G Model
QUAECO-626; No. of Pages 12
ARTICLE IN PRESS

The Quarterly Review of Economics and Finance xxx (2010) xxx–xxx

Contents lists available at ScienceDirect

The Quarterly Review of Economics and Finance

journal homepage: www.elsevier.com/locate/qref

Was there a U.S. house price bubble? An econometric analysis using national and regional panel data

Steven P. Clark[a], T. Daniel Coggin[b,*]

[a] Belk College of Business, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
[b] Horizon Investments, LLC, Charlotte, NC 28277, USA

ARTICLE INFO

Article history:
Received 9 December 2009
Received in revised form 2 September 2010
Accepted 7 December 2010
Available online xxx

JEL classification:
C32
C33
R11

Keywords:
Error correction models
Cointegration
U.S. house price bubble
Regional U.S. house prices

ABSTRACT

The purpose of this study is to examine the existence of a U.S. house price bubble. Specifically, we focus on the time series statistical relationship between real U.S. and regional house prices and a number of fundamental economic variables related to house prices using quarterly data from the first quarter of 1975 through the second-quarter of 2005, the approximate end of the recent house price rise. We find that U.S. house prices and our fundamental economic variables are unit root variables that are not cointegrated, even after allowing for structural breaks. Thus our analysis confirms the existence of an interesting and important anomaly suggested by some prior research on this period. We then discuss the implications of our results for the common practice of using error correction models of house prices, and for the current policy debate regarding the causes of the recent U.S. housing market collapse.

1. Introduction

> "... Although we certainly cannot rule out home price declines, especially in some local markets, these declines, were they to occur, likely would not have substantial macroeconomic implications."
> Alan Greenspan, Fed Chairman, testimony before the Joint Economic Committee, U.S. Congress, June 9, 2005.

> "... You're not going to see the collapse [of the housing market] that you see when people talk about a bubble ... we have, I think, an excessive degree of concern right now about homeownership and its role in the economy ... there is not the degree of leverage that we have seen elsewhere. This is not the dot com situation."
> Barney Frank (D, MA), Chairman of U.S. House Financial Services Committee, on the floor of the House recognizing National Homeownership Month, June 27, 2005.

Bubble Trouble? Not Likely
Chris Mayer and Todd Sinai, *Wall Street Journal*, September 19, 2005.

> "... House prices have risen by nearly 25 percent over the past two years. Although speculative activity has increased in some areas, at a national level these price increases largely reflect strong economic fundamentals, including robust growth in jobs and incomes, low mortgage rates, steady rates of household formation, and factors that limit the expansion of housing supply in some areas. House prices are unlikely to continue rising at current rates. However, as reflected in many private-sector forecasts such as the Blue Chip forecast mentioned earlier, a moderate cooling in the housing market, should one occur, would not be inconsistent with the economy continuing to grow at or near its potential next year."
> Ben Bernanke, Chairman, President's Council of Economic Advisers, testimony before the Joint Economic Committee, U.S. Congress, October 20, 2005.

> "... There's too much money chasing too few good deals, with too much debt and too few brains ... That's why I am getting out."

* Corresponding author. Tel.: +1 704 926 2071.
E-mail addresses: dcoggin@horizoninvestments.com, dcoggin@bellsouth.net (T.D. Coggin).

1062-9769/$ – see front matter © 2010 The Board of Trustees of the University of Illinois. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.qref.2010.12.001

Please cite this article in press as: Clark, S. P., & Coggin, T.D. Was there a U.S. house price bubble? An econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance* (2011), doi:10.1016/j.qref.2010.12.001

Tom Barrack, real estate investor, *Fortune* magazine, October 31, 2005.

> "...There is virtually no risk of a national housing price bubble, based on the fundamental demand for housing and predictable economic factors."
> *Real Estate Insights*, December 2005, National Association of Realtors.

John Paulson Bet Big on Drop in Housing Values
Gregory Zuckerman, Trader Made Millions on Subprime, *Wall Street Journal*, January 15, 2008.

> "... We had a housing bubble; that's now obvious."
> Lessons From the Housing Bubble, *Wall Street Journal*, May 29, 2008.

> "... It is the biggest housing bubble in U.S. history."
> Robert Shiller, video interview with Peter Coy of *BusinessWeek* posted at businessweek.com, September 26, 2008.

As the chronology of quotes and references listed above illustrates, not surprisingly, much has been written and commented upon regarding the existence (and consequences) of a U.S. "house price bubble." Some additional examples of academic and institutional research include studies and commentary of McCarthy and Peach (2004), Himmelberg, Mayer and Sinai (2005), Shiller (2005), Casse and Shiller (2003, 2006), Goodman and Thibodeau (2008), and Mikhed and Zemcik (2009). As suggested above, near the height of the current house price rise in December 2005, many analysts and commentators (and some academics and government officials) were suggesting the *absence* of a bubble. However, by early to mid 2008 the "media consensus" verdict was in. While there were *some* pockets of disagreement, there seemed to be general agreement in the financial media and elsewhere that the U.S. indeed experienced a house price bubble. Financial market "bubbles" can be defined in several different ways. Recent summaries are given in Brunnermeier (2008) and Pollock (2008). In the econophysics literature, Zhou and Sornette (2006) have proposed and tested a definition that defines a bubble as "faster-than-exponential" growth in price. In the economics literature Homm and Breitung (2010) and Phillips, Wu and Yu (in press) define a bubble as an "explosive root process," meaning a time series process that is integrated of order greater than 1.0. In this paper, we shall focus on the existence of what we will refer to as the "classic financial bubble model."

As the beginning quotes of our study suggest, the idea of a U.S. house price bubble is not new. However, the academic literature on this subject *is* relatively new. In studies based on approximately the same time period as ours, some have questioned the existence of a bubble [see, e.g., Himmelberg et al. (2005), and Smith and Smith (2006)]; while some other studies have presented evidence supporting a bubble [see, e.g., Zhou and Sornette (2006), Glaeser, Gyuourko and Saiz (2008) and Mikhed and Zemcik (2009)]. Mixed results are found by Goodman and Thibodeau (2008), Holly, Pesaran and Yamagata (2010), and Clark and Coggin (2009) who estimated a simple error correction model to test if U.S. regional house prices are converging. Interestingly, Zhou and Sornette (2003) find evidence of a British house price bubble; while Cameron, Muellbauer and Murphy (2006) (using a more traditional economic fundamentals-based model) find no such evidence.

Our study adds to this literature in several ways. First, we focus on the time period we now know includes the start of the rise in U.S. house prices up until its peak. Specifically, our data begin with the first quarter of 1975 and end with the second-quarter of 2005. If there were a bubble, then it would be observed during this period. Second, our study uses one of the largest sets of potential explanatory variables for the existence of a U.S. house price bubble of any study we have yet seen. Third, our analysis specifically accounts for structural breaks when testing for both unit roots and cointegration. Structural breaks are well known to confound the results of statistical tests that ignore them. Fourth, we use panel data. This fourth point requires some elaboration.

Another recent study using U.S. housing market panel data is Mikhed and Zemcik (2009), who confirm the existence of a bubble. We chose not to apply their panel econometric tests in our data for 3 main reasons. One, while the panel unit root and cointegration tests in Mikhed and Zemcik (2009) do allow cross-sectional dependence, they do *not* allow structural breaks which are known to potentially weaken or invalidate such tests. While there are constant improvements in this area, at this writing we know of *no* panel data unit root or cointegration tests that explicitly allow *both* cross-sectional dependence *and* structural breaks. Our individual-level tests are technically immune to the cross-sectional dependence problem and explicitly allow structural breaks. Two, our sample data include panel data (e.g., regional house prices) *and* non-panel data (e.g., national mortgage rates) data, which would potentially complicate a panel data test. Three, the null hypothesis of most existing panel unit root/cointegration tests is "*all* panel members of the panel contain a unit root/are cointegrated" versus the alternative of "at least *one* panel member/pair does *not* have a unit root/is not cointegrated." Thus if the null of panel cointegration *were* rejected (as we suggest it is likely to be in these data), then *individual* (non-panel) tests would be *still* be required. We therefore feel that the individual (regional) tests we provide will be more informative in our data at this time.

Our discussion is organized as follows. Section 2 presents the classic financial bubble model and suggests 2 varieties. Section 3 describes our data and time series statistical methodology, including the unit root and cointegration tests. Section 4 presents our results suggesting that U.S. house prices and our fundamental economic variables are unit root variables that are not cointegrated, even after allowing for structural breaks. Section 5 summarizes our findings and discusses some regulatory factors that likely contributed to this anomalous result.

2. The classic financial bubble model in 2 varieties

2.1. The fundamental value model

The classic (economic) fundamental value model of a financial bubble seeks to *directly* estimate the *fundamental value* of the asset and then compare it to the *observed value* in the market. Recent examples of this approach to the U.S. housing market include Himmelberg et al. (2005), and Smith and Smith (2006). Himmelberg, Mayer and Sinai (hereafter HMS) present a model of house prices that departs from conventional wisdom in that it focuses on the *user cost of housing*. Specifically, HMS argue that "... The key mistake committed by the conventional measures of overheating in housing markets is that they erroneously treat the purchase of a house as if it were the same as the annual cost of owning" (HMS, p. 74). They then proceed to develop a model of the annual cost of (home) ownership that assumes:

> "... A correct calculation of the financial return associated with an owner-occupied property compares the value of living in that property for a year—the 'imputed rent,' or what it would have cost to rent an equivalent property—with the lost income that one would have received if the owner had invested in an alternative capital investment—the 'opportunity cost of capital.' This comparison should take into account differences in risk, tax benefits from owner-occupancy, property taxes, maintenance

expenses and any anticipated capital gains from owning the home" (p. 74).

As HMS note, using this line of reasoning one can compare the true annual cost of owning a house (also known as the "imputed rent") to the observed ("actual") rental cost to judge whether the cost of owning a home is excessive (or cheap) relative to rental cost. A "house price bubble" thus occurs when homeowners (or potential homeowners) have unreasonably high expectations about the future value of a house relative to the current cost, which causes them to "pay too much" for the house today. HMS then propose a formula for the annual cost of homeownership (imputed rent), based on the prior work of Hendershott and Slemrod (1983), and Poterba (1991). The applications of the fundamental value model to the U.S. housing market for the most recent period [including the papers by HMS, and Smith and Smith (2006)] have tended to question the existence of a *nationwide* house price bubble, while suggesting one may exist in a few *local* markets.

2.2. The cointegrating regression model

A recent front page article in the *Wall Street Journal* illustrates the idea of a variation on the classic financial bubble model we will refer to as "the cointegrating regression model." In this *WSJ* article, there is the following passage concerning real estate investor John Paulson, who made an estimated $15 billion betting that home prices were overvalued in 2007 (Zuckerman, 2008):

> "[Mr. Paulson] . . . 'This is crazy . . . Where is the bubble we can short?' . . . Analyzing reams of data late at night in his office, Mr. Paulson became convinced investors were far underestimating the risk in the mortgage market. In betting on it to crumble, 'I've never been involved in a trade that had such unlimited upside with a very limited downside . . ..''

Mr. Paulson noted he had analyzed "reams of data" to come to the conclusion that house prices were way ahead of economic fundamentals. That is, Paulson seemed to be suggesting that any prior relationship between economic fundamental data and house prices had disappeared. Following this basic logic, we will present a generally accepted model of house prices that suggests what fundamental economic variables might reasonably be considered to "determine" the level of prices.

A number of previous studies have suggested a simple "supply and demand model" for housing. We follow here the concise description of this model given in Gallin (2006) in his earlier study of house prices and income. Specifically, the demand for housing is a function of several primary variables such as income (Y), population (N), wealth (W), the user cost of housing (UC),[1] mortgage interest rates (M), unemployment (U) and other "demand shifters" (θ_d). The supply of housing depends on the price of housing (P), the cost of construction (C), cost of land (L) and the other "supply shifters" (θ_s):

$$Q_d = D(Y, N, W, UC, M, U; \theta_d) \quad (1)$$

$$Q_s = S(P, C, L; \theta_s) \quad (2)$$

Thus the price of housing (P) can thus be rewritten as a function of all variables in the model:

$$P = f(Y, N, W, UC, M, U, C, L; \theta_d, \theta_s) \quad (3)$$

[1] User cost of housing is defined as: UC = P[$(1 - \tau_y)(M + \tau_p) + \delta$-cg], where P = price of housing, M = mortgage interest, τ_y and τ_p are income and property taxes, δ is maintenance and depreciation, and cg is capital gain.

A log-linear specification of this model relates the log of house price to the log of all the driving variables. Note that in this specification, all variables are measured in *levels* (or *rates* for mortgage interest and unemployment). Assuming demand shocks (such as population and income) are nonstationary, we can reasonably assume that house prices will be nonstationary as well. Further assuming constant coefficients and stationary unobserved components (such as trends and cycles), house prices will be *cointegrated* with (unit root) fundamentals, and the strength of the relationship will depend on the elasticities of supply and demand. In this study, we will actually *test* a cointegrating model more fully described below.

Following the logic above, one can test for the existence of a bubble by specifying a model such as described in Eq. (3) above and then conduct an empirical test of the predicted relationships. This is commonly done using a multiple regression model, although other specifications are possible. As noted by Gallin (2006), whether or not house prices are in fact related to *any* theorized fundamental variables in *any* way is an empirical question. In this study we specifically focus on a few of the most important and timely of them, and look for the existence of a (possibly) cointegrating relationship with real house prices.

If 2 (or more) unit root time series variables are cointegrated, then there exists a linear combination of them that is stationary, implying convergence to a (stationary) equilibrium relationship. Nonlinear specifications are also possible. The existence of such a cointegrating relationship implies the *absence* of a bubble. If 2 (or more) unit root time series variables are not cointegrated, then they will continue on their nonstationary time path of divergence. Thus in the cointegrating regression model, *the absence of an equilibrium relationship* would imply the *presence* of a bubble. Examples of this approach include the classic paper by Diba and Grossman (1988) for stock prices, and a more recent paper by Mikhed and Zemcik (2009) for U.S. house prices.

3. Data and methodology

3.1. Data

3.1.1. U.S. house price indexes

The (nominal) U.S house price index data are taken from the U.S. Federal Housing Finance Agency (FHFA). The FHFA was formed by a merger of the Office of Federal Housing Enterprise Oversight (OFHEO) with the Federal Housing Finance Board (FHFB) in 2008. The index data are quarterly [not seasonally adjusted, 1980 = 100, calculated using the Case–Shiller geometric-weighted repeat sales procedure fully described in Calhoun (1996)]. Our sample data span the period 1975(Q1) through 2005(Q2) [n = 122 quarters], are expressed in June 2005 dollars using the Consumer Price Index (all urban consumers, U.S. city average, *all items less shelter*, not seasonally adjusted, 1982–84 = 100) and converted to natural logarithms for the statistical analysis. By design, our sample data period ends very near the end of the current U.S. house price rise.

We deliberately focus on this *specific* time period for 4 main reasons. One, it is beyond dispute that the period leading up and through the end of the dramatic rise in U.S. house prices is a topic of vital concern to academics and policy makers trying to understand the causes of the rise and the economic carnage that followed the collapse. Two, consistent with this concern is an ongoing discussion of the specific relationship between house prices and fundamental economic variables during this period in trying to explain why the housing market subsequently collapsed. Simply put, if there were a U.S. house price bubble, then *this* is when it happened. Three, while the *aftermath* of the bubble period is of immense importance; that

Please cite this article in press as: Clark, S. P., & Coggin, T.D. Was there a U.S. house price bubble? An econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance* (2011), doi:10.1016/j.qref.2010.12.001

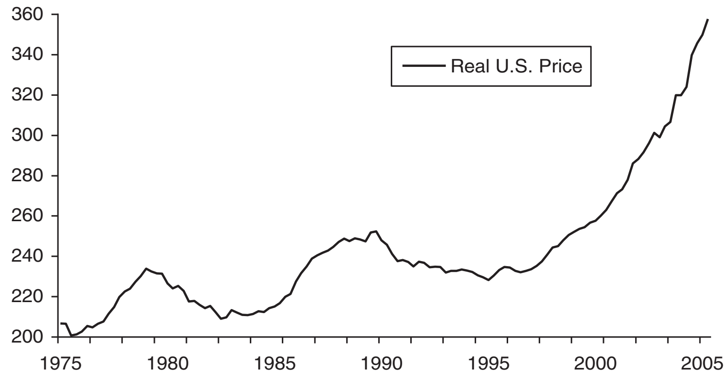

Fig. 1. Real U.S. house price index (in $ thousands).

Table 1
Definitions of super-regions.

Super-regions	FHFA regions
Northeast	New England
	Middle Atlantic
Midwest	East North Central
	West North Central
South	South Atlantic
	East South Central
	West South Central
West	Mountain
	Pacific

is *not* the focus of this study. We focus here on the existence of a bubble and the period leading up to it. Four, there is a purely statistical argument for focusing on this specific period. As noted by Homm and Breitung (2010), extending data past the end point of a financial bubble decreases the power of statistical tests used to detect them.

The FHFA (panel) house price data are supplied in the form of 9 regions plus the U.S. The states within each FHFA region are presented in Appendix A. As an informational note, the Census Bureau refers to these 9 regions as "divisions." In addition to the U.S., we further grouped the FHFA regions into 4 "super-regions" based on the 4 regions used by the U.S. Census Bureau to collect and analyze income and other demographic variables. The U.S. Census Bureau does not classify demographic data according to the FHFA regions. However they do classify data into 4 regional categories into which the FHFA 9 regions can be grouped. Thus the home price indexes, median family income and other economic data can be based on exactly the same geographical units. We calculated simple averages of the individual FHFA regional house price indexes to create the super-regions defined in Table 1.[2] We will hereafter refer to them as "regions."

The FHFA indexes (and their predecessor, the OFHEO indexes) have been used in the majority of prior studies of U.S. house prices and should be distinguished from the S&P/Case–Shiller (panel) house price indexes on which futures are traded on the Chicago Mercantile Exchange. The FHFA indexes rely on data collected by Fannie Mae and Freddie Mac, which were placed into a conservatorship run by the FHFA in 2008. Thus they exclude loans that are either too big (at this writing >$417,000) or too risky for Fannie Mae and Freddie Mac to guarantee. The Case–Shiller indexes include these data, but are currently limited to 20 major U.S. housing markets. As we will note in our Discussion, the definition of "too risky" used by Fannie and Freddie is *not* unproblematic, and may have actually contributed to the creation of a housing bubble. Up until December 2008 the FHFA indexes were included in data used by the Federal Reserve to estimate household wealth, while the Case–Shiller indexes are increasingly used as the basis for futures contracts on the U.S. housing market. For a more complete discussion of these indexes, see Drieman and Pennington-Cross (2004), and Wessel (2008).

3.1.2. U.S. and regional house price growth rates

We acknowledge that U.S. house prices are known to vary by region, state, city and even neighborhood.[3] We shall focus in this study on *regional* U.S. house prices. Figs. 1 and 2 present graphs of the (non-logged) real U.S. and regional house price indexes. The dramatic and sustained price rise beginning in the mid-1990s is clearly evident in these regional graphs. As a prelude to our main analysis and to gain more insight into the nature of the rise in U.S house prices, we also calculated both U.S. and regional growth rates for the entire period of our study and for 2 subperiods. The graphs in Figs. 1 and 2 and our unit root tests (discussed later) suggest a structural break in our data at about 1996(Q1). Therefore we used this breakpoint to define the 2 subperiods: Q1/1975–Q4/1995 and Q1/1996–Q2/2005. Table 2 presents the results of the growth rate calculations.

We can summarize the key features of Table 2 as follows. First of all, growth rates in real house prices rise uniformly from subperiod 1 to subperiod 2, and the increase is substantial. Second, with regard to real regional house prices specifically, the West has the highest growth rates in both subperiods, while the South has the lowest. In fact, real house prices in the South actually *declined* in the first subperiod.

3.1.3. Additional economic variables for classic bubble model

The list of fundamental economic variables that can reasonably be included as additional "supply and demand shifters" related to house prices as in Eq. (3) above is large indeed. For recent comprehensive summaries of the U.S. and international literature on this topic, see Labonte (2003), and Claessens, Kose and Terrones (2008). Our primary purpose here is *not* to comprehensively examine such a list. In fact, there appears to be *no* single generally agreed upon set of variables used in testing models of house prices in the literature. We selected several additional fundamental economic variables which are included in this literature *and* for which we

[2] We experimented with various methods to combine the regional house prices into the "super-regions" and found that the simple average performed as well as any of them.

[3] As a representative sample of studies documenting this assertion, see Abraham and Hendershott (1996), Johnes and Hyclack (1999), and Rapach and Strauss (2009).

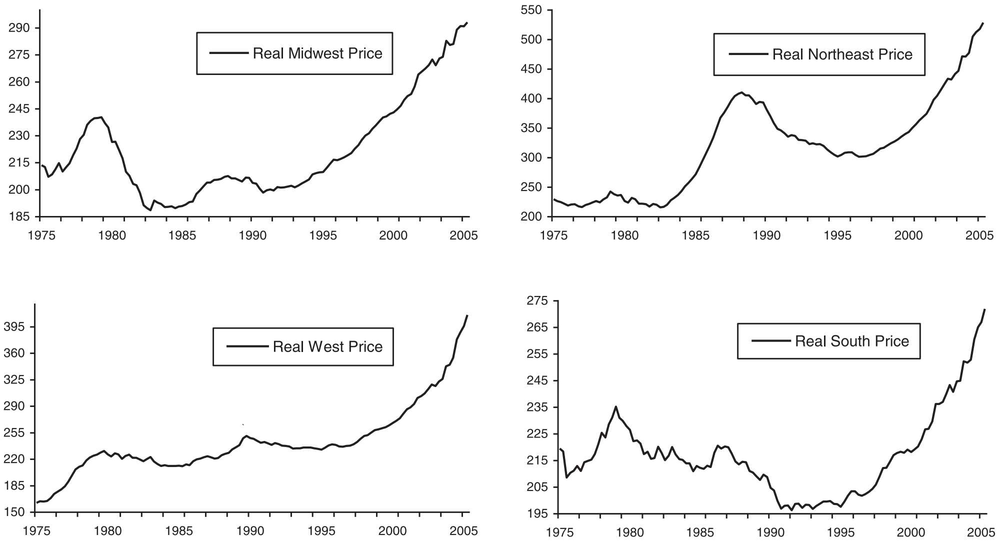

Fig. 2. Real regional U.S. house price indexes (in $ thousands).

Table 2
Quarterly growth rates, by region (in %).

	Real house price indexes				
	U.S.	Northeast	South	Midwest	West
Full period	0.4548	0.6923	0.1773	0.2619	0.7708
Q1/75–Q4/95	0.1536	0.3590	−0.0915	0.0174	0.4720
Q1/96–Q2/05	1.1498	1.4640	0.7881	0.8226	1.4733

could gather both national and regional data (although in some cases, no regional data are available).

Since it is possible that real house prices and *individual* economic variables might *not* be cointegrated in a bivariate model, but *might* be cointegrated with *other* variables in a *multivariate* model, we also grouped some of the additional fundamental economic variables in 2 smaller subgroups. The first group of 3 variables is generally considered to be *demand* related and include variables related to the actual decision to *buy* a house:

Ratio of 30-year fixed mortgage rate to one-year adjustable rate (30yrMortgage/Adjustable)
Ratio of the one-year adjustable rate to the 12-month LIBOR rate (Adjustable/LIBOR)
Ratio of median new house price/owner's equivalent rent (RealHousePrice/Rent).

We will call this group "housing market variables."

The second group includes 2 variables representing the *capital markets* related to house prices:

S&P500 stock index homebuilders industry composite (RealHomebuilders)
Lehman Brothers Mortgage-backed government bond index (RealLBMortgage-backed).

We will call this group "financial market variables." The complete set of additional variables and the rationale for choosing each of them are presented in Table 3. The sources and more detailed descriptions of *all* variables are given in Appendix B.

3.2. *Time series statistical methodology*

3.2.1. *Unit root tests*

In our empirical analysis, we apply 2 types of time series statistical tests: unit root tests and cointegration tests. We use the augmented Dickey–Fuller (ADF) test of Dickey and Fuller (1979) and Said and Dickey (1984) to test the null hypothesis that a time series y_t contains a unit root after allowing for autocorrelation in the time series regression errors with *j*-lags of the changes in y_t:

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum \gamma_j \Delta y_{t-j} + \varepsilon_t \quad (4)$$

In our analysis the maximum number of lags allowed is 4, chosen by the AIC criterion. A time trend is included in the structural break tests discussed below. The ADF test is a *t*-test on ($\alpha_1 = 0$) the coefficient for the lag of *y*. A significant ADF *t*-test indicates that y_t is stationary (i.e., rejects the presence of a unit root).

If a time series contains a unit root, then it is nonstationary, non-mean reverting, and essentially nonpredictable using only past values of the series itself. However, it is now well known that *structural breaks* such as economic depressions, tax cuts/hikes, or energy price shocks can spuriously cause the appearance of a unit root. For this reason, we also apply the LM (Lagrange Multiplier) unit root *t*-test of Lee and Strazicich (2003, 2004), which allows a single endogenous (unknown) break. We apply their Model C, the most general model, which allows a single break in both the test regression intercept and slope at the same point in time. Lee and Strazicich note that failure to allow for a structural break under the null [as in the test of Zivot and Andrews (1992)] can distort the size of the unit root *t*-test. Their LM unit root test avoids this problem by allowing a structural break under both the unit root null and the alternative.

Our study focuses on the existence of a classic financial bubble commencing with the dramatic price rise in the mid-1990s. This implies the existence of a *single* structural break in our 122 quar-

Please cite this article in press as: Clark, S. P., & Coggin, T.D. Was there a U.S. house price bubble? An econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance* (2011), doi:10.1016/j.qref.2010.12.001

Table 3
Additional economic variables.

Variable	Availability	Rationale
Real house price index	U.S. and regional	Real U.S. and regional house price indexes are indicators of the level of local house prices
30-year mortgage rate	U.S.	Long-term mortgage rate is an indicator of the cost of borrowing to finance a house
30-Year Mortgage Rate/Adjustable rate	U.S.	Ratio of 30-year mortgage rate to adjustable rate is an indicator of level of attractiveness of adjustable rate mortgages relative to fixed rate
Adjustable Mortgage Rate/LIBOR	U.S.	Ratio of adjustable mortgage rate to 12-month LIBOR rate is an indicator of the effect of LIBOR rate on adjustable rate. Adjustable rate mortgages are often pegged to the LIBOR rate
Unemployment	U.S. and regional	Unemployment is an indicator of strength of the local job market and economy
Household Debt/Income	U.S.	Household debt/median family income is an indicator of level of household debt and thus strength of family balance sheets
Real Median Family Income	U.S. and regional	Real median family income is an indicator of wealth available to purchase a house
New House Price/Median Family Income	U.S. and regional	Ratio of new house price (first time buyer) to median family income is an indicator of local housing affordability
New houses for sale/New houses sold	U.S. and regional	Ratio of new houses for sale/new houses sold is an indicator of strength of local housing market
Homeowner rate	U.S. and regional	Homeowner rate is an indicator of strength of local housing market and economy
Real house price/Rent	U.S. and regional	Real house price/rent is an indicator of the attractiveness of home-ownership versus renting
Real residential fixed investment/GDP	U.S.	Level of residential investment to GDP suggested by Shiller (2006) is an indicator of level of residential investment to total investment and thus strength of the residential housing market
Real homebuilders stock index	U.S.	Real S&P Homebuilders stock index is an indicator of the level and strength of homebuilder stocks
Real Lehman mortgage-backed securities index	U.S.	Real Lehman mortgage-backed securities index is an indicator of the level and strength of the mortgage bond market

ters of data. It is possible to test for 2 or more breaks. However, in a dataset of this size, the size of the separate time periods then potentially becomes less than 40 quarters. We believe testing for unit roots and cointegration with such small sub-periods is not advisable. Furthermore, we feel the single break in the mid-1990s is clearly of greatest importance in our data.

3.2.2. Cointegration tests

As noted above, it is well known that what appears to be a real over-time relationship between 2 unit root (or more) variables can be entirely spurious. The Engle–Granger (1987) regression model shows that if 2 unit root time series variables are *cointegrated*, then there exists a linear combination of the 2 that is stationary with mean, variance and autocorrelation that is independent of time:

$$y_1 = \alpha_0 + \alpha_1 y_2 + \varepsilon_t \tag{5}$$

That is, if 2 unit root variables are cointegrated (y_1 and y_2 here), then they cannot move independently of each other for long. The null hypothesis of the Engle–Granger model is no cointegration (i.e., a unit root in the regression model residuals) versus the alternative that the 2 variables are cointegrated (i.e., the residuals are stationary). The Engle–Granger statistical test for cointegration applies a unit root test (the Dickey–Fuller ADF *t*-test here) to the regression model residuals. We note that except in special cases that are beyond the scope of this paper, cointegration tests assume that *all* variables contain a unit root.[4] We shall make that standard assumption here.

Just as we need to be concerned about the presence of structural breaks in unit root tests, we should also be concerned in cointegration tests. The original, single-equation (2-step) residual-based Engle–Granger cointegration model did not include a structural break. Gregory and Hansen (1996) derived a version of their model that allows a single endogenous structural break in which the original (no break) Engle–Granger specification is a special case. We apply their Model 4, the most general case, which allows a single break in the model regression intercept and slope occurring at the same point in time. The Gregory–Hansen model tests a null hypothesis of no cointegration (i.e., a unit root in the model regression residuals) versus an alternative hypothesis of cointegration with a single structural break, and applies the Dickey–Fuller ADF *t*-test to the residuals from the cointegrating regression model.

In addition to the Gregory and Hansen (1996) single-equation model, we also apply the recent *system equation model* of Qu (2007). It is well known that single-equation models of cointegrating relationships can be sensitive to the choice of dependent and independent variables. This is known in econometrics as the *normalization problem*. The vector-autoregression (VAR), system-of-equations model of Johansen (1991) effectively solves this problem by estimating *all* possible cointegrating relationships *at once*. Qu (2007) extends this basic model to allow an unknown number of structural breaks at unknown times. Specifically, Qu (2007) presents a series of nonparametric tests for a null hypothesis of r_0 cointegrating vectors (where r_0 can be zero) against the alternative of more than r_0 cointegrating vectors existing in some sub-sample of the time series data. In our example, we will test the null of zero cointegrating vectors against an alternative of segmented cointegration. When the number of breaks is unknown, Qu recommends using the SQ and WQ tests which are based on the eigenvalues and ranks of the sub-sample moment matrices. These tests are fully described in Qu (2007). In our case, in order to err on the side of being conservative in our tests, we will assume the there are up to 2 endogenous breaks in our data.[5]

[4] For example, Phillips (1995) presents the FM-VAR cointegration model that allows both I(1) and I(0) variables.

[5] As noted by Qu (2007, p. 586), "... constructing tests using $m = 2$ is sufficient to generate consistency for any unknown number of regimes."

Table 4
Unit root tests for all variables, by region.

	U.S.		Midwest		Northeast		South		West	
Variable	DF-ADF	LS-ADF	DF-ADF	LS-ADF	DF-ADF	LS-ADF	DF-ADF	LS-ADF	DF-ADF	LS-ADF
Real house price index	0.62	−4.08	−0.68	−3.94	−1.29	−3.15	0.35	−4.26	0.20	−3.18
Real 30yrMortgage	−1.96	**−4.74***	na	na	na	na	na	na	na	na
Nominal 30yrMortgage	−0.67	−4.30	na	na	na	na	na	na	na	na
30yrMortgage/Adjustable	**−3.56***	−4.46	na	na	na	na	na	na	na	na
Adjustable/LIBOR	−2.48	−4.50	na	na	na	na	na	na	na	na
Unemployment	−2.36	−3.67	−1.60	−2.97	−2.86	−3.78	**−3.00***	−3.91	**−2.97***	−4.47
House Holddebt/Income	0.95	−3.06	na	na	na	na	na	na	na	na
Real median family income	−1.56	−4.33	−1.15	−3.37	−0.66	−2.52	−0.98	−3.64	−1.14	−2.75
New house price/Income	0.18	−2.62	−3.51	**−5.41***	0.31	−3.18	0.62	−3.61	−0.14	−3.62
ForSale/Sold	−1.82	−4.37	−2.44	−3.54	−2.01	−3.37	−1.69	**−4.64***	−1.68	−4.28
Homeowner rate	0.07	−2.50	−0.46	−2.86	−0.84	**−4.68***	−0.38	−2.33	−0.23	−3.73
Real house price/Rent	−2.45	−2.72	**−3.37***	−2.84	**−3.53***	−3.29	**−3.39***	−2.17	**−3.06***	−3.29
Real residential investment/GDP	−2.82	−4.19	na	na	na	na	na	na	na	na
Real homebuilders index	−0.10	−4.23	na	na	na	na	na	na	na	na
Real LBMortgage-backed index	−0.17	−4.22	na	na	na	na	na	na	na	na
95% C.V.	−2.90	−4.50	−2.90	−4.50	−2.90	−4.50	−2.90	−4.50	−2.90	−4.50

Note: "na" denotes regional data not available, so national variable used in all subsequent tables.
* Denotes reject unit root null at 0.05 level in bold type. DF-ADF test regression has constant term only. Maximum number of lags = 4. LS-ADF tests use their Model C (single break in slope and intercept at same time).

3.2.3. *Cointegration and error correction models*

A number of early studies suggested the presence of both serial correlation and mean reversion in house prices. This literature is summarized in Capozza, Mack and Mayer (1997). Specifically, as derived and discussed by Capozza et al. (1997), the interaction of the 2 allow the existence of convergence and cycles in regional house prices. To facilitate their discussion, they introduced a simple *error correction* model:

$$\Delta P_t = \alpha\Delta P_{t-1} + \beta(P_{t-1} - P^*_{t-1}) + \gamma\Delta P^*_t, \quad (6)$$

where P_t denotes log of real housing price at time t, α denotes the degree of serial correlation, β denotes the degree of mean reversion, P^* denotes the fundamental value determined by economic conditions, and γ is the contemporaneous adjustment of prices to current shocks. This model is typical and broadly representative of many error correction models for house prices in this literature. As discussed by Meen (2002), models of the U.S. housing market have used both the error correction model [e.g., Abraham and Hendershott (1996), and Malpezzi (1999)] and the standard time series regression model [e.g., DiPasquale and Wheaton (1994)] to specify the relationship between real house prices and key micro- and macroeconomic variables.

In applying an error correction model, one must *first* ascertain whether or not key variables in the model are stationary or nonstationary (unit root). This is an important question resulting from the *Granger Representation Theorem* [Engle and Granger (1987), and Johansen (1991)]. This theorem can be summarized as stating that, when dealing with 2 (or more) nonstationary variables, cointegration and error correction are statistically *equivalent* (i.e., *if* one *then* the other). Thus, when dealing with time series variables, specifying and testing an error correction model should generally be preceded by testing the variables for nonstationarity. As we noted earlier, time series regressions of *any* type that include nonstationary regressors may be *spurious*. Error correction models are not immune to this problem. That is, even though error correction models are specified in first differences, the error correction term is *only* valid for variables that are stationary (in levels) *or* cointegrated (in levels).

As first noted by Gallin (2006), this result has *not* always been recognized in the housing literature. We will first test all variables for stationarity. We will then apply tests for cointegration where appropriate. If we find cointegration, then we can apply an error correction model to estimate the magnitude of the error correction term. If not, then error correction models would *not* be appropriate for these nonstationary, non-cointegrated variables.

4. Results

4.1. *Unit root tests*

Table 4 presents the results of the unit root tests for all national and regional variables. With respect to the U.S. variables, the DF-ADF and LS-ADF tests suggest they all contain unit roots except the LS-ADF test for Real 30yr Mortgage Rate and the DF-ADF test for 30yr Mortgage/Adjustable Rate Mortgage. As a result we will include Nominal 30yr Mortgage Rate in all subsequent cointegration statistical tests and omit Real 30yr Mortgage. We will also omit 30yr Mortgage/Adjustable Rate Mortgage from subsequent cointegration tests that do not allow structural breaks. With respect to the regional variables, we will omit Midwest New House Price/Income and Northeast Homeowner Rate from all subsequent cointegration tests. We will also omit Midwest, Northeast, South and West Unemployment, Midwest-Real House Price/Rent. Variables for which no regional data are available or do not have a unit root are denoted "na" in all subsequent tables.

4.2. *Cointegration tests*

4.2.1. *Bivariate tests*

Table 5 presents the results of the bivariate cointegrating regression tests including all the economic variables separately. In these tests, in each case the dependent variable is Real (U.S. or regional) House Price Index. In Table 5 we see that for both for the classic Engle–Granger cointegrating regression model and the Gregory–Hansen model (which allows a single endogenous break in both the regression intercept and slope) we cannot reject the null hypothesis of "no cointegration" in *any* test.

4.2.2. *Multivariate tests using housing market variables*

As we noted earlier, it is possible that while individual variables might not be cointegrated with real house prices, linear combinations of them might be. For this reason, we also combined the individual economic variables into 2 subgroups described earlier. In Table 6 we present the multivariate cointegration test

Table 5
Bivariate cointegration tests for all variables, by region.

Variable	U.S. Engle–Granger	Gregory–Hansen	Midwest Engle–Granger	Gregory–Hansen	Northeast Engle–Granger	Gregory–Hansen	South Engle–Granger	Gregory–Hansen	West Engle–Granger	Gregory–Hansen
Nominal 30yr/Mortgage	−1.32	−3.91	−1.42	−2.80	−1.51	−2.58	−0.43	−3.71	−1.89	−3.97
30yr/Adjustable	na	−3.98	−0.30	−2.85	−0.69	−2.31	0.46	−2.55	1.05	−3.41
Adjustable/LIBOR	−1.08	−3.42	−0.23	−2.62	−1.40	−2.44	−0.74	−2.95	−1.29	−3.12
Unemployment	−0.08	−4.22	−0.35	−2.54	−1.41	−1.96	na	−2.14	na	−4.06
Householddebt/Income	−0.97	−4.36	na	na	na	na	na	na	na	na
Realmedian family income	−1.44	−3.14	−1.41	−2.56	−1.27	−1.99	−0.14	−2.52	−1.28	−2.62
Newhouse price/Income	−1.27	−3.58	na	na	−1.68	−3.32	−1.11	−2.60	−2.76	−4.04
For sale/Sold	−1.42	−3.34	−0.77	−3.02	−1.54	−2.38	−0.02	NA	−2.24	−3.93
Homeowner rate	−1.34	−2.87	−2.54	−3.20	na	na	−1.05	−3.80	−2.65	−3.40
Real house price/Rent	0.93	−3.39	na	−3.26	na	−4.89	na	−3.34	na	−2.82
RealResI nvestm ent/GDP	0.20	−3.34	0.31	−3.05	−1.40	−2.42	0.46	−3.24	−0.06	−4.58
RealHomeBuilders	−2.51	−2.93	−1.81	−3.63	−1.58	−2.80	−0.81	−4.00	−3.81*	−4.36
RealLehman Mort-backed	−1.00	−4.40	−1.41	−3.34	−2.60	−3.60	−0.19	−3.81	−1.25	−4.50
95% C.V.	−3.39	−4.95	−3.39	−4.95	−3.39	−4.95	−3.39	−4.95	−3.39	−4.95

Note: Dependent variable in Engle-Granger and Gregory-Hansen tests is *Real House Price Index*. “na” denotes regional data not available (*HouseholdDebt/Income*) or cointegration test not appropriate due to absence of a unit root for independent variable. All tests in this table fail to reject no cointegration null.

Table 6
Cointegration tests for housing market variables, by region.

Variables		U.S		Midwest		Northeast		South		West	
		Engle–Granger	Gregory–Hansen	Engle–Granger	Gregory–Hansen	Engle–Granger	Gregory–Hansen	Engle–Granger	Gregory–Hansen	Engle–Granger	Gregory–Hansen
Panel 1: Engle–Granger and Gregory–Hansen tests											
30yr/Adjustable											
Adjustable/LIBOR											
Real HousePrice/Rent											
	Test	−1.04	−4.67	−1.63	−3.91	−1.69	−5.04	−1.16	−3.99	−0.99	−3.73
	95%C.V.	−4.19	−6.00	−4.19	−6.00	−4.19	−6.00	−4.19	−6.00	−4.19	−6.00
Panel 2. Qu tests											
Variables		SQ	WQ	SQ	WQ	SQ	WQ	SQ	WQ	SQ	WQ
RealHousePriceIndex											
30yr/Adjustable											
Adjustable/LIBOR											
Real HousePrice/Rent											
	Test	8.81	4.67	8.90	4.47	9.29	4.73	10.52	5.55	8.21	4.14
	95% C.V.	15.47	8.35	15.47	8.35	15.47	8.35	15.47	8.35	15.47	8.35

Note: Dependent variable in Engle–Granger and Gregory-HanSen tests is *Real House Price Index.* All tests in this table fail to reject no cointegration null.

Table 7
Cointegration tests for financial market variables, by region.

Variables	U.S Engle–Granger	U.S Gregory–Hansen	Midwest Engle–Granger	Midwest Gregory–Hansen	Northeast Engle–Granger	Northeast Gregory–Hansen	South Engle–Granger	South Gregory–Hansen	West Engle–Granger	West Gregory–Hansen
Panel 1. Engle–Granger and Gregory–Hansen tests										
RealHomeBuilders										
RealLBMortgage-backed										
	−1.93	−3.78	−1.74	−4.10	−2.06	−3.84	−2.37	−4.12	−3.44	−4.33
95% C.V.	−3.81	−5.50	−3.81	−5.50	−3.81	−5.50	−3.81	−5.50	−3.81	−5.50
Panel 2. Qu tests										
Variables	SQ	WQ	SQ	WQ	SQ	WQ	SQ	WQ	SQ	WQ
RealHomeBuilders										
RealLBMortgage-backed										
	7.22	3.78	6.40	3.42	6.49	3.51	6.72	3.36	7.55	3.95
95% C.V.	11.70	6.41	11.70	6.41	11.70	6.41	11.70	6.41	11.70	6.41

Note: Dependent variable in Engle–Granger and Gregory–Hansen tests is *Real House Price Index*. All tests in this table fail to reject no cointegration null.

results for real (U.S. and regional) house price indexes and the *housing market variables*: 30yr/Adjustable, Adjustable/LIBOR and RealHousePrice/Rent. In Panel 1, where Real (U.S. and regional) House Price is the dependent variable and the set of housing market variables are the independent variables, we see that in *no* case do we reject the null of no cointegration for the Engle–Granger and Gregory–Hansen models. In Panel 2, where we present the results of the Qu (2007) cointegrating VAR model including all variables simultaneously and allowing up to 2 endogenous breaks, we again see that in *no* case can we reject the null of no cointegration.

4.2.3. Multivariate tests using financial market variables

In Table 7 we present the results for the *financial market variables*: RealHomebuilders and RealLBMortgage-backed. Just as with the results for the housing market variables in Table 6, in *no* case in Table 7 (in Panel 1 for the Engle–Granger and Gregory–Hansen tests, and in Panel 2 for the Qu tests) can we reject the null of no cointegration.

5. Summary and discussion

We have presented a detailed time series statistical analysis of the relationship between real U.S. and regional house prices and a number of fundamental economic variables related to house prices over the period 1975 through the second quarter of 2005. We believe this specific time period worthy of serious and intense concern. Hence we deliberately focus on this time period because of its relevance to the efforts of academics and policy makers trying to understand the causes of the precipitous rise leading to the subsequent decline of U.S. house prices, and the economic carnage that followed.

We divided our data into 2 subperiods on the basis of the dramatic rise in house prices that began in the mid-1990s. First, we showed that growth rates in real U.S. and regional house prices *uniformly* increase from the first subperiod to the second, and the increases across regions are substantial. Within regions, the West had the largest increase and the South had the smallest. Second, we showed that real U.S. and regional house prices and the fundamental economic variables we selected are nonstationary, unit root variables, thus supporting a test of the "cointegrating regression model" of a financial bubble for U.S. (national) and regional house prices. We then found that real house prices and our fundamental economic variables are not cointegrated, even after allowing for structural breaks. Thus our detailed quantitative analysis confirms the existence of an interesting and important anomaly suggested by the financial media and some prior academic research on this period. That is, consistent with testing the "cointegrating regression model" of a financial bubble, our finding of *no* cointegrating relationships (and thus *no* convergence to a stable equilibrium relationship) between house prices and *any* of our fundamental economic variables confirms the existence of a U.S. house price bubble for this period.

We also further confirm that, as noted by Gallin (2006), *error correction models* in this literature that attempt to relate house prices to fundamental economic variables are likely to be inappropriate. This is a result of the *Granger Representation Theorem* which states that when dealing with 2 (or more) nonstationary variables, cointegration and error correction are statistically *equivalent* (i.e., *if* one *then* the other). Thus when dealing with time series variables, specifying and testing an error correction model for nonstationary variables in the absence of evidence of cointegration will produce a spurious result.

One interpretation of our finding of a U.S. house price "bubble," combined with similar results of others, is that U.S. house prices during this period are *unrelated* to many commonly suggested variables in the literature. This result is somewhat disconcerting and raises an important question, "Is this result due to the fortuitous combination of social, economic and political conditions surrounding this particular time period, or is there something more fundamental going on?" We have acknowledged that our study does not include every potential, or even plausible, variable that might influence house prices. Therefore we feel it is reasonable to ask, "Do we need more comprehensive models of house prices that include additional potentially causal variables?" Put more descriptively, "Is is possible for us to formulate better substantive models using more appropriate econometric methods on more highly disaggregated data that might help policy makers better understand what happened and possibly avoid another devastating episode like this one?" We believe the results of our study and others on this important topic clearly suggest the answer is yes.

Thus an extension of our analysis might include additional variables to test the strength of our results after accounting for *additional* relevant underlying economic and regulatory factors. For example, while the apparent lack a statistically verifiable relationship between house prices and *any* of the fundamental economic variables is an intriguing anomaly, the lack of a relationship between house prices and family income [noted previously in Meen (2002) and Gallin (2006)] presents an *especially* interesting one. Economic theory (and common sense) clearly suggests that these 2 variables should indeed be related. Yet 2 recent summaries (arguably from opposite sides of the political spectrum) agree that the policy of *deliberately relaxing credit standards* contributed to the *disconnect* between house prices and family income during this period [see Baily, Litan and Johnson (2008), and Wallison (2009) for detailed discussions]. They argue that this policy effectively short-circuited the ability of lenders to adequately examine the *riskiness* of mortgage loans. Thus including and analyzing data on credit scores and mortgage applications (if such data can be obtained) in an analysis like the one conducted here might be a very interesting exercise that could shed some light on the policy debate currently surrounding this important issue.

As suggested above, another avenue for future research is to recreate our study using more disaggregated MSA-level panel data (such as the 20 cities in the Case–Shiller index). Finally, like all others on this topic of which we are aware, this study uses *linear* models for both the unit root and cointegration tests. As *nonlinear* unit root and cointegration econometric tests become available, it will be possible to specify and test such models where appropriate.

Acknowledgements

The authors thank the Editor and an anonymous reviewer for helpful and constructive comments. Gauss computer code for some of the statistical calculations was graciously provided by Zhongjun Qu. The usual disclaimer applies.

Appendix A. U.S. states by FHFA region

Region	State	Region	State
East	Kentucky	New England	Maine
South Central	Tennessee		Vermont
	Mississippi		New Hampshire
	Alabama		Massachusetts
			Rhode Island
West	Oklahoma		Connecticut
South Central	Texas		
	Arkansas	Middle Atlantic	New York
	Louisiana		New Jersey
			Pennsylvania
East	Wisconsin		
North Central	Michigan	South Atlantic	Delaware
	Illinois		Maryland
	Indiana		District of Columbia
	Ohio		Virginia
			West Virginia
West	North Dakota		North Carolina
North Central	South Dakota		South Carolina
	Minnesota		Georgia
	Nebraska		Florida
	Iowa		
	Kansas	Mountain	Montana
	Missouri		Idaho
			Wyoming
Pacific	Alaska		Nevada
	Washington		Utah
	Oregon		Colorado
	California		Arizona
	Hawaii		New Mexico

Appendix B.

Variable definitions, 1975Q1–2005Q2 ($N = 122$ quarters, except where noted).

Real data in $2005Q1 (except where noted). Data are seasonally adjusted where appropriate.*House Price Indexes for U.S. and Regions – (RealHousePriceIndex)*

Source: U.S. Federal Housing Finance Association (FHFA).

Quarterly index data, not seasonally adjusted, 1980 = 100, calculated using the Case–Shiller geometric-weighted repeat sales procedure fully described in Calhoun (1996). Regional data based on U.S. Census Bureau divisions (U.S. +9): United States, East North Central, East South Central, Middle Atlantic, Mountain, New England, Pacific, South Atlantic, West North Central, West South Central. The 9 divisions are used to create 4 regions (as defined by the U.S. Census Bureau) by summing and averaging the index values for the divisions within each region. The 4 regions are: West, Midwest, Northeast and South.

12-month LIBOR Rate – (LIBOR)

Source: British Bankers Association.

Interest rate on 12-month LIBOR deposits. Monthly data in $U.S. from 1986.

30-year Mortgage Rate – (30yrMortgage)

Source: Board of Governors of the U.S. Federal Reserve System.

30-year conventional mortgage rate (fixed-rate, first mortgages), not seasonally adjusted. Monthly data converted to quarterly mean by simple average. *Real 30-year Mortgage Rate* is the nominal rate adjusted for inflation using the annual (4-quarter) change in the CPI-less Shelter.

Adjustable Mortgage Rate – (Adjustable)

Source: Federal Housing Finance Board.

Effective adjustable interest rate on conventional single-family mortgages, not seasonally adjusted, all houses, monthly data from 1986.

Civilian Unemployment Rate for U.S. and Regions – (Unemployment)

Source: U.S. Department of Labor, Bureau of Labor Statistics.

Percent civilian unemployment. Monthly data are seasonally adjusted and converted to quarterly mean by simple average. Regional data from 1976.

CPI less Shelter

Source: U.S. Department of Labor, Bureau of Labor Statistics.

Consumer Price Index (CPI), all urban consumers, U.S. and Regional city averages, all items less shelter, monthly data, not sea-

Please cite this article in press as: Clark, S. P., & Coggin, T.D. Was there a U.S. house price bubble? An econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance* (2011), doi:10.1016/j.qref.2010.12.001

sonally adjusted, 1982–84 = 100. This variable is used to convert nominal data to real $2005Q1 equivalent.

CPI-Owner's Equivalent Rent of Primary Residence for U.S. and Regions – (Rent)

Source: U.S. Department of Labor, Bureau of Labor Statistics.

Consumer Price Index (CPI), all urban consumers, U.S. and Regional city averages, owner's equivalent rent of primary residence, monthly data from 1982, not seasonally adjusted, 1982 = 100.

Homeownership Rates for U.S. and Regions – (HomeownerRate)

Source: U.S. Census Bureau.

The proportion of households that are owners is termed the homeownership rate. It is computed by dividing the number of households that are owners by the total number of households. Quarterly data, not seasonally adjusted.

Household Debt Ratio – (HouseholdDebt/Income)

Source: Board of Governors of the U.S. Federal Reserve System.

This variable is the quarterly ratio of household credit market debt outstanding to annualized disposable personal income (individual components are seasonally adjusted). Data not available for regions.

Homebuilders Common Stock Index – (RealHomebuildersIndex)

Source: Merrill Lynch.

Capitalization-weighted, price-level index of homebuilding stocks based on homebuilding stocks included in the S&P 500 stock index, compiled and maintained by Merrill Lynch. Stocks included in the index for second quarter 2005: Centex Corp., D.R. Horton, KB Home, Lennar Corp. and Pulte Homes.

Lehman Brothers Mortgage-backed Securities Index – (RealLBMortgage-backedIndex)

Source: Lehman Brothers.

A price-level index made up of mortgage-backed securities that is usedfor benchmarking purposes. The Lehman Brothers MBS Index consists of fixed-rate securities, such as mortgage pools created by the Government National Mortgage Association (GNMA), Federal Home Loan Mortgage Corp (FHLMC) and Federal National Mortgage Association (FNMA). This index serves as a performance benchmark for many mortgage-backed securities funds.

Median Family Income for U.S. and Regions – (RealMedianFamily-Income)

Source: U.S. Census Bureau, *Historical Income Tables-Families* (all races).

Annual data from 1975 through 2005, not seasonally adjusted. Data are gross income (before taxes) as used by mortgage lenders in making the lending decision. Annual data converted to quarterly data by linear interpolation of annual changes in income. Pretesting verified that our basic results are not sensitive to using the original annual income data. Converting to quarterly data strengthens tests for a structural break by increasing the number of data points.

Median New House Price for U.S. and Regions – (NewHousePrice)

Source: U.S. Census Bureau.

Median sales price of new houses sold. Annual data from 1975 through 2005, not seasonally adjusted. Annual data converted to quarterly data by linear interpolation of each annual change in price.

New Houses for Sale for U.S. and Regions – (ForSale)

Source: U.S. Census Bureau.

Number of new houses for sale at end of period. Monthly data, not seasonally adjusted.

New Houses Sold for U.S. and Regions – (Sold)

Source: U.S. Census Bureau.

Number of new houses sold during period. Monthly data, not seasonally adjusted, quarterly data are 3-month sums of monthly data.

Ratio of Real Private Residential Fixed Investment to Real GDP – (RealResInvestment/GDP)

Source: U.S. Department of Commerce, Bureau of Economic Analysis.

Private residential fixed investment divided by GDP. Both individual variables are quarterly in billions of chained 2000 dollars, and converted to seasonally adjusted annual rate. Data not available for regions. This variable was suggested by Shiller (2006).

References

Abraham, J., & Hendershott, P. H. (1996). Bubbles in metropolitan housing markets. *Journal of Housing Research*, *7*, 191–207.

Baily, M. N., Litan, R. E., & Johnson, M. S. (2008). *The origins of the financial crisis, Brookings Institution Initiative on Business and Public Policy*. Washington, DC: The Brookings Institution.

Brunnermeier, M. K. (2008). Bubbles. In S. N. Durlauf, & L. E. Blume (Eds.), *The new Palgrave dictionary of economics*. New York: Palgrave Macmillan.

Calhoun, C. A. (March 1996). *OFHEO house price indexes: HPI technical description*. Washington, DC: Office of Federal Housing Enterprise Oversight.

Case, K. E., & Shiller, R. J. (2003). Is there a bubble in the housing market? (with comments and discussion). *Brookings Papers on Economic Activity*, *2*, 299–342.

Case, K. E., & Shiller, R. J. (August 30, 2006). Full house. *Wall Street Journal*, A10.

Cameron, G., Muellbauer, J., & Murphy, A. (2006). *Was there a British house price bubble? Discussion Paper #5619*. Centre for Economic Policy Research, London.

Capozza, D. R., Mack, C., & Mayer, C. J. (1997). *The dynamic structure of housing markets. Working Paper*. University of Michigan Business School.

Claessens, S., Kose, M. A., & Terrones, M. E. (2008). *What happens during recessions, crunches and busts?* Washington, DC: Research Department, International Monetary Fund.

Clark, S. P., & Coggin, T. D. (2009). Trends, cycles and convergence in U.S. regional house prices. *Journal of Real Estate Finance and Economics*, *39*, 264–283.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, *74*, 427–431.

Diba, B. T., & Grossman, H. I. (1988). The theory of rational bubbles in stock prices. *Economic Journal*, *98*, 746–754.

DiPasquale, D., & Wheaton, W. (1994). Housing market dynamics and the future of housing. *Journal of Urban Economics*, *35*, 1–27.

Drieman, M. H., & Pennington-Cross, A. (2004). Alternative methods of increasing the precision of weighted repeat sales house prices indices. *Journal of Real Estate Finance and Economics*, *28*, 299–317.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, *55*, 251–276.

Gallin, J. (2006). The long-run relationship between house prices and income. *Real Estate Economics*, *34*, 417–438.

Glaeser, E. L., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, *64*, 198–217.

Goodman, A. C., & Thibodeau, T. G. (2008). Where are the speculative bubbles in US housing markets? *Journal of Housing Economics*, *17*, 117–137.

Gregory, A. W., & Hansen, B. E. (1996). Residual-based tests for cointegration in models with regime shifts. *Journal of Econometrics*, *70*, 99–126.

Hendershott, P., & Slemrod, J. (1983). Taxes and the user cost of capital for owner-occupied housing. *Journal of the American Real Estate and Urban Economics Association*, *10*, 375–393.

Himmelberg, C. P., Mayer, C. J., & Sinai, T. M. (2005). Assessing high house prices: Bubbles, fundamentals, and misperceptions. *Journal of Economic Perspectives*, *19*, 67–92.

Holly, S, Pesaran, M. H., & Yamagata, T. (2010). A spatio-temporal model of house prices in the USA. *Journal of Econometrics*, *158*, 160–173.

Homm, U., & Breitung, J. (2010). *Testing for speculative bubbles in stock markets, a comparison of alternative methods, Working paper*. Department of Economics, University of Bonn. http://www.ect.uni-bonn.de/mitarbeiter/joerg-breitung/bubbles_feb10.pdf.

Johansen, S. (1991). Estimation and hypothesis testing on cointegrating vectors in Gaussian vector autoregressive models. *Econometrica*, *59*, 1551–1580.

Johnes, G., & Hyclack, T. (1999). House prices and regional labor markets. *Annals of Regional Science*, *33*, 33–49.

Labonte, M. (2003). *U.S. housing prices: Is there a bubble? Report for Congress, Number RL31918*. Congressional Research Service, Library of Congress, Washington, DC.

Lee, J., & Strazicich, M. C. (2003). Minimum Lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics*, *85*, 1082–1089.

Lee, J., Strazicich, M. C. (2004). *Minimum LM unit root test with one structural break. Working paper*. Department of Economics, Appalachian State University. http://econ.appstate.edu/RePEc/pdf/wp0417.pdf.

Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics*, *8*, 27–62.

Meen, G. (2002). The time-series behavior of house prices: A transatlantic divide? *Journal of Housing Economics*, *11*, 1–23.

McCarthy, J., & Peach, R. W. (2004). Are home prices the next "bubble"? *Economic Policy Review (Federal Reserve Bank of New York)*, *10*, 1–17.
Mikhed, V., & Zemcik, P. (2009). Testing for bubbles in housing markets: A panel data approach. *Journal of Real Estate Finance and Economics*, *38*, 366–386.
Phillips, P. C. B. (1995). Fully modified least squares and vector autoregression. *Econometrica*, *63*, 1023–1078.
Phillips, P. C. B., Wu, Y., & Yu, J. Explosive behavior in the NASDAQ: When did exuberance escalate asset values? *International Economic Review 51*, in press.
Pollock, A. J. (2008). *The human foundations of financial risk, Financial Services Outlook*. Washington, DC: American Enterprise Institute for Public Policy.
Poterba, J. (1991). House price dynamics. *Brookings Papers on Economic Activity*, 143–183.
Qu, Z. (2007). Searching for cointegration in a dynamic system. *Econometrics Journal*, *10*, 580–604.
Rapach, D. E., & Strauss, J. K. (2009). Differences in housing price forecastability across U.S. states. *International Journal of Forecasting*, *25*, 351–372.
Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive moving average models of unknown order. *Biometrika*, *71*, 599–607.
Shiller, R. J. (June 2, 2005). People are talking . . .. *Wall Street Journal*, A12.
Shiller, R. J. (2006). Long-term perspectives on the current boom in house prices. *Economists' Voice*, *3*, 1–11.
Smith, M. H., & Smith, G. (2006). Bubble, bubble, where's the housing bubble? *Brookings Papers on Economic Activity*, 1–50.
Wallison, P. J. (2009). *The true origins of this financial crisis*. Washington, DC: American Enterprise Institute for Public Policy Research.
Wessel, D. (2008). Home prices, by what measure? *Wall Street Journal*, A2.
Zivot, E., & Andrews, D. W. K. (1992). Further evidence on the great crash, the oil-price shock, and the unit root hypothesis. *Journal of Business & Economic Statistics*, *10*, 251–270.
Zhou, W. X., & Sornette, D. (2003). 2000–2003 real estate bubble in the UK but not in the USA. *Physica A*, *329*, 249–263.
Zhou, W. X., & Sornette, D. (2006). Is there a real-estate bubble in the USA? *Physica A*, *361*, 297–308.
Zuckerman, G. (January 1, 2008). Trader made billions on subprime. *Wall Street Journal*, front page

Please cite this article in press as: Clark, S. P., & Coggin, T.D. Was there a U.S. house price bubble? An econometric analysis using national and regional panel data. *The Quarterly Review of Economics and Finance* (2011), doi:10.1016/j.qref.2010.12.001