

The Long-Run Relationship between House Prices and Income: Evidence from Local Housing Markets

Joshua Gallin, Federal Reserve Board *

April, 2003

Abstract

The proposition that “housing prices can’t continue to outpace growth in household income” (*Wall Street Journal*; July 25, 2002) is the received wisdom among many housing-market observers. More formally, many in the housing literature argue that house prices and income are cointegrated. In this paper, I show that the data do not support this view. Standard tests using 27 years of national-level data do not find evidence of cointegration. However, it is known that tests for cointegration have low power, especially in small samples. I use panel-data tests for cointegration that have been shown to be more powerful than their standard time-series counterparts to test for cointegration in a panel of 95 metro areas over 23 years. Using a bootstrap approach to allow for cross-correlations in city-level house-price shocks, I show that even these more powerful tests do not reject the hypothesis of no cointegration. Thus the error-correction specification for house prices and income commonly found in the literature may be inappropriate.

*Thanks to Doug Elmendorf, Steve Oliner, Jeremy Rudd, Dan Sichel, Bill Wascher, participants at the 2002 Federal Reserve System Conference on Regional Economics and the 2002 Regional Science Association International Conference, and seminar participants at the University of Georgia. Special thanks to Norm Morin and Peter Pedroni for their help and comments. The views presented are solely those of the author and do not represent those of the Federal Reserve Board or its staff.

1 Introduction

In the second half of the 1980s, real house prices rose about 3 percent per year.¹ Then in 1990 alone, prices tumbled almost 5 percent. Prices continued to fall, on balance, through the end of 1994, reversing more than half of the gains posted in the late 1980s; prices did not return to their 1989 level until almost ten years later. Many coastal cities experienced even wilder swings: Real house prices rose about 65 percent in Los Angeles and about 45 percent in Boston in the second half of the 1980s, only to fall 30 percent in L.A. and 20 percent in Boston during the next five years.

Many real-estate market observers think that housing prices got too far ahead of fundamentals in the 1980s—especially in many coastal cities—and that the poor performance of house prices during the first half of the 1990s was the inevitable aftermath. During the past five years, real house prices have moved up almost 5 percent per year, out-pacing the gains from a decade ago and sparking fears that the housing market is once again over-valued.

Of particular concern to many is the fact that house-price gains have dwarfed per capita income gains in recent years. As can be seen in Chart 1, the ratio of house prices to per capita personal income moved up recently after trending down for most of the previous 20 years.² More specifically, from the middle of 1997 to the middle of 2002, real house prices rose about 28 percent while real per capita personal income rose about 15 percent. In contrast, during the previous 20-year period, real house prices rose only 8 percent while real per capita income rose 35 percent. The recent performance of house prices relative to income is taken as evidence by some that house prices are out of line with “fundamentals,” and that prices must stagnate or fall to allow income to catch up.

This idea is commonly formalized in the housing literature by positing a cointegrating relationship between house prices and fundamentals such as income, and then estimating an error-correction specification (Abraham and Hendershott, 1996; Malpezzi, 1999; Capozza et al., 2002; Meen, 2002). That is, house prices and income are thought to be linked by a stable long-run relationship;

¹All house price measures in this paper are from the Office of Federal Housing Enterprise Oversight’s weighted repeat-sales price index unless otherwise noted. Nominal prices are deflated by the personal consumption deflator from the National Income and Product Accounts.

²In this paper, I focus on per capita income rather than household income for two reasons. First, household size—the link between per capita and household income—is endogenous to the household formation problem, and therefore to housing demand. Second, I do not have city-level data on the number of households.

they may drift apart temporarily, but their tendency is to return to their long-run equilibrium.

The purpose of this paper is to test this view of the housing market. If prices and income are cointegrated, then the gap between the two may be a useful indicator of when house prices are above or below their equilibrium values, and therefore a useful predictor of future house-price changes. Conversely, if prices and income are not cointegrated, then the error-correction specifications common in the literature are inappropriate, and house prices need not stagnate or fall just because they have grown more quickly than has income of late.³

Why might such a relationship exist? Imagine a simple supply-and-demand framework. Suppose that the supply of housing, including land, slopes up either because there is a limited supply of attractive land or because of zoning restrictions (Glaeser and Gyourko, 2002). Because demand-side shocks such as population and income are not stationary, we should expect house prices to be non-stationary as well. However, if other shocks to the demand curve and all shocks to the supply curve are stationary, then population, income, and price will be cointegrated in a way that depends on the elasticities of supply and demand.

Many researchers simply assume that house prices and fundamentals are cointegrated (Abraham and Hendershott, 1996; Capozza et al., 2002). Others implicitly assume that they are not (Poterba, 1991). Meen (2002) is the only paper I am aware of that tests for cointegration of prices and fundamentals using national-level data. His reported tests do not find evidence for cointegration at conventional significance levels. However, Meen argued that the test statistics are “near” their critical values, and therefore concluded that prices and fundamentals are cointegrated. One contribution of this paper is to show that using 27 years of national-level data, one does not find evidence that prices, income, and other fundamentals are cointegrated. Thus my results, if not my interpretation, are in accord with Meen, and suggest that it is inappropriate to model house-price dynamics using an error-correction specification.

However, cointegration tests are known to have low power, particularly in small samples (Banerjee, 1999). A time span of 27 years may be too short to estimate what may be a genuine long-run relationship with slow adjustment. Starting with Quah (1990) and Levin and Lin (1992), researchers have developed panel tests for unit roots and cointegration that are more powerful than their

³The absence of cointegration does not preclude house-price bubbles. For example, one might not find cointegration, which is a linear relationship, if the housing market is beset by non-linear rational bubbles.

standard time-series counterparts. The second, and main, contribution of this paper is to apply recently developed tests of Pedroni (1999) and Maddala and Wu (1999) to a panel of 95 U.S. cities over 23 years. I show that even these more powerful tests cannot reject the hypothesis that prices and income are not cointegrated. My results contradict those of Malpezzi (1999), who found that one *can* reject the null of no cointegration in a similar panel. However, Malpezzi used a panel unit root test, which overstates the likelihood of cointegration because it ignores the first-stage estimation in his residuals-based cointegration test.

The rest of this paper is organized as follows. In the following section I briefly describe a simple model of housing supply and demand. The purpose of the model is to motivate why prices, per capita income, and perhaps other variables might be cointegrated. In Section 3, I describe the national-level data and show that there is little evidence for cointegration using Engle and Granger’s (1987) Augmented Dickey-Fuller (ADF) test. Section 4 is the heart of the paper. In it, I describe several tests for cointegration in panel data. Three are from Pedroni (1999): one is a panel version of Phillips and Ouliaris’ (1990) variance ratio test, one is a panel version of their Z_α test, and one is a panel version of Engle and Granger’s (1987) ADF test. My fourth test is a Maddala and Wu (1999) version of Engle and Granger’s ADF test. I also describe a bootstrap approach that allows for arbitrary cross correlations of the city-level shocks. I describe the city-level data and show that, using the bootstrapped critical values, none of the tests rejects the null of no cointegration. A secondary result is that the correlations among local housing markets can have a large effect the tests. I conclude in Section 5.

2 Housing Prices and Fundamentals

House prices and fundamentals like income—or some transformation of them—will be cointegrated if they are linked by a long-run equilibrium relationship. The interaction of housing supply and demand offers the most obvious and simple way to characterize such a relationship. Consider a simple supply and demand model of housing. In it, the demand for owner-occupied housing depends on income, Y ; population, N ; wealth, W ; the user cost of housing, UC ; and other demand shifters, θ_d . The supply of housing depends on the price of housing, P ; the cost of new construction, C ; and other supply shifters, θ_s :

$$Q_d = D(Y, N, W, UC; \theta_d)$$

$$Q_s = S(P, C; \theta_s).$$

The user cost of capital, in turn, depends on the price of housing; mortgage rates, m ; income and property taxes, τ_y and τ_p ; maintenance and depreciation, δ ; and expected capital gains, cg :

$$UC = P[(1 - \tau_y)(m + \tau_p) + \delta - cg] = P \cdot A$$

where A represents the term in brackets. Then the price of housing can be written as a function of all the other variables:

$$P = F(Y, N, W, C, A; \theta_d, \theta_s)$$

A log-linearized solution to the model would relate the log house price to the logs of all the driving variables. Under the assumption that the unobserved components of the model are stationary and the coefficients of the log-linearization do not change, house prices will be cointegrated with those fundamentals that also have a unit root, and the relationship will depend on the elasticities of supply and demand.

The point is not that house prices *must* be cointegrated with fundamentals. Indeed this simple model illustrates that there are many reasons why such a cointegrating relationship need not exist. For instance, the price elasticity of supply may not be stable over time because of changes in regulatory conditions, the price elasticity of demand may not be stable because of changing demographics, or demand shifters such as local taxes may not be stationary. The model does show what kind of assumptions are needed to generate a cointegrating relationship. Whether one exists is an empirical question.

3 National-Level Tests for Cointegration

In this section I present tests of cointegration of national-level house prices and various fundamentals. Suppose that the hypothesized cointegrating regression is given by

$$x_{0,t} = \alpha + \delta t + \sum_{m=1}^M \beta_m x_{m,t} + e_t, \quad (1)$$

where $m = 1, \dots, M$ indexes $I(1)$ variables and $t = 1, \dots, T$ indexes time. If the residual e_t is stationary, then we say that the x 's are cointegrated. Here, I use the common two-step procedure for testing for cointegration suggested by Engle and Granger (1987), sometimes called an augmented Engle-Granger

(AEG) τ -test. In the first stage, I estimate Equation (1) by OLS to get \hat{e}_t . In the second stage, I conduct an augmented Dickey-Fuller (ADF) τ -test on the residuals. The critical values differ from those of the standard ADF test because the residuals are estimated in the first stage.

My source for house-price data is the repeat-sales price index for existing homes, which is published by the Office of Federal Housing Enterprise Oversight (OFHEO).⁴ The index is based on price changes for homes that are resold or refinanced, but does not control for changes to the house through improvement or neglect; that is, while it does hold some characteristics constant, it is not a true quality-adjusted price. In addition, the repeat-sales sample excludes homes with jumbo, FHA, or VA mortgages (Calhoun, 1996). I used the BEA's measure of total personal income, the Census Bureau's measure of population, and the BLS's measure of average hourly wages for construction workers. I used the Standard and Poor's 500 stock index to measure stock-market wealth. I included the level of the personal consumption deflator from the BEA to control for inflation. All data are quarterly.

To calculate the user cost of housing, I used a weighted average of the rates on fixed-rate and 1-year adjustable-rate contracts for 30-year loans; the weights were the origination shares.⁵ I set expected capital gains to the average percentage increase in the house-price index during the previous three years.

The standard unit root tests, which are available upon request, do not reject the hypotheses that house prices, per capita income, population, the stock market, and construction wages all have a unit root, but that the non-price component of the user-cost, A , does not. The results were the same when I assumed rational and myopic, as opposed to backward-looking, expectations. I therefore did not include non-price part of the user cost in the regressions.⁶

Table 1 displays AEG τ -tests for cointegration of several sets of variables using quarterly data from 1975:Q1 to 2002:Q2. The first-stage levels equation yields an estimate of the cointegrating vectors, which are shown in the upper panel of the table. The coefficient estimates indicate that per capita income and the construction wage have positive and statistically significant effects on house prices, but that population has no effect and the price level and stock market have negative effects. Of course, if there is no cointegrating relationship, these

⁴Alternative tests based on the average existing house price from the National Association of Realtors yielded similar results.

⁵My measures of Federal and state and local tax rates are from the FRB/US model. See Reifschneider, Tetlow, and Williams (1999) for more information about the FRB/US model.

⁶The inclusion of the non-price part of the user cost or the simple mortgage rates does not affect the results.

Table 1
National-Level Tests for Cointegration of
House Prices and Fundamentals

null of no cointegration, 1975:Q1 to 2002:Q1

First-Stage Levels Regressions				
dependent variable is log(price)				
Independent variables				
(log values)	1	2	3	4
per capita income	.70** (.05)	1.45** (.20)	1.57** (.15)	1.71** (.17)
population	—	-2.94** (.76)	—	-1.57 (1.26)
stock market	—	—	—	-.13** (.02)
construction wage	—	—	—	.25** (.13)
PCE deflator	.03 (.10)	-.38** (.14)	-.76** (.15)	-1.04** (.14)
trend ×10	—	—	-.74** (.12)	-.10 (.24)
Second-Stage Test Results				
AEG τ -stat	-2.2	-2.0	-1.9	-3.4
critical value (10%)	-3.5	-3.8	-3.8	-4.7

Notes: ** - Significant at .05. Standard errors are in parentheses.

levels regressions are spurious. The lower panel presents the second-stage tests with 10 percent critical values that are based on the dimension of the proposed cointegrating vector and the presence or absence of a time trend (Davidson and MacKinnon, 1993).⁷ The lower panel of Table 1 shows that in none of the cases can we find strong evidence for the cointegration of house prices and fundamentals. In other words, the national-level data do not support the view that the log levels of house prices and various fundamentals are linked by a long-run stationary relationship.⁸

One criticism of these tests is that they are known to have low power against the alternative of cointegration, particularly when, as is the case here, the sample size is small. Thus the evidence in Table 1 may not be convincing to those who have strong priors that house prices and fundamentals, particularly income, are cointegrated.

4 City-Level Tests for Cointegration

Quah (1990, 1994) and Levin and Lin (1992) were among the first to devise panel tests for unit roots and to show that they can offer a substantial improvement in power relative to separate tests for each cross-sectional unit of the panel.⁹ The literature has blossomed since then. Some, like Im, Pesaran, and Shin (1997) have devised tests that impose fewer restrictions than did Levin and Lin. Others, like Pedroni (1997, 1999, 2001) developed related tests for cointegration. Banerjee (1999) provides an overview of the literature.

In this section I briefly describe three panel cointegration tests of Pedroni (1999) and of Maddala and Wu (1999). One of the underlying assumptions of all the tests is that shocks are either independent across cross sections or that the cross-sectional dependence can be modeled as an aggregate time effect. For the purposes of this paper, that assumption implies that shocks to housing markets in, say, San Francisco and Seattle have the same correlation as shocks to housing markets in Philadelphia and New York. As this is an unattractive feature, I discuss bootstrapped versions of the tests that relax this assumption by allowing for the cross-sectional dependence among cities evident in the data. I then describe the city-level data and present the test results.

⁷The time trend is not statistically significant in Column 4. The stationarity results did not change when I excluded the time trend.

⁸Meen (2002) conducted similar tests and concluded that prices and fundamentals are cointegrated. However, his reported test statistics were quite far from conventional critical values.

⁹See also Levin, Lin, and Chu (2002).

4.1 Some Panel Cointegration Tests

Suppose that the hypothesized cointegrating regression for each city is given by

$$x_{0,i,t} = \phi_i + \delta_i t + \sum_{m=1}^M \beta_{m,i} x_{m,i,t} + e_{it}, \quad (2)$$

where $i = 1, \dots, N$ indexes the city, $m = 1, \dots, M$ indexes variables, and $t = 1, \dots, T$ indexes time. Notice that this specification admits city-specific intercepts and time trends.

In this paper, I use Pedroni's panel-data versions of Phillips and Ouliaris' (1990) variance ratio test, $P-v$; Phillips and Ouliaris' Z_α test, $P-Z_\alpha$; and Engle and Granger's (1987) ADF test, $P-AEG$. These residual-based panel tests have the same structure as do their time-series counterparts, but are constructed by pooling information from the cross sectional units.¹⁰ See Appendix A for details on how the pooling is done.

Pedroni (1999) shows that the appropriately standardized version of his test statistics are asymptotically normal.¹¹ As with the standard time-series cases, under the alternative, the $P-v$ test diverges to positive infinity, so the right tail of the normal distribution is used for rejection, and the $P-Z_\alpha$ and $P-AEG$ tests diverge to negative infinity, so the left tail is used to reject.

Maddala and Wu (1999) present an alternative, and very general, test for cointegration based on Fisher (1932). Maddala and Wu's test is based on averaging the p -values for *any* test from each cross-sectional unit.

Suppose that we implement any cointegration test for each cross-sectional unit. Under the null, the significance levels for each test, p_i are distributed uniformly over $(0, 1)$. This implies that

$$-2 \log p_i \sim \chi^2(2).$$

Under the assumption that the tests are independent, the Maddala-Wu (MW) test statistic is

$$MW = -2 \sum_{i=1}^N \log p_i \sim \chi^2(2N),$$

¹⁰Pedroni also presents a panel version of Phillips and Ouliaris' Z_t test. The results from the $P-Z_\alpha$ test and $P-Z_t$ were almost identical, so I only report the $P-Z_\alpha$ test results in this paper. In addition, the tests I use here require that the AR(1) terms in the second stage residual regressions for each city be equal to each other (and less than 1) under the alternative hypotheses. Pedroni also proposes versions of his tests that do not impose this equality restriction. The results are not sensitive to the choice of tests.

¹¹Pedroni calculated each test statistics' mean and variance by simulation. Pedroni generously gave me the RATS code to construct the test statistics.

and the right tail of the distribution is used for rejection. This test can be applied using any underlying test. However, one must simulate an approximation to the entire distribution of whatever test statistic one is using for each cross-sectional unit in order to calculate the p_i 's. The Maddala-Wu tests I present in this paper are based on AEG τ -tests, where I simulated the AEG τ -test distributions for a 23-year time series. That is, I did not use the asymptotic distribution of the τ statistic.

The validity of Pedroni's tests depends in part on the assumption that any cross-sectional correlations are adequately captured by an aggregate time effect; Maddala and Wu's test requires complete independence. The assumption of cross-sectional independence is clearly violated in the data, and local housing market shocks will likely be correlated in ways that are not captured by a simple time effect. One way to address this problem is to bootstrap empirical distributions of the test statistics under the null to calculate critical values for the test. The key is to maintain the cross-sectional dependence while resampling.

More specifically, following Maddala and Wu (1999), one can get the bootstrap sample by estimating

$$\Delta x_{m,i,t} = \eta_{m,i}(L)\Delta x_{m,i,t-1} + u_{m,i,t} \quad t = 1, \dots, T \quad (3)$$

for each series m in city i , and then calculating the residuals, $\hat{u}_{m,i,t}$. Then, let $\psi(t) \sim \text{uniform}(1, T)$ index the random resampling and construct

$$\tilde{u}_{m,i,t} = \hat{u}_{m,i,\psi(t)} \quad \forall m, i, t.$$

Note that the same $\psi(t)$ is used for each city i , thereby maintaining the cross-correlation that exists in the data.

Given each set of resampled u 's, construct

$$\begin{aligned} \tilde{v}_{m,i,0} &= \sum_{j=0}^J \hat{\eta}_{m,i}(L)\tilde{u}_{m,i,-j} \\ \tilde{v}_{m,i,t} &= \hat{\eta}_{m,i}(L)\tilde{v}_{m,i,t-1} + \tilde{u}_{m,i,t} \\ \tilde{x}_{m,i,0} &= 0 \\ \tilde{x}_{m,i,t} &= \tilde{x}_{m,i,t-1} + \tilde{v}_{m,i,t} \end{aligned}$$

where the $\tilde{u}_{m,i,-j}$ are drawn as a separate bootstrap sample of size J .¹²

The null of no cointegration is true by construction for the resampled data, so one can build up the empirical distribution of the test statistic under the

¹²This is equivalent to setting $\tilde{u}_{m,i,0} = 0$ and simulating $J+T$ observations, but only using the last T observations.

null by replicating the process of resampling and re-estimation enough times. The bootstrapped test results presented later in this paper are based on 20,000 replications of 23 “years” of data for each city, where $J = 100$. These bootstrapped distributions are, by construction, the small-sample distributions and are therefore not directly comparable to the asymptotically normal distributions of Pedroni’s tests, but are comparable to the Maddala-Wu test. To help with comparisons, I also constructed the small sample distributions for Pedroni’s test statistics under the assumption of cross-sectional independence. That is, for each city I simulated 20,000 replications of 23 “years” of data for three unit-root processes under the assumptions of no cointegration and no cross-city correlations, and then tabulated the the values of Pedroni’s test statistics.

4.2 The City-Level Data

My data sources at the city and national levels are the same. OFHEO publishes a quarterly repeat-sales price index for over 300 metropolitan areas. However, only a short time sample exists for many of these cities. For the purposes of this paper, I restricted my attention to a sample of 95 cities for which the price data begin in 1978. See Appendix B for a list of the cities. Total personal income at the city level is available only at an annual frequency, and only through 2000.¹³

Charts 2 through 6 display house prices and per capita income on a log scale for 15 cities, all in current dollars.¹⁴ I chose these cities because they provide a good picture of the behavior of prices and income in most U.S. cities. The house-price run up of the late 1980’s is apparent in most of the cities, with Los Angeles providing the most dramatic example. The average house price in L.A. rose 29 percent in the two years leading up to the third quarter of 1990, reaching a peak of about \$255,000. During the same period, per capita income rose more than 12 percent, a rapid increase, but well short of house-price gains. During the next four years, nominal prices fell more than 20 percent and per capita income was about flat. But the charts also show that not all cities saw such dramatic swings. Indeed, Midwestern cities saw little, if any, price appreciation during the late 1980s, and essentially no depreciation after.

Turning to the more recent period, house prices have risen in all 15 cities

¹³Shiller and Perron (1985) showed that the power of unit root tests depends more on the number of years covered by the dataset than on the number of observations. Thus, the cost in terms of power of using annual data for the city-level analysis is likely small.

¹⁴I calculated the current-dollar values by multiplying the median house price in 2001:Q4 published by the National Association of realtors by the OFHEO index, re-indexed to that quarter, for each city.

shown, and indeed in just about every U.S. city, during the past few years. For example, during the two years leading up to the first quarter of this year, house prices rose about 25 percent in Boston, New York, and Washington, D.C.; prices rose only a bit less in San Francisco and Los Angeles and even slower-growth cities like Chicago, Cleveland, and Saint Louis saw double-digit price growth. Although I only have per capita income figures for cities through 2000, it seems likely that income growth did not keep pace in 2001.¹⁵

Taking a longer view in these cities, one can see that although both prices and per capita income have trended up over time, their levels can diverge for many years. It is not at all clear from these pictures that house prices and per capita income are cointegrated.

4.3 Results and Interpretation

The city-level cointegration tests are based on first-stage cointegrating relationships of house prices on per capita income and population. I estimated the first stage regressions, which all include a time trend, separately for each city. It is important to note that the bootstrap approach can account for any aggregate effect due to factors such as monetary or fiscal policy, national economic conditions, changes to the housing-finance industry, or cost shocks. Thus, I am really testing for the cointegration of prices, income, and population conditional on these other, potentially non-stationary, factors.

Table 2 displays the main results of the paper. The first column contains the value of each test statistic, the second column contains their asymptotic p-values, the third column contains their small-sample p-values, and the fourth column contains their bootstrapped p-values. The conclusions one might draw from the asymptotic p-values are in severe conflict. The $P-v$ and $P-AEG$ tests alone (lines 1 and 3) seem to provide strong evidence for cointegration, while the $P-Z_\alpha$ test (line 2) seems to provide none.

Using the small-sample p-values, the case for cointegration looks much weaker. One can no longer reject the null using the $P-AEG$ test, and one cannot reject using the MW statistic. However, one can still reject using the $P-v$ statistic. Notice that the $P-AEG$ test clearly suffers from extreme size distortion in small samples. Pedroni (2001) found a less extreme, but still significant distortion in the case where $N = 20$ and $T = 40$. My results indicate that the distortions become much worse as T gets even smaller.

Column 4 shows the bootstrapped p-values. Note that allowing for arbi-

¹⁵The available state-level figures for per capita income support this view.

Table 2
City-Level Test for Cointegration of
House Prices, Per Capita Income, and Population

null of no cointegration; $N = 95; T = 23$

	Test	Asymptotic	Small-Sample	Bootstrapped
	Statistic	p-value	p-value	p-value
1. $P-v$	2.7	.00	.01	.15
2. $P-Z_\alpha$	4.1	1.0	1.0	.31
3. $P-AEG$	-3.0	.00	.95	.65
4. MW	158.4	—	.95	.14

Notes: $P-v$ is asymptotically normal and the right tail is used for rejection. $P-Z_\alpha$ and $P-AEG$ are asymptotically normal and the left tail is used for rejection. The Maddala-Wu test is distributed chi-squared and the right tail is used for rejection (Maddala and Wu, 1999). The bootstrapped Maddala Wu test has a non-standard distribution that must be estimated by simulation; the right tail is used for rejection. The Pedroni tests include aggregate time effects.

trary cross correlations has a significant effect on the simulated distribution of the test statistics. The p-values for the $P-Z_\alpha$, $P-AEG$, and MW statistics are all smaller when one uses the bootstrapped distributions instead of the small-sample distributions, but one would still not reject the null of no cointegration using conventional significance levels. In contrast, the p-value for the $P-v$ statistic is larger using the bootstrapped distribution, and the difference is enough to change the results of the test from rejection to no rejection of the null.

Thus, using conventional significance levels, the tests are all in accordance: House prices, income, and population do not appear to be cointegrated at the national or local levels. These results contradict those of Malpezzi (1999), who found that one *can* reject the null of no cointegration in a similar panel. Malpezzi was able to reject the null of a unit root in the residuals of city-level cointegrating regressions of prices and income using a Levin-Lin test. His interpretation was that prices and income are therefore cointegrated. However, the Levin-Lin test is inappropriate in this context because the critical values tabulated by Levin and Lin (1992) are not adjusted for the fact that Malpezzi's residuals are estimated. In other words, the Levin-Lin tests is a unit root test, not a residuals-based

cointegration test. While the two types of tests are clearly related, the Levin-Lin test is only applicable to the case of cointegration when the researcher imposes the cointegrating relationship. In addition, Malpezzi did not correct for the fact that the Levin-Lin test requires independence across local housing markets.

If house prices and fundamentals are not cointegrated, how should we interpret the error-correction type models of Abraham and Hendershott (1996), Malpezzi (1999), and Capozza et al. (2002)? All three find that house prices increase more slowly when actual house prices are above a measure of the long-run equilibrium price level, and all three base their equilibrium measure on a first-stage levels regression. A strict interpretation of the results in this paper is that an error-correction model is a mis-specification, and that results from such a model are spurious. A looser interpretation is that even if prices, income, and population are cointegrated, we cannot verify this relationship. Our inability to verify the relationship implies an inability to accurately estimate it. In other words, even if we think a long-run relationship exists, we cannot say with much certainty when prices are in line with fundamentals and when they are not. Forecasts based on our best guess as to the degree of “disequilibrium,” such as those from an error-correction model, are therefore highly suspect.

5 Conclusion

Many housing market observers have become concerned that house prices have grown too quickly of late, and that prices are now too high relative to per capita incomes. Prices will likely stagnate or fall, the argument goes, until they are better aligned with income. This idea is often formalized in the housing literature by asserting a long-run equilibrium relationship between house prices and fundamentals such as income, population, and user cost. The validity of this assumption has important implications for how we model house price dynamics. If the assumption is accurate—so that house prices and fundamentals are cointegrated—then the error-correction specifications of Abraham and Hendershott (1996), Malpezzi (1999), and Capozza et al. (2002) are appropriate.

In this paper I have used standard tests to show that there is little evidence for cointegration of house prices and various fundamentals at the national level. I have also shown that bootstrapped versions of more powerful panel-data tests, applied to a panel of 95 U.S. metropolitan areas over 23 years, also do not find evidence for cointegration. This does not mean that fundamentals do not affect house prices, but it does mean that the *level* of house prices does not appear to

be tied to the *level* of fundamentals. Thus, the levels regressions found in the literature are likely spurious, and the associated error-correction models may be inappropriate.

A Statistical Appendix

This appendix provides definitions of the panel cointegration tests of Pedroni (1999). Suppose that the hypothesized cointegrating relationship for each city is given by

$$x_{0,i,t} = \alpha_i + \delta_i t + \sum_{m=1}^M \beta_{m,i} x_{m,i,t} + e_{it}, \quad (4)$$

Pedroni suggests the following test statistics:

$$\begin{aligned} \text{P-v} &= T^2 N^{\frac{3}{2}} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^2 \right)^{-1} \\ \text{P-} Z_\alpha &= T\sqrt{N} \left(\sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^2 \right)^{-1} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} (\hat{e}_{i,t-1} \Delta \hat{e}_{it} - \hat{\lambda}_i) \\ \text{P-AEG} &= \left(\tilde{s}_{N,T}^{*2} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t-1}^{*2} \right)^{-\frac{1}{2}} \sum_{i=1}^N \sum_{t=1}^T \hat{L}_{11i}^{-2} \hat{e}_{i,t}^* \Delta \hat{e}_{it}^* \end{aligned}$$

where

$$\begin{aligned} \hat{\lambda}_i &= \frac{1}{T} \sum_{j=1}^{k_i} \left(1 - \frac{j}{k_i + 1} \right) \sum_{t=j+1}^T \hat{\mu}_{i,t} \hat{\mu}_{i,t-j} \\ \hat{s}_i^2 &= \frac{1}{T} \sum_{t=1}^T \hat{\mu}_{i,t}^2 \\ \hat{s}_i^{*2} &= \frac{1}{T} \sum_{t=1}^T \hat{\mu}_{i,t}^{*2} \\ \tilde{s}_{N,T}^{*2} &= \frac{1}{N} \sum_{i=1}^N \hat{s}_i^{*2} \\ \hat{L}_{11i}^{-2} &= \frac{1}{T} \sum_{t=1}^T \hat{\eta}_{i,t}^2 + \frac{2}{T} \sum_{j=1}^{k_i} \left(1 - \frac{j}{k_i + 1} \right) \sum_{t=j+1}^T \hat{\eta}_{i,t} \hat{\eta}_{i,t-j} \end{aligned}$$

and where the residuals $\hat{\mu}_{i,t}$, $\hat{\mu}_{i,t}^*$, and $\hat{\eta}_{i,t}$ are from the following regressions:

$$\begin{aligned} \hat{e}_{i,t} &= \hat{\gamma}_i \hat{e}_{i,t-1} + \hat{\mu}_{i,t} \\ \hat{e}_{i,t} &= \hat{\gamma}_i \hat{e}_{i,t-1} + \sum_{k=1}^{K_i} \hat{\gamma}_{i,k} \Delta \hat{e}_{i,t-k} + \hat{u}_{i,t}^* \\ \Delta y_{i,t} &= \sum_{m=1}^M \hat{b}_{mi} \Delta x_{mi,t} + \hat{\eta}_{i,t}. \end{aligned}$$

B Data Appendix

Metropolitan Statistical Areas

1978 to 2000

Akron, OH	Minneapolis-St. Paul, MN-WI
Albuquerque, NM	Modesto, CA
Ann Arbor, MI	Monmouth-Ocean, NJ
Atlanta, GA	Nassau-Suffolk, NY
Austin-San Marcos, TX	New Orleans, LA
Bakersfield, CA	New York, NY
Baltimore, MD	Newark, NJ
Baton Rouge, LA	Norfolk-Virginia Beach-Newport News, VA-NC
Bergen-Passaic, NJ	Oakland, CA
Birmingham, AL	Oklahoma City, OK
Boston, MA-NH	Omaha, NE-IA
Boulder-Longmont, CO	Orange County, CA
Buffalo-Niagara Falls, NY	Orlando, FL
Canton-Massillon, OH	Philadelphia, PA-NJ
Charlotte-Gastonia-Rock Hill, NC-SC	Phoenix-Mesa, AZ
Chicago, IL	Pittsburgh, PA
Cincinnati, OH-KY-IN	Portland-Vancouver, OR-WA
Cleveland-Lorain-Elyria, OH	Raleigh-Durham-Chapel Hill, NC
Columbus, OH	Richmond-Petersburg, VA
Dallas, TX	Riverside-San Bernardino, CA
Dayton-Springfield, OH	Rockford, IL
Denver, CO	Sacramento, CA
Des Moines, IA	St. Louis, MO-IL
Detroit, MI	Salinas, CA
Eugene-Springfield, OR	Salt Lake City-Ogden, UT
Flint, MI	San Diego, CA
Fort Collins-Loveland, CO	San Francisco, CA
Fort Lauderdale, FL	San Jose, CA
Fort Wayne, IN	San Luis Obispo-Atascadero-Paso Robles, CA
Fort Worth-Arlington, TX	Santa Barbara-Santa Maria-Lompoc, CA
Fresno, CA	Santa Cruz-Watsonville, CA
Grand Rapids-Muskegon-Holland, MI	Santa Rosa, CA
Greensboro-Winston-Salem-High Point, NC	Sarasota-Bradenton, FL
Hamilton-Middletown, OH	Seattle-Bellevue-Everett, WA
Honolulu, HI	Spokane, WA
Houston, TX	Stockton-Lodi, CA
Indianapolis, IN	Syracuse, NY
Kalamazoo-Battle Creek, MI	Tacoma, WA
Kansas City, MO-KS	Tampa-St. Petersburg-Clearwater, FL
Lansing-East Lansing, MI	Tucson, AZ
Las Vegas, NV-AZ	Tulsa, OK
Los Angeles-Long Beach, CA	Vallejo-Fairfield-Napa, CA
Louisville, KY-IN	Ventura, CA
Madison, WI	Visalia-Tulare-Porterville, CA
Memphis, TN-AR-MS	Washington, DC-MD-VA-WV
Miami, FL	West Palm Beach-Boca Raton, FL
Middlesex-Somerset-Hunterdon, NJ	Wichita, KS
Milwaukee-Waukesha, WI	

References

- [1] Banerjee, A. (1999). "Panel Data Unit Roots and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics*, 61, 607-629
- [2] Calhoun, Charles, A. (1996). "OFHEO House Price Indexes: HPI Technical Description." Office of Federal Housing Enterprise Oversight. <http://www.ofheo.gov/house/download.html>.
- [3] Capozza, Dennis R., Hendershott, Partic H., Mack, Charlotte, and Mayer, Christopher, J. (2002). "Determinants of Real House Price Dynamics." NBER Working Paper 9262.
- [4] Davidson, R. and MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*, Oxford University Press.
- [5] Engle, R. F. and Granger, C. W. J. (1987). "Co-integration and Error Correction: Representation, Estimation, and Testing", *Econometrica*, 55, 251-276.
- [6] Fisher, R.A. (1932). *Statistical Methods for Research Workers*, Oliver & Boyd, Edinburgh, 12th Edition.
- [7] Glaeser E. L. and Gyourko, J. (2002). "The Impact of Zoning on Housing Affordability", Mimeo.
- [8] Hamilton, J.D. (1994). *Time Series Analysis*, Princeton University Press, Princeton, NJ.
- [9] Hilsenrath, J. E. (2002). "A 91-Year-Old Who Foresaw Selloff is 'Dubious' of Stock-Market Rally", *Wall Street Journal* July 25, 2002.
- [10] Im, K., Pesaran, M.H., and Shin, Y. (1997). "Testing for Unit Roots in Heterogeneous Panels", Mimeo, Department of Applied Economics, University of Cambridge.
- [11] Levin, A. and Lin, C. F. (1992). "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties", Department of Economics, University of California at San Diego, Discussion Paper No. 92-93 (revised 1993).
- [12] Levin, A., Lin, C. F., and Chu, C. J. (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties", *Journal of Econometrics*, 108, 1-24.

- [13] Malpezzi, Stephen. (1999). "A Simple Error Correction Model of House Prices", *Journal of Housing Economics*, 8, 27-62
- [14] Maddala, G. S., and Wu, S. (1999). "A Comparative Study of Unit Root Tests with Panel Data and New Simple Test", *Oxford Bulletin of Economics and Statistics*, 61, 631-652
- [15] McCoskey, S. and Kao, C. (1998). "A Residual-Based Test of the Null of Cointegration in Panel Data" *Econometric Reviews*, 17, 57-84
- [16] Meen, Geoffrey. (2002). "The Time-Series Behavior of House Prices: A Transatlantic Divide?" *Journal of Housing Economics*, 11, 1-23
- [17] Pedroni, P. (1997). "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis", Mimeo, Department of Economics, Indiana University.
- [18] Pedroni, P. (1999). "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors", *Oxford Bulletin of Economics and Statistics*, 61, 653-670
- [19] Pedroni, P. (2001). "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis", Mimeo, Department of Economics, Indiana University.
- [20] Poterba, J. (1991). "House Price Dynamics: The Role of Tax Policy and Demography", *Brookings Papers on Economic Activity*, Issue 2, 143-183.
- [21] Quah, D. (1990). "International Patterns of Growth I: Persistence in Cross-Country Disparities", Working Paper, Massachusetts Institute of Technology.
- [22] Quah, D. (1994). "Exploiting cross section variation in unit root inferences in dynamic data", *Economic Letters*, 44, 9-19.
- [23] Reifschneider, D., Tetlow, R., Williams, J. (1999) "Aggregate disturbances, monetary policy, and the macroeconomy: The FRB/US perspective" *Federal Reserve Bulletin*, 85, Issue 1, 1-19.
- [24] Shiller, R. J. and Perron, P. (1985). "Testing the Random Walk Hypothesis: Power versus Frequency of Observation," *Economic Letters* 18, 381-386.
- [25] Stock, J. H. and Watson, M. W. (1988). "Testing for Common Trends", *Journal of the American Statistical Association*, 83, 1097-1107.