TESTING FOR BUBBLES IN HOUSING MARKETS: A PANEL DATA APPROACH

Vyacheslav Mikhed
Petr Zemčík

CERGE-EI

Charles University
Center for Economic Research and Graduate Education
Academy of Sciences of the Czech Republic
Economics Institute

WORKING PAPER SERIES (ISSN 1211-3298)
Electronic Version

338

Working Paper Series 338
(ISSN 1211-3298)

Testing for Bubbles in Housing Markets: A Panel Data Approach

Vyacheslav Mikhed
Petr Zemčík

CERGE-EI
Prague, October 2007

ISBN 978-80-7343-137-2 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-126-5 (Národohospodářský ústav AV ČR, v.v.i.)

Testing for Bubbles in Housing Markets: A Panel Data Approach

Vyacheslav Mikhed and Petr Zemčík*

CERGE–EI†

October 2007

Abstract

We employ recently developed cross-sectionally robust panel data tests for unit roots and cointegration to find whether house prices reflect house-related earnings. We use U.S. data for Metropolitan Statistical Areas, with house price measured by the weighted-repeated-sales index, and cash flows either by market tenant rents or estimates of a fair market rent. In our full sample periods, an error-correction model is not appropriate, i.e. there is a bubble. We then combine overlapping ten-year periods, price-rent ratios, and the panel data tests to construct a bubble indicator. The indicator is high for the late 1980s, early 1990s and since the late 1990s for both panels. Finally, evidence based on panel data Granger causality tests suggests that house price changes are helpful in predicting changes in rents and vice versa.

Abstrakt

Testujeme, zda ceny domů odpovídají očekávanému výnosu z jejich vlastnictví. K tomuto účelu využíváme teprve nedávno vyvinuté testy pro přítomnost jednotkových kořenu a kointegrace v panelových datech. Tyto nové testy jsou robustní vůči závislosti daných proměnných mezi regiony. Používáme data z Metropolitních statistických oblastí ve Spojených Státech. Ceny domů měříme váženým indexem opakovaných prodejů a tok peněz buď skutečným tržním nájemným nebo jeho odhadem. Naše výsledky ukazují, že model opravy chyb není adekvátní pro celá časová rozmezí našich dvou databází. Jinými slovy, existuje cenová bublina. Následně definujeme indikátor cenové bubliny, jež je založený na desetiletých časových periodách, které se překrývají, na poměru cen vůči nájemnému a na testech v panelových datech. Indikátor je v obou databázích zvýšený koncem 80.let, počátkem 90.let a od konce 90.let. Závěrem uvádíme výsledky testu Grangerovy kauzality v panelových datech, které dokumentují, že změny v cenách domu jsou užitečné při predikci změn v nájmech a naopak.

Keywords: cointegration, panel data, unit root, bubble, house prices, rents
JEL classification: G12 - Asset Pricing; R21 - Housing Demand; R31 – Housing Supply and Markets; C33 - Models with Panel Data.

***Email:** petr.zemcik@cerge-ei.cz
†A joint workplace of the Center for Economic Research and Graduate Education, Charles University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic.
Address: CERGE–EI, P.O. Box 882, Politických vezňů 7, Prague 1, 111 21, Czech Republic

1 Introduction

A bubble is typically characterized by a discrepancy between house prices and fundamentals. The definition of fundamental variables reflects an underlying structural or present-value model. Both types of models lead to stationarity between house prices and selected factors. The former is a simple demand and supply model where supply determinants include depreciation, construction costs, etc., and demand determinants are, among other things, income, housing cost, and the user costs of owning a house. For example, Gallin (2006) uses the structural model to justify stationarity between house prices and income. The present-value model ties together asset prices with a stream of earnings related to a particular asset. Campbell and Shiller (1987) derive implications of this model for stationarity between financial assets and their cash flows: (i) they should be of the same order of integration, and (ii) if they are both non-stationary in levels but stationary in first differences, the two series should be cointegrated. Wang (2000) implements this methodology in the U.K. property market for house prices and rents.

McCarthy and Peach (2004) argue using aggregate data that there is no bubble in the U.S. market. Their arguments refer to both sides of the bubble definition. First, they conclude that standard aggregate house price measures often exaggerate price increases. Second, typical fundamentals such as income and rents are not appropriate measures of affordability and earnings associated with owning a house, respectively. Gallin (2004) addresses both issues and suggests that the national-level rent-price ratio may be stationary (i.e. there is no bubble) using standard error-correction models though evidence based on long horizon regression models is inconclusive. On the other hand, Shiller (2005) views the situation on the U.S. housing market as indicative of a bubble. He looks at the behavior of the house

University, Prague, and the Economics Institute of the Academy of Sciences of the Czech Republic. Address: CERGE–EI, P.O. Box 882, Politických vězňů 7, Prague 1, 111 21, Czech Republic

price index, construction costs, population, and interest rates in 1890-2004 and concludes that although there is a steady growth of population and a decline in the interest rates, they hardly justify the surge of housing prices after 1998.

A large number of studies analyzes regional data to assess potential bubble occurrence. These studies fall into two categories. The first one makes use of additional information contained in panel data. Malpezzi (1999) uses panel data unit root tests to study the long-run relationship between house prices and income in 133 U.S. metropolitan areas and rejects the no-cointegration hypothesis but does not account for the first-stage estimation of a cointegrating parameter. Gallin (2006) remedies this shortcoming and reverses the conclusion. The second category focuses on local markets with attention to fundamental variables being more closely attached to house characteristics. Smith and Smith (2006) estimate the fundamental values of a house using rent and price data for matched single-family homes in ten metropolitan areas. Their results indicate there is in fact no bubble in most of the considered cities. Himmelberg, Mayer, and Sinai (2005) calculate the user cost of housing in 46 metropolitan areas and also conclude that conventional metrics such as price-rent and price-income ratios are misleading in a test for bubble occurrence.

Studies based on the nationwide aggregates for both house prices and fundamentals typically employ standard univariate unit root methodology and require relatively long time series. However, these tests are notoriously known to have low power, in which case a discrepancy between house prices and fundamental factors can be wrongly viewed as non-stationary and hence interpreted as a bubble. In the 1990s and early 2000s, many researchers have focused on panel data to increase the power of unit root tests. A widely used and intuitive test is Im, Pesaran, and Shin (IPS henceforth) test proposed in Im, Pesaran, and Shin (2003), which is based on averaging single-series unit root tests and allows for an alternative hypothesis of stationarity only in some series in a panel. Pesaran (2007) develops a cross-sectionally augmented IPS test (CIPS), which is robust to the cross-sectional dependence often observed in panel data. A popular panel test for cointegration developed by Pedroni (1999, 2004) builds on a panel analogue of a regression of the type suggested in Engle and Granger (1987) for a simple time-series.

The panel data stationarity tests have been used to some extent in the context of the real estate market. Liow and Li (2006) investigate whether real estate company stock prices differ from their net asset values, which represent the underlying value of the real estate assets of a property stock. They use panel data from eight Asian-Pacific securitized real estate markets and find a long-run equilibrium relationship between the stock prices and net asset values. Closer to our approach are Malpezzi (1999) and Gallin (2006). Malpezzi (1999) uses the IPS test and Gallin (2006) complements this test with Pedroni (1999, 2004) cointegration tests. As already mentioned above, both of these studies analyze the relation between income and house prices.

Our study focuses on rent as a measure of the cash-flow variable. The choice of our fundamental factor is mainly data driven. We are able to collect two different panel datasets for rents, which enables us to achieve robustness with respect to the definition of a rent as well as to the dimensions of a given panel. The Bureau of Labor Statistics (BLS) is the source of the rent index computed as a part of a CPI calculation and defined as actual tenants' rent. The second measure of the rent is the fair market rent from the U.S. Department of Housing and Urban Development (HUD). Both of the rent series are combined with the house price index from the Office of Federal Housing Enterprise Oversight (OFHEO). The time series and cross-sectional dimensions of the data on rents determine the respective sizes of panels of data. Our first panel, based on the BLS rent index, covers 23 Metropolitan Areas from 1978 to 2006 semi-annually. The second panel uses data on 273 areas from 1986 to 2006 at a yearly frequency. In other words, the former dataset spans a greater time period for a smaller cross-section, while the latter spans a shorter period for a large number of regions.

The concentration on cash-flows in our analysis naturally leads to the present-value model as the appropriate theoretical framework, which is combined with the up-to-date panel data stationarity methodology. We first test for cross-section dependence using a test from Pesaran (2004), which indicates the existence of a strong mutual correlation among regions for both prices and rents. We then conduct the CIPS test to examine stationarity in levels and first differences for both the house prices and rents. In all datasets, both the house prices and rents are non-stationary in levels but stationary in first differences. Consequently, we test for cointegration using the Pedroni (1999, 2004) tests. The critical values for this test are calculated by bootstrap to account for the cross-sectional dependence. The house prices and rents are not cointegrated in the whole sample period in either dataset. In such a case, the price-to-rent ratio should be non-stationary as well. As expected, the CIPS test cannot reject the null hypothesis of the unit root in this variable. We conclude that an error-correction model is not appropriate for modelling house prices and rents. A broader interpretation is that it may take more than three decades for house prices to return to fundamentals.

While our full-sample evidence suggests long swings of house prices from their fundamental values, we would like to have a measure of how far away they are at a given point in time. Hence we conduct our tests using ten-period overlapping data windows. Based on the results, we define a “bubble indicator." We set it equal to unity if house prices are non-stationary while rents are stationary, and to zero if prices are stationary. For the other possibilities we test for the stationarity of the price-rent ratio, which is more convenient than a cointegration test that is often not applicable due to the different order of integration of the involved series. The bubble indicator is then set equal to the p-value of the CIPS test. Based on the 23 Metropolitan Statistical Areas (MSA), there are several periods in the U.S., when the bubble indicator is close to 1 or 1: the late 1980s and early 1990s plus a period since the year 2000. The second two are also confirmed by the 273 MSA dataset, only the bubble starts already in 1998. In both datasets, the bubble indicator decreases in 2006.

The non-stationarity of the price-to-rent ratio also has implications for research studying its predictive power with respect to either rents or house prices. For example, Capozza and Seguin (1996) argue that a rent-price ratio predicts capital appreciation in the housing market, Clark (1995) concludes that the rent-price ratio reflects the expectations of future rent growth, and Campbell, Davis, Gallin, and Martin (2006) decompose the rent-price ratio into the expected present value of rental growth, real interest rates, and future housing premia. To analyze this issue further, we first test for the stationarity of a simple average of price-rent and rent-price ratios across regions using standard univariate tests. In both datasets, all the considered aggregate series have unit roots. Panel data stationarity tests confirm this result.

The presence of unit roots prevents us from using standard regressions to investigate the predictive power of the price-to-rent ratios. Instead, we focus on a looser interpretation of the present-value formula, which suggests that house prices should have predictive power with respect to changes in rents, and vice versa. These ideas translate directly into testing for Granger causality in house prices and rents. This is only plausible for stationary series, in our case the differences in prices and rents. The recently developed methodology in Hurlin (2004) and Hurlin and Venet (2004) enables us to test for Granger causality in the panel data context. Our results suggest that changes in prices are helpful in predicting changes in rents and vice versa. This conclusion is more strongly supported using the 273 MSA.

The rest of this paper is organized as follows. We explain the idea behind bubbles in the housing market using the present-value model for house prices as a discounted stream of future rents in Section 2. We describe in detail our regional panel data in Section 3. Section 4 gives a survey of the panel data stationarity tests, introduces the bubble indicator and reports our results for their application to our data. Section 5 analyzes the mutual predictive power of changes in rents with respect to changes in prices and vice versa. Section 6 summarizes our findings.

2 Bubbles in Housing Markets

So far, studies on the non-stationarity of house prices have only used regionally aggregated data. Two notable exceptions are Malpezzi (1999) and Galin (2006) who investigate the plausibility of error correction models for house prices and income. This paper considers rents in Metropolitan Statistical Areas and augments the existing econometric framework by constructing a bubble indicator and by implementing Granger causality panel tests in the context of housing markets. We view a house as an investment vehicle, use a standard present-value formula to derive implications for the relationship between house prices and cash flows, and illustrate the consequences of a bubble presence in an economy. These consequences are later employed to test for rational bubbles using U.S. panel data on house prices and rents.

The standard present-value formula is:

$$P_{i,t} = E_t \left[\frac{C_{i,t+1} + P_{i,t+1}}{1+D} \right], \quad i = 1, \dots, N, \tag{1}$$

where $P_{i,t}$ is the price of a house i or, more in the light of our subsequent analysis, a regional house price index and E_t is mathematical expectation conditional on information at time t. $C_{i,t}$ is a cash-flow associated with owning a house, i.e. a rent

r_{it}. The formula (1) can be viewed as an implication of a Lucas (1978) endowment economy with risk neutral investors. In the equilibrium of this economy, income coincides with the cash-flow, which suggests that a study of the relationship between house prices and income is also appropriate.[1] D denotes a constant discount rate. This formula holds for all periods t. Invoking the law of iterated expectations results in the following formula:

$$P_{i,t} = E_t \left[\frac{C_{i,t+1}}{1+D} + \frac{C_{i,t+2}}{(1+D)^2} + ... + \frac{C_{i,t+k}}{(1+D)^k} + \frac{P_{i,t+k}}{(1+D)^k} \right]. \tag{2}$$

We impose for a moment the no-bubbles condition

$$\lim_{k \to \infty} E_t \left[\frac{P_{i,t+k}}{(1+D)^k} \right] = 0, \tag{3}$$

which yields

$$P^F_{i,t} = \sum_{j=1}^{\infty} \frac{1}{(1+D)^j} E_t[C_{i,t+j}], \tag{4}$$

which is often referred to as price reflecting fundamentals.

Following Campbell and Shiller (1987) and Wang (2000), we define the spread between the house price and cash flows as $S_{i,t} \equiv P_{i,t} - \frac{1}{D} C_{i,t}$. If the cash-flow process is $I(1)$ and (3) holds, then $P_{i,t}$ is also $I(1)$ (i.e. $\triangle P_{i,t}$ is stationary) and $S_{i,t}$ is stationary (i.e. house prices and cash flows are cointegrated). To illustrate this result we rewrite $S_{i,t}$ as:

$$S_{i,t} = \frac{1}{D} E_t \sum_{j=1}^{\infty} \frac{\triangle C_{i,t+j+1}}{(1+D)^j} = \frac{1}{D} E_t \left[\triangle P_{i,t+1} \right] \tag{5}$$

$$S_{i,t} = \frac{1+D}{D} E_t \sum_{j=1}^{\infty} \frac{\triangle C_{i,t+j}}{(1+D)^j} = \frac{1+D}{D} E_t \left[\triangle P_{i,t} \right]. \tag{6}$$

[1]Fundamentals (including income) can also be motivated by considering omitted variables in the present value formula as in Hamilton (1986) or via a general supply-demand model as in Gallin (2006).

The first equality stems from the fact that the conditional expected value of the future cash-flow is given by its current value. The second equality follows from equation (4). Note also that the stationarity of $S_{i,t}$ implies the stationarity of $P_{i,t}/C_{i,t}$ (and its inverse) since if $S_{i,t} = 0$ then $P_{i,t}/C_{i,t} = 1/D$.

Let us assume that the no-bubbles condition (3) is violated. In this case, both the house price $P_{i,t}$ and the spread $S_{i,t}$ are non-stationary.[2] Our previous discussion then suggests a strategy to determine empirically whether there is a bubble or not. A first natural step is to test for unit roots in series for house prices and cash-flows. There are four possible results of this test:

Case 1: $P_{i,t}$ stationary and $C_{i,t}$ stationary,

Case 2: $P_{i,t}$ stationary and $C_{i,t}$ non-stationary,

Case 3: $P_{i,t}$ non-stationary and $C_{i,t}$ stationary,

Case 4: $P_{i,t}$ non-stationary and $C_{i,t}$ non-stationary.

In Case 1, equation (1) passes a basic empirical test. While it can still fail to explain the behavior of house prices and cash-flows for other reasons, it is unlikely that the failure is due to the presence of a bubble. Case 2 indicates the failure of the present value model, since explosive cash-flows should be reflected in house prices. However, we are only interested in the failure of the model due to run-away prices and hence focus mainly on the two remaining cases. In Case 3, there is clearly a bubble. Case 4 calls for a test of the cointegration between house prices and cash-flows, assuming that first differences are stationary. Alternatively, one can test for the stationarity of P/r. In Section 4, we explicitly discuss panel data unit root

[2]An example of such a violation is a solution of the stochastic differential equation (1), which contains a “bubble" term that satisfies

$$B_{i,t} = \frac{1}{1+D} E_t B_{i,t+1}. \tag{7}$$

Consequently, the house price with a bubble may be written as:

$$P_{i,t} = P^F_{i,t} + B_{i,t}. \tag{8}$$

It is easy to show that the price obtained in equation (8) satisfies equality (1) and in this sense this bubble is “rational". See for example Hamilton (1986) for a survey of speculative bubbles.

and cointegration tests, and formulate a "bubble" indicator, which summarizes the results of these tests in a simple manner.

The present value formula is often used to justify the use of the rent-to-price ratio as a predictor for either the expected capitalization of investment in a house or the growth rate of rents (see Capozza and Seguin 1996 and Clark 1995). To illustrate, let us assume that a house is sold after one period. We can re-write (1) as

$$E_t \frac{\triangle P_{i,t+1}}{P_{i,t}} = D - E_t \frac{C_{i,t+1}}{P_{i,t}}. \tag{9}$$

One can regress the capitalization on the rent-to-price ratio. Capitalization can be replaced by the growth rate of rents since the two should be closely related. Statistically, regression (9) can be run if either both price differences and the rent-to-price are stationary or if they are of the same order of integration but cointegrated. Theory restricts the relationship further: Under the no bubbles condition and unit root in cash flows, both the price differences and rent-to-price ratio are stationary. Therefore, different means should be used to study predictability for a non-stationary rent-to-price ratio. If first differences in prices and rents are stationary we can test for Granger-causality in panel data. If the changes in prices Granger-cause changes in rents, price differences are useful in predicting the rent differences. The same principle applies in the opposite direction. We investigate this issue in Section 5.

3 Panel Data

The empirical analysis carried out in this study utilizes two datasets for the house price and rent indices. The first dataset is comprised of the house price index (HPI) from the Office of Federal Housing Enterprise Oversight and the rent of primary residence index (RI) estimated by the Bureau of Labor Statistics.[3] According to OFHEO (Calhoun 1996), the HPI is computed quarterly for the period 1975-2006 as the weighted repeat sales index based on the data on mortgage contracts recorded by the Federal Home Loan Mortgage Corporation (Freddie Mac) and the Federal National Mortgage Association (Fannie Mae).[4] The RI data for 25 Metropolitan Statistical Areas is available from the BLS for the period 1975-2006 in monthly, semi-annual, and annual frequency.

In order to harmonize the data in the frequency dimension and leave as many data points as possible, both the HPI and RI are recalculated to a semi-annual frequency with the first half of 1995 as the base period. Particularly, the HPI is computed as an arithmetic average of two quarterly values of this index, and the RI is estimated as the average of all monthly values of the index if monthly values are available or as a semi-annual value if such a value is available. After all recalculations, the HPI and RI are matched based on the names of the largest cities in the MSA because definitions of these areas are slightly different in OFHEO and BLS databases. The resulting dataset consists of 23 MSA and covers the period from the first half of 1978 to the second half of 2006. Both house prices and rents are adjusted for inflation using regional BLS CPI's.

The second dataset contains the HPI again from OFHEO and the fair market rent (FRM) from the U.S. Department of Housing and Urban Development. The HPI and FRM are matched using MSA in accordance with the 2000 U.S. census classification. The HUD defines FRM as the 40th, 45th or 50th percentile rent of the distribution of gross rents (i.e. including utilities) for standard quality dwellings. The rent distribution is derived based on data from the American Housing Sur-

[3] A description of the method of gathering data and calculating the rent index is presented in Fact Sheet No. BLS 96-5 “How BLS Uses Rent Data in the Consumer Price Index."

[4] The repeat sales approach to real estate index calculation was described first by Bailey, Muth, and Nourse (1963) and further advanced by Case and Shiller (1987, 1989). For more information on the HPI see Calhoun (1996).

veys, the 1990 Decennial Census, and Random Digit Dialing Telephone Surveys.[5] Compared to the BLS data, this dataset covers a wider range of regions but the covered period is shorter and the data collection frequency is yearly.

After matching data on the HPI and FRM, we drop MSA with missing values in either of the two series. Since data on the FRM for the sample of MSA used have annual frequency and the HPI for the same sample has quarterly frequency, the HPI is recalculated for each year by taking an arithmetic average of quarterly values for a corresponding year. In addition, all time series for the FRM that are given in nominal terms are transformed in the index with a base year 1995, since the base for the HPI is also 1995. This also allows us to use different percentiles of the FMR since all we need is a comparison with respect to the base year. As a result of this exercise, both the annual HPI and FRM indices are obtained for 273 MSA in the period 1986-2006. Again, we adjust the two series for inflation. Since regional CPI's are not available in this case, we use the U.S. aggregate CPI produced by BLS (Consumer Price Index for All Urban Consumers: All Items).

The price-rent ratios calculated for the 23 MSA using the first dataset are graphed in Figures 1 to 4. Based on the path of these ratios for 1978:h1-2006:h2, it is possible to divide metropolitan areas into four groups. The dynamics of the price-rent ratio in the first group (Figure 1) may hint that this group has experienced a discrepancy between home prices and rents in the early 1980s, at the end of the 1980s, and in the late 1990s. These are the "usual suspects," large cities such as New York or San Francisco, with cycles of booms and busts on the real estate market. The second group (Figure 2) consists mainly of Midwestern cities and has experienced a peak in the price-rent ratio again around 1980 and also since 1999-2000. The third group (Figure 3) are cities that have experienced three increases similar to the first group but less pronounced. Finally, the fourth group (depicted

[5]For additional details, see the description of FRM by the U.S. Department of Housing and Urban Development (1995).

in Figure 4) are cities that are difficult to categorize but with the exception of Portland have been experiencing a rise in the price-rent ratio. The average price-rent ratio for the 23 MSA presented in Figure 5 also has the three peaks with a substantial increase since 1999-2000. Interestingly, in all the above pictures, one can identify a slight decrease of the price-rent ratio at the end of the sample.

Table 1 presents summary statistics for the price-rent ratio, denoted by P/r, computed using the HPI and FRM for 273 MSA from the second database. The patterns found in the second database are very similar to those displayed in Figures 1 to 4 but without a decrease at the end of the sample. In particular, there are two potential bubbles hinted at by the data: in 1986-1991 and in 2000-2005. However, the behavior of the U.S. housing market differs somewhat in these two episodes. In the first episode the number of MSA with a very high price-rent ratio is relatively small and a high price-rent ratio in these areas is accompanied by a low ratio in many other areas. Hence, a break-down of the price-rent relationship does not have to be present on the aggregate level but only in some local markets. On the other hand, the second overall increase of the price-rent ratio in 2000-2006 is characterized by a large number of regions with very a high price-rent ratio: 137 out of 273 MSA have a ratio of more than 1.3. In addition, it is noteworthy that none of the MSA have experienced a decline in the price-rent ratio. These observations are roughly confirmed by Figure 6, which shows the average price-rent ratio for 273 Metropolitan Statistical Areas.[6]

[6]Using data on the median two-bedroom house price from the National Association of Realtors for 2006 and the house price index from OFHEO, we also calculated the house price for 1986-2006 as the HPI multiplied by the median house price. This enabled us to calculate price-to-rent ratios in absolute values, with the number of highest values above 22.5 increasing in the same periods as when using the relative number for P/r.

4 Panel Data Stationarity Tests

In this section, we investigate whether there is a long-run equilibrium relationship between house prices and rents corresponding to the present value formula (1). To do so, we conduct a battery of only recently developed panel data tests for unit roots and cointegration. Our test results are then interpreted in accordance with Section 2 and we also formulate a simple procedure designed to detect a bubble using a moving ten-year data window.

4.1 Unit Roots

Assume that the law of motion for panel data is the following AR(1) process:

$$y_{it} = \mu_i + \omega_i \, t + \rho_i y_{i,t-1} + \epsilon_{it}, \tag{10}$$

where $i = 1, \ldots, N$ is the cross-sectional dimension of the data and $t = 1, \ldots, T$ is the number of observed periods. μ_i is a fixed effect, $\omega_i t$ is an individual trend and ρ_i is an autoregressive coefficient. ϵ_{it} denotes an i.i.d. error term. The dependent variable y_i is said to contain a unit root if $|\rho_i| = 1$. We will consider two dependent variables, the house price $P_{i,t}$ and $r_{i,t}$, and their first differences.

There is an additional dimension here not present in univariate time series. ρ_i can be the same across cross-sections (i.e. $\rho_i = \rho$) or it can differ. Tests in Levin, Lin, and Chu (2002), Breitung (2000) and Hardi (2000) rely on the former assumption, while the Im, et al. (2003), Maddala and Wu (1999) and Choi (2001) tests rely on the latter. Assuming that autoregressive parameters cannot vary across individual series means the only alternative to a common unit root is the stationarity of all the series. This may be fairly restrictive in the context of our study since property prices (or rents for that matter) can rise substantially in some places while they stagnate or even decline in others. The tests which are based on the assumption of individual persistence parameters allows one to test the null hypothesis of a unit root in all series with the alternative hypothesis of unit roots in some (but not necessarily all) of the series. Therefore we decided to employ these tests in our empirical investigation.

Specifically, the Im, et al. (2003), Maddala and Wu (1999), and Choi (2001) tests combine the results of individual unit root tests. To see how, let us consider a standard specification, the augmented Dickey Fuller test (see Hamilton 1994, Ch. 17 for a textbook treatment; ADF henceforth):

$$\Delta y_{it} = \mu_i + \omega_i \, t + \alpha_i y_{i,t-1} + \sum_{j=1}^{p_i} \lambda_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \tag{11}$$

where ε_{it} is an error term. Note that $\alpha_i = \rho_i - 1$ and the lag order p_i vary across cross-sections. The respective null and alternative hypotheses for this test can be expressed as:

$$H_0: \ \alpha_i = 0, \tag{12}$$

$$H_1: \begin{cases} \alpha_i = 0 & \text{for } i = 1, 2, \ldots, N_1 \\ \alpha_i < 0 & \text{for } i = N_1 + 1, N_1 + 2, \ldots, N. \end{cases} \tag{13}$$

The ordering of regions may be changed as needed. H_1 states that at least one (a non-zero fraction) series is stationary.

Im, et al. (2003) first calculate the t-statistics for the α_i's in the individual ADF regressions (denoted as $t_{iT_i}(p_i)$) and then compute their average:

$$\bar{t}_{NT} = \frac{\sum_{i=1}^{N} t_{iT_i}(p_i)}{N}. \tag{14}$$

For the general case with a non-zero p_i for some cross-sections, the following statistic is asymptotically normally distributed:

$$W_{\bar{t}_{NT}} = \frac{\sqrt(N)\left(\bar{t}_{NT} - N^{-1}\sum_{i=1}^{N} E[t_{iT}(p_i)]\right)}{\sqrt{N^{-1}\sum_{i=1}^{N} Var[t_{iT}(p_i)]}} \rightarrow N(0,1). \tag{15}$$

Im, et al. (2003) (see Table 3 of the paper) provide $E[t_{iT}(p_i)]$ and $Var[t_{iT}(p_i)]$ for various combinations of T and p_i. To calculate the statistic, one needs to specify the deterministic components and the number of lags for each ADF regression. The set of choices for exogenous regressors consists of no regressors, an individual constant (a fixed effect) or an individual constant with a linear trend. As indicated in our specification (11), we opt for the most general case with the number of lags set to unity in each case. A complementary approach to the IPS test is used in Maddala and Wu (1999) and Choi (2001). They define a test based on functions of the p-value associated with the ADF test in individual regressions. As the results based on these tests are similar to the IPS test results, we do not report them here.

The IPS test is valid under the assumption of no cross-section dependence in the data. In other words, the residuals in the ADF regression equation (11) are not correlated. This may be a very strong assumption and Pesaran (2007) demonstrates that its violation often leads to undesirable finite sample properties of the IPS test. Therefore, we use the general diagnostic test for cross section dependence in panels proposed by Pesaran (2004) to find whether the dependence is present in the data. The test statistic

$$CD = \sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \mathrm{Corr}(\hat{\epsilon}_i, \hat{\epsilon}_j)\right) \Rightarrow N(0,1), \tag{16}$$

where $\hat{\epsilon}_i, i = 1, \ldots, N$ is a $(T \times 1)$ vector of estimated residuals from equation (11). This test exhibits much less size distortions then the standard Lagrange multiplier test based on squared correlation coefficients. Our test results are reported in

Table 2 and clearly indicate strong dependence in our data, in levels for both prices and rents, and in first differences for prices for the 273 MSA dataset.

Pesaran (2007) suggests a way of constructing a test robust to the presence of cross-section dependence in a panel. The test uses a cross-sectionally augmented Dickey-Fuller regression (CADF):

$$\Delta y_{it} = \mu_i + \omega_i\, t + \alpha_i y_{i,t-1} + \upsilon_i \bar{y}_{t-1} + \sum_{j=1}^{p_i} \lambda_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{p_i} \varpi_{ij} \Delta \bar{y}_{i,t-j} + \varepsilon_{it}, \quad (17)$$

where $\bar{y}_t$ is a cross-section mean. The presence of the lagged cross-section mean and its differences suffices to filter out the effect of an unobserved common factor. Let us denote $\tilde{t}_{i,T_i,N}(p_i)$ as the t-statistic for α_i in the CADF regression. Note that the t-statistic depends on the N-dimension here as well, reflecting the cross-sectional dependence. Pesaran (2007) shows that in a standard ADF regression, the t-test has a high empirical size in the presence of cross-sectional dependence, while it does not have this in the CADF regression. Assuming a balanced panel, $T_i = T$ for all i's. We also set $p_i = p = 1$ for all i's. Our notation then simplifies to $\tilde{t}_{i,T_i,N}(p_i) = \tilde{t}_i(T, N)$. We employ a truncated version of this statistic, restricting it to the interval between -6.42 and 1.70, which improves its finite sample properties. The CADF t-statistic is then used to construct a cross-sectionally augmented version of the IPS test:

$$\bar{t}^{\dagger} = \frac{1}{N} \sum_{i=1}^{N} \tilde{t}_i(N, T). \quad (18)$$

Critical values for this statistic are available in Pesaran (2007).

We use Gauss code to conduct our empirical analysis. The IPS test is implemented using Nonstationary Panel Time Series Module 1.3 for Gauss (NPT 1.3) written by Chihwa Kao. We programmed the CADF and CIPS tests ourselves.[7] We report our results for the CADF t-tests in Tables 3 and 4-5. The MSA are in

[7]Our code was cross-checked with the Gauss procedures kindly provided by professor Pesaran and his research assistant Takashi Yamagata. They yielded identical results.

alphabetical order and we calculate $\tilde{t}_i(T, N)$ for prices, rents, and price-rent ratios. As expected, the t-CADF statistic rejects much less often than the standard t-ADF (not reported). For the 23 MSA, only two price-rent ratios are deemed stationary for the whole sample period. For the 273 MSA, we report the top and bottom 25 based on P/r. Table 6 gives a summary of all conducted panel data unit root tests for both 23 MSA and 273 MSA. The tests indicate that using the whole sample period in both cases, prices and rents are integrated of order one and the price-rent ratio is non-stationary. P/r inverse (results not reported here) is not stationary either. Results based on the IPS test, while quantitatively different, are qualitatively the same as the results based on the CIPS test. Given that prices and rents series have the same order of integration, the natural next step is testing for cointegration in panel data. The cointegration test can indicate whether it is appropriate to formulate a dynamic model of the dependence of house prices on rents in terms of first differences or whether we should formulate an error-correction model.

4.2 Cointegration

The cointegration tests employed in this paper rely on the results of Pedroni (1999, 2004). Pedroni (1999) describes the framework for testing for cointegration in panel datasets with $m = 2, ..., M$ explanatory variables and Pedroni (2004) covers the case for just one regressor. The hypothesized cointegrating regression is

$$y_{i,t} = \mu_i + \omega_i\, t + \psi_i\, x_{i,t} + \zeta_{i,t} \text{ for } t = 1, ..., T,\ i = 1, ..., N. \tag{19}$$

Again, T is the time dimension and N the cross-sectional dimension. The slope coefficient ψ_i and the fixed-effects parameter μ_i are allowed to vary across individual panel members. Also included is an individual time trend with a coefficient ω_i. We substitute house prices for y and rents for x in the regression equation.

There are seven residual-based statistics proposed by Pedroni (1999).[8] The first four are based on pooling along the within-dimension and test the null hypothesis of no cointegration: $H_0 : \gamma_i = 1$ for all i where γ_i is the autoregressive coefficient of the residual $\hat{\zeta}_i$ extracted from estimating the regression equation (19). The alternative hypothesis is $H_1 : \gamma_i = \gamma < 1$ for all i's, i.e. it assumes a common value for γ_i's. These four statistics are a non-parametric variance ratio statistic, non-parametric statistics similar to the Phillips and Perron rho- and t-statistics, and a parametric statistic similar to the augmented Dickey-Fuller t-statistic.

The remaining three statistics are based on pooling along the between-dimension and again test the null hypothesis of no cointegration: $H_0 : \gamma_i = 1$ for all i, this time versus the alternative hypothesis $H_1 : \gamma_i < 1$ for all i, i.e. no common value for the autoregressive coefficient is presumed in this case. The statistics use a group mean approach and are again respectively analogous to the Phillips and Perron rho- and t-statistic, and to the augmented Dickey-Fuller t-statistic. Similarly to testing for unit roots in panel data, the group statistics are of main interest since a potential failure to reject the null hypothesis of a unit root in residuals from the panel data regression may be due to heterogenous autoregressive coefficients rather than due to the presence of a unit root. Pedroni (2004) also conducts some Monte Carlo experiments that show that the finite sample properties of the group ADF t-statistic dominate the properties of the other two group tests and hence we employ this one in our analysis.

To calculate the values of the Pedroni test statistic, we used Gauss code from Wagner and Hlouskova (2007). Since a panel data cointegration test robust to the presence of cross-correlation is not yet available, we used the bootstrapping methodology from Gallin (2006) and Maddala and Wu (1999) to calculate critical values for the cointegration test. The bootstrapping methodology preserves the

[8] For explicit formulae, see Table 1 in Pedroni (1999).

cross-sectional dependance in the residuals of (19) observed in the data. The results reported in Table 7 imply that prices and rents are clearly not cointegrated and hence the use of an error-correction model is not appropriate.

4.3 Bubble Indicator

Here we combine the previously described methodology of testing for unit roots and cointegration in panel data and further analyze the relationship between house prices and rents in levels. We propose an indicator summarizing the implications of the present value model. The theory suggests that there is a bubble if either (i) the price-level is non-stationary while the rent-level is stationary, or (ii) both series are of first order of integration but they are not cointegrated. In both cases, the relationship between the two variables breaks down and there is a bubble on the house market. The latter case prevails using the whole sample for both 23 and 273 MSA. However, we would also like to be able to assess how the likelihood of a bubble changes over time. In both datasets, we define overlapping ten-year intervals covering the two sample periods. In accordance with the theory, we define a bubble indicator to be 0 for stationary prices and one for non-stationary prices and stationary rents. For cases in between, one would ideally use a cointegration test. However, this is potentially problematic since the test is often not well defined in a given sub-sample due to a different order of integration of the two series and it is cumbersome to check (one has to conduct unit root tests for higher order differences in many cases). Therefore we propose to replace the cointegration test with a test for the stationarity of the price-rent ratio, which has an intuitive appeal. We equate the bubble indicator to the p-value of the CIPS test for P/r if it is not already 1 or 0.

We calculate the Pedroni tests and the CIPS unit root tests, which respectively allow for the possibility of different autocorrelation coefficients in the residuals of the cointegrating regression and in a given time series. For the CIPS tests, we extend the table for critical values in Pesaran (2007) by calculating the numbers for N=273 and T=10, which also enables us to assign p-values for P/r. We present our results in Tables 9 and 10. Looking at the CIPS unit root tests for 23 MSA, we can see that house prices were non-stationary and rents stationary from 1987 to 1988 when we use ten prior years of data in each year. This results in the indicator being 1. In 1990, the rents become non-stationary but the price-rent ratio is stationary, resulting in a bubble indicator of 0.03, below the 5% level of significance. Overall, the periods when the indicator is higher correspond to rising average price-rent ratios in Figure 5 prior to 1990 and prior to 2000. The indicator decreases by the end of a sample, which reflects a potential stagnation of the housing market. The rising indicator for 273 MSA also coincides with the corresponding rises in the price-rent ratios in Figure 6. Here the indicators can be calculated only since 1995, which is also a period with the tests suggesting a bubble. The indicator remains high until 2005 with a brief trough in 2002 and then again decreases at the end of the sample.

While the number of observations for our two bubble indicators is fairly small for a thorough time series analysis, we at least provide an illustrative comparison to major house market indicators. First, we use existing home sales, sale price of existing homes, and the housing affordability index from the National Association of Realtors. Second, we use new home sales sold and for sale, total construction spending, and total housing units started from the Census Bureau, Department of Commerce. Finally, we use the Housing Market Index (HMI) from the National Association of Home Builders. A survey of these indicators is given in Bauhmol (2005).

We split the indicators into two groups and depict their values together with both of our indicators in graphs 7 and 8 (all indicators are normalized to 1 in 1995). Interestingly, all of the indicators show a cooling of the housing market at the end of 2006. Initially, the bubble indicator for 273 MSA leads the bubble indicator for 23 MSA but they move together in the last five years or so. The HMI (see Figure 7) and the value of construction index (see Figure 8) have similar patterns as the bubble indicators. These observations are roughly confirmed in Table 11, which also indicates that patterns of correlations with housing indicators are similar for both bubble indicators.

5 Predictability of House Prices and Rents

The aggregate average price-to-rent ratio and its inverse are non-stationary and the same conclusion holds in our panel datasets. This prevents us from using the standard regression methodology and brings us to statistical predictability formulated in terms of Granger causality. First differences for both house prices and rents are stationary according to our panel data tests for unit roots. Therefore, we are in a position to test for causality between the two. Testing for causality gives an indication of whether changes in prices predict rents and vice versa. Similarly to recently developed panel data unit root and causality tests, there exists an analogous test for Granger causality in panel data with a short time-series dimension. This test is described in Hurlin (2004) and applied in Hurlin and Venet (2004).

Let y_i and x_j be two stationary variables. Consider the following linear model:

$$y_{it} = \mu_i + \sum_{l=1}^{L} \varphi_i^{(l)} y_{i,t-l} + \sum_{l=1}^{L} \delta_i^{(l)} x_{i,j,t-l} + \xi_{it}. \tag{20}$$

ξ_{it} are normally i.i.d. with zero mean and finite heterogeneous variances and $\xi_i = (\xi_{i1}, ..., \xi_{iT})'$ are independently distributed across groups. The null hypothesis assumes that x does not help in predicting y for any of the N individual units in the panel. It is referred to as Homogeneous Non Causality (HNC) and can be formally stated as:

$$H_0: \ \delta_i = 0, \ \ \forall i = 1, ..., N, \tag{21}$$

where $\delta_i = (\delta_i^{(1)}, ..., \delta_i^{(L)})'$. The alternative hypothesis encompasses the possibility that there are N_1 individual units with no causality and is defined as:

$$\begin{aligned} H_1: \ \ & \delta_i = 0, \ \ \forall i = 1, ..., N_1, \\ & \delta_i \neq 0, \ \ \forall i = N_1 + 1, ..., N, \end{aligned} \tag{22}$$

where $N_1 \in [0, N)$ is not known. Let $W_{NT}^{HNC} = (1/N) \sum_{i=1}^{N} W_{iT}$ where W_{it} denotes the Wald statistic associated with the individual test of H_0 for each $i = 1, \ldots, N$. Hurlin (2004) shows that the approximated standardized statistic

$$Z_{NT}^{HNC} = \sqrt{\frac{N}{2 \times L} \times \frac{(T - 2L - 5)}{(T - L - 3)}} \times \left[\frac{(T - 2L - 3)}{(T - 2L - 1)} W_{NT}^{HNC} - L \right] \tag{23}$$

converges in distribution to $N(0, 1)$ as $N \rightarrow \infty$ for a fixed $T > 5 + 2L$.

We first verify that there is no cross-sectional dependence in the residuals from regression (20). See the results of the CD test in Table 8. We then use Intercooled Stata 9.2. to run panel data Granger causality tests with $L = 1$. Results in Table 8 show that the null of HNC can only be rejected for the 23 MSA in the direction from differences in rents to differences in prices. Even in this case, the p-value is barely above the 10% level of significance. In other words, there is a (statistically speaking) causal two-way relationship between changes in house prices and changes in rents, in spite of the fact that the connection between the levels of the two variables breaks down due to the presence of a bubble.

6 Summary

We study the implications of the standard present value formula for the order of integration, cointegration and Granger-causality between house prices and rents, with and without a bubble term. We analyze two panels of data using recent advances in panel data econometrics. House prices are from one source only but rents are either calculated as a part of the consumer price index or estimated as a fair market rent. Utilization of the former results in a dataset with a longer time span and a smaller cross-sectional dimension. Using the latter implies a large cross-section and a shorter time series.

Over the whole sample period, the house prices and rents either have a different order of integration or are not cointegrated. This conclusion is confirmed using the price-rent ratio. This is consistent with the presence of a bubble term in our asset pricing model and it implies that one cannot use an error correction model for house prices with rents as the fundamental factor. We proceed a step further with our analysis and formulate a simple procedure that can help to determine the extent to which there is a (statistical) discrepancy between house prices and their fundamentals. At each point in time, we investigate a panel of house prices and rents for the last ten years. If the prices are non-stationary but rents are not, we view that as an indication of a bubble. If prices are stationary, the bubble indicator is zero. In all other situations it is equal to the p-value of the panel data unit root test for the price-rent ratio. Our bubble indicator coincides fairly well with the pattern of price-rent ratios, the HMI and the value of construction index. However, it has the advantage of being able to determine formally whether the price-rent ratio is "too high." In both regional data sets, price-rent ratios are suggestive of rational bubbles in several periods but mainly in the late 1980s and the early 2000s up to 2005.

Non-stationarity of the price-to-rent and rent-to-price ratios is documented for both the aggregate series and in our panels. In such a case, standard regression techniques cannot be used to investigate the predictive properties of the price-to-rent ratios for the growth rates of house prices or rents. However, since prices and rents are both $I(1)$, we can test for panel Granger causality. In spite of the fact that it can take over three decades for house prices to revert to a fundamental value corresponding to earnings, the first differences of house prices have predictive power with respect to rents and vice versa.

There are several possible extensions of our study, mainly considering different measures of property prices or variables providing information about fundamentals. On the price side, one can for instance attempt to use a quality adjusted index or correct for an upward bias present in the repeated-sales index used in this paper. On the side of fundamentals, yet another measure for rents can be employed, such as the owners' equivalent rent series. Also, income and interest rates can be included as explanatory variables. However, all these improvements come at a cost in the terms of reduced cross-sectional and/or time dimension of the data, a more complex structural theoretical framework, and a greater computational burden.

References

Bailey, M.J., Muth, R.F., and H.O. Nourse (1963).“A Regression Method for Real Estate Price Index Construction," Journal of American Statistical Association 58, 933-942.

Baumohl, B. (2005). The Secrets of Economic Indicators. Wharton School Publishing, New Jersey.

Breitung, J. (2000). “The Local Power of Some Unit Root Tests for Panel Data," in B. Baltagi (ed.) *Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels,* Amsterdam: JAI Press, 161-178.

Calhoun, C.A. (1996). “OFHEO House Price Indexes: HPI Technical Description," Office of Federal Housing Enterprise Oversight.

Campbell, J. Y. and R. J. Shiller (1987). “Cointegration and Tests of Present Value Models," The Journal of Political Economy, 95:5, 1062-1088.

Campbell, S.D., Davis, M.A. , Gallin, J., and R.F. Martin (2006). “What Moves Housing Markets: A Trend and Variance Decomposition of the Rent-Price Ratio." Federal Reserve Board Working Paper.

Capozza, D.R. and P.J. Seguin (1996). “Expectations, Efficiency, and Euphoria in the Housing Market." Regional Science and Urban Economics 26, 369-386.

Case, K.E. and R.J. Shiller (1987). “Prices of Single Family Homes since 1970: New Indexes for Four Cities," NBER Working Paper 2393.

Case, K.E. and R.J. Shiller (1989). “The Efficiency of the Market for Single-Family Homes," The American Economic Review 79, 152-137.

Clark, T.E. (1995). “Rents and Prices of Housing across Areas of the United States: A Cross-section Examination of the Present Value Model." Regional Science and Urban Economics 25, 237-247.

Choi, I. (2001). “Unit Root Tests for Panel Data," Journal of International Money and Finance 20, 249-272.

Engle, R.F. and C.W.J. Granger (1987). “Co-integration and Error Correction: Representation, Estimation, and Testing," Econometrica 55, 251-276.

Gallin, J. (2004). “The Long-Run Relationship between House Prices and Rents," Federal Reserve Board Working Paper.

Gallin, J. (2006). “The Long-Run Relationship between House Prices and Income: Evidence from Local Housing Markets," Real Estate Economics 34:3, 417-438.

Hamilton, J. D. (1986). “On Testing for Self-Fulfilling Speculative Price Bubbles," International Economic Review 27(3), 545-552.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton, New Jersey.

Hardi, K. (2000). “Testing for Stationarity in Heterogeneous Panel Data," Econometric Journal 3, 148-161.

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

Hurlin, Ch. (2004). “Testing Granger Causality in Hetergoeneous Panel Data Models with Fixed Coefficients," Working Paper 2004-05, Laboratoire d'Economie d'Orléans.

Hurlin, Ch. and B. Venet (2004). “Financial Development and Growth: A Reexamination Using a Panel Granger Causality Test," Working Paper 2004-18, Laboratoire d'Economie d'Orléans.

Im, K.S., Pesaran, M.H., and Y. Shin (2003). “Testing for Unit Roots in Heterogeneous Panels," Journal of Econometrics 115, 53-74.

Levin, A., Lin., C.F., and C. Chu (2002). “Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," Journal of Econometrics 101, 1-24.

Liow, K.H. and Y. Li (2006). “Net Asset Value Discounts for Asian-Pacific Real Estate Companies: Long-run Relationships and Short-term Dynamics," The Journal of Real Estate Finance and Economics, 33:4, 363-388.

Lucas, R.E., (1978),“Asset prices in an Exchange Economy," Econometrica 46, 1429-1445.

Maddala, G.S. 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, 631-52.

Malpezzi, S. (1999). “A Simple Error Correction Model of House Prices," Journal of Housing Economics 8, 27-62.

McCarthy, J. and R. Peach (2004). “Are Home Prices the Next Bubble?" Federal Reserve Bank of New York Economic Policy Review 10:3, 1-17.

Pedroni, P. (1999). “Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors," Oxford Bulletin of Economics and Statistics 61, 653-70.

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, 597-625.

Pesaran M.H. (2004). "General Diagnostic Tests for Cross Section Dependence in Panels," CESIFO Working Paper 1229.

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.

Posen, A. (1998). *Restoring Japan's Economic Growth.* Washington: Institute for International Economics.

Shiller, R.J. (2005). *Irrational Exuberance* 2nd ed. Princeton: Princeton University Press.

Smith, M. and G. Smith (2006). "Bubble, Bubble, Where's the Housing Bubble?" The Brookings Papers on Economic Activity 1, 1-50.

U.S. Department of Housing and Urban Development. (1995)."Fair Market Rents: Overview."

Wang, P. (2000). "Market Efficiency and Rationality in Property Investment," Journal of Real Estate Finance and Economics, 21:2, 185-201.

Wagner, M. and J. Hlouskova (2007). "The Performance of Panel Cointegration Methods: Results from a Large Scale Simulation Study," Institute for Advanced Studies Vienna, Working Paper.

Table 1: Price-rent ratios using indices, 273 MSA

Note:

This table specifies the number of regions where a given P/r condition is satisfied.

	P/r>1.5	P/r>1.3	P/r<0.9	P/r<0.8
1986	2	11	142	76
1987	5	15	145	72
1988	7	20	141	76
1989	3	18	130	72
1990	2	16	119	60
1991	1	5	116	50
1992	0	1	116	47
1993	0	0	105	36
1994	0	0	5	0
1995	0	0	0	0
1996	0	0	1	0
1997	0	0	1	0
1998	0	0	2	0
1999	0	5	4	0
2000	2	14	1	0
2001	7	28	3	0
2002	10	39	3	0
2003	16	54	4	0
2004	42	84	3	0
2005	74	118	2	0
2006	97	137	3	1

Table 2: Diagnostic tests for cross section dependence in panels

Notes:
(1) ADF regression: intercept, trend, and the first lag of the dependent variable.
(2) Under the null of no cross section dependence: $CD \Rightarrow N(0,1)$.
(3) *** - significant at the 1 % level, ** - significant at the 5 % level, * - significant at the 10 % level.

price-level		price-diff.		rent-level		rent-diff.
CD		CD		CD		CD
23 MSAs, 1978:01-2006:02, semi-annual data						
75.12	***	-0.69		29.13	***	-0.64
273 MSA, 1986-2006, annual data						
566.00	***	-1.99	**	209.01	***	1.43

Table 3: Cross-sectionally augmented Dickey-Fuller tests, 23 MSA

Notes:
(1) Sample: 1978:01-2006:02, semi-anual data, levels only.
(2) CADF regression: intercept; trend; the first lags of the difference of the dependent variable, the difference of the cross section mean, and the cross-section mean; the difference of the cross-section mean.
(3) Critical values for the CADF t-statistic are from Pesaran (2007), Table Ic: 1% -4.52 (denoted ***), 5% -3.79 (denoted **), and 10% -3.44 (denoted *).
(4) MSA are in alphabetical order.

MSA	price CADF	rent CADF	P/r CADF	P/r rank
Atlanta-Sandy Springs-Marietta, GA	-0.37	-0.41	-1.15	7
Boston-Quincy, MA	-1.90	-3.21	-1.69	8
Chicago-Naperville-Joliet, IL	-6.42 ***	-6.42 ***	-4.42 **	23
Cincinnati-Middletown, OH-KY-IN	-4.73 ***	-1.32	-2.75	19
Cleveland-Arkon, OH	-4.73 ***	-4.53 ***	-2.23	12
Dallas-Plano-Irving, TX	-2.95	-1.68	-2.93	21
Denver-Aurora, CO	0.89	-0.12	-0.78	5
Detroit-Livonia-Dearborn, MI	-1.08	-3.03	-0.72	3
Honolulu, HI	-0.13	-0.79	-0.49	2
Houston-Sugar Land-Baytown, TX	-3.05	-5.18 ***	-2.52	17
Kansas City, MO-KS	-0.01	-2.50	-0.76	4
Los Angeles-Long Beach-Glendale, CA	-1.78	-1.79	-2.5	16
Miami-Miami Beach-Kendall, FL	-0.45	-0.03	-0.91	6
Milwaukee-Waukesha-West Allis, WI	-4.40 **	-2.36	-2.33	15
Minneapolis-St. Paul-Bloomington, MN-WI	0.65	-5.06 ***	-0.42	1
New York-White Plains-Wayne, NY-NJ	-3.39	-3.90 **	-2.82	20
Philadelphia, PA	-1.50	-1.70	-1.87	9
Pittsburgh, PA	-3.32	-2.76	-2.27	13
Portland-Vancouver-Beaverton, OR-WA	-2.08	-2.29	-1.89	10
San Diego-Carlsbad-San Marcos, CA	-2.34	-3.18	-2.63	18
San Francisco-San Mateo-Redwood City, CA	-2.80	-2.00	-2.28	14
Seattle-Bellevue-Everett, WA	-2.96	-2.95	-3.51 *	22
St. Louis, MO-IL	-3.04	-2.08	-2.17	11

Table 4: Cross-sectionally augmented Dickey-Fuller tests, 273 MSA, top 25 by price-rent ratio

Notes:
(1) Sample: 1986-2006, anual data, levels only.
(2) CADF regression: intercept; trend; the first lags of the difference of the dependent variable, the difference of the cross section mean, and the cross-section mean; the difference of the cross-section mean.
(3) Critical values for the CADF t-statistic are from Pesaran (2007), Table Ic: 1% -5.41 (denoted ***), 5% -4.17 (denoted **), and 10% -3.64 (denoted *).
(4) MSA are in alphabetical order.

MSA	price CADF	rent CADF	P/r CADF	P/r rank
Ann Arbor, MI	-0.82	-1.37	0.20	8
Athens-Clarke County, GA	1.48	-1.44	-0.11	12
Barnstable Town, MA	1.44	-1.71	1.70	1
Bend, OR	-0.11	-1.56	0.53	3
Bridgeport-Stamford-Norwalk, CT	-0.06	-1.51	-0.56	19
Canton-Massillon, OH	-1.46	-2.20	-0.46	17
Deltona-Daytona Beach-Ormond Beach, FL	1.16	-1.46	-0.59	20
Des Moines-West Des Moines, IA	-2.68	0.39	0.08	9
El Paso, TX	-2.10	-3.34	-0.52	18
Flint, MI	-0.60	-2.11	-0.59	21
Jackson, MI	0.72	-3.15	0.00	10
Lakeland, FL	0.17	-1.52	0.52	4
Las Vegas-Paradise, NV	-0.96	-2.89	0.23	7
Lexington-Fayette, KY	-1.15	-2.18	0.47	5
Mansfield, OH	-0.90	-2.17	-0.64	23
Monroe, LA	-4.84 **	-1.50	-0.32	14
New Haven-Milford, CT	-1.54	-1.68	-0.70	25
Olympia, WA	-0.19	-1.29	-0.32	15
Poughkeepsie-Newburgh-Middletown, NY	-0.50	-1.36	-0.06	11
Rochester, MN	-1.28	0.98	-0.27	13
Saginaw-Saginaw Township North, MI	0.34	-3.01	1.02	2
Sherman-Denison, TX	-2.81	-1.31	-0.68	24
Spokane, WA	-0.63	-1.00	-0.61	22
St. Joseph, MO-KS	-1.42	-2.42	-0.43	16
Toledo, OH	0.63	-2.17	0.29	6

Table 5: Cross-sectionally augmented Dickey-Fuller tests, 273 MSA, bottom 25 by price-rent ratio

Notes:
(1) Sample: 1986-2006, anual data, levels only.
(2) CADF regression: intercept; trend; the first lags of the difference of the dependent variable, the difference of the cross section mean, and the cross-section mean; the difference of the cross-section mean.
(3) Critical values for the CADF t-statistic are from Pesaran (2007), Table Ic: 1% -5.41 (denoted ***), 5% -4.17 (denoted **), and 10% -3.64 (denoted *).
(4) MSA are in alphabetical order.

MSA	price CADF	rent CADF	P/r CADF	P/r rank
Abilene, TX	-6.42 ***	-2.13	-4.84 **	271
Amarillo, TX	-6.34 ***	-0.18	-4.42 **	268
Atlantic City, NJ	-5.26 **	-2.28	-3.60	255
Beaumont-Port Arthur, TX	-3.62	-2.20	-3.99 *	262
Boulder, CO	-2.97	-1.80	-4.53 **	269
Burlington-South Burlington, VT	-5.79 ***	-1.97	-3.83 *	259
Colorado Springs, CO	-4.27 **	-1.66	-3.47	249
Corpus Christi, TX	-6.07 ***	-1.45	-3.74 *	256
Dayton, OH	-3.77 *	-3.17	-3.47	250
F. Lauderd.-Pompano B.-Deerfield B., FL	0.53	-1.56	-3.58	253
Greeley, CO	-3.51	0.19	-3.97 *	261
Houma-Bayou Cane-Thibodaux, LA	-5.97 ***	-2.22	-3.56	252
Knoxville, TN	-2.87	-2.07	-3.91 *	260
Los Angeles-Long Beach-Glendale, CA	-6.42 ***	-0.82	-3.55	251
Madison, WI	-2.89	-3.02	-4.77 **	270
Ocean City, NJ	-3.90 *	-2.43	-4.10 *	264
Oxnard-Thousand Oaks-Ventura, CA	-5.95 ***	-2.74	-4.10 *	265
Philadelphia, PA (MSAD)	-4.93 **	-1.56	-6.42 ***	273
Phoenix-Mesa-Scottdale, AZ	-3.50	-1.84	-4.02 *	263
Reading, PA	-6.17 ***	-1.42	-4.20 **	267
San Diego-Carlsbad-San Marcos, CA	-3.84 *	-1.03	-3.80 *	258
Santa Cruz-Watsonville, CA	-3.88 *	-2.33	-4.11 *	266
Scranton-Wilkes-Barre, PA	-3.21	-0.28	-5.49 ***	272
Shreveport-Bossier City, LA	-6.42 ***	-1.31	-3.78 *	257
Waterloo-Cedar Falls, IA	-1.91	-2.66	-3.58	254

Table 6: Panel data unit root tests

Notes:
(1) The IPS test is based on the individual ADF regressions with an intercept, trend, and th first lag of the dependent variable. The test statistic has an asymptotic standardized normal distribution.
(2) The CIPS test is based on the individual CADF regressions with an intercept; a trend; the first lags of the difference of the dependent variable, the difference of the cross section mean, and the cross-section mean; and the difference of the cross-section mean. Critical values for the CIPS statistic are from Pesaran (2007), Table IIc. For the 23 MSA, these are: 1% -2.85, 5% -2.71, and 10% -2.63. For the 273 MSA, these are: 1% -2.70, 5% -2.57, and 10% -2.50.
(3) Significance at the 1%, 5%, and 10% levels is denoted as ***, **, and *, respectively, for both tests.
(4) For both tests, the null hypothesis is that of a unit root (it assumes individual an unit root process).

23 MSA, 1978:01-2006:02, semi-annual data							
Method	price-level	price-diff.		rent-level	rent-diff.		p/r
IPS	1.63	-20.29	***	-2.60	-13.00	***	5.48
CIPS	-2.26	-5.02	***	-2.58	-4.16	***	-2.00
273 MSAs, 1986-2006, annual data							
Method	price-level	price-diff.		rent-level	rent-diff.		p/r
IPS	6.17	-19.23	***	5.84	-22.16	***	12.69
CIPS	-2.35	-2.74	***	-1.90	-3.00	***	-2.12

Table 7: Pedroni cointegration tests with one regressor

Notes:
(1) 23 MSA: Sample 1978:01-2006:02, semi-anual data. The critical values were generated using 50,000 simulations by bootstrapping to preserve cross-sectional dependence and including autocorrelation. They are 1% -3.60, 5% -2.56, and 10% -2.03.
(2) 273 MSA: Sample 1986-2006, annual data. The critical values were generated using 50,000 simulations by bootstrapping to preserve cross-sectional dependence and including autocorrelation. They are 1% -9.17, 5% -5.97, and 10% -4.47.
(3) Significance at the 1%, 5%, and 10% levels is denoted as ***, **, and *, respectively, for both tests.
(4) The Pedroni test included a constant and a trend as deterministic variables.

Stat.	223 MSA	273 MSA
group adf-stat	1.82	28.69

Table 8: Hurlin Tests for Homogeneous Non-Causality in Panel Data

Notes:
Both CD and Z_{NT}^{HNC} asymptotically follow the standardized normal distribution. P-values for two-sided tests are reported.

H_0	CD	P-value	Z_{NT}^{HNC}	P-value
23 MSA, 1978:01-2006:02, semi-annual data				
price-diff. does not Granger cause rent-diff.	-0.70	> 0.50	-1.70	0.09
rent-diff. does not Granger cause price-diff.	1.35	0.18	-1.58	0.11
273 MSA, 1986-2006, annual data				
price-diff. does not Granger cause rent-diff.	0.96	0.34	-5.62	0.00
rent-diff. does not Granger cause price-diff.	-0.39	> 0.50	-5.75	0.00

Table 9: Bubble indicator, 23 MSA

Notes:
(1) Sample: 1978:01-2006:02, semi-anual data, ten-year data windows end in a given year.
(2) The CIPS test is based on the individual CADF regressions with an intercept; a trend; the first lags of the difference of the dependent variable, the difference of the cross section mean, and the cross-section mean; and the difference of the cross-section mean. Critical values for the CIPS statistic are from Pesaran (2007), Table IIc. 1% -2.92, 5% -2.73, and 10% -2.63.
(3) The Pedroni test included a constant and a trend as deterministic variables. The critical values were generated using 50,000 simulations by bootstrapping to preserve cross-sectional dependence and including autocorrelation. They are 1% -5.97, 5% -4.78, and 10% -4.13.
(4) Significance at the 1%, 5%, and 10% levels is denoted as ***, **, and *, respectively, for both tests.
(5) The (price) bubble indicator is 1 if the price level is non-stationary while the rent is stationary at 10% significance level; it is 0 if the price level is stationary; otherwise, it is equal to the p-value of the CIPS test conducted for the price-rent ratio.

Year	price CIPS	price-dif. CIPS	rent CIPS	rent-dif. CIPS	coint. Pedroni	P/r CIPS	bubble indicator
1987	-2.21		-2.71 *		-1.32	-2.62	1.00
1988	-1.75	-2.72 *	-2.65 *	-2.42	1.64	-2.62	1.00
1989	-2.71 *	-2.46	-3.22 ***	-2.35	0.56	-2.89 **	0.00
1990	-2.41	-2.67 *	-2.14	-2.60	0.95	-2.82 **	0.03
1991	-1.51	-2.75 **	-1.89	-2.56	0.66	-2.18	0.51
1992	-1.52	-3.02 ***	-1.55	-2.64 *	-3.11	-1.63	0.97
1993	-2.11	-2.56	-1.64	-2.63 *	-2.68	-2.31	0.35
1994	-2.21	-2.61	-1.53	-2.55	-1.74	-2.32	0.34
1995	-2.71 *	-2.95 ***	-1.83	-2.44	-2.81	-2.53	0.00
1996	-2.23	-2.55	-1.83	-2.51	-2.05	-2.22	0.46
1997	-2.79 **	-3.29 ***	-1.72	-3.11 ***	-4.08	-2.44	0.00
1998	-2.96 ***	-3.15 ***	-2.08	-3.15 ***	-6.44 ***	-2.70 *	0.00
1999	-1.91	-3.07 ***	-2.25	-2.79 **	-2.35	-1.85	0.87
2000	-1.69	-2.90 **	-2.50	-2.84 **	-0.97	-1.80	0.90
2001	-2.07	-3.59 ***	-2.66 *	-2.87 **	0.16	-1.88	1.00
2002	-1.83	-3.48 ***	-1.89	-2.51	0.39	-2.13	0.58
2003	-1.46	-3.59 ***	-1.37	-2.54	0.16	-1.80	0.90
2004	-1.58	-3.23 ***	-0.81	-2.41	0.69	-1.85	0.87
2005	-1.71	-3.14 ***	-0.71	-2.69 *	-0.81	-1.75	0.93
2006	-2.28	-3.03 ***	-0.94	-3.07 ***	-3.86	-2.34	0.32

Table 10: Bubble indicator, 273 MSA

Notes:
(1) Sample: 1976-2006, anual data, ten-year data windows end in a given year.
(2) The CIPS test is based on the individual CADF regressions with an intercept; a trend; the first lags of the difference of the dependent variable, the difference of the cross section mean, and the cross-section mean; and the difference of the cross-section mean. Critical values for the CIPS statistic were generated by replicating simulations from Pesaran (2007) leading to Table IIc for N=273 and T=10. They are 1% -3.18, 5% -2.94, and 10% -2.80.
(3) The Pedroni test included a constant and a trend as deterministic variables. The critical values were generated using 50,000 simulations by bootstrapping to preserve cross-sectional dependence and are 1% -31.92, 2% -26.42, 3% -23.89.
(4) Significance at the 1%, 5%, and 10% levels is respectively denoted as ***, **, and *, respectively, for both tests.
(5) The (price) bubble indicator is 1 if the price level is non-stationary while the rent is stationary at 10% significance level; it is 0 if the price level is stationary; otherwise, it is equal to the p-value of the CIPS test conducted for the price-rent ratio.

Year	price CIPS	price-dif. CIPS	rent CIPS	rent-dif. CIPS	coint. Pedroni	P/r CIPS	bubble indicator
1995	-2.03		-1.48		-18.05	-1.88	0.92
1996	-2.62	-3.19 ***	-1.57	-2.05	-16.32	-4.14 ***	0.00
1997	-2.39	-3.09 **	-2.35	-1.93	-14.06	-2.54	0.28
1998	-2.56	-2.85 *	-2.31	-2.73	-2.86	-0.85	1.00
1999	-2.74	-2.92 *	-3.33 ***	-2.67	-3.85	-1.18	1.00
2000	-2.75	-2.21	-1.94	-2.68	-6.92	-0.71	1.00
2001	-2.75	-2.28	-3.39 ***	-2.97 **	4.88	-2.74	1.00
2002	-2.56	-2.11	-2.22	-2.49	9.12	-2.50	0.31
2003	-1.42	-2.05	-2.21	-2.29	-0.46	-2.04	0.80
2004	-2.12	-2.11	-1.81	-2.65	-3.45	-1.94	0.88
2005	-2.40	-1.95	-1.40	-2.16	7.11	-1.78	0.96
2006	-2.37	-1.85	-1.26	-2.23	8.23	-2.15	0.69

Table 11: Correlation of bubble indicators with housing indicators

	Bub. ind. 273	Bub. ind. 23	Exist. home sales	Exist. home price	Afford. index	HMI	Housing starts	New home sales	Houses for sale	Value of constr.
Buble ind. 273	1.00	0.34	0.34	0.23	0.00	0.21	0.29	0.28	0.15	0.28
Bubble ind. 23	0.34	1.00	0.37	0.24	-0.03	0.29	0.40	0.35	0.12	0.52

Figure 1: Price-rent ratios in 23 MSA, Part 1/4

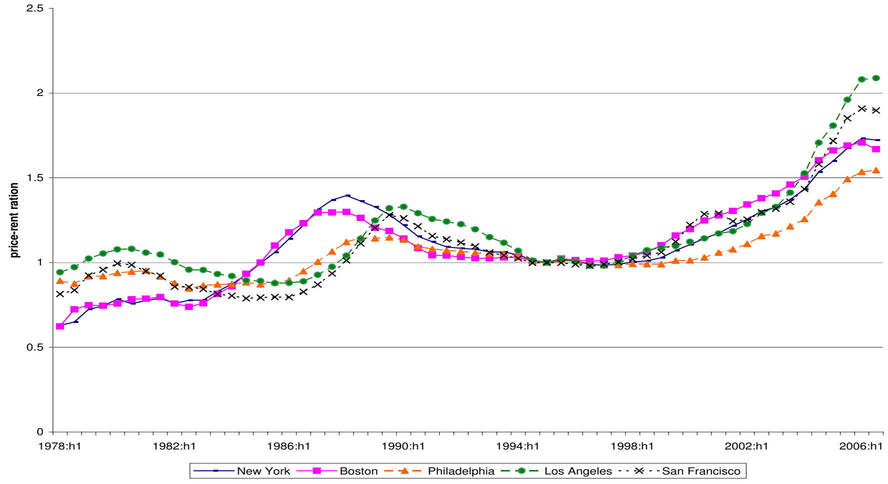


Figure 2: Price-rent ratios in 23 MSA, Part 2/4

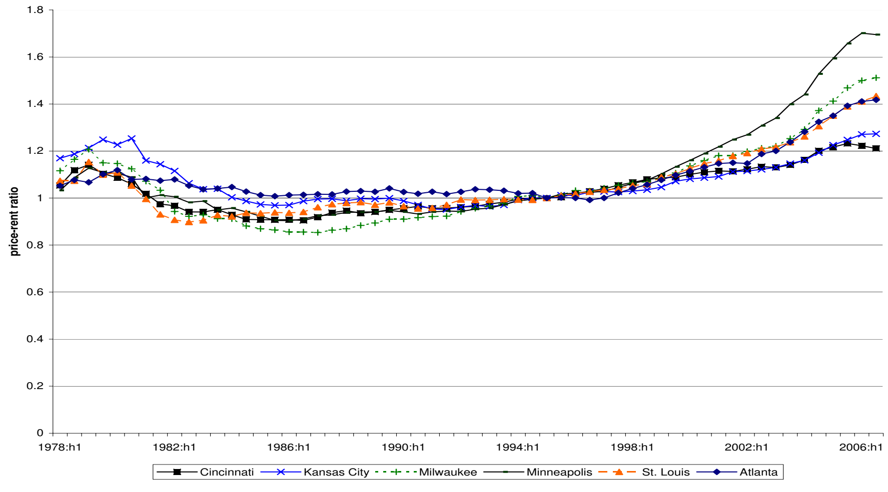


Figure 3: Price-rent ratios in 23 MSA, Part 3/4

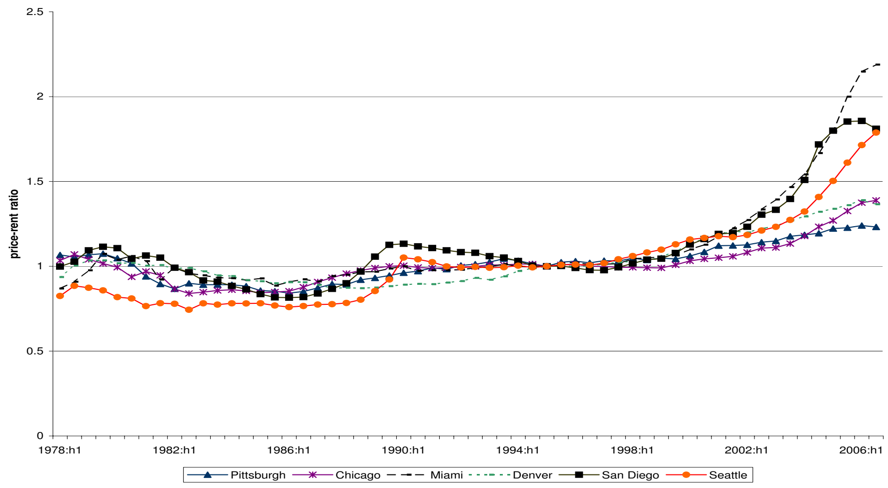


Figure 4: Price-rent ratios in 23 MSA, Part 4/4

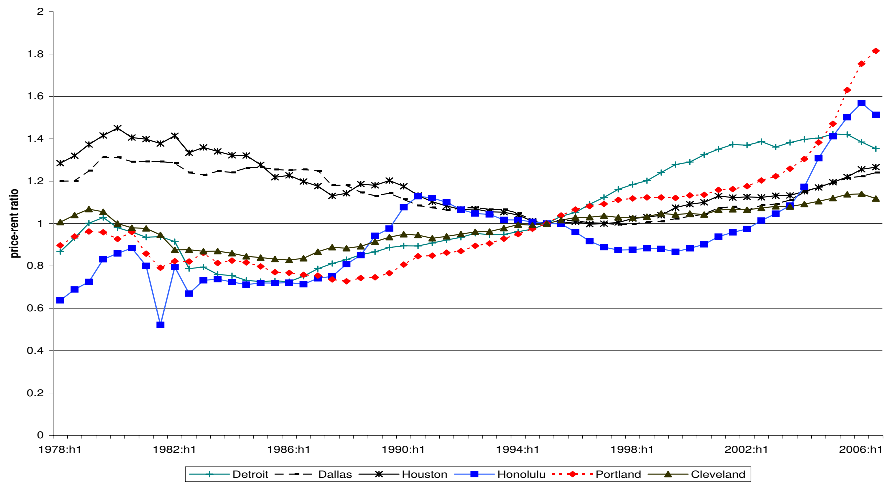


Figure 5: Average price-rent ratios in 23 MSA

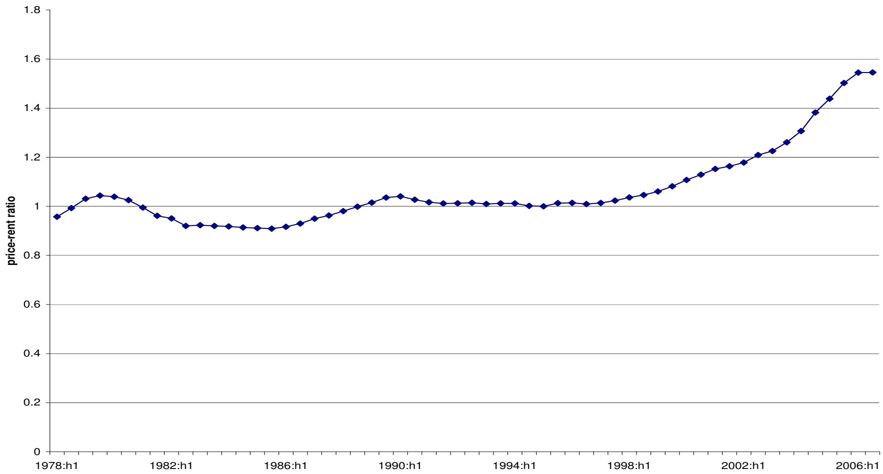


Figure 6: Average Price-rent Ratios in 273 MSA

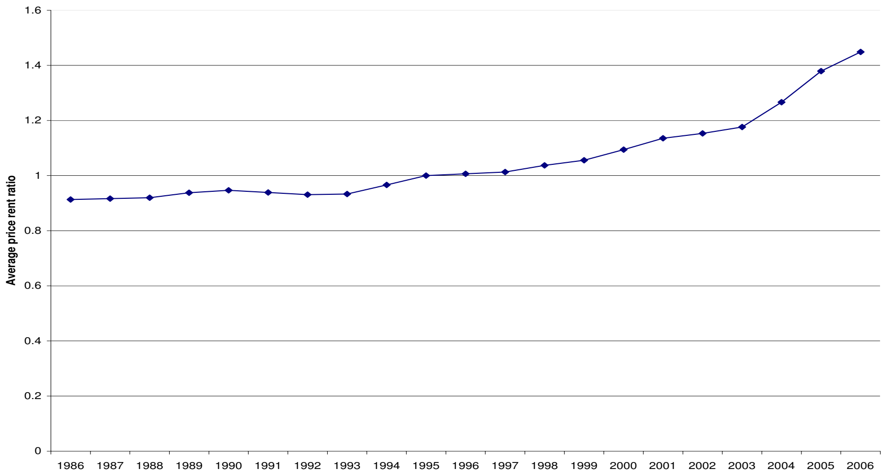


Figure 7: Bubble indicators, existing homes sales and prices, HMI, and affordability

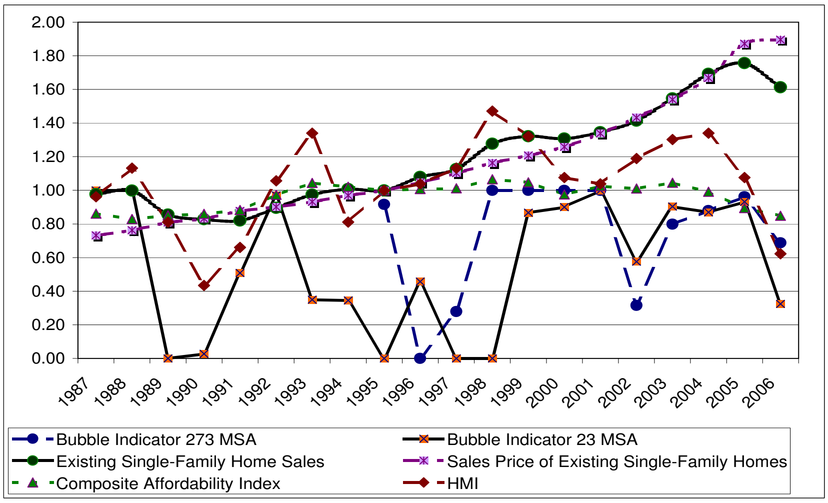
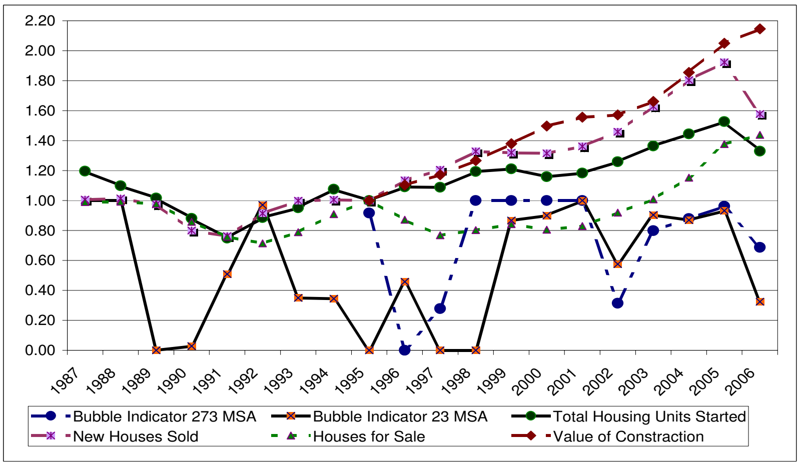


Figure 8: Bubble indicators, new houses sold and for sale, started units, and value of construction



Individual researchers, as well as the on-line and printed versions of the CERGE-EI Working Papers (including their dissemination) were supported from the following institutional grants:

- Center of Advanced Political Economy Research [Centrum pro pokročilá politicko-ekonomická studia], No. LC542, (2005-2009),
- Economic Aspects of EU and EMU Entry [Ekonomické aspekty vstupu do Evropské unie a Evropské měnové unie], No. AVOZ70850503, (2005-2010);
- Economic Impact of European Integration on the Czech Republic [Ekonomické dopady evropské integrace na ČR], No. MSM0021620846, (2005-2011);

Specific research support and/or other grants the researchers/publications benefited from are acknowledged at the beginning of the Paper.

Published by
Charles University in Prague, Center for Economic Research and Graduate Education (CERGE)
and
Economics Institute ASCR, v. v. i. (EI)
CERGE-EI, Politických vězňů 7, 111 21 Prague 1, tel.: +420 224 005 153, Czech Republic.
Printed by CERGE-EI, Prague
Subscription: CERGE-EI homepage: http://www.cerge-ei.cz

Editors: Directors of CERGE and EI
Managing editors: Deputy Directors for Research of CERGE and EI

ISSN 1211-3298
ISBN 978-80-7343-137-2 (Univerzita Karlova. Centrum pro ekonomický výzkum a doktorské studium)
ISBN 978-80-7344-126-5 (Národohospodářský ústav AV ČR, v. v. i.)

CERGE-EI
P.O.BOX 882
Politických vězňů 7
111 21 Praha 1
Czech Republic
http://www.cerge-ei.cz