NBER WORKING PAPER SERIES

EXPECTATIONS, EFFICIENCY, AND EUPHORIA IN THE HOUSING MARKET

Dennis R. Capozza
Paul J. Seguin

Working Paper No. 5179

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1995

The authors thank James Berkovic, Timothy R. Burch, Patric Hendershott, Roger Kormendi, Nejat Seyhun, Nancy Wallace and seminar participants at Arizona State University, Georgia State University, Penn State University, the University of British Columbia, the University of Michigan, and the NBER conference on Public Policy and the Housing Market for their helpful comments and Phil Maguire and Alastair McFarlane for their expert research assistance. This paper was prepared as part of the NBER Study on Housing Dynamics and was presented at the NBER Study Conference. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

NBER Working Paper #5179
July 1995

EXPECTATIONS, EFFICIENCY, AND EUPHORIA IN THE HOUSING MARKET

ABSTRACT

This paper studies expectations of capital appreciation in the housing market. We show that expectations impounded in the rent/price ratio at the beginning of the decade successfully predict appreciation rates, but only if we first control for cross-sectional differences in the quality of rental versus owner-occupied housing. We also demonstrate that observed rent/price ratios contain a disequilibrium component that also has power to forecast subsequent appreciation rates. Finally, we provide evidence consistent with euphoria: participants in housing markets appear to overreact to income growth.

Dennis R. Capozza
School of Business Administration
University of Michigan
Ann Arbor, MI 48109-1234

Paul J. Seguin
School of Business Administration
University of Michigan
Ann Arbor, MI 48109-1234

Expectations, Efficiency, and Euphoria in the Housing Market

Abstract

This paper studies expectations of capital appreciation in the housing market. We show that expectations impounded in the rent/price ratio at the beginning of the decade successfully predict appreciation rates, but only if we first control for cross-sectional differences in the quality of rental versus owner-occupied housing. We also demonstrate that observed rent/price ratios contain a disequilibrium component that also has power to forecast subsequent appreciation rates. Finally, we provide evidence consistent with euphoria: participants in housing markets appear to overreact to income growth

Interest among academics in long run trends in housing prices was recently spurred by a controversial[1] study by Mankiw and Weil (1989). Mankiw and Weil note that the surge in demand for housing caused by the postwar baby boom could be easily predicted 20 years in advance because the eventual population in high demand cohorts (over 20 years old) can be estimated from the cohort data of earlier years. Mankiw and Weil hypothesize that if expectations are rational then home buyers should anticipate the effects of a population bulge like the postwar baby boom on the price of housing and bid prices up in anticipation. Their empirical results reject efficiency, i.e. home prices appear to rise contemporaneously rather than in advance of predictable cohort shifts, suggesting that the housing market is informationally inefficient, at least with respect to this class of information.

---

[1]The MW article sparked an unusually large number of rebuttals which appeared in a subsequent issue of this journal.

In this study, we also address the issue of informational efficiency in the housing market. However, unlike Mankiw and Weil, we use census data disaggregated by metropolitan areas to analyze decadal appreciation rates. Exploiting disaggregated data on cross-sectional variations of appreciation rates is beneficial for a number of reasons. First, as with labor markets, supply and demand factors in real estate markets vary from locale to locale. By using aggregated data, many of these market specific factors are canceled out (diversified away) in the aggregation process. For example, while high household formation rates are easily predicted in the aggregate, decadal population growth at the metro area level varied from -9% to +130% during the post-war period and was primarily due to intercity migration rather than the baby boom. Thus, metro area rates on which local expectations are likely to be based are highly volatile, difficult to predict, and swamp the variation of the postwar baby boom.

Second, the use of cross-sectional disaggregated data provides at least two econometric benefits. By using disaggregated data, the number of usable observations, and hence the statistical power of our tests, are increased. Also, by using cross-sectional data, we circumvent a number of potentially troublesome time-series problems encountered in the literature investigating the predictive power of dividend yields.[2]

---

[2]A number of studies in the finance literature have tested the predictive power of the dividend yield to some aggregate stock index. Though some studies have reported success, including Fama and French (1988), a number of recent studies have argued that inference is difficult given the existence of only one time series and the relative variability of prices compared to dividends. See Hodrick (1992) and Goetzmann and Jorion (1993) for examples.

The primary objective of this study is to examine the efficiency of real estate markets, in general, and the rationality of expectations, in particular.[3] The specific proposition we examine is whether expected risk–adjusted total returns on housing are equal across metro areas. To derive empirical predictions from this proposition, we borrow from the stock dividend literature and show that if this proposition is true, then rent/price ratios should predict future appreciation rates. This empirical prediction is the basis of our tests.

However, due to high transactions and information costs, asymmetric information, and differences in the quality of rental versus owner occupied housing, appreciation rates and rent/price ratios calculated from census data will be measured with considerable error. Further, because some of this measurement error is common to both the dependent variable (appreciation rates calculated from prices) and the independent variable (rent/price ratios), results from standard econometric techniques will be biased.

To mitigate this bias, we adopt an instrumental variable technique. We first regress the observed rent/price levels on a number of variables previously identified as being related to property value. The fitted values from this cross-sectional regression incorporate less observation error than observed values.

---

[3]Meese and Wallace (1994) and Hamilton and Schwab (1985) also examine expectation of price appreciation in the housing market using dividend ratio or present value models. An alternative approach to explaining house values has been to base estimates on the asset pricing approach to valuing real estate developed in Capozza and Helsley (1990), Capozza and Sick (1994), and Capozza and Li (1994). Under this approach, housing must compete with other financial assets and therefore must satisfy capital market equilibrium conditions. Growth of population and income as well as both systematic and unsystematic risk are important in this approach while amenities are not. Examples of this approach include Capozza and Schwann (1989, 1990) and Abraham and Hendershott (1992)

Further, any remaining observation error is, presumably, uncorrelated with the observation error in the dependent variable. Thus, we use these fitted values as the independent variable in our main specification. When we regress appreciation rates on these first stage fitted values, we find that the rent/price ratio is highly significant and of the correct sign. However, we find that the residuals from the first-step regression also contain predictive power. We interpret this finding as evidence that observed rent/price ratios contain a disequilibrium component, indicating some mean reversion in the relative prices of rental versus owner occupied housing. To explore the source of this mean reversion, we add lagged income and population growth rates to the specification. We find that lagged income growth enters the equation significantly, which is at least consistent with "euphoria," i.e., an overreaction by market participants to recent economic activity.

In the following section, we review dividend theory and explicitly state our testable hypotheses. The third section discusses the data while the fourth presents the results of tests of expectations. The next two sections describe our work on efficiency and euphoria. Some brief conclusions and ongoing work are presented in the final section.

A Brief Review of Dividend Theory

A fundamental tenet in corporate finance is the dividend irrelevance proposition of Miller and Modigliani (1961). They point out that in the absence of taxes and transactions costs, a corporation's dividend policy does not affect the value of its shares. Further, even in the presence of taxes and transaction costs dividend policy does not affect share value. For example, suppose that, for tax reasons, investors prefer capital gains to dividends. Then firms would:

> adjust their dividend policies to take advantage of the negative effects of dividends by adjusting their dividend policies to supply the levels of yield that are most in demand. As a result the supply of shares at each level of yield will come to match the demand for shares at that level of yield, and investors as a group will be happy with the available range of yields. After equilibrium is reached, no corporation will be able to affect its share price by changing its dividend policy. (Black and Scholes, 1974)

Empirically, the finance literature has generally shown that equity prices are efficient with respect to this form of information, and it is difficult to detect any difference in risk adjusted returns between high and low dividend securities (see Black and Scholes (1974) and Miller and Scholes (1984)).

The focus of this study is not to test the dividend irrelevance proposition in the housing market. However we do make use of the concept to test the role of expectations. With competitive markets total risk-adjusted expected returns will be equal across urban areas. If there are differences in expected total returns across urban areas, then capital should flow to those areas with higher expected returns, increasing current price levels, and decreasing future expected total returns. Because expected total returns are the sum of the dividend or rent yield and an expected appreciation rate, urban areas where rent/price ratios are high should have lower expected appreciation. More formally, because:

$$E\{TR_{it}\} = \frac{R_{it}}{P_{it}} + \frac{E\{\Delta P_{it}\}}{P_{it}} \qquad (1)$$

where TR_{it} is the total return to housing in area i over time period t, R is the level of rent in area i over time period t, P_{it} is the price of housing, and E is the expectation operator, then:

$$\frac{E\{\Delta P_{it}\}}{P_{it}} = E\{TR_{it}\} - \frac{R_{it}}{P_{it}}. \qquad (2)$$

This identity[4], which is the basis for our empirical tests, indicates that expected capital gains should be negatively related to the rent/price ratio. Therefore, if information about existing rent ratios have been efficiently impounded into housing prices, then the rent/price ratio should have significant predictive power for future capital gains. Further, the relation is exact: investors should expect an area with a one-percent larger rent/price ratio to experience a one percent smaller appreciation rate per period.

Econometric Issues

Because expected appreciation rates are not observable[5], we need to use realized, rather than expected, capital gains. Thus, we examine whether expectations are rational by testing whether an area with a one-percent larger rent/price ratio, on average, experiences a one percent smaller appreciation rate per period, or a ten percent smaller appreciation rate per decade.

---

[4]Equation (2) follows directly from the present value relation when the expected growth of rents and the expected growth of prices are equal.

[5]An alternative approach to expectations in real estate markets has been to survey owners and renters directly (Case and Shiller (1988), Collins, Lipman, and Groeneman (1992)). This approach is helpful for assessing the average expectations of owners and renters and for understanding the differences among cohorts. However markets reflect only the expectations of the marginal buyer and seller. Because we cannot identify these marginal traders from a survey, we cannot infer from surveys how the expectations that influence prices are formed.

Several problems arise from using observed changes. First, the observed rents are not on the same houses as the observed prices. Because owned and rental units tend to be of different quality, there is measurement error in the data on the rent/price ratio. We do attempt to control for average quality between urban areas using housing characteristics available in the census data, but this does not fully correct for quality differences within each area. As a result, we expect the coefficient on rent/price to be biased towards zero.

Second, though arbitrage is normally assumed to keep prices close to long run equilibrium, all assets trade within a band determined by transaction and information costs. In real estate markets these costs are large relative to securities markets, so real estate trades occur at prices within a wide price band. This variation in trading prices within the band adds noise to the observations, with a concomitant reduction in statistical power. Further, this error in observed capital appreciation is positively correlated with the observation error inherent in the rent/price ratio at the beginning of the period, which we use as the predictive variable. Therefore, regressing the observed appreciation rate on the observed beginning of period rent/price ratio will result

in positively biased coefficient estimates[6]. Because the expected sign of the coefficient is negative, these biases reinforce each other.

Data

Our unique sample is a pooled, cross-sectional time series of 64 Standardized Metropolitan Areas (SMAs) in the US from 1960 to 1990. The data were collected primarily from the decennial census but supplemented with series from other sources. A complete description of the data and its sources appear in the Data Appendix.

Two features of our database enable us to test for informational efficiency with greater statistical power than earlier studies. First, we have a larger number of SMAs in our sample than in Hamilton and Schwab (1985) who examine 21 cities from the American Housing Survey, or Meese and Wallace (1994) who focus on one city. Second, while Hamilton and Schwab's investigation covers three years, our data covers four decades.

---

[6]To illustrate, assume that we want to regress y on x, but can only observe X = x + u and Y = y + v. In the standard error-in-variable case, where $E\{x'u\} = E\{y'v\} = E\{u'v\} = 0$, then OLS estimates of the slope parameter are biased towards zero by a factor proportional to $\frac{E\{x'x\}}{E\{x'x\} + E\{u'u\}}$. However, in this case, $E\{u'v\} > 0$, so OLS slope estimates will suffer from both a positive bias, and a bias towards zero. For example, in the case of a single regressor (with zero mean, for ease of exposition only), it is easy to show that the expected value of the OLS slope estimate is $(\beta + \frac{E\{u'v\}}{\sigma_x^2})(1 + \frac{\sigma_u^2}{\sigma_x^2})^{-1}$. The first term captures the positive bias while the second captures the bias towards zero.

The key variables for our analyses are the decadal percentage change in real house values, Δln(VALUE), calculated as differences in the logs of reported prices deflated by the level of the CPI, and the rent/price ratio (R/P), calculated as the ratio of the median annual rental rate deflated by the median value of owner occupied housing. Because our data is decadal, subscripts relative to "t" refer to decades, so "t-1" indicates an observation from the previous decade.

Expectations

To illustrate the magnitude of the biases outlined above, we initially estimated a simple one-step specification by regressing the real percentage change in housing values over the decade, $\Delta\ln(VALUE)_t$ against the gross rental rate measured as of the start of the decade, R/P_{t-1}, and separate intercepts for each of the three decades over which the real decadal appreciation was calculated, Y_{60s}, Y_{70s} and Y_{80s}. To the extent that R/P_{-1} has a coefficient of -10, these intercepts estimate total real return to housing over that decade. Separate decadal intercepts are included to mitigate the effect of macro-wide factors in the specification including changes in aggregate demographics, especially changes in age cohorts (Mankiw and Weil (1989)). Data from three decades of price appreciation for each of the 64 SMAs was used, yielding (with t-statistics in parentheses):

$$\begin{array}{lccccccc} \Delta\ln(VALUE) = & 0.07\ Y_{60s} & + & 0.30\ Y_{70s} & + & 0.09\ Y_{80s} & - & 0.10\ R/P_{t-1} \\ & (0.7) & & (2.8) & & (1.0) & & (-0.1) \end{array}$$

Intercepts associated with appreciation over the 1960s (Y_{60s}) and the 1980s (Y_{80s}) are small and insignificant indicating little average real housing total return over these decades. In contrast, real housing returned more than 30% over the 1960s, and this increase is significant. Of primary importance,

however, is the coefficient associated with the rent ratio as of the beginning of the decade. Because we are using annual rental rates and decadal capital appreciation rates, the estimated coefficient on the rent/price coefficient should be approximately -10, which is roughly one hundred times the estimated coefficient. Further, though the estimate is of the correct sign, it is insignificantly different from zero. We attribute this finding to reinforcing effects of both a bias towards zero attributable to measurement error, and a positive bias of a negative coefficient estimate due to correlated measurement errors, as detailed in footnote 6.

To mitigate the measurement error problems, we use an Instrumental Variables or Two Stage Least Squares technique. In the first stage, we regress the observed rent/price ratio on a set of variables suggested by Capozza and Helsley (1990) and Capozza and Sick (1994) in their investigations of urban growth with uncertainty.[7] These variables include income, income growth, population, population growth rate, property tax rate, developable area in the city, and construction costs. We also add variables specifically designed to help accommodate cross-sectional differences in quality, including utility cost rates and the number of baths, as well as lagged growth rates in population and income. The results are not sensitive to alternative specifications.

---

[7]A number of commentators have suggested that we run the first stage regression on the level of housing prices and then employ the fitted values to calculate the rent/price ratio. Though feasible, this procedure results in a generated regressor (the rent/price ratio) that is a non-linear transformation of the fitted value from the first stage regression (the price level). As a result, OLS standard errors would be biased. Under our approach, unadjusted standard errors are correct, and the nominal test size is accurate under the null that the parameter equals zero since the variable measured with error is linear in the observation error. See Murphy and Topel (1985, especially pg. 375), Pagan (1984) and Schwert and Seguin (1990, pg. 1133) for details.

Results of the cross-sectional regression on rent/price ratios on these instruments with intercepts that vary by decade appear in table 1. The high level of the R^2 (=0.52) suggests that the model is successful in capturing much of the cross-sectional dispersion in rent/price ratios.

The primary motivation for the first stage is not to provide a model of rent/price ratios, *per se*, but to partition these values into expected components and residuals. From a purely econometric view, the expected component or fitted value is an estimate of the rent/price ratio that either contains no measurement error, or, at least contains measurement error that is uncorrelated with measurement error in appreciation rates. In that technical respect, use of the expected component is desirable because specifications employing it would, presumably, suffer from less measurement error bias. Use of fitted values is also desirable from an economic perspective. The fitted values can be interpreted as quality adjusted rent/price ratios, or, the portion of the observed rent/price ratio that can be attributed to cross-sectional differences in the relative quality of rental versus owner occupied housing.

Next, we use the fitted values, denoted as $\hat{R/P}$, in the model linking capital appreciation over the decade to the rent/price ratio as of the beginning of the decade. With the same 192 observations, estimation via OLS yields:

$$\Delta \ln(\text{VALUE}) = \underset{(3.8)}{0.48 Y_{60s}} + \underset{(5.4)}{0.78\, Y_{70s}} + \underset{(4.0)}{0.50\, Y_{80s}} - \underset{(-3.4)}{6.13\, \hat{R/P}_{t-1}}$$

The estimates of all coefficients change dramatically. Unlike the original specification, intercepts estimating total returns over each of these decades are large, significant positive numbers reflecting total annualized real returns to

housing of about 5%, 8% and 5% per year for the decades, respectively. The movement in these estimates away from zero is consistent with a mitigation of the errors-in-variables problem[8].

Of primary importance, however, is the fact that the rent/price ratio as of the beginning of the decade is now significantly negative and is much closer to the value predicted by theory of -10[9]. Thus, current rent/price ratios appear to have power to predict subsequent capital appreciation.

The fact that the estimated coefficient is still less than its hypothesized value could arise from a number of sources that are not necessarily mutually exclusive. First, if homeowners have a strong preference for capital gains, perhaps due to asymmetric tax treatment, then a 1% decrease in rent yield will result in a less than one percent increase in required capital gains. However, implicit rents are never taxed while capital gains may be taxed on realization. Thus, we believe that housing prices and rents will be determined by marginal investors for whom rents are preferred to capital gains and therefore view this explanation as being unlikely.

A second possibility is that our results are indicative of a rejection of one of the fundamental assumptions unpinning our investigation, namely, that expected total returns are the same across SMAs for a given period. Specifically, if

---

[8]Since estimates of an intercept and estimates of a slope coefficient are negatively correlated, a negative movement away from zero for a slope will tend to occur with a positive movement in the intercept.

[9]If we use a standard t-test, the coefficient is significantly different from -10. However, it should be noted that reported OLS standard errors understate true standard errors when using a generated regressor. Using the correct standard error from Murphy and Toepel (1985, eq. 17, p 375) the coefficient is not significantly different from -10 at the 5% level.

expected total returns are not constant across SMAs, but systematically vary in the cross-section-- perhaps due to variation in risks or even risk tolerances, and this variation is positively correlated with the beginning of the decade rent-to-value ratio, then the coefficient estimate on the ratio would suffer from a positive bias.

The final possibility stems from our use of gross rents. Ideally, net rents, which include only cash flows to the owner, rather than gross rents, which also include utility costs, taxes and other costs, should be used. However, such data are available from the census only for 1980. Thus, our rent measures are too large, and to the extent that expenses are proportional to gross rents, the coefficient is biased towards zero. Using only the data for 1980 where the tax and utility cost data are available we reestimated our model but found only slight movement in the coefficient towards -10.

Efficiency

As a preliminary test of efficiency, we test for the information content of the residuals from the first stage regression reported in Table 1. These residuals are constructed as the differences between the actual rent/price ratios and the expected rent/price ratio conditional on many factors chosen to capture cross-sectional differences in quality between the two types of housing. Therefore, we interpret the residuals as deviations of rent/price ratios from their theoretic or equilibrium values. When these residuals, denoted as $\hat{e}_{t-1}$ are included in our main specification, estimation via OLS using the same 192 data points yields[10]:

---

[10]The residuals were created in an OLS regression, so they should be orthogonal to each of the independent variables in the first stage regression and to the fitted values from that regression. However, that first stage regression was estimated on levels using all 256 data points, while this regression uses

$$\Delta \ln(\text{VALUE}) = \underset{(3.6)}{0.46\ Y_{60s}} + \underset{(5.1)}{0.76\ Y_{70s}} + \underset{(3.8)}{0.48\ Y_{80s}} + \underset{(2.4)}{3.88\ \hat{e}_{t-1}} - \underset{(-3.2)}{5.81\ \hat{e}_{t-1}}$$

The results indicate that the residuals have significant predictive power for subsequent capital appreciation.

The implications of this finding for statements about market efficiency depend crucially on the determinants of the deviations themselves. For example, if some important, but perhaps unobservable, variables were omitted from our first stage specification, like the total or systematic risk of housing or risk tolerances, then the residuals would be correlated with these omitted variables. If one or more of these omitted variables were also correlated with cross-sectional variations in equilibrium required rates of return, then the significance of the residuals in the above specification is merely capturing cross-sectional differences in equilibrium total required returns to housing. Under this scenario, the significance of the residuals does not imply any inefficiency.

Other interpretations lead to conclusions of inefficiencies. As mentioned above, all assets trade within a band determined by transaction and information costs. Because these costs are notoriously large in real estate markets, these bands are expected to be wide. A negative first stage regression residual may indicate that house values are close to the upper bound defined by transactions costs. In the next decade, prices will appear, on average, lower (towards the center of the transactions band). Thus a negative residual at the beginning of

---

only those 192 data points for which first differences of value are available. For this subset of data, the residuals are not perfectly orthogonal to the fitted values, so their inclusion in the specification slightly alters the estimates of the coeffients for the other variables.

the decade should be related to a lower level of observed capital appreciation over the decade. Under this scenario, the residual is inversely related to a transitory or disequilibrium component of house prices, and the predictive ability of the residual is due to the reversion of housing prices back to their equilibrium value.

An alternative hypothesis is based on the notion of a rental market and an owner occupied market with relative prices that wander from their true equilibrium values. For example, a positive residual may reflect quality-adjusted rental rates that are "too high." The rational strategy would be for agents to substitute out of rental units and into owner occupied housing, and, in so doing, drive the relative prices of the two forms of housing back towards their long-run equilibrium values. In this case, substitution into housing would increase the value of housing. Of course, such moves are expensive, so the predicted increase in housing values would be far from instantaneous. Thus a positive residual observed today does not predict an instantaneous jump in house values, but higher subsequent appreciation rates.

Euphoria

In this section we use decadal data on the growth of income and population to test for cross-sectional variation in expected total required returns. Both income growth and population growth are highly autocorrelated (0.78 and 0.52 respectively), so there is much predictability in these time series and much dispersion in cross-sectional forecasted values. In table 2 we add lagged population growth and lagged income growth to our specification.

The estimated coefficient associated with the first-step residual is virtually identical to its earlier value, while the fitted rent/price ratio as of the beginning of the decade moves much closer to (and is insignificantly different from) its

theoretic value of -10. Of considerable interest, however, is the result that lagged income growth enters the model significantly. This is strong evidence against the hypothesis that required total returns to residential real estate are constant in the cross-section and that observed returns vary randomly around an aggregate average. Instead, total returns vary systematically with recently experienced income growth.

A number of possible conclusions can be drawn from this result. First, it might be argued that systematic cross-sectional variations in expected total returns are rational, in that they reflect cross-sectional differences in the risk characteristics of the underlying real estate assets. This argument is valid only under the unlikely scenario that cross-sectional differences in *lagged* income growth rates are somehow correlated with cross-sectional exposures to *contemporaneous* risk factors and thus, risk premia. A second argument is that the growth rates affect the location of observed prices within the transaction band. Thus, the inclusion of these variables merely reduces the errors-in-variables problem. We agree that this would be a valid conclusion if contemporaneous growth rates were employed, but lagged rates are used.

Thus, we conclude that agents have not correctly processed the information content of the lagged values of these variables into the rent-to-value ratios. More specifically, the sign indicates that owners are euphoric in that they appear to overreact to income changes. This finding is consistent with the theoretical conclusions of matching models, for example Wheaton (1990), and with empirical evidence from Abraham and Hendershott (1992) who use annual data on repeat sales and find a lag structure on income that produces cyclical movements in house prices. Our results suggest that the cycle from annual data is also reflected in decadal data.

Conclusion

In this study, we analyze single family house price appreciation for 64 metro areas over four census periods. Data on house prices, rents, household income, population, housing characteristics, property taxes (1980 only) and utility costs (1980 only) were collected from the decennial censuses for 1960 to 1990. We use these data to test whether equilibrium total returns (the service flow or rent from the property plus price appreciation) are approximately equal across markets. Specifically, we test whether areas with low rent/price ratios have higher expected appreciation rates.

However, we argue that observed rent/price ratios are comprised of two components: an equilibrium component that varies across SMAs due to cross-sectional differences in the relative quality of rental and owner occupied housing, and a disequilibrium component. We identify these two components in a first stage regression and then use this decomposition in estimating the relation between rent/price ratios and subsequent rates of appreciation.

The quality-adjusted or equilibrium component of the rent/price ratio is a valuable predictor of subsequent house price movements. Further, consistent with rational expectations, each percent increase in the annual rental flow is offset by a reduction in annual appreciation by 0.6- 0.85%. Because the equilibrium component of the rent/price ratio ranges from 4% to 8%, cities with the lowest rent/price ratios can be expected to appreciate 27-40% more per decade than those with the highest rent/price ratios.

The disequilibrium component of the rent/price ratio also has significant predictive power. When rent/price ratios are higher than predicted by our first stage regression, subsequent appreciation rates are typically above average. Though we cannot dismiss the possibility that the disequilibrium component is

somehow proxying for cross-sectional differences in risk, we believe that the disequilibrium component reflects non-permanent deviations in the prices of housing and/or rental properties.

We next test whether expectations incorporate all available information by adding lagged population and lagged income growth to the specification. Lagged income growth enters significantly with a negative sign, indicating that, while investors and homeowners appear to process past experience into an expected appreciation rate, there appears to be systematic bias to expectations. Specifically, the negative sign associated with lagged income growth suggests possible euphoria: when income growth has been high, owners set rent/price ratios as if they systematically overestimate subsequent appreciation rates.

Like most studies of the housing market our results are not fully consistent with asset market efficiency. However, because information and transactions costs are unusually large for this asset class, this conclusion may not be surprising. Further, given the capital constraints of residential real estate owners and short sale restrictions, the forces of arbitrage usually in place to eliminate inefficiencies are highly impaired.

References

Abraham, J. and P. Hendershott, "Patterns and Determinants of Metropolitan House Prices, 1977-91," mimeo, Ohio State University, 1992.

Black, F. and M. Scholes, "The Effect of Dividend Yield and Dividend Policy on Common Stock Prices and Returns", *Journal of Financial Economics*, 1974, 1, 1-22.

Capozza, D.R. and R. Helsley, "The Stochastic City", *Journal of Urban Economics*, 1990, 28, 187-203.

Capozza, D.R. and Y. Li, "The Intensty and Timing of Investment: The Case of Land", *American Economic Review*, 1994, 84, 4, 889-904.

Capozza D. R. and G. Schwann, "The Asset Approach to Pricing Urban Land: Empirical Evidence", *Journal of the American Real Estate and Urban Economics Association*, 1989, 17, 2, 161-174.

Capozza D. R. and G. Schwann, "The Value of Risk in Real Estate Markets", *Journal of Real Estate Finance and Economics*, 1990, 3, 117-140.

Capozza, D. R. and G. Sick, "The Risk Structure of Land Markets", *Journal of Urban Economics*, 1994, 35, 297-319.

Case, K. and R. Shiller, "The Behavior of Home Buyers in Boom and Post-Boom Markets," *New England Economic Review*, 1988, 29-46.

Collins, S., B. Lipman, and S. Groeneman, "Experience and Expectation: The Home Price Predictions of Owners and Renters," National Association of Realtors, 1992.

Fama, E. F. and K. R. French, "Dividend Yields and Expected Stock Returns", *J. of Financial Economics*, 1988, 22, 3-25.

Goetzmann, W. N. and P. Jorion, "Testing the Predictive Power of Dividend Yields", *Journal of Finance*, 1993, 48, 663-679.

Hamilton, B. W. and R. M. Schwab, "Expected Appreciation in Urban Housing Markets," *Journal of Urban Economics,* 1985, 18, 103-118.

Hendershott, P. and T. Thibodeau, "The Relationship between Median and Constant Quality House Prices," *J. of the American Real Estate and Urban Economics Association,* 1990, 18, 323-334.

Hodrick, R. J., "Dividend Yields and Expected Stock Retruns: Alternative Procedures for Inference and Measurement", *Review of Financial Studies*, 1992, 5, 243-280.

Mankiw G. and A. Weil, "The Baby Boom, the Baby Bust, and the Housing Market", *Regional Science and Urban Economics,* 1989, 19, 235-258.

Meese, Richard and Nancy Wallace, "Testing the Present Value Relation for Housing Prices: Should I Leave my House in San Francisco," *Journal of Urban Economics,* 1994, 35, 245-266.

Miller, M. and F. Modigliani, "Dividend Policy, Growth, and the Valuation of Shares", *Journal of Business*, 1961, 34, 411-433.

Miller, M. and M. Scholes, "Dividend and Taxes: Some Empirical Evidence", *Journal of Political Economy*, 1982, 90, 1118-1141.

Murphy, Kevin M., and Robet H. Topel, 1985, Estimation and inference in two-step econometric models, *Journal of Business and Economics Statistics* 3, 370-379.

Ozanne, L. and T. Thibodeau, "Explaining Metropolitan Housing Price Differences," *Journal of Urban Economics*, 1983, 13, 51-66.

Pagan, A. "Econometric issues in the analysis of regressions with generated regressors." *International Economic Review*, 25 (Feb. 1984), 221-247.

Rose, L. "Topographical Constraints and Urban Land Supply Indexes," *J. of Urban Economics*, 1989, 26, 335-347.

Schwert, G. William, and Paul J. Seguin, 1990, Heteroskedasticity in stock returns, *Journal of Finance* 44, 1129-1155.

Wheaton, W., 1990, "Vacancy, Search and Prices in a Housing Market Matching Model," *J. of Political Economy*, 98, 1270-92.

Table 1

Instrumental Variable Regression for Rent-to-Price Ratio

For each of the 64 SMA's and each of four usable census years, we regress the ratio of rental income to housing price levels on the set of explanatory variables suggested by Capozza and Helsley (1990) and Capozza and Sick (1994) in their investigation of urban growth with uncertainty. These variables include income, population, population growth rate, property tax rate, utility cost rate, number of baths, developable area in the city, and construction cost. Alternate formulations including building age and / or number of rooms yield equivalent results. Detailed descriptions of the source and/or computation of these variables appear in the data appendix. The notation (x100) means the estimated coefficient has been scaled up by a factor of 100. There are 256 usable observations. The R^2= 0.52 and the regression F-statistic is 24.3 which is significant at any standard significance level.

Explanatory Variable	Coefficient Estimate	T-statistic
Y_{60}: *Intercept for 1960 Census Observations*	0.22	4.1***
Y_{70}: *Intercept for 1970 Census Observations*	0.24	4.3***
Y_{80}: *Intercept for 1980 Census Observations*	0.22	3.9***
Y_{90}: *Intercept for 1990 Census Observations*	0.22	3.9***
Δln(Real Income) (x100)	-1.79	-2.0*
ln(Real Income) (x100)	-1.30	-2.2**
ln(Population) (x100)	-0.24	-2.9***
Δln(Population) (x100)	0.27	0.7
Tax Rate: *Median value fo property taxes (Monthly)*	0.87	6.5***
Utility Rate: *Median heating, water, gas, electricity*	-0.53	-2.2**
Median number of baths (x100)	-0.32	-1.9*
Conditional Land Supply Index (x100)	1.42	3.1***
ln(Local construction cost index/CPI level)(x100)		

Table 2

Second Step Regression using Estimated Rent-to-Price Ratio

For each of the 64 usable SMA's and each of three observation intervals (1960 to 1970, 1970 to 1980, and 1980 to 1990), we regress the appreciation in housing price value over the interval (measured as the first difference in the log of the value of the median house price) on a set of explanatory variables including lagged income growth, lagged population growth, separate intercepts for each of the three final years over which the real decadal appreciation was calculated, and fitted values and residuals from a regression with beginning-of-period rent-to-value ratio as the dependent variable. Detailed descriptions of the source and/or computation of these variables appear in the data appendix. There are 192 usable observations. The R^2= .33 and the regression F-statistic is 16.7 which is significant at any standard significance level.

Explanitory Variable	Coefficient Estimate	T-statistic
Y_{60s}: *Intercept for 1960-1970 Observations*	0.96	5.3***
Y_{70s}: *Intercept for 1970-1980 Observations*	1.17	6.5***
Y_{80s}: *Intercept for 1980-1990 Observations*	0.64	4.9***
Fitted Rent-to-price Ratio$_{t-1}$	-8.46	-4.4***
Δln(Population)$_{t-1}$	0.07	1.0
Δln(Income/CPI deflator)$_{t-1}$	-0.79	-4.1***
$\hat{e}_{t-1}$: Residual from First Step Regression (Table 1)	-3.84	2.5**

Appendix A
DESCRIPTION OF DATA AND COMPUTATIONS

CITIES STUDIED: The following SMSAs are included in the sample.

1 Akron
2 Albany NY
3 Albuquerque
4 Anaheim-Santa Ana
5 Atlanta
6 Baltimore
7 Birmingham
8 Boston (includes Brockton MA)
9 Buffalo-Niagara Falls
10 Charleston SC
11 Charlotte
12 Chattanooga
13 Chicago (not including Gary-Hammond IN or Kenosha WI)
14 Cincinnati
15 Cleveland
16 Columbus OH
17 Dallas-Fort Worth @
18 Denver
19 Des Moines
20 Detroit
21 El Paso
22 Fort Lauderdale
23 Grand Rapids
24 Hartford-New Britain @
22 Honolulu
26 Houston
27 Indianapolis
28 Jacksonville FL
29 Kansas City
30 Knoxville
31 Lansing MI
32 Las Vegas
33 Los Angeles
34 Louisville
35 Memphis
36 Miami
37 Milwaukee
38 Minneapolis-St. Paul
39 Nashville
40 New Orleans
41 New York-Northern New Jersey
42 Oklahoma City
43 Omaha
44 Orlando
45 Philadelphia
46 Phoenix
47 Pittsburgh
48 Portland OR
49 Providence RI
50 Riverside-San Bernardino
51 Rochester NY
52 Saint Louis
53 Salt Lake City-Ogden @
54 San Antonio
55 San Diego
56 San Francisco-Oakland
57 San Jose
58 Seattle-Tacoma @
59 Syracuse
60 Tampa-St. Petersburg
61 Toledo
62 Tulsa
63 Washington DC
64 West Palm Beach FL

@ = SMSAs formed in 1980 from 2 prior distinct SMSAs. Pre-1980 data are a weighted average of data for each component SMSA.

New York-Northern New Jersey includes the following SMSAs: New York (which includes Westchester and Rockland counties), Nassau-Suffolk, Jersey City, Newark, and Paterson-Passaic. In some cases, the value for 41 is a weighted average of reported values for each of component SMSAs.

SOURCES

The following abbreviations are used:

CCDB	City and County Data Book, a Supplement to the Statistical Abstract.
CHMHC	Census of Housing, Vol. II, Metropolitan Housing Characteristics.
CHUS	Census of Housing, Vol. II, General Characteristics. Part I, US Summary.
ERP	Economic Report of the President to the Congress.
PUMS	Public Use Microdata Sample. Computer file of household micro data, consisting of a 0.1% stratified random sample from the 1970 and 1980 Censuses of Housing and Population, and a 1% stratified random sample from the 1940 and 1950 Censuses of Population and Housing. The 1960 PUMS contains no SMSA identifier and so is unusable.
REIS	Machine readable (DOS) files included in the Regional Economic Information System, prepared and distributed by the Bureau of Economic Analysis, US Department of Commerce.
SA	Statistical Abstract of the United States.

DATA

Data on the following variables are used in this study. Unless otherwise indicated, all data are for metro areas and are taken from the decennial Census of Housing and Population. Owner-occupied dwellings are defined by the Census Bureau as being sited on 10 acres or less of land, and consist largely of 1 unit structures, as condominiums were rare before the 1980 Census. All dollar amounts are not adjusted for inflation unless otherwise indicated.

Baths	Median number of baths in owner-occupied dwellings. Available only on the 1980 PUMS.
CLSI	Conditional Land Supply Index. Ranges from .5 (for a city that completely occupies an island) to 1.0 (for a city on a featureless plain). Source: for 38 cities, values are given in Table 2 of

	Table 2 of Rose [1989]. For remaining 26 cities, values are informal subjective estimates by Capozza and Meguire.
CPI	Consumer Price Index, all urban residents, average of monthly values. Source: 1940-60: Ibbotson Associates, Stocks, Bonds, Bills and Inflation, Table B-10, last column. 1960-90: 1992 ERP
Gross_Rent	Median of the sum of monthly contract rent on rented unfurnished dwellings, plus estimated monthly cost of utilities. Data for rented 1 unit detached houses were reported only for the 1970 Census. Data for other years is for all rented dwellings. 1940: CHUS, Table 104, last column. 1950: 1956 CCDB, Table 2 (Anaheim, Fort Lauderdale, Las Vegas, Orlando and West Palm Beach) or Table 3 (all other), col. 15. 1960: 1967 CCDB, Table 3. 1960-80: CHMHC, Table A-2. 1990: CD-ROM put out by Census Bureau.
HCCI	Historical Construction Cost Index, all US, 1/1/75 = 1.0. Measured as of July 1 of each year. Source: R.S. Means & Co., Building Construction Cost Data, 1983, 1992 editions.
NCON	Local construction cost, expressed in terms of US average=100. Source: 1980-90: R. S. Means & Co., Building Construction Cost Data, quarterly publication.
Pers Income	1970-90: = GNP+Transfer Payments+Interest on Government Debt-Retained Earnings of Corporations. In per capita terms. Source: REIS, File CA-25, series 030. 1940-50: Same as 1970-90, except computed for principal state surrounding SMSA. Source: REIS, Table SA-52.
Population	Total population by SMSA. 1970-90: REIS, File CA5, series 020. 1960: SA, 1983-4, Table 20, col. 1.
Rooms	Median number of rooms in owner-occupied dwellings. Source: CHMHC, Table A-6.
Tax&Insur	Median value of the sum of property taxes and property/casualty insurance owed on owner-occupied dwellings, at monthly rates. Only 1980 data are available. Available only on the 1980 PUMS.
Utility_Cost	Median of the sum of the monthly cost of heating, water, gas, and + electricity. For owner-occupied dwellings only. Source: computed from estimated monthly cost of each of component appearing on the 1980 PUMS.

Value	Value of owner-occupied dwellings, estimated by owner-occupant. Source: 1960: 1967 CCDB, Table 3, col. nn. 1970: 1972 CCDB, Table n, col. nn. 1980: 1987 CCDB Table n, col. nn. 1990: CD-ROM put out by Census Bureau.
Yearbuilt	Median coded value for year in which owner-occupied dwellings were built. Available only from the 1970 and 1980 PUMS. Coding is as follows: 1=1979-80, 2=1975-78, 3=1970-74, 4=1960-69, 5=1950-59, 6=1940-49, 7=1939 or earlier

COMPUTED VARIABLES

The following variables are computed from the above data as follows:

Age:	if Yearbuilt=1, =1; if 1<Yearbuilt<7, = Census year - midpoint of YearBuilt; if Yearbuilt=7, = Census year - 1925.
Rent	= 12*Gross_Rent/Value
ln(Value)	= LOG(VALUE/CPI)
Δln(Value)	= ln(Value)(t) - ln(Value)(t-10)
ln(Income)	= LOG(FAMERN/CPI)
Δln(Income)	= ln(Income)(t) - ln(Income)(t-10)
Δln(pop)	= LOG(Population(t)/Population(t-10))
Taxrate	= (Taxes & Insurance)/Value
Utilityrate	= Utility_Cost/Value

Pre-1980 values of TAXRATE, UTILITYRATE and NCON are set to their 1980 values.