Jing Zhang*, Robert de Jong and Donald Haurin

Are US real house prices stationary? New evidence from univariate and panel data

DOI 10.1515/snde-2013-0134

Abstract: Many papers in the housing literature treat the intertemporal evolution of the logarithm of US real house prices as a unit root process. They also study the cointegration relationship among the logarithm of real house prices and fundamental economic variables such as income and they apply an error correction specification for modeling and forecasting real house prices. This paper argues that the logarithm of US real house price is not a unit root process. Instead, the evidence from a 120-year national dataset and metro area level and state level panel data sets supports the notion that US house prices are trend stationary. One result of this conclusion is that the validity of analyses of US house prices based on cointegration and error correction models needs to be reconsidered.

Keywords: cointegration; error correction; house price; stationarity; unit root.

1 Introduction

Understanding real house price dynamics is an important issue because the housing market is an important economic sector. Many papers in the housing literature apply cointegration analysis and error correction specifications for modeling US real house prices (Abraham and Hendershott 1996; Malpezzi 1999; Capozza, Hendershott, and Mack 2004; Gallin 2006; Mikhed and Zemcik 2009; Holly, Pesaran, and Yamagata 2010). For cointegration and error correction analysis to be valid, real house price should be a unit root process.[1]

In a regression equation the stationarity of the variables ensures that hypothesis tests are valid. If a series is not stationary but trend-stationary, it can be transformed into a stationary series by subtracting the deterministic trend. If a nonstationary series has a unit root, a stationary series can be generated by differencing. Moreover, in a context with multiple series with unit roots, if they have a long run equilibrium relationship (or more formally, if there exists a linear combination among them that is stationary), they are said to be cointegrated. In the cointegration case, variables adjust to discrepancies from the long run relationship, and hence an error correction specification is appropriate to capture the impact of the deviation from the long run equilibrium on the short-run dynamics. Therefore, a prerequisite for applying the error correction model is the existence of a cointegrating relationship, and a prerequisite for the cointegration analysis is that the variables contain unit roots.

Determining whether there is a unit root in US real house prices also sheds light on the appropriateness of theoretical urban models that explain real house prices. If real income has a unit root and real house price is trend stationary as our results suggest, then the models such as the one by Capozza and Helsley (1989, 1990) that suggest an equilibrium relationship between real house price and real income are puzzling.

1 Following conventions in the literature, all series under consideration in this paper are in logarithms, and we take statements such as “unit root in real house prices” to mean “unit root in the log of real house prices.”

***Corresponding author: Jing Zhang,** ECCA, Moody’s Analytics, 121 N Walnut St, West Chester, PA 19380, USA, e-mail: zhang.728@osu.edu
Robert de Jong and Donald Haurin: Economics, The Ohio State University, Columbus, OH, USA

Our paper studies a fundamental and important question: do US real house prices contain a unit root? Surprisingly, the literature has not taken a careful look at this question. To answer the question of whether US real house prices have a unit root, we apply unit root tests to national data, to state level panel data, and to MSA (Metropolitan Statistical Area) panel data. We first study a national real home price index over 120 years constructed by Robert Shiller. This is a long time series that is of particular interest because the literature typically uses data sets that begin around 1975. A central feature of the data is that there appear to be structural breaks, and therefore we apply unit root tests that explicitly allow for such breaks. Whether there is a national housing market has been debated and thus we next apply the Pesaran (2007) panel unit root test to a data set of 48 states and Washington, DC and to one that contains 363 MSAs for the 1975–2011 period. Because areas with an inelastic supply of housing usually have more severe house price cycles, we conduct separate tests in set of MSAs where supply is most inelastic and most elastic.

To understand why our results differ from results found in the literature, we restrict our data to the areas and years in the samples used in the most cited papers that apply panel unit root tests and cointegration tests to US real house prices. We find that these paper's results regarding whether real house price has a unit root change when the sample period is extended to the most recent data.

We also employ nonlinear unit root tests to consider the possibility that house prices might be nonlinear, and a sequential panel selection method to select house price series with stronger evidence of stationarity from a panel.

The remainder of this paper is organized as follows. Section 2 discusses the related literature. Section 3 provides unit root tests results for Shiller's 120-year home price index. Section 4 reports the panel unit root test results for the state and MSA panels as well as for supply elastic and inelastic subgroups. The second half of Section 4 returns to the literature and determines what happens to the results in the literature when the time period of analysis is extended to include recent data, and then discusses the implications of our results for theoretical house price models. Section 5 presents the results for nonlinear unit root tests. Section 6 provides the Sequential Panel Selection Method and the results. Section 7 concludes.

2 Literature review

Error correction models are widely used in the housing literature to model US house price dynamics. Examples include Abraham and Hendershott (1996), Malpezzi (1999), and Capozza, Hendershott, and Mack (2004). Abraham and Hendershott (1996) use annual data for 30 metropolitan areas over the period 1977–1992 to investigate the determinants of real house price appreciation. The explanatory variables consist of three parts: the change in fundamental price, the lagged real house price appreciation with its coefficient called the "bubble builder," and the deviation of house price from its equilibrium level in the previous period with its coefficient called the "bubble burster." They find a positive bubble builder coefficient and a negative bubble burster coefficient. Moreover, the absolute values of these coefficients are higher for the coastal city group than for the inland city group.

Capozza, Hendershott, and Mack (2004) empirically study which variables explain the significant difference in the geographic patterns of the bubble builder and the bubble burster coefficients found in Abraham and Hendershott's (1996) error correction specification. Using a panel dataset of 62 MSAs from 1979 to 1995, they find that higher real income growth, higher population growth, and a higher level of real construction costs increase the bubble builder coefficient, while higher real income growth, a larger population, and a lower level of real construction costs increase the absolute value of the bubble burster coefficient.

Abraham and Hendershott (1996) and Capozza, Hendershott, and Mack (2004) do not provide formal unit root and cointegration tests results to justify the existence of a long run equilibrium relationship among house prices and fundamental variables, which should be a prerequisite for the validity of error correction models.

Malpezzi (1999) uses a dataset which includes 133 MSAs and covers 18 years from 1979 through 1996 and states that short run real house price changes are well modeled by an error correction formulation. The panel unit root test of Levin, Lin, and Chu (2002, LLC test) is applied to real house price changes, the house-price-to-income ratio, and the residuals of the regression of real house prices on real per capita incomes. The first two are the dependent variables in Malpezzi's error correction model. A unit root is rejected for price changes, but cannot be rejected for the price-to-income ratio. Moreover, a unit root is rejected for the residuals of the regression of real house prices on real per capita incomes, and hence Malpezzi concludes that real house prices and real incomes are cointegrated. This cointegration test procedure suffers from several shortcomings. First, before applying the cointegration test, Malpezzi does not examine if real house prices and income have a unit root, respectively. Second, critical values of the LLC panel unit root test have not been shown to work for residuals from the first stage regression, so the claim that house prices and incomes are cointegrated based on the LLC critical values could be misleading. Third, the LLC test does not allow cross-sectional dependence in the regression errors, hence the test result may be biased.

The 2000s' US housing boom, which reached its peak in 2006, raises the question of whether US real house prices are supported by fundamentals; that is, whether real house prices and the fundamental economic variables such as income have a long run equilibrium relationship. The housing literature formalizes this argument by discussing the cointegration relationship among real house prices and the fundamental variables. Thus far, the results are mixed. Some papers such as Gallin (2006) and Mikhed and Zemcik (2009) apply cointegration tests and claim that there is no long run equilibrium relationship, which cast doubts on the validity of applying error correction models to US real house prices. Other papers argue that there is a cointegration relationship, such as Holly, Pesaran, and Yamagata (2010).

Gallin (2006) tests for the existence of a long run relationship among US house prices and economic fundamental variables by applying cointegration tests to both national level data and city level panel data. The augmented Engle-Granger cointegration test is applied to national level house prices, per capita income, population, construction wage, user cost of housing, and the Standard and Poor's 500 stock index. No cointegration relationship is found. He also applies panel cointegration tests to city level house prices, per capita income, and population, for 95 MSAs over 23 years from 1978 to 2000. The panel cointegration tests he uses are Pedroni (1999) and Maddala and Wu (1999). He also applies a bootstrapped version of the tests to take into account cross-sectional dependence. The null hypothesis of no cointegration cannot be rejected, neither by the original tests nor the bootstrapped version. To test for a unit root, Gallin applies the ADF unit root test to the national level real house prices and a unit root is not rejected. But he does not provide panel unit root test result for the city level panel data.

Mikhed and Zemcik (2009) also examines if US house price and fundamental factors are cointegrated. The innovation in their paper is that they include more fundamental variables to avoid the possibility that the omission of potential demand and supply shifters cause the lack of cointegration relationships. The fundamentals included are house rent, a building cost index, per capita income, population, mortgage rate, and the Standard and Poor's 500 stock index. Their sample includes 22 MSAs over 1978–2007, and they examine several different time periods (1978–2007, 1978–2006, 1978–2005, 1997–2007, 1978–1996). They apply the CIPS panel unit root tests to real house prices and the fundamental variables for all periods, setting the time lag in the CIPS test to 1 year. For real house prices, a unit root is rejected at the 5% level for 1978–2007 and rejected at the 10% level for 1978–2006, but cannot be rejected for other periods. The authors interpret this as a correction of the house price bubble around 2006. They further investigate the cointegration relationship of house prices and fundamentals for periods prior to 2005 when a unit root cannot be rejected in house prices. They apply the Pedroni (1999, 2004) panel cointegration test and bootstrap the critical values for possible cross-sectional dependence. No evidence of cointegration relationships is found in any of the cases. Hence they claim that US real house price dynamics are not explained by fundamentals, and this is evidence of housing bubbles in those subsamples.

Holly, Pesaran, and Yamagata (2010) study the determination of US real house prices using a panel of state level data (48 states, excluding Alaska and Hawaii and they include the District of Columbia) over 29 years from 1975 to 2003. Unlike Gallin (2006) and Mikhed and Zemcik (2009), Holly et al. find that real house prices and real income are cointegrated. The innovation of Holly et al. is that they apply the common correlated effects (CCE) estimators of Pesaran (2006) to study a panel of real house prices. They first use an asset non-arbitrage model to show that the ratio of real house price and real income should be stationary, hence the log of real house price and the log of real income should be cointegrated with a cointegration vector of (1, –1). Then they empirically find such a cointegration relationship and estimate a panel error correction model for the dynamics of adjustment of real house prices to real incomes.

Most of the above papers use time periods that end before 2006 and thus they do not include the recent US housing bust period in their sample. An exception is Mikhed and Zemcik (2009), which reports unit root tests for periods ending after 2006, and a unit root is rejected at the 5% level over 1978–2007 for real house prices. The authors then apply cointegration tests to the subperiods ending before 2006 where a unit root is not rejected. This is an ad hoc solution that does not address the issue of whether real house prices have a unit root or are stationary.

3 Evidence from Shiller's 120-year national real home price index

3.1 Data

The periods examined in the literature usually start at or after 1975 because reliable US house price data are available since then, both for national and regional house prices. However, if the periods of the recent house price bubble and bust are omitted, as occurs in many papers in the literature, then only 20–30 years of data remain. Therefore, we examine the full period to increase the power of the unit root tests.

We study the US real home price index constructed by Shiller (2005).[2] It is a national level index that begins in 1890. Shiller constructed this index "by linking together various available series that were designed to provide estimates of the price of a standard, unchanging, house…" He also created an index for the period 1934–1953. Before 1953 there are only annual data, and therefore we use the logarithm of the annual data for this analysis. Figure 1 shows the path of this index. Real house prices appear to be relatively stable except for the period during the two world wars and the Great Depression, and a sharp spike during the 2000s.

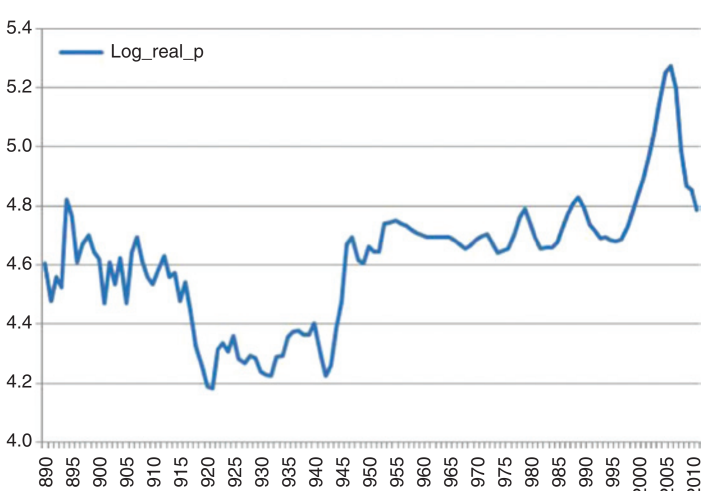

Figure 1: Log of Shiller's Real Home Price Index, 1890–2011.
Note: The contruction of Shiller's home price index is described in Shiller (2005).

2 Shiller's home price index can be found at http://www.econ.yale.edu/~shiller/data.htm. See the Appendix for a more thorough description of the construction of the series and a discussion of the reliability of it. Detailed descriptions of each of the sources for constructing the index are also available in Shiller's book Irrational Exuberance, 2nd ed.: 234–235.

3.2 Methodology: univariate unit root tests

We apply the Lee and Strazicich (2003) (L-S) minimum LM endogenous two breaks unit root test, the multiple breaks unit root tests in Carrion-i-Silvestre, Kim, and Perron (2009) (CKP), the augmented Dickey-Fuller (ADF) unit root test, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test to the 120 year time series of house prices.

Because two breaks clearly appear in Figure 1, with one break around World War I (1914–1918) and the second around World War II (1939–1945), we apply the Lee-Strazicich minimum LM endogenous two structural breaks unit root test. The L-S test estimates the points of structural break by searching for the two break points where the unit root t-statistic is minimized, and tests the null of a unit root against the alternative of trend stationarity. It allows for breaks under both the null and alternative hypotheses. If breaks are not included in the null, as in Lumsdaine and Papell (1997), then a rejection of the null may imply a unit root process with breaks and not necessarily a trend stationary series. Model C in the Lee and Strazicich paper is adopted here, which is the most general case that allows for changes in both level and slope.

We apply unit root test statistic in Lee and Strazicich (2003), which is obtained from the following regression:

$$\Delta y_t = \delta' \Delta Z_t + \phi \tilde{S}_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta \tilde{S}_{t-i} + u_t, \tag{1}$$

where $\tilde{S}_t = y_t - \tilde{\Psi}_x - Z_t\tilde{\delta}$, $t=2, \ldots, T$, $\tilde{\Psi}_x = y_1 - Z_1\tilde{\delta}$; $\tilde{\delta}$, are the regression coefficients from a regression of Δy_t on ΔZ_t; $Z_t=[1, t, D_{1t}, D_{2t}, DT_{1t}, DT_{2t}]'$, where $D_{jt}=1$ and $DT_{jt}=t-T_{B_j}$ for $t \geq T_{B_j}+1$ and 0 otherwise, $j=1, 2$, and T_{B_j} is the time of a break point. The lagged terms $\Delta\tilde{S}_{t-i}$ are included to correct for serial correlation, and the number of lags k is selected following the general-to-specific procedure described in their paper. The LM test statistic is given by

$$\tilde{\tau} = t \text{ statistic testing the null hypothesis } \phi = 0. \tag{2}$$

The two breaks ($\lambda_j = T_{B_j}/T$, $j=1, 2$) are determined endogenously by a grid search over the time span [0.1T, 0.9T] (to avoid end-point issues).

Because of the potential existence of three breaks during 1890–2011 due to the two world wars and the 2000s' housing bubble, we also apply Carrion-i-Silvestre, Kim, and Perron's (2009) unit root tests that allow for multiple structural breaks. They extend the unit root tests of Elliott, Rothenberg, and Stock (1996) and Ng and Perron (2001) to allow for multiple breaks (we use their tests with three breaks), and provide five test statistics $(P_T^{GLS}, MZ_a^{GLS}, MSB^{GLS}, MZ_t^{GLS}, MP_T^{GLS})$ to test the null of a unit root against the alternative of trend stationarity. The tests considered in Carrion-i-Silvestre, Kim, and Perron (2009) estimate the break points, allow for breaks under both the null and the alternative hypotheses, and allow for breaks in both the level and the slope.

We also apply the ADF and KPSS tests because they are popular in the literature that tests the stationarity of a series. The ADF test tests the null of a unit root against the alternative of trend stationarity, while the KPSS test tests the null of trend stationarity against the alternative of a unit root.

3.3 Empirical results

Table 1 reports the results of the Lee and Strazicich (2003) two breaks test. A unit root cannot be rejected for the period 1890–2011. Break points are found in 1916 and 1949. These values closely correspond to the visual inspection of Figure 1. However, it is possible that a third break occurs during the most recent housing bubble, hence biasing the test towards non-rejection of a unit root. If the period of the recent housing bubble is excluded, a unit root is rejected at the 1% level over 1890–1998. The estimated break points are 1916 and 1947 for this time span. We also apply this two-break test to the 1920–2011 period to avoid the structural break at World War I, thus allowing for the break at World War II and a possible break during the recent housing

Table 1: Lee and Strazicich (2003) minimum LM endogenous two breaks unit root test for Shiller's 120-year log real home price index.

	Break points	Test stat.	Lags
1890–2011	1916, 1949	−5.02	2
1890–1998	1916, 1947	−6.54***	2
1920–2011	1946, 2000	−6.53***	3

Model C which allows for changes in both level and trend is adopted here. The number of lags (k) is selected following the general-to-specific procedure described in their paper. Critical values of the unit root statistic, which depend upon the location of the breaks, are tabulated in Lee and Strazicich (2003 Table 2). For T=100, the 1%, 5%, 10% critical values are −6.16, −5.59, −5.27 for λ_1=0.2, λ_2=0.4, are −6.41, −5.74, −5.32 for λ_1=0.2, λ_2=0.6, are −6.32, −5.73, −5.32 for λ_1=0.6, λ_2=0.8, and are −6.33, −5.71, −5.33 for λ_1=0.2, λ_2=0.8. *** denotes significance at the 10%, 5%, and 1% level, respectively.

bubble. A unit root again is rejected at the 1% level. The estimated break points are 1946 and 2000. Therefore, a unit root is rejected for both 1890–1998 and 1920–2011.

Table 2 reports the tests statistics for three breaks in Carrion-i-Silvestre, Kim, and Perron (2009), for different numbers of lags. The estimated break points are 1917, 1944, and 1998, which are similar to the above results. The rejection of a unit root is very robust to the number of lags. None of the statistics can reject a unit root when the number of lags equals zero; three of them reject a unit root at the 10% significance level when one lag is included; all of them reject a unit root at the 5% or 1% level when more than one lag is included. The selected number of lags is two using the Bayesian information criterion (BIC) and is zero using the modified Akaike information criterion (MAIC) by Ng and Perron (2001) and Perron and Qu (2007). Because real house prices and real house price changes are documented in the literature as having strong serial correlation, we think the number of lags selected based on BIC is more reliable. Therefore, the three breaks tests provide strong evidence of trend stationarity with three breaks over 1890–2011.

Results for the ADF and the KPSS tests are reported in Table 3. For the period 1890–2011, the ADF test cannot reject a unit root and the KPSS test rejects stationarity, which is not surprising because ignoring the possibility of structural breaks will bias these test statistics towards the unit root hypothesis. To avoid the effects of dramatic changes due to the world wars and the Great Depression, we restrict the test period to 1950–2011, and the unit root hypothesis is rejected at the 5% level by the ADF test, suggesting that real house price is trend stationary during this 60-year period. Moreover, to examine the possible effect of the recent bubble on the unit root test results, we examine the period 1950–2006. The ADF test cannot reject a unit root and the KPSS test rejects trend stationarity at the 5% level. When we look at the series that ends at 1998 (excluding the recent bubble), the ADF test rejects a unit root at the 1% level and KPSS test cannot reject trend stationarity, implying strong evidence of trend stationarity over 1950–1998. Therefore, including or excluding

Table 2: Carrion-i-Silvestre, Kim, and Perron (2009) three breaks unit root tests for Shiller's 120-year log real home price index.

Number of lags	0	1	2	3	4	5
P_T^{GLS}	12.34	9.58	6.19**	6.14**	4.55***	1.63***
MZ_a^{GLS}	−23.77	−30.89*	−48.37***	−48.72***	−66.11***	−186.44***
MSB^{GLS}	0.14	0.13*	0.10***	0.10***	0.09***	0.05***
MZ_t^{GLS}	−3.38	−3.87*	−4.87***	−4.89***	−5.71***	−9.63***
MP_T^{GLS}	11.53	8.95	5.78***	5.74***	4.25***	1.52***

The estimated break points are 1917, 1944 and 1998. Breaks are allowed in both level and trend. The test statistics for different number of lags in CKP(2009) are reported. The 1%, 5% and 10% critical values are 5.78, 7.59 and 8.80 for P_T^{GLS}, −43.82, −34.31 and −30.01 for MZ_a^{GLS}, 0.11, 0.12 and 0.13 for MSB^{GLS}, −4.68, −4.14 and −3.87 for MZ_t^{GLS}, and 5.78, 7.59 and 8.80 for MP_T^{GLS}. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively.

Table 3: ADF and KPSS unit root test for Shiller's 120-year log real home price index.

	1890–2011	1950–2011	1950–2006			1950–1998
ADF *t*-stat.	−2.46	−3.82**	−1.07			−4.93***
1%, 5% and 10% critical values	(−4.04, −3.45, −3.15)	(−4.12, −3.49, −3.17)	(−4.14, −3.50, −3.18)			(−4.17, −3.51, −3.18)
KPSS LM-stat.	0.18**	0.14*	0.17**			0.07
1%, 5% and 10% critical values	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)			(0.22, 0.15, 0.12)
	1975–2011		1975–2006		1975–2002	
	Annually	Quarterly	Annually	Quarterly	Annually	Quarterly
ADF *t*-stat.	−3.48*	−3.35*	−2.50	−3.03	−3.21	−2.78
1%, 5% and 10% critical values	(−4.24, −3.54, −3.20)	(−4.03, −3.44, −3.15)	(−4.30, −3.57, −3.22)	(−4.03, −3.45, −3.15)	(−4.36, −3.60, −3.23)	(−4.05, −3.45, −3.15)
KPSS LM-stat.	0.09	0.12*	0.14*	0.23**	0.08	0.09
1%, 5% and 10% critical values	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)	(0.22, 0.15, 0.12)

An intercept and a linear time trend are included. The null hypothesis for the Augmented Dickey-Fuller test is that the series has a unit root. The null hypothesis for the Kwiatkowski-Phillips-Schmidt-Shin test is that the series is trend stationary. The number of lags included in the ADF test are based on the Schwarz Info Criterion. For 1950–2006, the SIC selected lag length is 2, and SIC selected lag lengths for all other annual periods are 1. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

the entire recent bubble-bust cycle results in trend stationarity, but including only the boom period of this bubble will result in a unit root conclusion.[3] To further emphasize the potential problem in the literature when only a subperiod is studied, we apply the ADF test and the KPSS test to periods starting at 1975 for both annual data and quarterly data. A unit root is weakly rejected at the 10% level for 1975–2011, but cannot be rejected for 1975–2006 and 1975–2002. This result coincides with the findings in the literature which generally claim that real house prices have a unit root.

Table 4 presents a summary of the findings of the ADF test, the L-S two breaks test and the CKP three breaks test for different time periods. Our preferred method is the CKP three breaks test for the full sample period, 1890–2011, resulting in a conclusion that house prices are trend stationary with breaks.

Table 4: Summary of unit root test results for log of Shiller's real home price index.

	ADF	L-S two breaks	CKP three breaks
1890–2011	Unit root	Unit root	Trend stationary
1890–1998	–	Trend stationary	–
1920–2011	–	Trend stationary	–
1950–2011	Trend stationary	–	–
1950–2006	Unit root	–	–
1950–1998	Trend stationary	–	–

3 As a robustness check, we also apply unit root tests to the period 1950–1989 and 1950–1982, which, respectively, ends at the peak and the start of the housing boom prior to the 2000s bubble, and find evidence of trend stationarity for both periods: a unit root is rejected by the ADF test at the 5% level and the 1% level, respectively. Therefore, of the examined five periods since 1950, 1950–2006 is the only one for which we cannot reject a unit root.

4 Evidence from MSA and state level panel data

4.1 Data

We consider two panel data sets: the first includes house prices for 363 MSAs in the US and the second reports data for 48 states and Washington, DC. Both data sets cover the 1975–2011 period. We use the Freddie Mac House Price Index (FMHPI), which is a monthly nominal repeated-sales index estimated using data on house price transactions (including refinanced) on one-family detached and townhome properties whose mortgage has been purchased by Freddie Mac or Fannie Mae.[4] We deflate the nominal house price index by the CPI-U for each month and then average over a year to obtain annual data and then take logarithms.[5]

4.2 Methodology: panel unit root test

In the literature, early specifications of panel unit root tests did not allow for cross-sectional dependence (Levin, Lin, and Chu 2002; Im, Pesaran, and Shin 2003). More recently, panel unit root tests allow for cross-sectional dependence (Bai and Ng 2004; Moon and Perron 2004; Pesaran 2007).

We apply the Pesaran CIPS test (2007). There are several reasons for this choice. First, the CIPS test allows for cross-sectional dependence, which is an important consideration for house prices because house prices in different areas are very likely to be affected by common effects such as changes in interest rates or technology. Second, it allows for cross-sectional heterogeneity in the intercept, trend, and autoregressive coefficients. The third reason is its favorable size and power for large N and small T, which is the case for our sample. In particular, Moon and Perron (2004) and Bai and Ng (2004) both assume that $N/T \to 0$ as N and $T \to \infty$ when deriving the asymptotic properties of the test, while Pesaran (2007) assumes $N/T \to k$ (where k is a finite positive constant) as N and $T \to \infty$.

To justify the use of Pesaran's (2007) CIPS test, we need to establish the presence of cross-sectional dependence. We use the CD (cross-section dependence) test proposed in Pesaran (2004) for this purpose. This test statistic asymptotically converges in distribution to a standard normal distribution as T and $N \to \infty$ in any order under the null of no cross-sectional dependence,[6] and is defined as

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right), \tag{3}$$

where $\hat{\rho}_{ij} = Corr(\hat{\varepsilon}_i, \hat{\varepsilon}_j)$ and $\hat{\varepsilon}_i$ is the estimated residual from the ADF(p) regression that estimates the model

$$\Delta y_{it} = a_{i0} + a_{i1} t + a_{i2} y_{i,t-1} + \sum_{j=1}^{p} \delta_{ij} \Delta y_{i,t-j} + \varepsilon_{it}. \tag{4}$$

Pesaran's (2007) CIPS panel unit root test is based on a cross-sectionally augmented ADF (CADF) regression, which filters out the cross-sectional dependence by augmenting the ADF regressions with the lagged cross-section mean and the lagged first differences of the cross-sectional mean. The CADF regression estimates the model

$$\Delta y_{it} = a_{i0} + a_{i1} t + a_{i2} y_{i,t-1} + a_{i3} \bar{y}_{t-1} + \sum_{j=0}^{p} d_{ij} \Delta \bar{y}_{t-j} + \sum_{j=1}^{p} \delta_{ij} \Delta y_{i,t-j} + v_{it}. \tag{5}$$

4 The FMHPI series can be found at http://www.freddiemac.com/finance/fmhpi/.
5 The CPI-U series can be found at http://www.bls.gov/cpi/#data.
6 Pesaran (2015) shows robustness of the CD test to the null of weak cross-sectional dependence.

Let $\tilde{t}_i$ denote the t-ratio for a_{i2} in the above regression. Then the CIPS statistic is

$$\text{CIPS}=N^{-1}\sum_{i=1}^{N}\tilde{t}_i. \tag{6}$$

The null and alternative hypotheses for the CIPS test are

$$H_0: a_{i2}=0 \quad \text{for } i=1, 2, \ldots, N, \tag{7}$$

and

$$H_1: \begin{cases} a_{i2}=0 & \text{for } i=1, 2, \ldots, N_1 \\ a_{i2}<0 & \text{for } i=N_1+1, N_1+2, \ldots, N \end{cases}. \tag{8}$$

4.3 Empirical results

To confirm the presence of cross-sectional dependence, we carry out CD tests using 1, 2 and 3 lags and using both the MSA and state data sets. We also consider subsamples of the top 20 and 50 most supply inelastic and elastic MSAs. These MSAs are selected based on the supply elasticity estimated in Saiz (2010), who provides housing supply elasticity measures for 269 metro areas. After matching his data with the MSA sample used here, supply elasticity measures for 254 MSAs are available.

The cross-sectional dependence is confirmed by the CD test statistics in Table 5, which are statistically highly significant. This implies that the panel unit root tests that do not allow for cross-sectional dependence are inappropriate.

Table 6 reports the CIPS statistics with an intercept and a linear time trend included for varying augmented orders and for the state and MSA sample and the MSA subsamples. We examine the periods 1975–2011, 1975–2006, 1975–2002, and 1975–1998 to determine if and how the recent housing bubble affects unit root test results. For the 363 MSAs group, the period 1975–2011 strongly rejects a unit root at the 1% level, but a unit root cannot be rejected for all other periods except 1975–1998 which rejects a unit root at the 10% level when the augmented order is one. This result indicates that we can reject a unit root if we examine the period that includes both the boom and bust years of the recent bubble, but a unit root cannot be rejected if we exclude the bust. For the state sample, the periods 1975–2011 and 1975–2006 both reject a unit root, but 1975–2002 and 1975–1998 cannot reject.[7]

The CIPS test statistics for the supply elastic and inelastic MSA groups also are reported in Table 6. The recent literature has identified that housing supply elasticity plays an important role in house price dynamics. For example, Glaeser, Gyourko, and Saiz (2008) find, both theoretically and empirically, that places with

Table 5: Pesaran (2004) cross-section dependence test (CD test) for log real house prices, 1975–2011.

	ADF(1)	ADF(2)	ADF(3)
All 363 MSAs	403.44	405.37	358.66
Top 50 supply inelastic MSAs	85.90	85.49	80.71
Top 20 supply inelastic MSAs	38.21	38.12	36.95
Top 50 supply elastic MSAs	69.16	66.70	57.39
Top 20 supply elastic MSAs	26.53	26.69	24.84
States+DC	94.96	91.92	81.55

An intercept and a trend are included. The CD statistic tends to $N(0, 1)$ under the null hypothesis of no error cross-sectional dependence. The 2-sided 5% critical value is 1.96.

7 We also apply the CIPS test to monthly Freddie Mac house price data for the state panel, and set the maximum number of lags to 24 months. A unit root is rejected when the number of lags equals 11, 13, or is greater than 14. This result is consistent with the finding from the annual data.

Table 6: Pesaran (2007) CIPS panel unit root test for log real house prices.

		CADF(1)	CADF(2)	CADF(3)
All 363 MSAs	1975–2011	−2.786***	−2.686***	−2.822***
	1975–2006	−2.306	−2.165	−2.282
	1975–2002	−2.410	−2.222	−2.278
	1975–1998	−2.510*	−2.353	−2.364
Top 50 supply inelastic MSAs	1975–2011	−3.280***	−2.974***	−3.351***
	1975–2006	−2.507	−2.103	−2.178
	1975–2002	−2.415	−1.926	−1.741
	1975–1998	−2.358	−1.698	−1.339
Top 20 supply inelastic MSAs	1975–2011	−3.306***	−2.621	−3.264***
	1975–2006	−2.796**	−1.816	−1.947
	1975–2002	−3.028***	−1.768	−1.610
	1975–1998	−3.319***	−2.103	−1.789
Top 50 supply elastic MSAs	1975–2011	−2.548*	−2.651**	−2.766***
	1975–2006	−2.424	−2.379	−2.500
	1975–2002	−2.561*	−2.437	−2.528
	1975–1998	−2.364	−2.189	−2.237
Top 20 supply elastic MSAs	1975–2011	−2.540	−2.722**	−2.549
	1975–2006	−2.438	−2.414	−2.153
	1975–2002	−2.567	−2.476	−2.183
	1975–1998	−2.296	−2.174	−1.995
States+DC	1975–2011	−3.093***	−2.970***	−2.995***
	1975–2006	−2.697**	−2.699**	−2.923***
	1975–2002	−2.312	−2.387	−2.608*
	1975–1998	−2.051	−2.079	−2.386

An intercept and a trend are included. Reported are truncated CIPS statistics. 1%, 5%, 10% critical values for N=363, T=30 and T=20 are −2.609, −2.532, −2.485 and −2.627, −2.534, −2.482. For N=50, T=30 and T=20, they are −2.73, −2.61, −2.54 and −2.76, −2.62, −2.54. For N=20, T=30 and T=20, they are −2.88, −2.72, −2.63 and −2.92, −2.73, −2.63. Critical values for N=50 and N=20 are given in Table II(c) in Pesaran (2007). Critical values for N=363 are not reported in Pesaran's paper and we generate them by ourselves. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively.

more elastic housing supply have smaller house price increases and shorter house price bubbles. Also, we can see from Figure 2 that the average of the log real house price index for the top 20 supply inelastic MSAs is much more volatile and has much bigger and prolonged cycles than that for the top 20 supply elastic MSAs.

Table 6 shows that the top 50 supply inelastic MSA group strongly rejects a unit root over 1975–2011 with the test statistics much higher in absolute value than that for the 363 MSA group, while all other periods cannot reject a unit root. The top 20 supply inelastic MSA group reveals even stronger evidence for trend stationarity – a unit root is very strongly rejected for all time periods when a 1-year lag is selected. In contrast, the supply elastic groups reveal little evidence of trend stationarity. The top 50 supply elastic MSA group cannot reject a unit root at the 5% level for all periods when we select a 1-year lag, and the top 20 supply elastic group cannot reject a unit root even at the 10% level. Therefore, it appears that the test statistics are impacted by the house supply elasticity of the areas under consideration.[8]

8 Pesaran, Smith, and Yamagata (2013) propose a CIPSM test which is an extension of the CIPS test. The CIPS test allows for only one common factor affecting the cross-sectional dependence, while the CIPSM test allows for multiple common factors. In the CIPSM test, an additional variable is included when there are two common factors, and the lagged cross-sectional mean and the lagged first differences of the cross-sectional mean of that additional variable are used to filter out the cross-sectional dependence created by the second common factor. Similarly, in the case of three common factors, two additional variables are included in the CIPSM test. We carry out the CIPSM test with lag order being one year, and consider three cases of the additional variables included: income, population, and income plus population. (By income we mean log real per capita income, and population means

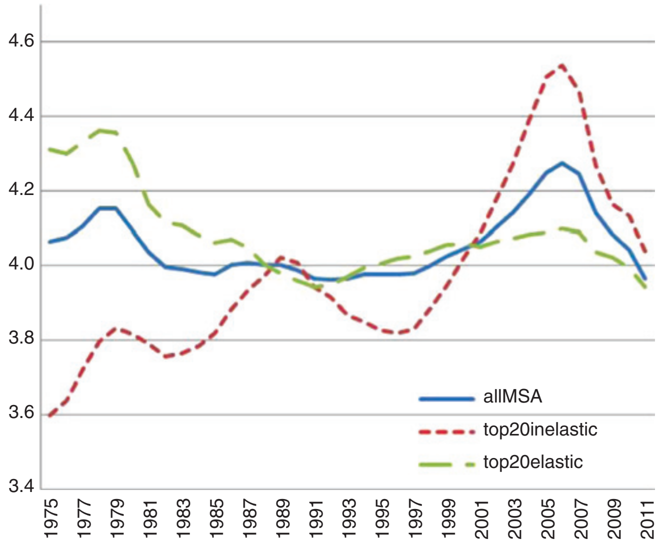

Figure 2: Average log real house price index, 1975–2011.
Note: House prices are based on the Freddie Mac house price index. Housing supply elasticity measure is provided in Saiz (2010).

4.4 Panel unit root tests for samples studied in the literature

Our panel unit root test results indicate trend stationarity if the data extends through 2011. This result differs from many studies in the literature, which find that real house prices have a unit root process. This raises the question as to the causes of the difference in results. Possible explanations are (1) the different time periods covered, (2) the different areas included, and (3) the different unit root tests that are applied. To answer this question, we apply the CIPS panel unit root test to the area samples studied in the most cited papers that examine panel unit root tests and cointegration tests for US real house prices.

4.4.1 Malpezzi (1999)

We apply the CIPS panel unit root test to a subsample that matches the MSAs and years in Malpezzi's paper. Of the 133 MSAs studied by Malpezzi, 124 of them are available in our dataset. The first row of Table 7 presents

Table 7: Panel unit root test for log real house prices of the MSAs in Malpezzi (1999).

	CADF(1)	CADF(2)	CADF(3)
1979–1996	−1.831	−1.894	−1.459
1975–2011	−2.855***	−2.802***	−3.001***

Pesaran (2007) truncated CIPS statistics are reported. An intercept and a trend are included. 1%, 5%, 10% critical values for N=100, T=30 and T=20 are −2.66, −2.56, −2.51 and −2.70, −2.57, −2.51. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively.

log population.) For the 363 MSAs group over 1975–2011, a unit root is rejected at the 5% level when income is the additional variable included, but it cannot be rejected if the additional variable(s) included is(are) population or income plus population. For the top 20 supply inelastic MSAs group and the to 20 supply elastic MSAs group, results are very similar to the one common factor CIPS test: the former group strongly rejects a unit root in all the cases of the additional variables included, and the latter group has little evidence of rejecting a unit root.

Table 8: Panel unit root test for log real house prices of the MSAs in Gallin (2006).

	CADF(1)	CADF(2)	CADF(3)
1978–2000	−2.244	−2.090	−2.206
1975–2011	2.921***	−2.720***	−2.941***

Pesaran (2007) truncated CIPS statistics are reported. An intercept and a trend are included. 1%, 5%, 10% critical values for N=70, T=30 and T=20 are −2.69, −2.58, −2.52 and −2.72, −2.59, −2.53. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively.

the CIPS statistics. A unit root cannot be rejected over the years studied in Malpezzi's sample 1979–1996, and this result is robust to the varying augmented lags, and thus we replicate his result. The second row of Table 7 presents the CIPS statistics for the same group of MSAs over 1975–2011, and a unit root is rejected at the 1% level, regardless of the number of the augmented lags. This change of the test results is clearly a result of the different time periods available in his and our samples.

4.4.2 Gallin (2006)

Of the 95 MSAs in Gallin's sample, 83 of them can be matched with our data and we apply the CIPS test to this group of MSAs. The first row of Table 8 reports the CIPS statistics over 1978–2000, the years examined in Gallin's sample. A unit root is not rejected. But if the data are extended to cover the 1975–2011 period, a unit root is rejected as reported in the second row. The results are robust to changing the number of lags. Therefore, the 1975–2011 period data contain new information that may invalidate the cointegration analysis in Gallin's paper.

4.4.3 Mikhed and Zemcik (2009)

We apply the CIPS test to the 22 MSAs in Mikhed and Zemcik's paper that are in our sample. The first and second rows of Table 9 report the CIPS statistics over 1978–2005 and 1978–2007, their sample periods. Our results are very similar to theirs when one augmented lag is selected (which is the case in their paper). Specifically, a unit root is rejected at the 5% level for 1978–2007, but cannot be rejected for 1978–2005. The third row of Table 9 reports the test statistics for the entire period 1975–2011 and a unit root is rejected at the 1% level when the lag is set to one year. The non-rejection of a unit root for periods before 2005 could be an indication of less test power caused by the limited number of time periods. As before, the application of cointegration methodology in Mikhed and Zemcik's paper may well have been inappropriate.

4.4.4 Holly, Pesaran, and Yamagata (2010)

Table 10 reports the CIPS statistics using our sample for 48 states and Washington, DC. The house price data in Holly et al. are the house price index from the Office of Federal Housing Enterprise Oversight and

Table 9: Panel unit root test for log real house prices of the MSAs in Mikhed and Zemcik (2009).

	CADF(1)	CADF(2)	CADF(3)
1978–2005	−2.567	−2.142	−1.866
1978–2007	−2.838***	−2.301	−2.054
1975–2011	−3.046***	−2.506	−2.534

Pesaran (2007) truncated CIPS statistics are reported. An intercept and a trend are included. 1%, 5%, 10% critical values for N=20, T=30 and T=20 are −2.88, −2.72, −2.63 and −2.92, −2.73, −2.63. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively.

Table 10: Panel unit root test for log real house prices of the states in Holly, Pesaran, and Yamagata (2010).

	CADF(1)	CADF(2)	CADF(3)
1975–2003	−2.358	−2.455	−2.614*
1975–2011	−3.093***	−2.970***	−2.995***

Pesaran (2007) truncated CIPS statistics are reported. An intercept and a trend are included. 1%, 5%, 10% critical values for N=50, T=30 are −2.73, −2.61, −2.54. *, **, and *** denotes significance at the 10%, 5%, and 1% level, respectively.

are deflated by a state level consumer price index, while we use the Freddie Mac house price index deflated by the national consumer price index. The first row shows the test statistics for the time period 1975–2003 studied in Holly et al., and a unit root cannot be rejected. This agrees with their unit root test result. However, the second row shows that, if the data are extended to also cover the 2004–2011 period, a unit root is strongly rejected. Again, this change in unit root test results indicates that the cointegration analysis and the error correction model discussed in Holly et al.'s paper may well have been inappropriate.

For each of these studies, we find that a unit root cannot be rejected if we consider samples similar to those in the original study and thus we confirm their results. However, if these samples are extended to 1975–2011, a unit root is always rejected. Therefore, the reason for our disagreement with the literature is the difference in the time span covered in the analysis. In particular, stopping an analysis midway through the 1998–2011 housing cycle yields different results than including the complete cycle.

4.5 Implications of the results

Both the panel unit root tests and the univariate unit root tests provide supportive evidence that real house price is a trend stationary series with structural breaks, rather than a unit root process. This result presents a challenge to urban models that aim to find the determinants of real house prices. As mentioned in the introduction, house price models usually identify real income as one of the most important factors determining real house prices. But our study suggests some income-house-price "puzzles." Real house prices and real income are very likely two processes with different stationarity properties. Real house prices reject a unit root based on findings in our paper, while there is strong evidence that real income is a unit root

Table 11: Pesaran (2007) CIPS panel unit root test for log real per capita income.

		CADF(1)	CADF(2)	CADF(3)
All 363 MSAs	1975–2009	−1.740	−1.657	−1.705
	1975–2006	−1.788	−1.708	−1.704
	1975–1998	−1.754	−1.539	−1.372
Top 20 supply inelastic MSAs	1975–2009	−1.785	−1.786	−1.585
	1975–2006	−1.589	−1.712	−1.491
	1975–1998	−1.512	−1.478	−1.308
Top 20 supply elastic MSAs	1975–2009	−1.897	−1.671	−1.720
	1975–2006	−2.144	−2.068	−2.057
	1975–1998	−2.042	−1.967	−1.385

An intercept and a trend are included. Reported are truncated CIPS statistics. 1%, 5%, 10% critical values for N=363, T=30 and T=20 are −2.609, −2.532, −2.485 and −2.627, −2.534, −2.482. For N=20, T=30 and T=20, they are −2.88, −2.72, −2.63 and −2.92, −2.73, −2.63. Critical values for N=20 are given in Table II(c) in Pesaran (2007). Critical values for N=363 is not reported in Pesarans paper and we generate them by ourselves.

9 The per capita personal income is annual data from Bureau of Economic Analysis, and we convert this nominal data to real terms by deflating it by CPI-U.

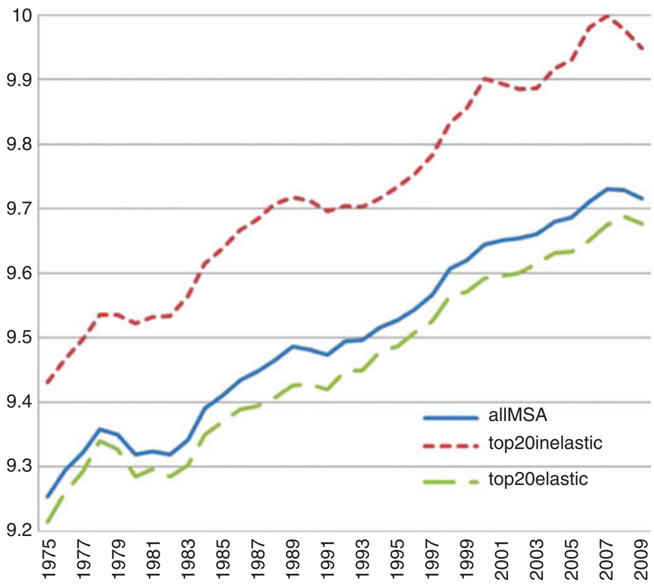

Figure 3: Average log real per capita personal income, 1975–2009.
Note: Personal income is from the Bureau of Economic Analysis. Housing supply elasticity measure is provided in Saiz (2010).

process. Table 11 reports the CIPS panel unit root test statistics when applied to log of real per capita personal income, for the all-MSA group, top-20-supply-inelastic-MSA group, and top-20-supply-elastic-MSA group, and covering different time periods: 1975–2009, 1975–2006, and 1975–1998.[9] None of the test statistics are significant. How a unit root real income process determines a trend stationary real house price is a challenge for urban models.

In addition to the stationarity puzzle, there is also an issue related to house price trends. Figure 1 indicates that there is hardly any trend for the national level real home price index, but Figure 3 shows that there is an obvious upward trend for real income. A possible explanation is to include housing supply because if housing supply is elastic, house price trends will follow trends in construction costs. As documented in Gyourko and Saiz (2006), national level real construction costs trended down over 1980–2003. Therefore, in supply elastic MSAs, one would expect real house prices to decline even when real income is increasing. In contrast, if housing supply is inelastic, then real house prices will rise as income increases. This explanation is supported by Figure 2. Real house prices have an upward trend for the top 20 supply inelastic MSAs, and a downward trend for the top 20 supply elastic MSAs. In contrast, Figure 3 shows that, for real income, there are obvious upward trends for all the three groups, and the three paths are almost parallel.

5 Nonlinear unit root tests

5.1 Unit root test with nonlinear ESTAR alternative

To consider the possibility that house price series might be nonlinear,[10] we implement a nonlinear unit root test proposed in Kapetanios, Shin, and Snell (2003) (hereafter, KSS test), which tests the null of a unit root

10 Canarella, Miller, and Pollard (2012) examine stationarity for the house price indices included in the S P/Case-Shiller Composite-10 index. They implement a number of unit root tests including linear tests, nonlinear tests, and tests with structural breaks. The series they examine are the log difference of city house prices, and the ratio of log city house price index to log Composite-10 house price index, but they do not examine any house price series itself. They find different results for different unit root test procedures applied and different series examined.

Table 12: Kapetanios, Shin, and Snell (2003) nonlinear unit root test.

Shiller series	lag=1	lag=2	lag=3
1950–2011	−4.363***	−3.580**	−3.886**
1890–2011	−3.099	−3.613**	−3.265*
The 15 states for which a unit root cannot be rejected are AR, CO, GA, LA, MI, MN, MS, ND, NE, OH, OK, SC, SD, TX, and WV.			

Asymptotic critical values: −3.93 for 1%, −3.40 for 5%, −3.13 for 10%. *, ** and *** denotes significance at the 10%, 5%, and 1% level, respectively.

process against the alternative of a globally stationary exponential smooth transition autoregressive (ESTAR) process. We carry out this test for Shiller's national house price series, as well as the 49 individual states.

Table 12 presents the results. For Shiller's series from 1950 to 2011, therefore excluding the two World Wars and the Great Depression, the KSS test rejects a unit root at the 1% or 5% level for lag lengths varying from one through three. If we include the time period prior to 1950 and examine the period of 1890–2011, the test does not reject a unit root with one lag, but rejects a unit root at the 5% level with two lags and at the 10% level with three lags. For the 49 individual states, the test with one lag rejects a unit root at the 10% or higher level for 34 states. The 15 states for which the test cannot reject a unit root are mostly inland areas.

The above results suggest that the evidence of stationarity for house prices is very strong under the KSS nonlinear unit root testing procedure: A unit root can be rejected for Shiller's series and almost 70% (34 out of 49) of the states.

5.2 Unit root test with smooth breaks

Arguably, structural breaks in house prices could be smooth and gradual. We consider this issue by implementing a unit root test proposed in Enders and Lee (2012). This test uses the low frequency components of a Fourier expansion to approximate smooth structural breaks of unknown forms. Enders and Lee suggest pretesting for a deterministic nonlinear trend using an F-statistic before carrying out their unit root test. They define the F-statistic in their equation (10) using a single frequency component, and suggest a grid-search method to select the frequency value k that minimizes the sum of squared residuals from the unit root test regression equation (and thus maximizes the F-statistic). We implement the F-test for Shiller's series from 1890 to 2011, for lag lengths varying from one through three, and experiment with one through five for the frequency value k. For each lag length, the k that maximizes the F-statistic is always k=1, and the corresponding F-statistic is always insignificant.[11] This result indicates that a single frequency component Fourier expansion is not a good approximation for the trend in Shiller's house price series. Therefore, Enders and Lee's (2012) unit root test may have low test power in our case, and we do not proceed in applying their unit root test.

6 Sequential panel selection method

To investigate which series in the panel have stronger evidence of stationarity, we employ the Sequential Panel Selection Method (SPSM) of Chortareas and Kapetanios (2009) in this section.

The SPSM separates a panel into two groups by the following steps. A panel unit root test is first implemented to the entire panel. If a unit root is not rejected, the null that all series are non-stationary is accepted. If a unit root is rejected, the most stationary series (measured as the minimum individual CADF statistic in our case of the CIPS test) is excluded from the panel, and the panel unit root test is re-run on the remaining

11 The F-statistic is 3.28 for lag=1, 3.77 for lag=2, and 3.37 for lag=3.

series. Repeat this process until the null is not rejected. Then the panel is separated into a group of series with evidence for stationarity and a group of series without evidence for stationarity.

We apply this SPSM to the panel of state house price series, using the CIPS panel unit root test with one lag. At the 10% significance level, 19 out of the 49 states are selected as areas with evidence for stationarity: FL MA NY AL PA UT OK NH NJ MD TN RI MO CT CO NC TX NV ID. A pattern observed is that 10 of these 19 states are along the Atlantic Ocean.

A related question is how the stationarity properties of the disaggregated house price series affect the stationarity of the aggregated national series. Intuitively, states with large population and expensive housing such as the big coastal states will have higher weights in constructing the national-level index, and the stronger evidence of stationarity in these areas contribute to the stationarity of national level index.

7 Conclusion

This paper examines a fundamental and important question regarding the time series properties of US real house prices; specifically, do US real house prices contain a unit root. This stationarity property is of vital importance for analyzing univariate and panel time series data, modeling and forecasting the dynamics of the series, and conducting tests such as cointegration. Choosing an inappropriate model due to misunderstanding of whether a unit root exists could invalidate the usual statistical tests, such as t-tests and F-tests.

We first examine Shiller's national level real home price index over a 120-year period. We apply the multiple breaks unit root tests in Carrion-i-Silvestre, Kim, and Perron (2009) (CKP), the Augmented Dickey-Fuller (ADF) unit root test, and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. The CKP three breaks test rejects a unit root for 1890–2011. When the ADF and the KPSS tests are applied to periods beginning at 1950, whether the recent bubble is included or excluded affects the results. Selecting a period that includes the recent house price boom, but not the bust, results in the ADF test being unable to reject a unit root. Moreover, if the period typically used in the literature (starting in 1975) is examined, the ADF test generally cannot reject a unit root at the 5% level. However, these results are reversed if the complete price cycle is included.

Arguably, the US housing market is not a single national market, but is composed of state or MSA level markets. Thus, we also conduct the Pesaran (2007) CIPS panel unit root test using samples of states and MSAS. We also conduct this test in supply elastic and inelastic MSAs given that house price cycles differ between them. We find that, first, real house prices are trend stationary if both the boom and bust periods of the recent price bubble are included, both for the 363 MSA group and the state group. But not including the bust period of this bubble results in non-rejection of a unit root for the 363 MSA group. For the state group, a unit root in real house prices can be rejected if the period ends at 2006, but a unit root cannot be rejected if the period ends at 2002. Second, housing supply elasticity affects unit root test results; specifically, there is stronger evidence of rejecting a unit root in supply inelastic MSAs than elastic MSAs.

We then apply the CIPS panel unit root test to the samples studied in the literature. When our data is restricted to the same areas and time periods, we verify these studies' results. In nearly all cases the result is that a unit root cannot be rejected. But if the dataset is extended to 2011, a unit root is always rejected at a highly significantly level. Thus, the difference in unit root test results between the literature and this paper occur because of the smaller number of time periods studied in the literature.

We also implement the nonlinear unit root test of Kapetanios, Shin and Snell (2003) to consider the case that house prices might be nonlinear, and the test in Enders and Lee (2012) which uses a Fourier expansion to approximate smooth breaks. The evidence of stationarity is still strong under the nonlinear unit root testing procedure. We then carry out the Sequential Panel Selection Method of Chortareas and Kapetanios (2009) to obtain more insights on which individual house price series have stronger evidence of stationarity in the panel.

The unit root tests results in this study have important implications, both empirically and theoretically. Empirically, if US real house prices are indeed trend stationary, then the cointegration analysis and the error correction models widely used in modeling and forecasting US house price dynamics need to be reconsidered. Theoretically, a model should be able to explain the following issues. First, the model should explain the relationships between real income, which has a unit root, and US real house prices, which does not have a unit root. Second, it should explain that the evidence in favor of trend stationarity in real house prices for supply inelastic cities is stronger than for supply elastic cities. Third, although areas with inelastic and elastic housing supply have similar trends in real income, the former group has an upward trend and the latter group has a negative trend in real house prices.

Appendix on Shiller's house price series

Shiller constructs the extended US time series of real house prices by linking together existing house price data from 1890 to 1934 that was derived from Grebler, Blank, and Winnick (1956) with the home-purchase component of the CPI-U from 1953 to 1975. The series next uses the Office of Federal Housing Enterprise Oversight (OFHEO) data from 1975 to 1987 and thereafter uses the Case-Shiller-Weiss index. In the above time series, there is a gap between 1934 and 1953. Shiller constructs an index of house prices for this period using data on the sales price of houses reported in newspapers from five large cities.

The data from Grebler, Blank, and Winnick (1956) is viewed as the best source of data for the 1890–1934 period. It has been used to construct the house price series reported in the Historical Statistics of the United States, Colonial Times to 1970, Part 2, Tables 259–260 (1975). Grebler, Blank, and Winnick (1956) used survey information from 22 cities in the U.S. Department of Commerce's 1937 *Financial Survey of Urban Housing*. At least two cities were located in each of the nine census regions except the East South Central. The survey data included the value of property in 1934, the year of acquisition by the current owner, and the original purchase price. Price increases (measured for the median property) were then determined for the sample period for single family owner-occupied properties. The results were then adjusted for annual depreciation and for structural additions (net depreciation was calculated to be 1.375 percent). Grebler, Blank, and Winnick (1956) compared their time series of house prices to construction costs and found a close conformation of the two indexes.

The most careful evaluation of the Shiller data is by Davis and Heathcote (2007). They construct house price indexes and compare their results with Shiller for the periods 1930–1950, 1950–1960, and 1960–1970. Only in 1970–1980 is there a disagreement as the differences in the earlier periods were only 0.1 percent through 1960 and 0.5 percent from 1960 to 1970. From 1980 on the differences in indexes again were relatively small. They argue that "Shiller's series underestimates true house price growth, especially during the 1970s." However, the estimates of structural breaks in the Shiller price series occur in periods many years from the period of disputed data.

Acknowledgment: We thank the editor and anonymous referee for their valuable comments and suggestions.
Disclaimer: The views expressed in this paper are those of the authors and should not be attributed to Moody's Analytics.

References

Abraham, J., and P. Hendershott. 1996. "Bubbles in Metropolitan Housing Markets." *Journal of Housing Research* 7 (2): 191–207.
Bai, J., and S. Ng. 2004. "A PANIC Attack on Unit Roots and Cointegration." *Econometrica* 72 (4): 1127–1177.
Canarella, G., S. M. Miller, and S. K. Pollard. 2012. "Unit Roots and Structural Change: An Application to US House-Price Indices." *Urban Studies* 49 (3): 757–776.
Capozza, D., and R. Helsley. 1989. "The Fundamentals of Land Prices and Urban Growth." *Journal of Urban Economics* 26 (3): 295–306.
Capozza, D., and R. Helsley. 1990. "The Stochastic City." *Journal of Urban Economics* 28 (2): 187–203.

Capozza, D., P. Hendershott, and C. Mack. 2004. “An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets.” *Real Estate Economics* 32 (1): 1–32.
Carrion-i-Silvestre, J. L., D. Kim, and P. Perron. 2009. “GLS-based Unit Root Tests with Multiple Structural Breaks both under the Null and the Alternative Hypotheses.” *Econometric Theory* 25 (06): 1754–1792.
Chortareas, G., and G. Kapetanios. 2009. “Getting PPP Right: Identifying mean-Reverting Real Exchange Rates in Panels.” *Journal of Banking and Finance* 33: 390–404.
Davis, Morris A., and Jonathan Heathcote. 2007. “The Price and Quantity of Residential Land in the United States.” *Journal of Monetary Economics* 54 (4): 2595–2620.
Elliott, G., T. J. Rothenberg, and J. H. Stock. 1996. “Efficient Tests for an Autoregressive Unit Root.” *Econometrica* 64 (4): 813–836.
Enders, W., and J. Lee. 2012. “A Unit Root Test Using a Fourier Series to Approximate Smooth Breaks.” *Oxford Bulletin of Economics and Statistics* 74 (4): 574–599.
Gallin, J. 2006. “The Long-run Relationship between House Prices and Income: Evidence from Local Housing Markets.” *Real Estate Economics* 34 (3): 417–438.
Glaeser, E., J. Gyourko, and A. Saiz. 2008. “Housing Supply and Housing Bubbles.” Journal of Urban Economics, 64 (2): 198–217.
Grebler, Leo, David M. Blank, and Louis Winnick. 1956. *Capital Formation in Residential Real Estate: Trends and Prospects*. Princeton, NJ: Princeton University Press.
Gyourko, J., and A. Saiz. 2006. “Construction Costs and the Supply of Housing Structure.” *Journal of Regional Science* 46 (4): 661–680.
Historical Statistics of the United States, Colonial Times to 1970, Part 2. 1975. U.S. Bureau of the Census: Washington, DC.
Holly, S., M. H. Pesaran, and T. Yamagata. 2010. “A Spatio-temporal Model of House Prices in the US.” *Journal of Econometrics* 158 (1): 160–173.
Im, K., M. H. Pesaran, and Y. Shin. 2003. “Testing for Unit Roots in Heterogeneous Panels.” *Journal of Econometrics* 115 (1): 53–74.
Kapetanios, G., Y. Shin, and A. Snell. 2003. “Testing for Cointegration in Nonlinear Smooth Transition Error Correction Models.” *Econometric Theory* 22: 279–303.
Lee, J., and M. Strazicich. 2003. “Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks.” *Review of Economics and Statistics* 85 (4):1082–1089.
Levin, A., C. Lin, and C. Chu. 2002. “Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties.” *Journal of Econometrics* 108 (1): 1–24.
Lumsdaine, R. L., and D. H. Papell. 1997. “Multiple Trend Breaks and the Unit Root Hypothesis.” *Review of Economics and Statistics* 79 (2): 212–218.
Maddala, G., and S. Wu. 1999. “A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test.” *Oxford Bulletin of Economics and Statistics* 61 (S1): 631–652.
Malpezzi, S. 1999. “A Simple Error Correction Model of House Prices.” *Journal of Housing Economics* 8 (1): 27–62.
Mikhed, V., and P. Zemcik. 2009. “Do House Prices Reflect Fundamentals? Aggregate and Panel Data Evidence.” *Journal of Housing Economics* 18 (2): 140–149.
Moon, H. R., and B. Perron. 2004. “Testing for a Unit Root in Panels with Dynamic Factors.” *Journal of Econometrics* 122 (1): 81–126.
Ng, S., and P. Perron. 2001. “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power.” *Econometrica* 69 (6): 1519–1554.
Pedroni, P. 1999. “Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors.” *Oxford Bulletin of Economics and Statistics* 61 (S1): 653–670.
Pedroni, P. 2004. “Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis.” *Econometric Theory* 20 (3): 597–625.
Perron, P., and Z. Qu. 2007. “A Simple Modification to Improve the Finite Sample Properties of Ng and Perron’s Unit Root Tests.” *Economics Letters* 94 (1): 12–19.
Pesaran, M. H. 2004. “General Diagnostic Tests for Cross Section Dependence in Panels.” CESifo Working Papers, no. 1233.
Pesaran, M. H. 2006. “Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure.” *Econometrica* 74 (4): 967–1012.
Pesaran, M. H. 2007. “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence.” *Journal of Applied Econometrics* 22 (2): 265–312.
Pesaran, M. H. 2015. “Testing Weak Cross-sectional Dependence in Large Panels.” *Econometric Reviews* 34 (6–10): 1089–1117.
Pesaran, M. H., L. V. Smith, and T. Yamagata. 2013. “Panel Unit Root Tests in the Presence of A Multifactor Error Structure.” *Journal of Econometrics* 175 (2): 94–115.
Saiz, A. 2010. “The Geographic Determinants of Housing Supply.” *The Quarterly Journal of Economics* 125 (3): 1253–1296.
Shiller, R. 2005. *Irrational Exuberance*, 2nd ed., Princeton University Press, Princeton, NJ, USA.

Supplemental Material: The online version of this article (DOI: 10.1515/snde-2013-0134) offers supplementary material, available to authorized users.