APTARA	REEC	reec12098	**Dispatch:** May 20, 2015	**CE:** AFL
	Journal	**MSP No.**	**No. of pages:** 25	**PE:** Charlotte Ching

2015 V00 0: pp. 1–25
DOI: 10.1111/1540-6229.12098

REAL ESTATE
ECONOMICS

Do House Price Levels Anticipate Subsequent Price Changes within Metropolitan Areas?

Q1 Nai Jia Lee,* Tracey N. Seslen** and William C. Wheaton***

This research examines the relationship between hedonically controlled housing price levels and subsequent changes in those prices across locations within MSAs. Are areas with a high price relative to an "imputed rent" paying for higher appreciation? In an efficient market (e.g. Gordon Growth Model), as fundamentals (impute rent) differ across locations and change over time, anticipation of these should generate a positive correlation between (residual) price levels and subsequent price changes. We undertake these tests in four different MSAs using a panel of repeat-sale house price indices that have been scaled to price levels with the hedonic attributes of the house and ZIP code. In three markets we find that identical houses in higher priced ZIP codes subsequently appreciate faster. In one market we find that there is little statistical difference.

This research examines the price efficiency of the housing market across locations within a metropolitan market. Our innovation is to combine temporal efficiency with the traditional attribute efficiency of "price hedonics." A Gordon Growth Model (Gordon 1959, henceforth GGM) would suggest that units in areas with growing fundamentals should have rising "rent," and if prices are the discounted value of future rent, prices should be rising commensurate with rent. Most importantly, such areas should have higher price/rental ratios—all relative to a location with stagnant or deteriorating fundamentals. In GGM there is no uncertainty about future rental growth. When there is uncertainty but forward-looking expectations, Campbell and Shiller (1988, 1998) propose testing whether price rent ratios are positively correlated with subsequent rent (and hence price) growth.

Our unique approach is not to try and measure actual contract "rent" (for an owner occupied unit), but rather to take the unit's predicted price in a

*National University of Singapore, Singapore 117566 or rstlnj@nus.edu.sg.
Q2 **University of Washington, Seattle, WA 98195 or seslen@uw.edu.
***MIT, Cambridge, MA 02139 or wheaton@mit.edu.

cross-section hedonic equation as a measure of imputed rent. Our test is then to examine the residuals from this hedonic equation, which includes most measured housing attributes as well as several ZIP-code variables. If the market is efficient, positive residuals (paying more for observationally equivalent units) should be compensated for by actual subsequent price appreciation (presumably due to improving fundamentals). Hence we ask if the difference between the actual price and "imputed rent"—a ZIP-code-level residual—is correlated with subsequent appreciation in those ZIPs. If ZIP-code residuals are just noise in the measurement of "rent" then there should be no correlation with subsequent appreciation. Positive correlation is strong evidence that subsequent changes in ZIP fundaments are anticipated in current prices—as would be called for in an efficient forward-looking market. If we were to find a negative correlation we would have strong evidence of market inefficiency. Why would households pay more for observationally equivalent houses in locations where fundamentals (and prices) subsequently decline?

While there has been much previous discussion about housing market efficiency, in our literature review virtually all of this research involves examining price efficiency either in theory, or empirically *between* metropolitan markets. There is only one attempt to use the Campbell–Shiller approach to test whether prices are efficient across locations *within* a metropolitan market (Meese and Wallace 1994). In this article, we test for such efficiency by creating a panel data base of house prices at the ZIP-code level. This is done by first obtaining ZIP-level *repeat-sale price indices* that span roughly 25 years. We then undertaking a cross-section (hedonic) adjustment for the attributes of the house and ZIP code using thousands of actual transactions for a year near the middle of the indices. Scaling the time-series indices by the predicted hedonic level, we have what a common house costs across both time and space (ZIP). We generate this unique data for four (quite different) MSAs (metropolitan statistical areas) and then undertake a separate panel analysis of the data in each market. In three of these markets we find that identical houses in ZIP codes that are priced higher appreciate faster (in the near future) than in lower priced ZIPs. In one market we find that there is little statistical difference. In the first three we infer that prices are forward looking, while in the latter we can say little.

Q3

In the following section, we begin with a review of theoretical models and empirical research on house price efficiency. The third section lays out our simple version of a hedonic price model that must be efficient inter-temporally as well as across heterogeneous properties. In the fourth section we discuss the creation of our panel data base, and in the fifth section we present several panel models that test whether locations with hedonically unexplained higher prices subsequently appreciate faster. The final section concludes with some caveats and suggestions about additional work.

 Q4

Literature Review

The literature on the efficiency of housing markets probably begins with a paper by Case and Shiller (1988), which reports that in surveys, the recent purchasers of houses have completely unrealistic expectations about the future appreciation of their newly acquired asset. This notion is consistent with many older studies on the behavioral origins of "bubbles" (originally published 1978, Kindleberger 2000). Q5

The idea that housing might not be efficiently priced then receives some support in another paper by Case and Shiller (1989), which demonstrates that there is considerable serial correlation and predictability in national house prices. The authors argue that such predictability should allow arbitrage purchases which would eliminate the predictability—if the market was truly efficient. Later work at the MSA level by Capozza, Hendershott and Mack (2004) reinforces the existence of serial correlation and so too the argument of Case and Shiller.

A series of more theoretical papers effectively deflates the notion that serial correlation *per se* implies irrational "bubbles." Poterba (1984) presents a perfectly informed forward model of house prices in which positive shocks are followed by declining prices and *vice versa*. This "mean reversion" is intrinsic to the market because of the existence of lags in the supply of new housing assets. Even if these lags are well known and fully incorporated into expectations, they still create serial correlation and (somewhat) predictable house price movements in reaction to shocks. Grossman and Laroque (1990) incorporate transactions costs into the adjustment of housing demand, and their model also creates predictable patterns to house prices. Then there is a series of papers by Stein (1995) in which liquidity constraints in financial markets create predictable and positive correlations between sales volumes and prices. In a series of follow up models, Ortalo-Magne and Rady (1999, 2001) show that such predictable patterns also arise as renters become owners in a market that again has financial constraints. In short, a wide range of theoretical models now suggest that those looking for evidence of inefficiency must look beyond the existence of serial correlation and price predictability for their evidence.

An alternative approach to testing price efficiency emerges following the work of Campbell and Shiller (1988, 1998). If markets are efficient, price/earnings ratios should be positively correlated with subsequent earnings (and price) growth. DiPasquale and Wheaton (1996) and Capozza and Helsley (1990) extend this argument to theoretical models of urban land markets. If there is anticipated growth in market-wide population, income or other economic "fundamentals," these will generate distinct spatial patterns to Ricardian rent growth, which when anticipated, create systematic variation in price/rent ratios across locations. Similarly if these fundamentals experience stochastic shocks as well as growth, risk premiums will likewise vary spatially and create further variation in price/rent ratios between locations.

The first attempt to use price/rent ratios to test for housing market efficiency occurs in Meese and Wallace (1994). These authors examine the ratio of average house prices-to-average rents across counties in California and do find some positive correlation to subsequent price appreciation. Himmelberg, Mayer and Sinai (2005) examine the rise in average price/rent ratios prior to the financial crises of 2008 but suggest that the implied appreciation in prices is only about the historic average. A recent paper by Campbell *et al.* (2009, hereafter CDGM) undertakes a comprehensive time series analysis of price/rent ratios and concludes that there is strong evidence of mean reversion. When ratios are high they tend to fall (and *vice versa*), with most of the adjustment being done by prices rather than rents.

The paper by CDGM has substantial implications for market efficiency, not well drawn out by the authors. If rents oscillate, or are subject to (somewhat) predictable shocks then the volatility of prices should actually be less than rents—when the prices are the present discounted value of expected future rents (Abel and Blanchard 1986). Contrary to the CDGM results, with price efficiency it should be rents rather than prices that do most of the adjusting. With this pattern, efficient price rent ratios are highest when oscillations have driven rents to a (temporary) bottom, and lowest when rents are at a (temporary) peak. Hence regardless of whether the rents in question are smoothly trending up or down, or oscillating in some manner, efficient price/rent ratios are positively correlated with near-term subsequent rental movements. When prices are the PDV of rents the positive correlation will extend to prices as well.

Most recently, however, the use of average housing price/ rent ratios has come under critical scrutiny in a paper by Smith and Smith (2006). These authors argue that the average rental unit in most metropolitan areas is dramatically different from the average owned unit. Furthermore these average differences vary between market and change between census years. As a test, they sample sales prices and rents across 10 metropolitan areas but use specific neighborhoods within which virtually identical houses on the same street are both rented and sold. Their derived price/rent ratios—now for truly the same asset—are dramatically different than MSA average price rent ratios. These differences also vary significantly as between the 10 metropolitan areas sampled. While not explicitly testing for price efficiency, these results would cast considerable doubt over any empirical tests that rely on using average prices and average rents.

In this research we stick with the idea that prices and rents are the correct comparison for a test of efficiency, but following the critique of the Smiths we do not try to use *measured rents*. Rather we assume that the imputed "rent" for any unit is a linear combination of its various attributes (a hedonic equation). In this article a high actual price relative to the weighted average of attributes (a predicted price) is assumed equivalent to a high price/rent ratio. To ascertain whether this approach has validity we will regress the subsequent price appreciation in an area against location-specific Hedonic residuals for those areas, in a panel extending almost 25 years.

Hedonic Pricing with Temporal Efficiency

In the standard model of efficient market inter-temporal optimization Poterba (1984) follows Hendershott and Slemrod (1983) and derives the "user cost" of consuming "housing services." At each time period t this has two components, the dollar price of a unit of housing services p_t, and then the opportunity cost of tying up that dollar of capital. If the consumption of housing services can be freely adjusted each period, and there are no constraints on savings or borrowing, then the "user cost" varies over time—each period time obeying (1) below, with 1-period forward expectations.[1]

$$UC_t = p_t[r - E(p_{t+1})/p_t + 1]. \tag{1}$$

Alternatively we can think of (2) as reflecting the inter-temporal shadow cost of spending *one dollar* on housing services:

$$uc_t = r - E(p_{t+1})/p_t + 1. \tag{2}$$

When the housing market is differentiated by a series of fixed location or housing attributes X, there has to be an added dimensionality to (1) and (2). Across the market, permutations of X will generate a Hedonic price function $p_t(X)$ which can move over time differently for each value of X. In this case, the utility that a representative agent derives from living at X for one period must incorporate not just the utility flow from the consumption of X but also must consider how the price of that particular permutation of X is expected to evolve over that period. Rather than try and directly model how the shadow prices of X or the distribution of X will evolve into the future (we give some examples below), we take the simple utility function $U(c, X)$, and incorporate uc_t into an implied inter-temporal budget constraint: $c_t = y - p_t(X)uc_t$. In (3)

[1] Here we simplify and ignore taxes and operating expenses, r is likewise assumed to be both the after-tax cost of both debt financing as well as the opportunity cost of equity funds.

below it must be remembered that uc_t incorporates the expectation of prices at a specific X.

$$U(y - p_t(X)uc_t, X) = U_t. \tag{3}$$

In equilibrium at each time period, with fixed utility across X, Expression (3) can be differentiated totally with respect to X, set to zero, and using a linear approximation together with Equation (2), the hedonic price function can be written as depending on attributes and the expected appreciation of the price of those particular combination of attributes in the future.

$$p_t(X) = (r - E(p_{t+1})/p_t + 1)\sum_i \frac{\partial U/\partial X_i}{\partial U/\partial c_t} X_i. \tag{4}$$

We might also rewrite (4) into an expression for the implied "price/rent" ratio of any particular permutation of X. In this model the imputed rent of X is simply the summation of the current willingness-to-pay utility terms on the right had side X. Hence the equilibrium marginal condition implies that for any X, the imputed price rent ratio should be negatively related to the future price appreciation associated with that X:

$$E(p_{t+1})/p_t = (r + 1) - \frac{p_t(X)}{\sum_i \frac{\partial U/\partial X_{i,t}}{\partial U/\partial c_t} X_{i,t}}. \tag{5}$$

Equations (4) or (5) deserve some further exploration because each has a price level being determined by expected future prices. An immediate question of course is what determines differences in expected price appreciation across locations or permutations of X? Here, many examples come to mind, remembering that while the X values themselves are fixed, the anticipated valuation of them in the future can certainly be expected to change. Two such are described below.

Example #1: With expected rising energy costs in a monocentric urban spatial model, peripheral locations face increases in commuting costs—far more than do more central locations. When these rises in energy costs are fully anticipated, it should be the case that central locations not only have higher rents (and prices) because of current commuting cost savings, but also higher price/rent ratios because of the expected increase in those cost savings over the future.

Example #2: Another possibility is that consumer preferences or their distribution (the marginal utility terms in (4) and (5)) are expected to change. So if a market has anticipated strong growth in income, and it is known that preferences for coastal locations within that market are highly income elastic, then coastal sites will not only have higher rents (and prices) today but also higher price-rent ratios in anticipation of the increasing future rents (and prices) for ocean frontage.

As a first approach to developing an estimating equation we could simply try to directly estimate Equation (4), a standard hedonic model, and see if adding actual *subsequent* appreciation in a unit's ZIP improves its explanatory power, of course with an expected positive sign. Such estimation however could raise important specification issues. If the dependent variable is price at time t, and if the independent variable is subsequent price change from t to $t+1$, any serial correlation in prices (as many authors have found) would generate an errors in variables problem. Appreciation between t and $t+1$ is correlated with that from $t-1$ and t which in turn must impact the measured price at t, by construction.

As an alternative, we could work with Equation (5). Here the dependent variable is price appreciation between from t to $t+1$ and the RHS variable is the ratio of price/rent at t. To actually implement this we might estimate "rent" as a linear combination of attributes (X) and then calculate its ratio to actual price—as in the right-hand side of Equation (5). Then following Campbell and Shiller (1988) this imputed price/rent ratio becomes the independent variable in a regression predicting actual subsequent growth in price across locations. Methodologically this is similar to the regression undertaken by Meese and Wallace (1994), but in our case we use an imputed "rent" for the actual unit whose price is being measured, rather than the average contract rent for apartments in the same location as single-family house sale prices. A final difference between our equation and (5) is that rather than create an imputed price rent ratio and track its movements, we use the difference between price and imputed rent (a residual) and observe it over time.

To actually implement his approach, requires that we examine how the value for identical units varies both across locations as well as over time. We estimate such a panel housing values effectively in two stages. First, with a large number of individual property transactions we estimate residuals from a standard hedonic equation of prices against X, and we then aggregate the value of the residuals at the ZIP level. These ZIP residuals *levels* capture some combination of current rent for omitted attributes (cross-section noise) as well as the expected future change-in-valuation of those attributes that we do measure. Secondly, we grow these residuals across time using a set of repeat sale price indexes provided to us at the ZIP level. These repeat-sale indexes are supposed to represent value *movement* for truly identical units. These data allow us to ask whether higher prices levels—controlling for the flow of housing services—have predictive power in explaining future price appreciation. If the residuals are exclusively error from omitted attributes (cross-sectional noise) they should not be positively correlated with actual subsequent price appreciation.

The equilibrium pattern of prices that emerges over time with forward-looking efficiency is well known. Suppose that fundamentals (imputed rent) are somehow fluctuating or moving in a cyclic pattern that is anticipated (albeit partially). In this case, if prices are the present discounted value of future rent, then prices will also fluctuate with similar timing to the underlying movements in fundamentals. As discussed in the second section, the price movement should be less volatile than that in rent so that price rent ratios will be high when rent is at a temporary bottom (and prices are starting to recover). The ratio will be low when rent is at a temporary high (and prices have already started to fall). With this pattern, at any time, the price rent ratio (or in our case linear residuals) should be positively associated with subsequent near term rent (and price) movement. An illustration of this is shown below (Figure 1).

As an alternative (and null hypothesis), consider prices which are not at all forward-looking and are simply (inefficiently) moving around a constant and nonchanging set of fundamentals (the CDGM result). Any form of mean reversion means that when prices are high (relative to fundamentals) then they have to subsequently decline and *vice versa*. This, however creates a negative correlation that is the exact opposite result from the efficient pattern (in Figure 1).

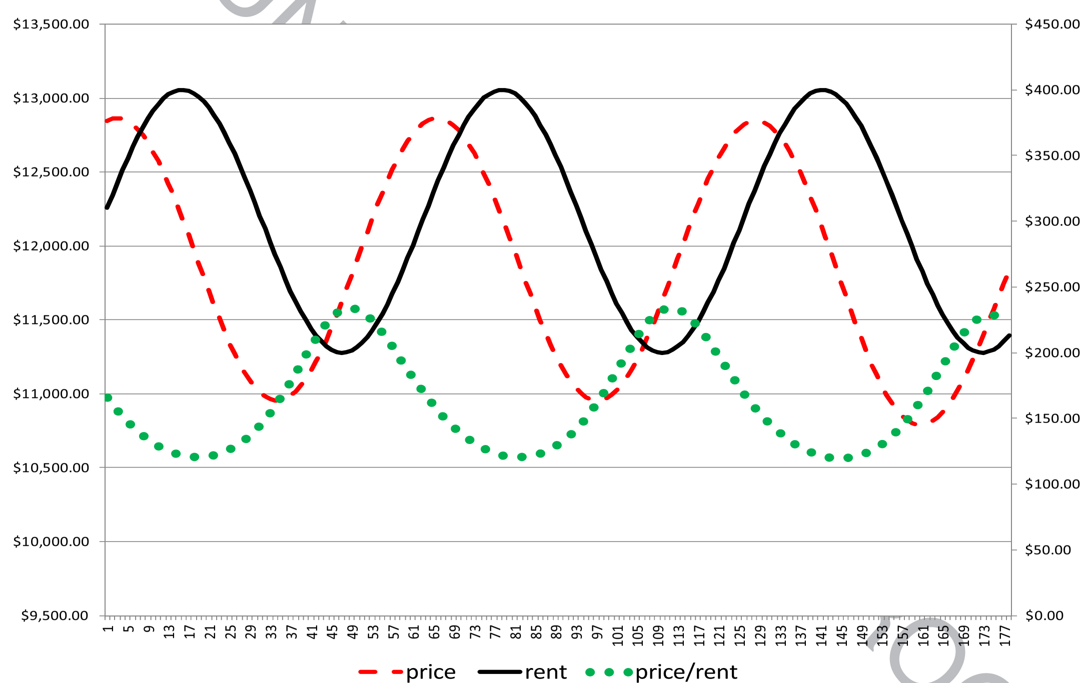
Over the long run, prices across locations can cris-cross, remain relatively stable, or diverge depending on whether the fundamentals (or their valuations) across locations are so behaving. In any of these cases, the key distinction is the correlation between price-to-imputed rent and subsequent price appreciation. With efficiency it is (at least somewhat) positive and with inefficiency it is (somewhat) negative.

ZIP-Level Panel Data

To test these ideas we put together a panel database of housing price *levels* by ZIP code for four MSAs. We do this in two stages. First we obtain repeat-sales housing price *indices* provided to us by Case Shiller Weiss/ FISERV at the ZIP-code level. The data cover four MSAs: Boston, Chicago, Phoenix and San Diego that we were allowed to select.[2] In choosing these MSAs, we attempt to create a sample representing a diverse set of demographic, geographic and

[2]FISERVE price indexes at the ZIP level are proprietary and generally only sold. The company graciously gave us four markets that we selected for geographic diversity out of a listed set of MSAs for which ZIP data was available.

Figure 1 ■ Rents, NPV prices, P/R ratio.[a]



[a]In Figure 1, Rents oscillate between 200 and 400, with 60 (quarter) periodicity. Prices are the NPV for 200 periods with 2.5% quarterly discounting.

housing market related conditions. The Boston metropolitan area included 249 ZIP codes from 1982 through 2004; Chicago comprises 152 ZIP codes from 1987 through 2004; Phoenix includes 164 ZIP codes spanning 1984 to 2004; and San Diego covers 86 ZIP codes starting from 1975 and going through 2004. After dropping ZIP series that did not contain observations for the full length of the MSA time series, the final sample consists of 109 ZIP observations for Boston, 51 for Chicago, 80 for Phoenix and 42 for San Diego. As repeat sales price indexes thoroughly control for individual unit or parcel attributes, if different ZIP codes move differently over time, we can rule out systematic changes in the ZIP housing stocks. In theory, different trends in repeat sales indices for ZIPs offer a clean measure of how the location fundamentals of different areas are changing over time.

The Repeat-Sale price indices however, have no measurement of levels and the crucial question for this research involves the relationship between trends and levels. To convert the repeat sale indexes into a full set of prices we need a cross-section in one year of what identical houses sold for across ZIPs. Ideally we would like to pick a year in the middle of our index time series (*e.g.*, 1990), but the vendors we sought transactions data bases from did not have data back that far. The year 1998 was the earliest for which we were able to acquire large data sets for each market of individual house transactions. The Warren Group provided approximately 15,000 sales for Boston, while Data Quick gave us 34,000 sales for Phoenix, 13,000 sales for Chicago and 14,000 for San Diego. With this data we estimated Hedonic equations using detailed unit attributes, as well as ZIP fixed effects and two ZIP characteristics: median household income and gross residential density. The hedonic results are presented in the Appendix A for each of our four MSA markets.

With the estimated Hedonic equations we specified an identical (benchmarked) house for each market generally consisting of a three-bedroom unit built in the 1960s with 1.5 bathrooms. We also used the mean lot size and floor area for each MSA as our benchmark and, similarly, used the MSA median ZIP income and density. From the hedonic regression, we derived the price of the benchmarked house in 1998 across all ZIP areas using the ZIP fixed effects. We then inflated/deflated these predicted price levels across time using the CSW indices (which in turn have been deflated using the national CPI). The result is a measure of what the same house sells for both (in constant 2005 dollars) across ZIPs and across time. These house prices, for each of the four MSAs, are plotted as Charts 1–4 in Appendix B.[3]

[3]If there was sufficient sales data, for sufficiently long, to estimate a hedonic equation for each period we could construct this panel from the ground up. If those hedonic

In Chart 1, Boston prices generally have been trending up, peaking once around 1986–1988 and then again in 2005. Cyclically, all ZIP codes seem to move together and there clearly is no independent cyclicality at the ZIP level. In longer term trends, however, there are differences. In many of the ZIP areas that area initially priced higher, prices appear on the surface to continue to grow faster and diverge over time from the prices of identical houses in the initially lower price ZIPs. There also are few if any cases of prices ever crossing so that the ranking in prices across ZIP does not change much from 1982 to 2005. The most expensive areas remain the same for the 24 years in the data.

The house price trends of San Diego and Phoenix are shown in Charts 2 and 3. Real prices generally trend up in San Diego, like Boston. In addition the prices in the most expensive ZIP again have a steeper gradient than prices in the cheaper neighborhoods. On the surface San Diego (and Boston) looks something like Figure 1 in that there is some degree of price divergence. In Phoenix (Chart 3), there are two peaks with prices falling to a trough in 1993 and then picking up again thereafter. The most expensive neighborhood in 1985, occupied the top position until 1998. Prices in several other expensive ZIPs surpass this top neighborhood around 1998. Again lower priced areas say at the bottom, and once again there is some appearance of divergence between top and bottom tier ZIP.

The price trends in Chicago (shown in Chart 4) appear a bit different from the price trends in the other three MSAs. While the very most expensive neighborhoods stay the same over this period, there are many changes in ranking from the bottom to mid-level priced neighborhoods and from mid-level to upper tier. The number of ZIP price crossings is much more noticeable than in Boston, San Diego or Phoenix.

Panel Empirical Tests of Pricing Efficiency

As our panel of prices across ZIP and time for each MSA controls carefully for all unit characteristics and two important location features (median ZIP income and density), our test of the relationship between change-in-price and price level can be reduced to a simple bivariate analysis with the panel data. Our model is a direct comparison between a ZIP's relative price level at time period $t - k$ and then the subsequent relative price appreciation from that period up to

equations yielded stable coefficients over time then for a given set of attributes, the constructed ZIP prices would be parallel over time and exhibit similar appreciation profiles. With changing coefficients (time variation in shadow values), the so-created ZIP indices could diverge, converge or cross. The patterns observed in the repeat-sales indexes implicitly reveal whether shadow prices are constant or changing over time.

period t. The notion of relative references the market average behavior of both prices and their change; hence we demean the variables across ZIP codes. This equation is (6).

$$\Delta P_{i,t,k} - \Delta \bar{P}_{t,k} = \alpha + \beta(P_{i,t-k} - \bar{P}_{t-k}) + e_{i,t}. \tag{6}$$

In (6), we denote $\Delta P_{i,t,k}$ as the cumulative change in price of ZIP i over a span of k periods that ends in period t and begins in period t – k. Similarly $\Delta \bar{P}_{t,k}$ represents the mean price change across all ZIP in the MSA—also for an interval of k periods ending in t. The variable $\bar{P}_{t-k}$ represents average ZIP price level at the start of the k period interval ending at t. We conduct the tests using yearly changes, and then nonoverlapping two-year, three-year and five-year windows. Using nonoverlapping intervals prevents auto correlation that occurs by construction with smoothing over time. As we lengthen the window, the number of observations drops. With yearly data we have approximately 20 (years) by 80 (ZIP) observations for each MSA. If we extended k to be the full length of the sample, the model would cease to be a panel and become a simple cross-section of 20-year cumulative growth rates regressed against initial levels. Such a simplified cross section analysis would preclude allowing ZIP prices ever to multiple-cross over time. Using the panel with yearly intervals provides a much more robust test of whether high prices *in any given year* predict higher price appreciation in the subsequent year.

As an alternative we might think of simply undertaking a traditional full panel model with both time and cross-section fixed effects—such as in (7). This equation, however, would be inappropriate, because with ZIP fixed effects we would be examining the prices of each ZIP code relative to their average over the sample period. Here we are strictly interested in relative performance across ZIPs—not across time for a given ZIP. An additional problem with the more general specification (7) is that our specific null hypothesis is whether the time fixed effects are the average (across ZIP areas) of appreciation-minus-price level in each period. For our particular question specification (6) is clearly more appropriate than (7).

$$\Delta P_{i,t,k} = \alpha + \beta P_{i,t-k} + \eta_t + \gamma_i + e_{i,t}. \tag{7}$$

Estimating the model in (6), our results are shown in Appendix C. We present results for each of the four markets using one-, two-, three- and five-year nonoverlapping windows. We find that changes in prices for San Diego, Boston and Phoenix are positively associated with price levels at the beginning of that yearly interval. The results are highly significant at one-, two- and three-year intervals, but are noticeably less so when a five-year interval is used. Chicago, alternatively, has price levels negatively correlated with the subsequent change in prices, but the results are insignificant at common thresholds for all of the year intervals tested.

We should also examine the magnitude of the estimated coefficients to see if they seem plausible with theory. We analyze these in the final columns of the tables in Appendix C. We begin by taking the average annual coefficient for each market—that is the coefficient for each window of years divided by the number of years in the window. We then multiply this by 100,000 dollars. This dollar value is about one standard deviation in the cross-section dispersion of ZIP residual prices in 1998 (averaged across the four cities). The result (labeled ΔP/100KP in the fourth column of the Tables) table gives the annual subsequent appreciation resulting in each market for a ZIP that has a $100,000 higher price residual value.

In the case of Boston, the yearly subsequent appreciation that results from a positive ($100,000) residual is virtually the same regardless of the number of subsequent years over which the appreciation is measured. To our thinking this infers that the differences in fundamental growth across Boston ZIPs are relatively constant over time. Prices between ZIP seem to be smoothly diverging with higher priced areas experiencing constantly higher appreciation rates. The magnitude of the coefficient however is a bit smaller than theory would suggest. If P/R ratios are 20, then a 0.5% higher growth rate in R, *ceteris paribus*, would warrant only a 10% higher price (the present discounted value of the 0.5% growth in R). A $100,000 deviation on average Boston Prices represents about a 25% price premium and, such a price premium would require a 1.25% annual increase in R (if discounting is done at 5%).

In San Diego, the coefficients get distinctly larger as the window lengthens. This would suggest that fundamentals may be increasingly diverging across ZIPs. Higher priced neighborhoods are not only appreciating faster, but the rate of appreciation increases when measured over longer periods. In Phoenix, on the other hand, the rate of appreciation attenuates when measured over longer periods. This would occur if higher priced ZIPs are experiencing only temporary improvements in their fundamentals. In both of these cities the average coefficient is about two times that in Boston, and 1.0% greater yearly appreciation is more consistent with the $100,000 higher price level, using on a P/R value of 20.

Chicago remains a puzzle. It is tempting to assert that fundamentals are deteriorating in higher priced areas, but this would be inconsistent with efficient (forward-looking) pricing. If the deteriorating fundamentals can be even partially foreseen, then the prices should not be higher. While the magnitude of the coefficients is just as large as the other three cities, the standard errors are far greater, so the negative impact is too imprecise to be meaningful. The 95% confidence band, for example, falls well into positive territory.

It is prudent to think about whether the errors in (6) are true "noise" or whether there might be some type of specification error. In the third section, we offered several examples for how one ZIP might appreciate faster or more slowly than others—as differences in the underlying growth in the area's fundamentals get capitalized into current period price. If, as is likely, these fundamentals, and their change are not observable, their impact will be potentially be captured by the error term. In this case the error term would be correlated across time for one ZIP, *i.e.*, autocorrelated. Here we try to correct for the possible impact of such omitted variables on the specification of the model error term.

When ZIP fundamentals are changing systematically over time, but are unobservable (as we hypothesize they might be), the error term should contain first-order autocorrelation across time (as rewritten in 8). In the discussion that follows, for simplicity we assume a yearly model ($k = 1$).

$$e_{i,t} = \theta_i e_{i,t-1} + \varepsilon_{i,t}. \tag{8}$$

In (8), θ_i is the correlation across time in the errors of each ZIP i. The term $\varepsilon_{\mathrm{i,t}}$ is the true random error for the observation. In this formulation, locations with systematic but unobservable changes in their fundamentals should have high positive autocorrelation while areas with little or no change in fundamentals should have zero autocorrelation. Hence it is important that the degree of autocorrelation be allowed to vary across the ZIPs. Following Parks (1967) there are a number of ways of devising an estimation strategy that will remove the correlations in the error term.

We begin by taking the lagged one period value of the LHS of (8), multiply it with an estimate of θ_i (which we label r_i) and then difference it with current period. Following the same differencing with the RHS we get (9) wherein the error is returned to random noise.

$$\begin{aligned}
&(\Delta P_{i,t} - \Delta \bar{P}_t) - r_i(\Delta P_{i,t-1} - \Delta \bar{P}_{t-1}) \\
&\quad = \alpha(1 - r_i) + \beta[(P_{i,t-1} - \bar{P}_{t-1}) - r_i(P_{i,t-2} - \bar{P}_{t-2})] + [e_t - r_i e_{t-1}] \\
&\quad = \alpha(1 - r_i) + \beta[(P_{i,t-1} - \bar{P}_{t-1}) - r_i(P_{i,t-2} - \bar{P}_{t-2})] + \varepsilon_{i,t}.
\end{aligned} \tag{9}$$

To estimate the parameters in this version, we adopt a two-step procedure. In the first step we find values for r_i by running a Prais-Winsten GLS for each ZIP in an MSA. In these equations we use the demeaned change in price as the dependent variable and the demeaned price level as the independent variable (again with $k = 1$). The correlations of the residuals are uniformly positive. In

Chicago 45 of 51 ZIP have positive autocorrelation, in Boston 91 of 109, in Phoenix 69 of 80 and San Diego 32 of 42. Thus there is considerable empirical evidence of auto correlated errors - which again offers evidence of persistence in the unobservable determinants of price appreciation. After obtaining these r_i we then transform the dependent variable (the demeaned change in price) and independent variable (demeaned price).

As $\varepsilon_{\mathrm{i,t}}$ is by definition noise, we can run a simple OLS in a second stage to obtain final estimates of alpha and beta in (9) above. The results are reported in Appendix D for the case of yearly intervals ($k = 1$). Boston, Phoenix and San Diego continue to have positive and significant coefficients for price. The magnitude of the coefficients is similar to the results in Appendix C, with Phoenix having the largest coefficient and San Diego has the smallest. Chicago, alternatively, has a negative coefficient for lagged price, but it still is not significant. The results are quite consistent with the first set of results in which there was no adjustment for autocorrelation.

Conclusion

In this article, we examine the movement of prices across locations within metropolitan areas, controlling for housing attributes. In particular, we construct a test of price "efficiency" that is a variant of the Campbell–Shiller test. These and subsequent authors have examined the predictive power of price/rent ratios on subsequent price and earnings growth. In housing, those attempting such a test previously have used average prices and average rents. Recent research, however, has documented that owned and rented assets are very different, and that these vary over time and space. We have proposed using an alternative comparison of price to an imputed rent measure for the same housing unit. To our thinking a unit's predicted price using its current attributes in a hedonic equation represents an imputed rent. The residuals from this hedonic equation should capture both the omitted determinants of rent as well as the expected future growth in rent determinants (which is observable only *ex post*).

With this notion we test if housing Hedonic equation residuals have a significant positive correlation with subsequent price growth. We do this in four MSA. In each we use a combination of detailed sales data in one year combined with repeat sales price indices at the ZIP-code level. With this combination of data we find in three of our four MSA that ZIPs with higher price level residuals in any given year have statistically significant greater price appreciation in subsequent years. In our fourth MSA there is an insignificant relationship. This provides at least some evidence that across locations within MSA, price levels reflect expected growth in location "fundamentals" and hence future price changes.

The authors are indebted to the MIT Center for Real Estate. They remain responsible for all results and conclusions derived there from.

References

Abel, A. and O. Blanchard. 1986. The Present Value of Profits and Cyclic Movements in Investment. *Econometrica* 54(2): 249–273.

Campbell, J. and R.J. Shiller. 1988. Stock Prices, Earnings, and Expected Dividends. *Journal of Finance* 43(3): 661–676.

—— and ——. 1998. Valuation Ratios and the Long Run Stock Market Outlook. *Journal of Portfolio Management*: 24(2): 11–26.

——, M. Davis, J. Gallin and R. Martin. 2009. What Moves Housing Markets: A Variance Decomposition of the Price-Rent Ratio. *Journal of Urban Economics* 66(2): 90–102.

Capozza, D. and R. Helsley. 1990. The Stochastic City. *Journal of Urban Economics* 28(2): 187–203.

——, P. Hendershott and C. Mack. 2004. An Anatomy of Price Dynamics in Illiquid Markets: Analysis and Evidence from Local Housing Markets. *Real Estate Economics* 32: 1–21.

Case, K.E. and R.J. Shiller. 1988. The Behavior of Home Prices in Boom and Post-Boom Markets. *New England Economic Review*: 29–46.

—— and ——. 1989. The Efficiency of Market for Single-Family Homes. *American Economic Review* 79(1): 125–137.

DiPasquale, D. and W.C. Wheaton. 1996. *Urban Economics and Real Estate Markets*. New Jersey: Prentice Hall.

Gordon, R. 1959. Dividend Earnings and Stock Prices. *Review of Economics and Statistics* 41(2): 99–105.

Grossman, S. and G. Laroque. 1990. Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods. *Econometrica* 58: 25–51.

Hendershott, P. and J. Slemrod. 1983. Taxes and the User Cost of Capital for Owner-Occupied Housing. *Journal of the American Real Estate and Urban Economics Association* 10(4): 375–393.

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

Q6 Kindleberger, C. 2000. *Manias, Panics and Crashes* (4th ed.). (1st ed. 1978). Wiley.

Meese, R. and N. Wallace. 1993. Testing the Present Value Relation for Housing Prices: Should I Leave My House in San Francisco? *Journal of Urban Economics* 35(3): 245–266.

Ortalo-Magne, F. and S. Rady. 1999. Boom In, Bust Out: Young Households and the Housing Price Cycle. *European Economic Review* 43(4): 755–766.

—— and ——. 2001. Housing Market Dynamics: On the Contribution of Income Shocks and Credit Constraints. Discussion Paper 470, CESIfo.

Parks, R. 1967. Efficient Estimates of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated. *Journal of American Statistical Association* 62: 500–509.

Poterba, J.M. 1984. Tax Subsidies to Owner Occupied Housing: An Asset Market Approach. *The Quarterly Journal of Economics* 99(4): 729–752.

Smith, M.H. and G. Smith. 2006. Bubble, Bubble, Where's the Housing Bubble? *Brooking Papers* 37(1): 1–68.

Stein, J. 1995. Prices and Trading Volume in the Housing Market: A Model with Down Payment Effects. *Quarterly Journal of Economics* 110: 379–406.

Appendix A: Hedonic Equations

(A) Boston Hedonic Price Regression (1998)

lnsale	Coef.	*T*	*P*>|*t*|	Sig. coef.
Built 1960–1980	−0.05026	−7.11	0	−0.05026
Built 1940–1960	−0.13523	−16.09	0	−0.13523
Built 1900–1940	−0.1662	−18.46	0	−0.1662
Built pre-1900	−0.19206	−15.89	0	−0.19206
1 bedroom	−0.12093	−3.78	0	−0.12093
2 bedroom	−0.06042	−4.29	0	−0.06042
3 bedroom	−0.00808	−0.66	0.507	0
4 bedroom	0.005488	0.46	0.643	0
1 bath	−0.20654	−16.47	0	−0.20654
1.5 bath	−0.14318	−12.71	0	−0.14318
2 bath	−0.16332	−13.98	0	−0.16332
2.5 bath	−0.06041	−6.19	0	−0.06041
Interior square feet	0.000298	21.11	0	0.000298
Sq. feet squared	−1.32E-08	−5.58	0	−1.3E-08
Lot size	0.14155	13.27	0	0.14155
Lot size sq.	−0.01635	−8.98	0	−0.01635
Regsco	0.015277	4.09	0	0.015277
Density	0.041415	8.51	0	0.041415
Ln median income	0.357986	2.16	0.031	0.357986
_cons	−2.95215	−2.95	0.003	−2.95215

Note: Fixed effects omitted, $R^2 = 0.7168$.

(B) Phoenix Hedonic Price Regression (1998)

Lnsale	Coef.	*P*>|*t*|	Sig. coef.
Built 1980–1990	−0.0886	0	−0.0886
Built 1970–1980	−0.14813	0	−0.14813
Built 1960–1970	−0.20073	0	−0.20073
Built before 1960	−0.2527	0	−0.2527
1 bath	−0.11637	0	−0.11637
1.5 bath	−0.12147	0	−0.12147
2 bath	−0.05861	0	−0.05861
2.5 bath	−0.02484	0	−0.02484
3 bath	−0.00131	0.814	0
5 bedroom	−0.00331	0.709	0
6 bedroom	−0.02291	0.037	−0.02291
7 bedroom	−0.0511	0	−0.0511
Room8-	−0.08753	0	−0.08753
Interior sq. feet	0.0006	0	0.0006
Interior sq. feet square	−4.42E-08	0	−4.4E-08
Lot size	0.260908	0	0.260908
Lot size sq.	−0.02199	0.001	−0.02199
Pool	0.052737	0	0.052737
Garage	0.057881	0	0.057881
Density	−0.00434	0.511	0
ln Median Income	0.011572	0.901	0
Constant	11.18304	0	11.18304

Note: Fixed effects omitted, $R^2 = 0.776$.

(C) San Diego Price Regression (1998)

lnsale	Coef.	$P>\|t\|$	Sig. coef.
Built 1980–1990	0.009864	0.45	0
Built 1970–1980	−0.09978	0	−0.09978
Built 1960–1970	−0.09502	0	−0.09502
Built before 1960	−0.0653	0.001	−0.0653
1 bedroom	0.076764	0.177	0
2 bedroom	0.033985	0.085	0.033985
3 bedroom	0.052283	0.001	0.052283
4 bedroom	0.032621	0.02	0.032621
5 bedroom	−0.0742	0.015	−0.0742
1.5 bath	−0.06265	0.059	−0.06265
2 bath	−0.01135	0.667	0
2.5 bath	−0.03225	0.188	0
3 bath	−0.02204	0.344	0
Interior sq. feet	0.000467	0	0.000467
Interior sq. feet square	−2.58E-08	0	−2.6E-08
Lot size sq.	−1.46E-06	0.847	0
Lot size	0.001165	0.734	0
Garage	0.135294	0	0.135294
Pool	0.076756	0	0.076756
Density	0.00358	0.284	0
lnmedinc	0.85806	0	0.85806
_cons	2.262603	0.001	2.262603

Note: Fixed effects omitted, $R^2 = 0.6994$

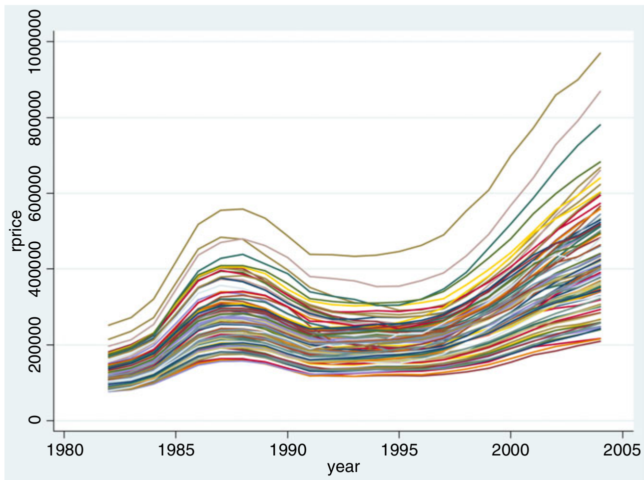
(D) Chicago Hedonic Regression (1998)

lnsale	Coef.	$P>\|t\|$	Sig. coef.
Built after 1980	0.195267	0	0.195267
Built 1970–1980	0.124359	0	0.124359
Built 1970-1960	0.078294	0	0.078294
Built 1960–1970	0.033203	0	0.033203
1 bath	−0.27247	0	−0.27247
1.5 bath	−0.21495	0	−0.21495
2 bath	−0.19741	0	−0.19741
2.5 bath	−0.14139	0	−0.14139
3 bath	−0.11663	0.004	−0.11663
Interior sq. feet	0.00036	0	0.00036
Interior sq. feet square	−2.60E-08	0	−2.6E-08
Lot size	0.381877	0	0.381877
Lot size sq.	−0.0505	0	−0.0505
Density	0.025079	0	0.025079
Nonwhite	−0.96735	0	−0.96735
_cons	11.77202	0	11.77202

Note: Fixed effects omitted, $R^2 = 0.6372$

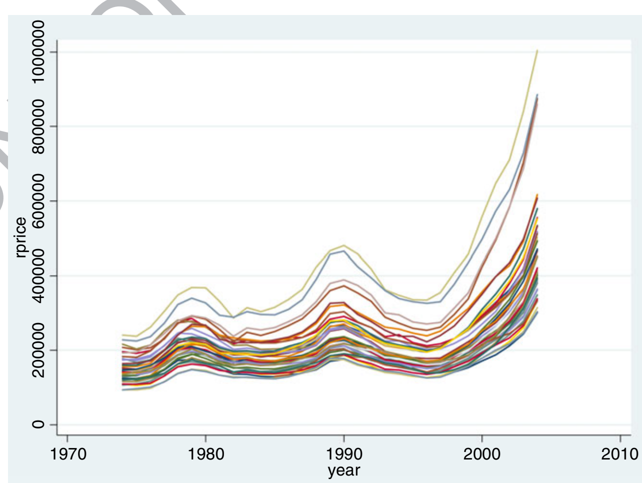
Appendix B: Panel Price Series

Chart 1 ■ Boston.



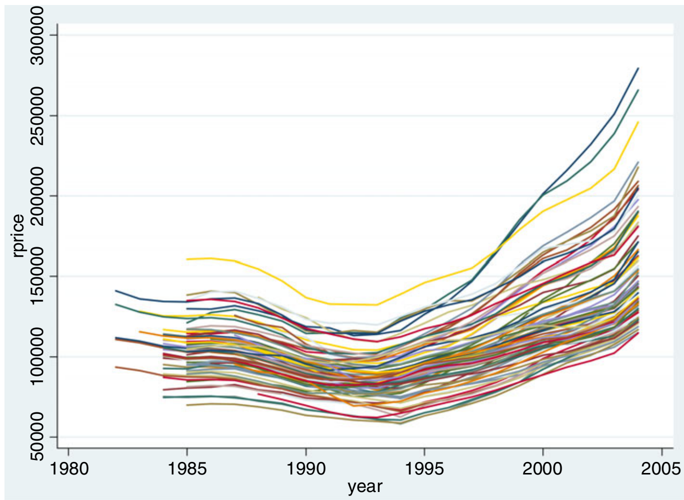
Note: rprice refers to the real price of identical house with specified features.

Chart 2 ■ San Diego.



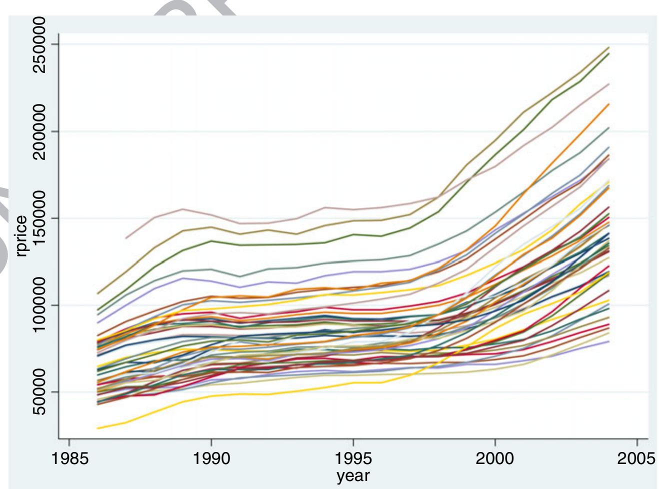
Note: rprice refers to the real price of identical house with specified features.

Chart 3 ■ Phoenix.



Note: rprice refers to the real price of identical house with specified features.

Chart 4 ■ Chicago (Cook County).



Note: rprice refers to the real price of identical unit with specified features.

Appendix C: Panel Regression Results

Demeaned Change in Price Regression against Demeaned Prices

C.1 Yearly Change in Price against One Year Lagged Price

Boston: Dependent variable: Yearly Change in Price (Demeaned)

	Coef.	Std. err.	T	ΔP / 100KP|
Price (Demeaned)				
Lag 1	5.31e-08	7.17e-09	7.40	0.53%
R^2	0.0230			
Adjusted R^2	0.0226			

Phoenix: Dependent variable: Yearly Change in Price (Demeaned)

	Coef.	Std. err.	T	ΔP / 100KP|
Price (Demeaned)				
Lag 1	1.48e-07	2.47e-08	6.00	1.48%
R^2	0.0222			
Adjusted R^2	0.0216			

San Diego: Dependent variable: Yearly Change in Price (Demeaned)

	Coef.	Std. err.	t	ΔP / 100KP|
Price (Demeaned)				
Lag 1	2.83e-08	1.29e-08	2.20	0.28%
R^2	0.0039			
Adjusted R^2	0.0031			

Chicago: Dependent variable: Yearly Change in Price (Demeaned)

	Coef.	Std. err.	t	
Price (Demeaned)				
Lag 1	−5.06e-08	3.89e-08	−1.30	−.51%(NS)
R^2	0.0019			
Adjusted R^2	0.0008			

C.2 2-Yearly Change in Price against Two-Year Lagged Price

Boston: Dependent variable: 2 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	t	ΔP / 100KP|
Price (Demeaned)				
Lag 2	1.07e-07	2.32e-08	4.61	0.52%
R^2	0.0197			
Adjusted R^2	0.0188			

Phoenix: Dependent variable: 2 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	t	ΔP / 100KP|
Price (Demeaned)				
Lag 2	2.62e-07	7.08e-08	3.69	1.31%
R^2	0.0182			
Adjusted R^2	0.0169			

San Diego: Dependent variable: 2 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	t	ΔP / 100KP|
Price (Demeaned)				
Lag 2	1.23e-07	3.92e-08	3.13	0.61%
R^2	0.0138			
Adjusted R^2	0.0123			

Chicago: Dependent variable: 2 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	t	ΔP / 100KP|
Price (Demeaned)				
Lag 2	−2.11e-07	1.29e-07	−1.64	−1.05%(NS)
R^2	0.0068			
Adjusted R^2	0.0043			

C.3 3-Yearly Change in Price against Three Year Lagged Price

Boston: Dependent variable: 3 Yr. Change in Price (Demeaned)

	Coef.	Std. Err.	t	ΔP / 100KP
Price (Demeaned)				
Lag 3	1.88e-07	4.22e-08	4.45	0.61%
R^2	0.0261			
Adjusted R^2	0.0248			

Phoenix: Dependent variable: 3 Yr. Change in Price (Demeaned)

	Coef.	Std. Err.	t	ΔP / 100KP
Price (Demeaned)				
Lag 3	4.67e-07	1.27e-07	3.68	1.55%
R^2	0.0268			
Adjusted R^2	0.0248			

San Diego: Dependent variable: 3 Yr. Change in Price (Demeaned)

	Coef.	Std. Err.	t	ΔP / 100KP
Price (Demeaned)				
Lag 3	7.10e-08	6.89e-08	1.03	0.24%(NS)
R^2	0.0020			
Adjusted R^2	0.0001			

Chicago: Dependent variable: 3 Yr. Change in Price (Demeaned)H

	Coef.	Std. Err.	t	ΔP / 100KP
Price (Demeaned)				
Lag 3	−3.64e-07	2.14e-07	−1.71	−1.19%(NS)
R^2	0.0099			
Adjusted R^2	0.0065			

C.4 5-Yearly Change in Price against Five-Year Lagged Price

Boston: Dependent variable: 5 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	*t*	Δ*P* / 100KP|
Price (Demeaned)				
Lag 5	2.86e-07	1.38e-07	2.07	0.57%
R^2	0.0101			
Adjusted R^2	0.0077			

Phoenix: Dependent variable: 5 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	*t*	Δ*P* / 100KP|
Price (Demeaned)				
Lag 5	4.85e-07	3.87e-07	1.25	0.97%(NS)
R^2	0.0063			
Adjusted R^2	0.0023			

San Diego: Dependent variable: 5 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	*t*	Δ*P* / 100KP|
Price (Demeaned)				
Lag 5	5.85e-07	2.27e-07	2.58	1.17%
R^2	0.0227			
Adjusted R^2	0.0192			

Chicago: Dependent variable: 5 Yr. Change in Price (Demeaned)

	Coef.	Std. err.	*t*	Δ*P* / 100KP|
Price (Demeaned)				
Lag 5	−8.75e-07	6.11e-07	−1.43	−1.73%(NS)
R^2	0.0139			
Adjusted R^2	0.0071			

Appendix D: Quasi-Differenced Tests

Boston: Dependent variable: Yearly Change in Price (Demeaned and Quasi-Differenced)

	Coef.	Std. err.	t	$P>\|t\|$	[95% Conf.	interval]
Price (Demeaned and quasi differenced)						
Lag 1	4.40e-08	6.96e-09	6.33	0.000	3.04e-08	5.77e-08
R^2	0.0177					
Adjusted R^2	0.0173					

Phoenix: Dependent variable: Yearly Change in Price (Demeaned and Quasi-Differenced)

	Coef.	Std. err.	t	$P>\|t\|$	[95% Conf.	interval]
Price (Demeaned and quasi differenced)						
Lag 1	1.33e-07	2.73e-08	4.88	0.000	7.97e-08	1.87e-07
R^2	0.0156					
Adjusted R^2	0.0149					

S. Diego: Dependent variable: Yearly Change in Price (Demeaned and Quasi-Differenced)

	Coef.	Std. err.	t	$P>\|t\|$	[95% Conf.	interval]
Price (Demeaned and quasi differenced)						
Lag 1	2.93e-08	1.28e-08	2.28	0.023	4.12e-09	5.44e-08
R^2	0.0043					
Adjusted R^2	0.0035					

Chicago: Dependent variable: Yearly Change in Price (Demeaned and Quasi-Differenced)

	Coef.	Std. err.	t	$P>\|t\|$	[95% Conf.	interval]
Price (Demeaned and quasi differenced)						
Lag 1	−8.19e-08	5.03e-08	−1.63	0.104	−1.81e-07	1.68e-08
R^2	0.0032					
Adjusted R^2	0.0020					

Author Queries

Q1: Author: Please confirm that given names (red) and surnames/family names (green) have been identified correctly.

Q2: Author: Please provide missing department/division details in all the affiliations.

Q3: Author: The reference "Meese and Wallace 1994" is cited in the text but is not listed in the references list. Please either delete in-text citation or provide full reference details following journal style.

Q4: Author: A running head short title was not supplied; please check if this one is suitable and, if not, please supply a short title that can be used instead.

Q5: Author: The reference "Kindleberger 1978" is cited in the text but is not listed in the references list. Please either delete in-text citation or provide full reference details following journal style.

Q6: Author: Please provide publisher location for reference Kindleberger (2000).