}) = \frac{\sum_{\mathrm{t=k+1}}^{\mathrm{T}} \mathrm{X}_{\mathrm{t}}^{2} * \mathrm{X}_{\mathrm{t-k}} - \sum_{\mathrm{t=k+1}}^{\mathrm{T}} \mathrm{X}_{\mathrm{t}} * \mathrm{X}_{\mathrm{t-k}}^{2}}{\mathrm{T-k}} \tag{10}$$

The standardized TR test statistic, $\frac{\widehat{\gamma}_{2.1}(\mathrm{k})}{\mathrm{VAR}[\widehat{\gamma}_{2.1}(\mathrm{k})]^{1/2}}$, which is the ratio of sample estimate of a bicovariance function to its standard deviation is next computed.

To further identify the critical values for the standardized TR test statistics, the Monte Carlo simulation is implemented. The first step in Monte Carlo simulation is to identify fitted ARIMA model for each housing market series. Based upon the fitted ARMA model, a Monte Carlo simulation is conducted to generate 1,000 values of $\widehat{\gamma}_{2.1}(k)$ along with its corresponding standard deviations for each lag k. The null hypothesis of time reversibility for each housing market is jointly tested for all lags 1 to 10 based on the following TR portmanteau statistic, which is distributed as χ^2 with $n\text{-}m+1$ degrees of freedom.

$$P_{m,n} = \sum_{k=m}^{n} \left(\frac{\widehat{\gamma}_{2,1}(k)}{\widehat{\sigma}_{\gamma}} \right)^2 \tag{11}$$

If the null hypothesis of time reversible or symmetry in housing price is rejected, we can further identify whether asymmetry is inherent in functional form or in the innovation of data generating process. Based on Ramsey and Rothman (1996), the sources of time irreversibility come in two different forms. The first form of irreversibility is induced by non-linear functional form as opposed to non-Gaussian innovations. This is referred to “Type I time irreversibility” in which a nonlinear model with Gaussian innovations is the appropriate model. The second form of irreversibility is caused by non-Gaussian innovations as opposed to non-linearity in functional form. This is called “Type II time irreversibility.” In such case, the suitable model for housing markets will be a linear model with non-Gaussian innovations. Differentiating between two types of asymmetries is often overlooked yet remains utmost important for developing a correct forecasting model for any time series. These two types of time irreversibility can be differentiated by performing TR test on the standardized TR test statistics on the ARIMA residuals. The TR test statistics, $\gamma_{2,1}(k)$ are calculated using residuals from fitted model and standardized by their standard deviation which is calculated as follows:

$$var\left[\widehat{\gamma}_{2,1}(k)\right]^{1/2} = 2\left(\mu_4\mu_2 - \mu_3^2\right)/(T-k) - 2\mu_2^3(T-2k)/(T-k)^2 \tag{12}$$

where $\mu_3 = E\left[X_t^2\right], \mu_3 = E\left[X_t^3\right]$, and $\mu_4 = E\left[X_t^4\right]$. When the null hypothesis of time reversibility is rejected under both the raw housing returns and ARMA residuals, one can conclude that asymmetric behavior is caused by the non-linearity in the

Table 5 The duration dependence test on positive runs on monthy returns on S&P Case-Shiller national and city home price index, seasonally adjusted

		Period	Obs	Total Runs	Total	Total	Log-Logistic Test				LRT	
					Negative Runs	Positive Runs	α	t-value	β	t-value	H_0: β = 0	(p-value)
S&P Case-Shiller Home Price Index												
United States		1987:01–2019:03	386	35	17	18	-2.0695***	(-4.5411)	-0.2502	(-1.4162)	2.0401	(0.1532)
Northeast Region												
1	Boston MA	1987:01–2019:03	386	69	34	35	-0.7497***	(-2.6858)	-0.6520***	(-4.2951)	23.8786***	(0.0000)
2	NewYork NY	1987:01–2019:03	386	65	32	33	-0.8541***	(-2.9895)	-0.6009***	(-3.9114)	90.6626***	(0.0000)
3	Washington DC	1987:01–2019:03	386	59	30	29	-1.0635***	(-3.4443)	-0.5441***	(-3.5182)	14.4004***	(0.0014)
Midwest Region												
4	Cleveland OH	1987:01–2019:03	386	95	48	47	-0.9995***	(-3.8285)	-0.4149***	(-2.6113)	7.1284***	(0.0075)
5	Detroit MI	1991:01–2019:03	338	51	26	25	-1.9616***	(-4.7386)	-0.1381	(-0.7088)	0.5027	(0.4728)
6	Illinois CH	1987:01–2019:03	386	69	37	32	-1.1509***	(-3.9960)	-0.4192***	(-2.9680)	8.0615***	(0.0045)
7	Minneapolis MN	1989:01–2019:03	362	63	31	32	-1.6516***	(-5.1245)	-0.2013	(-1.3123)	1.7893***	(0.1810)
Southeast Region												
8	Atlanta GA	1991:01–2019:03	338	36	18	18	-1.4636***	(-3.3797)	-0.5218***	(-2.7655)	7.9114***	(0.0049)
9	Charlotte NC	1987:01–2019:03	386	94	40	54	-0.8002***	(-3.1337)	-0.5806***	(-3.8580)	16.6219***	(0.0000)
10	Miami FL	1987:01–2019:03	386	63	32	31	-0.5755**	(-1.9682)	-0.8164***	(-5.1167)	31.7246***	(0.0000)
11	Tampa FL	1987:01–2019:03	386	72	35	37	-0.4143	(-1.5372)	-0.8208***	(-5.3361)	37.5825***	(0.0000)
Southwest Region												
12	Dallas TX	2000:02–2019:03	230	54	24	30	-0.9505***	(-2.7916)	-0.5166***	(-2.8618)	9.5016***	(0.0020)
13	Phoenix AZ	1989:01–2019:03	362	22	10	12	-0.1684***	(-0.3974)	-1.1098***	(-5.2711)	40.5402***	(0.000)
Western Region												
14	Denver CO	1987:01–2019:03	386	57	28	29	-0.3183	(-1.0479)	-0.9521***	(-5.7842)	57.9892***	(0.0000)
15	Las Vegas NV	1987:01–2019:03	386	68	34	34	-0.9736***	(-3.2643)	-0.5701***	(-3.5738)	13.9138***	(0.0000)
16	Los Angeles CA	1987:01–2019:03	386	34	16	18	-0.4546	(-1.1617)	-0.9599***	(-4.9404)	30.9064***	(0.0000)
17	Portland OR	1987:01–2019:03	386	52	26	26	-0.9134***	(-2.7232)	-0.7257***	(-4.3472)	20.7931***	(0.0000)

Table 5 (continued)

		Period	Obs	Total Runs	Total	Total	Log-Logistic Test				LRT	
					Negative Runs	Positive Runs	α	t-value	β	t-value	H_0: β=0	(p-value)
18	San Diego CA	1987:01–2019:03	386	54	27	27	-0.3827***	(-1.2039)	-0.9287***	(-5.2123)	36.3636***	(0.0000)
19	San Francisco CA	1987:01–2019:03	386	44	22	22	-1.2051***	(-3.1363)	-0.5848***	(-2.1037)	9.8588***	(0.0017)
20	Seattle WA	1987:01–2019:03	350	77	41	36	-1.0005***	(-3.5126)	-0.4348**	(-2.5337)	84.1341***	(0.0000)

The Duration Dependence test is performed on the monthly returns of S&P Case-Shiller National and City Home Price Index. The studied periods for all cities home price index are from January 1978 to June 2019 except the home price index for Detroit (1991:01), Minneapolis (1989:01), Atlanta (1991:01), Dallas (2000:01), and Phoenix (1989:01). Actual run counts do not include the partial runs which may occur at the beginning or at the end of the period investigated. Total runs are the numbers of total positive and negative runs. The sample hazard rate, $h_i = N_i/(M_i + N_i)$, indicates probability that a run ends at length *i* provided that it lasts until *i*. The likelihood ratio test (LRT) of the null hypothesis of no duration dependence or constant hazard rate (H_1: β=0) is asymptotically distributed χ^2 with one degree of freedom. The critical values are 6.635, 3.841, and 2.706 at 1%, 5%, and 10% significance levels. *P*-value is the marginal significance levels–the probability of obtaining the calculated value of LRT or higher under the null hypothesis. *, **, and *** Indicate significance at the 10%, 5%, and 1% levels, respectively

Table 6 Standardized time reversibility test statistics for S&P Case-Shiller U.S. national and city home price index, seasonally adjusted

		Period	Obs	Standardized TR Test Statistics at Lag k										$P_{1,10}$	Rejection Type
				k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10		
S&P Case-Shiller Home Price Index															
	United States	1987:01–2019:03	386	-0.2848	-0.5359	-0.6167	-0.5685	-0.5183	-0.4717	-0.3632	-0.2124	-0.0987	-0.0464	1.7518	Time Reversible
Northeast Region															
1	Boston MA	1987:01–2019:03	386	-0.9795	-0.6469	-0.1756	0.0170	-0.0654	-0.1104	-0.2012	-0.2012	-0.2582	-0.0312	1.5740	Time Reversible
2	NewYork NY	1987:01–2019:03	386	-0.4089	-0.2748	-0.2619	-0.3092	-0.2748	-0.2609	-0.1358	-0.0274	-0.0391	-0.0053	0.5712	Time Reversible
3	Washington DC	1987:01–2019:03	386	-0.3780	-0.4747	-0.2607	-0.4643	-0.4605	-0.5831	-0.4961	-0.5212	-0.4054	-0.4079	2.0520	Time Reversible
Midwest Region															
4	Cleveland OH	1987:01–2019:03	386	0.8567	-1.1479	-9.1520***	-10.9528***	-6.7801***	1.6663	-3.9775***	1.3717	-4.8483***	5.1975***	322.7464***	Type II
5	Detroit MI	1991:01–2019:03	338	-0.2437	-2.2605	-2.1033	-0.3933	-0.7202	-1.4678	-2.5670	-2.1357	-2.2657	-2.3761	34.3505***	Type II
6	Illinois CH	1987:01–2019:03	386	-1.1048	-7.1518***	-3.5625	-3.8525	-1.8668	-7.0828***	-5.0362***	-5.4259***	-2.2648	-1.4055	195.4631***	Type II
7	Minneapolis MN	1989:01–2019:03	362	-4.2177***	-3.5557	-4.6003	-8.5176***	-6.6001***	-7.7887***	-6.6261***	-5.3233***	-4.9807***	-4.1687***	342.8011***	Type II
Southeast Region															
8	Atlanta GA	1991:01–2019:03	338	-1.0451	-0.5196	-1.9122	-2.3393	-1.7945	-2.3225	-3.0142	-4.6116***	-6.0735***	-5.9073***	121.2424***	Type II
9	Charlotte NC	1987:01–2019:03	386	-5.2695***	-1.7724	-0.4264	1.0994	-0.1260	-1.5898	-0.9093	-0.3957	-0.2932	0.2676	35.9847***	Type II
10	Miami FL	1987:01–2019:03	386	-0.2386	-0.0768	-0.4632	-0.1733	-0.1393	-0.1671	-0.1219	-0.1033	-0.0405	-0.0367	0.3832	Time Reversible
11	Tampa FL	1987:01–2019:03	386	-0.0884	0.2598	-0.0585	-0.0802	-0.1451	-0.1133	-0.0975	0.0128	0.0446	0.0615	0.1345	Time Reversible
Southwest Region															
12	Dallas TX	2000:02–2019:03	230	0.0272	-2.3257	-0.4402	-1.3271	-2.5393	-2.5242	-0.9947	0.4337	0.4991	-0.8904	22.4037***	Type II
13	Phoenix AZ	1989:01–2019:03	362	-0.1066	-0.2507	-0.5737	-0.7806	-0.9815	-1.0849	-1.0625	-0.9997	-0.8622	-0.7399	6.5723	Time Reversible
Western Region															
14	Denver CO	1987:01–2019:03	386	-0.2194	-0.0614	0.0084	-0.1399	-0.2558	-0.1985	-0.0739	-0.0930	-0.0878	-0.0417	0.1217	Time Reversible
15	Las Vegas NV	1987:01–2019:03	386	0.0472	-0.3424	0.7879	0.7628	0.7714	0.7411	0.2476	0.0744	-0.2485	-0.3913	2.7482	Time Reversible

Table 6 (continued)

		Period	Obs	*Standardized TR Test Statistics at Lag k*										$P_{1,10}$	*Rejection Type*
				k=1	*k=2*	*k=3*	*k=4*	*k=5*	*k=6*	*k=7*	*k=8*	*k=9*	*k=10*		
16	Los Angeles CA	1987:01–2019:03	386	0.0273	-0.0189	-0.0097	0.0221	0.0372	0.0060	0.0039	-0.0280	-0.0416	-0.0517	0.0083	Time Reversible
17	Portland OR	1987:01–2019:03	386	-0.0812	-0.1822	-0.1021	0.0693	0.0437	0.0447	-0.0822	-0.0343	-0.0305	0.0175	0.6804	Time Reversible
18	San Diego CA	1987:01–2019:03	386	-5.6526	-1.4456	-1.9424	-1.0152	-1.0168	-0.7384	-0.8279	-0.9848	-1.1664	-1.2052	44.8923***	Type II
19	San Francisco CA	1987:01–2019:03	386	-0.2523	-0.2219	-1.1446	-1.1846	-1.5328	-1.5673	-1.5802	-2.1160	-2.1385	-2.1470	23.7894***	Type II
20	Seattle WA	1987:01–2019:03	350	0.1472	-1.1254	-0.2704	-0.7740	-0.1447	0.3192	0.6713	0.9904	1.0618	0.7469	5.2011	Time Reversible

Time reversibility test is implemented on the monthly changes to S&P Case-Shiller National and City Home Price Index. The studied periods for all cities home price index are from January 1978 to March 2019 except the home price index for Detroit (1991:01), Minneapolis (1989:01), Atlanta (1991:01), Dallas (2000:01), and Phoenix (1989:01). Standardized time reversibility test statistics are the ratio of TR test statistics to the standard deviation of the TR test staticties. $P_{1,10}$ is the time reversibility portmantaeu statistic which provides a joint test on a set of standardized TR test statistic values. With 10 degrees of freedom $\chi^2_{0.90}=15.99$, $\chi^2_{0.95}=18.31$, and $\chi^2_{0.99}=23.20$. Type I rejection is the time irreversbility due to non-linear but Gaussian innovation, Type II rejection is time irreversiblity due to linear but non-Gaussian innovation, and time reversibility indicates symmetry. *, **, and *** Indicate significance at the 10%, 5%, and 1% levels, respectively

Table 7 Standardized time reversibility test statistics for ARMA residuals for S&P Case-Shiller national and city home price index

		Period	Obs	*Standardized TR Test Statistics at Lag* k^2										$P_{1,10}$	*Rejection Type*
				k = 1	*k = 2*	*k = 3*	*k = 4*	*k = 5*	*k = 6*	*k = 7*	*k = 8*	*k = 9*	*k = 10*		
S&P Case-Shiller Home Price Index															
United States		1987:01–2019:03	386	-0.0734	0.0228	0.0474	-0.0162	0.0108	-0.0147	-0.0155	-0.0023	0.0676	0.0525	0.0163	Time Reversible
Northeast Region															
1	Boston MA	1987:01–2019:03	386	-0.0508	0.5589	-0.9803	0.2355	-0.1178	0.1359	0.3504	-0.2743	-0.4715	-0.5682	2.1070	Time Reversible
2	NewYork NY	1987:01–2019:03	386	-0.1480	0.0053	-0.0827	-0.1422	-0.3316	-0.2843	0.0688	-0.0222	-0.2399	0.0190	0.3029	Time Reversible
3	Washington DC	1987:01–2019:03	386	0.0038	0.0195	0.0626	0.0812	0.0138	0.0440	0.0189	0.0108	0.0063	0.0246	0.2855	Time Reversible
Midwest Region															
4	Cleveland OH	1987:01–2019:03	386	0.1315	-0.5190	-1.5022	-1.7005	-0.6144	0.0833	-0.2828	-0.0749	-0.5808	0.8877	7.0304	Type II
5	Detroit MI	1991:01–2019:03	338	-0.0348	-0.0058	0.0018	-0.0032	0.0034	-0.0035	-0.0063	0.0027	0.0049	-0.0002	0.0014	Type II
6	Illinois CH	1987:01–2019:03	386	0.0072	-0.4136	0.2537	0.1349	0.3462	-1.4601	-0.5290	-0.7318	0.1196	-0.1042	3.3459	Type II
7	Minneapolis MN	1989:01–2019:03	362	-1.0226	-0.3336	-0.1511	-1.5249	-0.3183	-1.1427	-0.4756	-0.4642	-0.4201	-0.4109	5.6994	Type II
Southeast Region															
8	Atlanta GA	1991:01–2019:03	338	-0.1416	-0.1515	-0.2149	0.0637	0.1526	-0.1430	-0.1893	-0.0966	-0.1266	-0.2731	0.2728	Type II
9	Charlotte NC	1987:01–2019:03	386	-2.4268	-0.4010	0.1278	-0.2250	0.2375	-0.2448	-0.4327	0.1854	0.3377	-0.0167	6.5696	Type II
10	Miami FL	1987:01–2019:03	386	-0.1950	-0.0014	-0.2582	0.1313	-0.0075	-0.0630	0.1001	0.0214	0.0060	-0.1335	0.1543	Time Reversible
11	Tampa FL	1987:01–2019:03	386	-0.0381	0.1297	0.1033	-0.0645	-0.0946	-0.2139	-0.2931	-0.0517	-0.1536	-0.2040	0.2416	Time Reversible
Southwest Region															
12	Dallas TX	2000:02–2019:03	230	-0.1915	-0.3880	0.0761	-1.0357	-0.2051	-0.3957	-0.0277	0.5285	0.6009	-0.5732	2.4342	Type II
13	Phoenix AZ	1989:01–2019:03	362	0.0539	0.2009	-0.4279	0.1320	-0.0606	-0.0151	-0.0480	-0.0563	0.1327	0.0607	0.2745	Time Reversible
Western Region															
14	Denver CO	1987:01–2019:03	386	-0.5732	0.8430	0.6176	0.8131	0.3243	-1.7125	0.1478	-0.4539	-0.4101	0.0787	5.5220	Time Reversible
15	Las Vegas NV	1987:01–2019:03	386	0.1637	-0.4422	0.1454	0.0838	0.1473	0.1489	0.0196	0.2098	-0.0018	0.0702	0.3437	Time Reversible
16	Los Angeles CA	1987:01–2019:03	386	0.0092	0.1084	0.0281	0.0166	-0.0314	-0.0046	-0.0165	0.1156	0.0098	0.0579	0.0000	Time Reversible
17	Portland OR	1987:01–2019:03	386	0.3806	-0.0218	-0.2015	0.1375	-0.0851	0.2745	0.4310	0.1014	0.2684	0.6864	1.0266	Time Reversible
18	San Diego CA	1987:01–2019:03	386	0.1071	0.3987	-0.8494	0.2620	-0.1204	-0.0300	-0.0953	-0.1117	0.2634	0.1204	1.0814	Type II

Table 7 (continued)

		Period	Obs	*Standardized TR Test Statistics at Lag* k^2										$P_{1,10}$	*Rejection Type*
				k = 1	*k = 2*	*k = 3*	*k = 4*	*k = 5*	*k = 6*	*k = 7*	*k = 8*	*k = 9*	*k = 10*		
19	San Francisco CA	1987:01–2019:03	386	0.0058	0.0772	-0.2282	0.1081	0.0098	0.0803	-0.0056	0.0334	-0.1670	-0.1243	0.1208	Type II
20	Seattle WA	1987:01–2019:03	350	0.0945	-0.1421	-0.0573	-0.1297	-0.0938	0.0548	-0.1155	0.0187	-0.0721	-0.0059	0.0799	Time Reversible

Time reversibility test is implemented on the residuals from S&P Case-Shiller National and City Home Price Index. The period studied for all cities home price index is from January 1978 to March 2019 except the home price index for Detroit (1991:01), Minneapolis (1989:01), Atlanta (1991:01), Dallas (2000:01), and Phoenix (1989:01). Standardized time reversibility test statistics are the ratio of TR test statistics to the standard deviation of the TR test statictics. $P_{1,10}$ is the time reversibility portmantaeu statistic which provides a joint test on a set of standardized TR test statistic values. With 10 degrees of freedom $\chi^2_{0.90} = 15.99$, $\chi^2_{0.95} = 18.31$, and $\chi^2_{0.99} = 23.20$. Type I rejection is the time irreversbility due to non-linear but Gaussian innovation, Type II rejection is time irreversiblity due to linear but non-Gaussian innovation, and time reversibility indicates symmetry. *, **, and *** Indicate significance at the 10%, 5%, and 1% levels, respectively functional form of the model. The logic is that fitting ARIMA to the housing model should remove the linear dependency in the housing return dynamics. On contrary, if pattern of housing returns is time irreversible, but the ARIMA residuals fail to do so, one can conclude that such asymmetric behavior is caused by asymmetry in non-Gaussian innovations.

The empirical TR test results reported in Table 6 reveal that housing markets in nine out of twenty cities exhibit a time irreversible or asymmetric pattern. The TR portmanteau test statistics reject the null hypothesis of time reversible for these housing markets at the 1% significance level. It is interesting to note that all housing markets in the Midwest region show strong evidence of time irreversible pattern. These cities are Cleveland, OH, Detroit, MI, Illinois, CH, and Minneapolis, MN. This is also the case for housing markets in Miami and Tampa, FL in Midwest region, Atlanta, GA and Charlotte, NC in Southeast region, Dallas, TX in Southwest region, and San Diego and San Francisco in Western region. However, when the TR test is performed on the residuals from the ARIMA model, the null hypothesis of time reversible cannot be rejected for the entire U.S. housing market and any city as reported in Table 7. The finding of asymmetric or time irreversible behavior in housing markets in nine cities, but not in the residuals suggests that such asymmetry is driven by non-Gaussian innovations of the housing returns, which is consistent with Type II time irreversibility. Therefore, the appropriate-designed housing market model for these nine cities will be a linear model with non-Gaussian error terms.

5 Conclusions

This study provides a comprehensive analysis of overall U.S. housing market dynamics in twenty cities across five census regions during 1987 and 2019. A battery of tests is employed to discover the nonlinear nature in U.S. housing markets and whether such nonlinearity is in the chaos form and consistent with a characteristic of bubbles. The source of nonlinearity is also identified in response to the need of designing an appropriate model for forecasting housing markets.

The overall U.S. housing markets including home markets in twenty cities exhibit some nonlinear serial dependence. The nonlinear dependence is driven by erratic and chaotic behavior, not by the movement of underlying macroeconomic variables. Persistence of positive or negative returns tends to characterize U.S. house price pattern in various cities except for Phoenix, AZ. Nonlinear behavior in housing markets in eighteen cities is consistent with bubbles, except for home markets in Detroit, MI, Minneapolis, MN as well as U.S. national home market. Housing price patterns in nine cities are time irreversible or asymmetric. Such asymmetric behavior is caused by asymmetry in innovations, not in the functional form. These housing markets are Cleveland, OH, Detroit, MI, Illinois, CH, Minneapolis, MN, Atlanta, GA Charlotte, NC, Dallas, TX, and San Diego and San Francisco, CA. The finding here suggests that the appropriate model for these markets is a linear model with non-Gaussian innovations.

With the overwhelming evidence of nonlinear chaotic process in the U.S. housing markets coupled with the predictable components and bubble episodes, the overall findings here imply that U.S. housing market is relatively inefficient. In addition, macroeconomic variables have little or no explanatory power in predicting house price movements. Similar nonlinear chaotic bubble-like characteristics are shown to explain the behavior of housing markets in most of the cities in five census regions suggesting that U.S. housing prices seem to behave uniformly across regions. As for researchers and policymakers, the consistent finding of nonlinear behavior warrants the use of nonlinear time series model to accurately forecast the house prices at the aggregate as well as city levels. The most notable finding here provides some warning for forecasters and policymakers against relying on the empirical finding solely on the aggregate index. While city home prices are driven by a bubble-like process, the national home index does not have a bubble due to diversification effect across cities. The next logical extension is to develop a model that has the features determined in this study to help successfully predict house prices and potential bubble occurrence.

Declarations

Conflict of interest No conflict of interest

References

Acemoglu D, Ozdaglar A, Tahbaz-Salehi A (2007) The growth of house prices in Australian capital cities: what do economic fundamentals explain? Aust Econ Hist 40(3):225–238

Arias M, Gascon C, Rapach D (2016) Metro business cycles. J Urban Econ 94:90–108

Asal M (2019) Is there a bubble in the swedish housing market? J Eur Real Estate Res 12(1):32–61

Belke A, Marcel W (2005) Boom or bubble in the US real estate market? Intereconomics 40(5):273–284

Besarria C, Paes N, Silva M (2018) Testing for bubbles in housing markets: some evidence for Brazil. Int J Hous Markets Anal 11(2):754–770

Broock W, Scheinkman J, Dechert W, LeBaron B (1996) A test for independence based on the correlation dimension. Economet Rev 15(3):197–235

Cannon S, Miller N, Pandher G (2006) Risk and return in the U.S. housing market: a cross-sectional asset-pricing approach. Real Estate Econ 34(4):519–552

Caraiani P (2014) What drives the nonlinearity of time series: a frequency perspective. Econ Lett 125(1):40–42

Devaney R (1989) An Introduction to Chaotic Dynamical Systems, 2nd edn. Addison-Wesley

Emekter R, Jirasakuldech B, Went P (2012) Rational speculative bubbles and commodities markets: application of duration dependence test. Appl Financ Econ 22(7):581–596

Enders W, Siklos P (2001) Cointegration and threshold adjustment. J Bus Econ Stat 19(2):166–176

Engelhardt G (2003) Nominal Loss aversion, housing equity constraints, and household mobility: evidence from the United States. J Urban Econ 53(1):171–195

Engsted T, Hviid S, Pedersen T (2016) Explosive bubbles in house prices? Evidence from the OECD countries. J Int Financ Markets Inst Money 40(C):14–25

Evans G (1991) Pitfalls in testing for explosive bubbles in assets price. Am Econ Rev 81(4):922–930

Gallagher E (2019) All recessions are not equal: The effect of sectoral shifts on unemployment using regional data. CMC Senior Theses. 2261. https://scholarship.claremont.edu/cmc_theses/2261

Genesove D, Mayer C (2001) Loss aversion and seller behavior: evidence from the housing market. Q J Econ 116(4):1233–1260

Hepsen A, Kalfa N (2009) Housing market activity and macroeconomic variables: an analysis of turkish dwelling market under new mortgage system. Istanbul Univ J School Bus Adm 38(1):38–46

Herring R, Wachter S (2002) Bubbles in Real Estate Markets. Zell/Lurie Center Working Papers #402, Wharton School Samuel Zell and Robert Lurie Real Estate Center, University of Pennsylvania

Himmelberg C, Mayer C, Sinai T (2005) Assessing high house prices: bubbles, fundamentals, and misperceptions. J Econ Perspect 19(4):67–92

Hott C (2012) The influence of herding behavior on house prices. J Eur Real Estate Res 5(3):177–198

Hsieh D (1991) Chaos and nonlinear dynamics: application to financial markets. J Financ 46(5):1839–1877

Huang M (2013) The role of people's expectation in the recent US housing boom and bust. J Real Estate Financ Econ 46(3):452–479

Jirasakuldech B, Emekter R, Went P (2006) Rational speculative bubbles and duration dependence in exchange rates: an analysis of five currencies. Appl Financ Econ 16:233–243

Kim S, Bhattacharya R (2009) Regional housing prices in the USA: an empirical investigation of nonlinearity. J Real Estate Financ Econ 38(4):443–460

Larrain M (1991) Testing chaos and nonlinearities in t-bill rates. Financ Anal J 47(5):51–62

Lehkonen H (2010) Bubbles in China. Int Rev Financ Anal 19(2):113–117

Leung C, Ng C (2019) Macroeconomic Aspects of Housing. *Oxford Research Encyclopedia of Economics and Finance* (Oxford Research Encyclopedias). Oxford University Press

Lim KP, Brooks R (2011) The evolution of stock market efficiency over time: a survey of the empirical literature. J Econ Surv 25(1):69–108

May R (1976) Simple mathematical models with very complicated dynamics. Nature 261(5560):459–467

Maynou L, Monfort M, Morley B, Ordóñez J (2021) Club convergence in European housing prices: the role of macroeconomic and housing market fundamentals. Econ Model 103:1–12

McCarthy J, Peach R (2004) are home prices the next 'bubble'? Fed Reserve Bank N Y Econ Policy Rev 10(3):1–17

McQueen G, Thorley S (1994) Bubbles, stock returns, and duration dependence. J Financ Quant Anal 29(3):379–401

McQuenn G, Thorley S (1991) Are stock returns predictable? A test using markov chains. J Financ 46(1):239–263

Mikhed V, Zemcik P (2007) Testing for bubbles in housing markets: a panel data approach. J Real Estate Financ Econ 38(4):366–386

Miles W (2008) Boom-bust cycles and the forecasting performance of linear and non-linear models of house prices. J Real Estate Financ Econ 36(3):249–264

Mill E (1989) Social returns to housing and other fixed capital. AREUEA J 17(Summer):197–211

Muellbauer J, Murphy A (1997) Booms and busts in the UK housing market. Economic Journal 107(445):1720–1746

Nartea GV, Cheema MA (2014) Bubble footprints in the Malaysian stock market: are they rational? Int J Account Inf Manag 22(3):223–236

Nartea GV, Cheema MA, Szulczyk K (2017) Searching for rational bubble footprints in the Singaporean and Indonesian stock markets. J Econ Financ 41:529–552

Otto G (2007) The growth of house prices in Australian capital cities: what do economic fundamentals explain? Aust Econ Rev 40(3):225–238

Peng T (2016) The effect of quality determinants on house prices of eight capital cities in australia a dynamic panel analysis. Int J Hous Mark Anal 9 (3): 355–375. https://doi.org/10.1108/IJHMA-06-2015-0028

Peng C, Tsai I (2019) The Influences of Housing Prices on Residential Mobility and Unemployment. *European Real Estate Society*. https://ideas.repec.org/p/arz/wpaper/eres2019_136.html. Accessed 15 May 2022

Ramsey J, Rothman P (1996) Time reversibility and business cycle asymmetry. J Money Credit Bank 28(1):1–21

Rosenthal S (1999) Residential building and the cost of construction: new evidence on the efficiency of the housing market. Rev Econ Stat 81(2):288–302

Seslen T (2004) Housing Price Dynamics and Household Mobility Decisions. Paper presented at the USC LUSK/FBE Real Estate Seminar
Shi S, Valadkhani A, Smyth R, Vahid F (2016) Dating the timeline of house price bubbles in australian capital cities. Econ Record 92(299):590–605
Shiller R (2005) Irrational Exuberance, 2nd edn. Princeton University Press, Princeton
Skalin J, Teräsvirta T (2002) Modelling asymmetries and moving equilibria in unemployment rates. Macroecon Dyn 6(2):202–241
Stevenson S (2008) Modeling housing market fundamentals: empirical evidence of extreme market conditions. Real Estate Econ 36(1):1–29
Stevenson S, Young J (2014) A Multiple error-correction model of housing supply. Hous Stud 29(3):362–379
Trindade A, Zhu Y, Andrews B (2010) Time series models with asymmetric laplace innovations. J Stat Comput Simul 80(12):1317–1333
Watanapalachaikul S (2021) The impact of COVID-19 Pandemic toward asian pacific stock markets during year 2020: an empirical study of logarithmic returns and duration dependence test model. J Pacific Inst Manag Sci 19(2):113–117

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.