IS THE SAN FRANCISCO HOUSING MARKET EFFICIENT?

May 10, 2006

Vineet Bhagwat

Department of Economics
Stanford University
Stanford, CA 94305
vbhagwat@stanford.edu

under the direction of:
Professor Jonathan Levin

ABSTRACT

With prices skyrocketing for housing, especially in the San Francisco Bay Area, there is much talk about a housing bubble. However, the rapid price increases could also be due to changes in fundamentals: implicit rents, interest rates, and property and income tax rates.

I use sales and implicit rental data from 1983 to 2005 for the City of San Francisco that is available through national accounting databases to examine the relationship between housing prices and fundamental factors. I find that the housing market is marked with inefficiencies in the short-run but is ultimately efficient in long-run. The short-run inefficiencies are most likely due to the high prevalence of transaction costs in the market.

Keywords: housing market efficiency, housing bubble, present value relation

1 I thank Jonathan Levin for his invaluable input and guidance over the past year, without his help this thesis would not have been possible. I also thank Geoffrey Rothwell for laying the foundations of thesis writing. Any errors are my own.
Introduction

The popular press is rife with speculation of a bubble in the housing market. After all, the median price of homes in San Francisco has more than doubled from first quarter of 1998 to the second quarter of 2005 (U.S. Bureau of Labor Statistics). During a housing bubble, prices rise as buyers pay high prices because they expect a further rapid increase in prices – more rapid than can be sustained over a long time period. Eventually, the market cools, leaving some buyers with homes for which they drastically overpaid. Understanding the determinants of a bubble would have tremendous policy implications not only in the housing market but also in the stock market, which is also bubble-prone.

The goal of this paper is to assess whether or not a bubble exists in the housing market by examining housing prices in city of San Francisco in relation to rental prices and other cost of capital indices. In order to assess the existence of a bubble, the definition of a bubble must be clear. In its most widespread use, a “bubble” is defined as “a situation in which excessive public expectations of future price increases cause prices to be temporarily elevated,” (Case and Shiller, 2003). The value of a home can increase over time and expectations of future price increases do play a role in the decision to buy a home.

However, the observation of a rapid price increase does not imply a bubble. If there are changes in fundamental factors, such as a long-term shift in supply and demand, there may be a substantial price increase without a housing bubble. For example, if suddenly for some reason all current homes were extended to include an extra bedroom and bathroom, the prices of homes would surely increase. However the price increase stems from a fundamental quality increase that raises buyer willingness to pay and it
would not be characterized as a bubble. So one must be careful in controlling for fundamental factors when assessing the existence of a bubble.

In this paper I control for fundamental changes in supply and demand, such as a change in the quality of the housing stock, by relating house purchase price to rental prices, or implicit rents. Implicit rents are defined as the potential rents were a homeowner to rent the home to a tenant. Implicit rents are seen as a return to the homeowner from owning the home, much like a dividend to a stockholder. The idea is that even if housing quality or demand for housing changes, housing prices and implicit rents should still share a common trend. If the rent-to-price ratio is low, this suggests that the return on owning the home is low relative to other assets. This relatively low return is unlikely to persist and in order for it to reach the level of competing assets, housing prices must fall, (McCarthy and Peach, 2004).

A comparison between rental rates and home purchase price is complicated by the fact that most people borrow money to purchase a house. Thus, the cost of owning a home must account for the cost of capital – in particular interest rates. Interest rates “influence home ownership affordability and represent the yield on a competing asset in a household’s portfolio,” (McCarthy and Peach, 2004). The fall in interest rates and mortgage rates during the 1990s significantly improved the affordability of homes and boosted demand for home ownership. Thus, one must control for interest rates since the rise in home prices could be due to the increased demand stemming from a drop in price people pay to borrow money. In addition, many people buy homes as an investment that will provide an income tax shelter. Although the recent tax breaks have reduced the amount of this shelter, it is still a factor in the demand for homes and so should also be
included in the analysis of the housing market. On the flip side, a homeowner will face property taxes on the home, which should work to decrease demand as the property tax rate rise, and vice versa.

There are many other fundamental factors, on the supply side. Many economists believe that in the real estate market, especially in urban areas, the supply of housing is relatively inelastic (Case and Schiller, 1989). The inelastic nature of supply implies that any increase in demand for housing, whether due to irrational expectations or not, cannot be met with an increase in supply and so prices must increase. Since this price increase could be due to the fact that supply is inelastic it cannot be attributed to a bubble. A study that properly controls for relative elasticity in various local markets around the country would be able to assess the existence of a bubble even in an inelastic supply situation.

Like McCarthy and Peach (2004), I present a descriptive analysis of the San Francisco housing market by looking at the rent-to-price ratio and adjusting this ratio to incorporate fundamental factors listed above. I augment my descriptive analysis by borrowing from the methodology of Meese and Wallace (1994). I construct a present value model that equates the real return to owning a home with the cost of owning a home, or the cost of capital. I then use data for the city of San Francisco to run short-run and long-run tests housing market efficiency. The present value model incorporates interest rates, property taxes, income taxes, and depreciation into the cost of capital series. I find that the present value model does not hold in the short-run, but the short-run inefficiencies are arbitraged away in the long-run. The short-run inefficiencies are most likely due to the high prevalence of transaction costs in the market.
This paper is organized as follows. In the next section I will discuss the existing literature on the topic of housing bubbles, after which I will outline my study. I then provide data on the present value tests in the housing market and a discussion of the results.

**Literature Review**

My analysis builds on a variety of methodological approaches developed by earlier authors. Here I discuss the findings and impact of four research papers in this area that all took slightly different approaches to the problem of identifying a bubble in the housing market.

I start with a recent research paper written by Jonathan McCarthy and Richard W. Peach, of the Federal Reserve Bank in New York. McCarthy and Peach (2004) examine

![Figure 1](image_url)

**Figure 1**

*Ratio of Median Price of Existing Homes Sold to Median Family Income*


Notes: The ratio is calculated by dividing the median sales price of an existing home by median family income. Median family income for 2002 is based on an estimate from the National Association of Realtors.

*Source: McCarthy and Peach (2004)*
national data from 1985 to 2002 to see if a housing bubble exists in the nation as a whole. While many of their methodological tools, such as rent-to-price ratio analysis, are useful, a national approach will gloss over the local variation in metropolitan areas that is key to the formation of a bubble. Furthermore, they conduct a completely descriptive analysis while a more concrete, quantitative method might prove useful in assessing the existence of a bubble. Nevertheless, it is useful to see how they obtain their results since I incorporate many of their methods in my analysis.

They obtain housing sales data from the OFHEO housing price index and the rental data from the shelter rent component of the CPI, sources that will be discussed below. They start with two pieces of evidence to indicate the presence of a national housing bubble: price-to-income ratio and rent-to-price ratio. A high price-to-income ratio is an indicator that home price affordability is eroding, that potential homebuyers are finding it more difficult to meet down payments and monthly mortgage payments. This, in turn, should reduce housing demand, resulting in downward pressure on prices in the market. From Figure 1, one can see that “the median home price…is now about three times median household income, surpassing peaks in the early 1980s, when there was arguably a bubble in the housing market,” (McCarthy and Peach, 2004). Also observe that real price appreciation decreased significantly over the few years following the previous peak in the price-to-income ratio.
The second piece of evidence McCarthy and Peach (2004) present is the rent-to-price ratio. If one views rents as the monthly return on owning a home, the rent-to-price ratio should be stable over time, as any deviations should be met by changes in demand. That is, if the rent-to-price ratio were low, this would indicate that the return on owning a home is low relative to other assets, which would yield the prediction that home prices should decline over the next few years. Indeed, years of low rent-to-price ratios are followed by years of significant price declines. From 2000 to 2003, the rent-to-price ratio is below its historical average, raising fears of a housing bubble that is about to burst.

**Figure 2**  
**Ratio of Owners’ Equivalent Rent to OFHEO Index**

Sources: Office of Federal Housing Enterprise Oversight (OFHEO); U.S. Department of Labor, Bureau of Labor Statistics.

Note: Shaded areas indicate periods designated national recessions by the National Bureau of Economic Research.  

Despite the above evidence, McCarthy and Peach (2004) argue that there is in fact no national housing bubble. Neither of the above measures account for interest rates. Surely, interest rates affect the cost of borrowing, which in turn affect affordability and demand for housing. To incorporate interest rates into the home price-to-income ratio,
the authors look at the maximum loan amount for families at the median income and the ratio of principal and interest payments to median family income. In both instances, the evidence overturns the initial evidence of a bubble from the simple price-to-income ratio. From 1990-2002 there has been a steady increase in the maximum loan available to a family at the median income level. In addition, the ratio of principal and interest payments to median family income has been falling steadily since 1980. Both pieces of evidence point to increasing affordability of homes. Despite rising prices in relation to median income, affordability is at an all time high due to rising loan amounts and relatively decreasing monthly payments on those loans (McCarthy and Peach, 2004).

Furthermore, the authors incorporate cost of capital measures into the rent-to-price ratio. The rent-to-price ratio is adjusted by “subtracting the interest rate and property tax rate, both on an after-tax basis, and the depreciation rate,” (McCarthy and Peach, 2004). The adjusted rent-to-price ratio shows no evidence of a bubble. In fact, since 2000, the adjusted ratio has been increasing steadily and is above its historical average. Thus, the authors find that while the initial data suggests that there was rapid price increases, the incorporation of the interest rate into the pricing models negates the existence of a bubble.

McCarthy and Peach (2004) focus on the national housing market. The study, however, does not provide much insight into the characteristics of local housing markets, which is presumably a higher concern to individual homeowners than the national market. National data inevitably has the effecting of smoothing local volatilities, and so in order to gain insight into local markets one must use local data. In addition, the paper provides descriptive data on the housing market, but not much in the form of regressions
and concrete hypothesis testing. So in order to get a more concrete picture of the housing market, I turn my attention to Abraham and Hendershott (1996), a paper that analyzes regional housing data to build a model of the determinants of the equilibrium price level and real housing price appreciation.

Abraham and Hendershott (1996) start with the premise that there is an equilibrium price level to which the market is constantly adjusting. This leads to two groups of fundamental factors. One group explains the changes in the equilibrium price and the other explains the adjustment dynamics and deviations from the equilibrium price. They include growth in real income, real construction costs, and changes in the real after-tax interest rate in the group that accounts for changes in the equilibrium price. Lagged real appreciation and the difference between the actual and equilibrium price levels are in the second group. Thus, forecasts of real house price appreciation depend on two factors: “forecasts of changes in fundamentals and forecasts of initial gaps between actual and equilibrium prices.”

A major component of the article is incorporating a proxy for the tendency for bubbles to build as well as burst. Their proxy is “the deviation between the actual metropolitan house price level and a “fundamental” price level,” (193) where the fundamental price level is determined based on their empirical estimates. Their model then becomes a multiple-stage estimation, where in the first stage they estimate the equilibrium price based on fundamentals and in the second stage estimate the bubble through deviations from the equilibrium price.

Abraham and Hendershott (1996) apply this analysis to data from 30 metropolitan statistical areas provided in the Freddie Mac-Fannie Mae repeat transaction database.
Using data from thirty metropolitan areas, the authors aim to capture the variation in the housing markets across the country and make comparisons across markets. For example, not only will they use their model to estimate differences in price appreciation across the usual regions of the country (Northeast, Midwest, Southeast, West) but they will also try to model differences in coastal and inland cities. The reason for splitting the country into coastal and inland cities is that coastal cities tend to be more supply constrained by limited land availability and local building requirements than inland cities.

They find that their model “can explain roughly three-quarters of real price swings in the West and about half the changes in the Northeast,” (205). They also find that there was a “35 percent above-market premium in Northeast prices, a 15 to 20 percent premium in West Coast prices, and probably significant under-pricing in Texas,” (205). While coastal and inland respond similarly to real income growth and to changes in real after-tax interest rates and local price deviation, they differ in their responses to construction costs and “disequilibrium variables” such as the lagged appreciation rate and deviation from equilibrium price. Inland cities were three times more responsive in price to construction costs and about two times less responsive to the lagged real appreciation rate than coastal cities. The evidence of price bubbles is much stronger for the coastal cities, and confirms the suspicion that coastal cities exhibit higher transaction costs in terms of relative supply inelasticity and stricter building requirements (Abraham and Hendershott, 1996).

Abraham and Hendershott (1996) provides a good framework for a model that describes an equilibrium housing price level by modeling changes in the equilibrium price and adjustment dynamics or changing deviations from the equilibrium price. While
a good methodological approach, I prefer one that keeps the investor at the heart of the analysis. Bubbles are driven by investor psychology and expectations. Investors expect further increase in prices and therefore are willing to pay exorbitant prices for a home that a year or two ago would have fetched half the price. Thus, a methodology that focuses on the investor will, in my opinion, yield more fruitful results as to the source of a housing bubble, if it exists. Meese and Wallace (1994) and Case and Schiller (2003), discussed below, take a more investor-driven approach.

Furthermore, a drawback to Abraham and Hendershott (1996) that is similar to one from McCarthy and Peach (2004) is the use of aggregate data at the region level. The regions used are those defined in the census estimates – namely the Northeast, West, Southeast, and Midwest regions of the United States. While to a lesser extent than national aggregation, data at the region level still does not fully account for local housing market swings. It is important to note, however, that the authors also do separate analysis of coastal and inland regions to account for differences in relative supply inelasticity. Another drawback to their model is that the price dynamic terms are capturing a market expectations mechanism in a blunt way. The model does not explain how these expectations are formed.

I now examine two papers that address these concerns. Case and Schiller (2003) use survey data to provide insight into the market expectations mechanism that was endogenous to the model in Abraham and Hendershott (1996). Meese and Wallace (1994) is the first study I examine that uses local metropolitan level data to assess housing market inefficiencies.
One critique of Abraham and Hendershott (1996) is that the model they present does not explain how the market expectations mechanism operates. Since even the slightest rumor can change expectations immediately and throw housing prices in a precipitous decline, developing a metric for market expectations is quite important. Karl Case and Robert Schiller write a very thorough and extensive Brookings Paper that addresses this very topic. In addition to conducting the standard econometric analysis of the housing market, the authors present the results of a questionnaire conducted in 2003 of people who bought homes in the previous year in four major cities: Los Angeles, San Francisco, Boston, and Milwaukee. The purpose of this questionnaire was to learn more about the thought process of homebuyers, to gain insight into the reasons they used to justify buying a home. The fact that the respondents to the survey reside in four different metropolitan areas allows the authors to compare responses from metropolitan areas that have reputedly gone through a bubble (San Francisco, Los Angeles, and Boston) with one that has not (Milwaukee). Some of the questions from the survey are listed below:

1. *In deciding to buy your property, did you think of the purchase as an investment?*
2. *Why did you buy the home that you did?*
3. *How much of a change do you expect there to be in the value of your home over the next 12 months?*
4. *On average over the next 10 years, how much do you expect the value of your property to change each year?*
5. *It’s a good time to buy because housing prices are likely to rise in the future.*

*Agree or Disagree?*
The results from these questions are quite striking, especially the responses to questions 3 and 4. On average, homeowners expected prices to increase by about 8%, which could be reasonable during a housing boom. However, the responses to question 4 reveal that homeowners expect a 12% increase in home prices per year over the next 10 years. This would imply a more than tripling of housing prices over the next 10 years, which one can assume is highly unlikely. This result is enhanced by the fact that for the vast majority of buyers, investment was a “major” or “partial” factor in their decision to buy a home. More importantly, there is a significant difference in the responses between those that reside in the “bubble cities” (Los Angeles, San Francisco, and Boston) in comparison to those that live in the “non-bubble city”, Milwaukee. The “bubble cities” displayed inflated expectations of future prices as compared to Milwaukee, and a higher percentage of respondents in the “bubble cities” bought into the theory that prices will continue to rise because “people want to live here,” (324).

While the survey does provide good insight into investor psychology, one must be cautious in interpreting the results. One problem with using surveys is that the wording of the question can yield starkly different responses. John Quigley, in a discussion section at the end of paper, wonders if changing the wording of a question on expected home prices increases could have yielded different results: “Might the responses have been different if an equivalent question had been asked,” (356). Despite the inevitable wording bias, the survey still provides a glimpse into the factors that affect market
expectations, an important element to understand if one truly wants to test for market inefficiencies or a bubble.

I now turn my attention to the second critique of Abraham and Hendershott (1996) – the use of aggregated regional data does not provide much insight into the volatilities of a local market. Meese and Wallace (1994) do not incorporate a direct test for the existence of a housing bubble, but rather assess it indirectly by testing the present value relation for housing prices in the San Francisco Bay Area. The key feature of this study is that instead of using datasets published by government agencies like the previous two studies, they self-construct their own home price and rental indices. They accomplish this for the price index by running a hedonic regression of house price on set of house attributes, like by square footage, number of bedrooms, bathrooms, etc. The rental index is constructed by extracting data from classifieds in the Oakland Tribune and the San Francisco Chronicle on the rental prices of two bedroom apartments. A third data series employed is the cost of capital series, which incorporates user costs of owning a home – namely, the mortgage rate, income and property tax rate, and depreciation. Using these three data series, the authors construct a present value relation that should hold in an efficient housing market and conduct a short-run and long-run test to verify this relation.

The present value relation is setup such that the price of a home today equals the sum of rental prices and expected appreciation in the sale divided by the cost of capital. An efficient housing market in the short-run will exhibit no serial correlation in returns, defined as the percentage difference between the sum of future prices and rents and current prices. However, their short-run test reveals that there is indeed serial correlation
in returns. In fact, there is even serial correlation in the deviations from the short-run present value relation. Thus, the authors find that the housing market is inefficient in the short-run. The housing market can still be efficient in the long-run if these forecast errors are stationary, that is, if any of the deviations from the present value relation are not permanent. The long-run test does in fact show that these deviations are stationary and that the housing market is efficient in the long-run. The authors explain the short-run violation and long-run verification of housing market efficiency with the argument that “the housing market does not conform with the usual assumptions of efficient markets theory,” (246). The housing market is characterized by many transaction costs – from real estate agent commissions to zoning restrictions – and information asymmetries that lead to friction in the short-run market. However, an efficient market will do away with these frictions in the long-run, and indeed this is what they find. Meese and Wallace (1994) prefer to explain the results of their test through high transaction costs rather than the existence of a bubble: “We find the argument for bubbles or irrational behavior in the Bay Area housing market less compelling than an explanation based on the high transactions costs that are known to characterize the market,” (263-64).

**Study Design**

I will combine the approaches of Meese and Wallace (1994) and McCarthy and Peach (2004). One problem with McCarthy and Peach (2004) is that they assess national data and present evidence on the existence of a national housing bubble. However, real estate is inevitably a local market. A single-family home for sale in New York is most likely not a viable option for a homebuyer in the San Francisco Bay Area. There is also
the question of elasticity of supply of housing. Housing is presumed to have inelastic supply in many metropolitan areas in the East Coast, for example New York City. However, there is a lot of undeveloped land in central United States, which means that the supply of housing should be more elastic in these places. When aggregating the data over the entire country, McCarthy and Peach lose much of the variation in the local markets. While they do analyze state-level differences, I believe the data need to be disaggregated even further, down to the metropolitan city level.

While Meese and Wallace (1994) do use local data and a constant quality adjusted price index, they did not obtain a municipality-specific rental series. Instead, they used rents for two-bedroom apartments obtained from the classifieds sections of The Oakland Tribune and The San Francisco Chronicle. While apartment rents might be the closest proxy for housing rental prices, a more precise approximation exists. There is an element of the Consumer Price Index called the owners-equivalent rent (REQ). From the documentation to the CPI: “REQ measures the change in the amount a homeowner would pay to rent, or would earn from renting, his or her home in a competitive market. It is a measure of the change in the price of the shelter service provided by owner-occupied housing.” The REQ is available for many metropolitan areas, including the San Francisco Bay Area, which comes in very handy for my paper.

I first present descriptive data similar to McCarthy and Peach (2004), namely, the adjusted and unadjusted rent-to-price ratios. I then use the present value model constructed by Meese and Wallace (1994) to test for efficiency in San Francisco housing market. I will run short-run and long-run tests – in the short-run an efficient market will exhibit uncorrelated returns and errors and in the long-run the forecast errors will be
stationary in an efficient market. However, instead of constructing data series from scratch and using apartment rents as a proxy for household rent, I apply the approach of McCarthy and Peach (2004) by using data published by government agencies, and in particular, by using the owners-equivalent rental series of the CPI.

**Data Description**

One must be careful in choosing the right data, however. There are several data series available on housing prices, each constructed in a slightly different manner. The results could differ depending on the selection of the index, and so I will discuss each in turn.

There are four main methods of constructing a housing price index: prices of repeat sales, constant-quality new home price, median home prices of existing homes, and median home prices of new homes. The median prices of existing homes sold is released by the National Association of Realtors, while the median prices of new homes sold is released by the Bureau of the Census of the U.S. Department of Commerce. From a simple observation of a time series graph of these two indices, it is clear that neither series is seasonally adjusted, as there are many volatile jumps at predictable frequencies. This volatility may also come from the fact that the product mix of sales varies from month to month (McCarthy and Peach 2004). Furthermore, the series might not accurately reflect housing stock values because the underlying price data reflect only recent sales. Thus, it seems that neither index is ideal for the task at hand.

The Office of Federal Housing Enterprise Oversight (OFHEO) releases an index that tracks repeat home sales in metropolitan areas around the country. The index is
constructed by “observing sales prices or appraised values of properties on which the mortgage loans have been purchased by Fannie Mae and Freddie Mac more than once,” (http://www.ofheo.gov). As mentioned earlier, one confounding factor in the increase in housing prices is the change in the quality of a home. Change in quality is something that skews all price indices and housing is no different. The OFHEO index is commonly referred to as a “constant-quality” index. However, there is still some debate as to how accurately it adjusts for quality. The OFHEO index attempts to adjust for quality by sampling repeat home sales instead of new homes sales. New homes tend to be more spacious and equipped with more amenities, and adjusting for this quality increase would be more complicated than when sampling new homes than when sampling repeat sales. One hitch in the OFHEO index, however, could be the fact that it does not take into account changes in the physical characteristics of those properties. So depending on the amount of alterations in the set of properties the index could be skewed upward.

The U.S. Bureau of the Census publishes an alternative index that explicitly attempts to control for physical and locational characteristics of homes. It tries to accomplish this by applying hedonic methods to data on new homes sales. Such methods include using “regressions of home prices on characteristics such as total square footage, number of bathrooms, number of bedrooms, and the presence of air conditioning and fireplaces,” (http://www.census.gov). The index samples from the four major census regions and constructs a national index by applying proper weights to the four areas based on homes sales activity in each region. While the idea of controlling for quality is an appealing one, the critics maintain that such an index will underestimate overall housing price inflation (Hulten 2003). In classic urban development models, construction
occurs at the fringe, where the supply of land is elastic. The supply of land is quite inelastic by definition in the central areas of the city, and thus *ceteris paribus*, prices will increase much more rapidly in the center areas of the city than on the fringes. Since the newer homes are on the fringe and U.S. Bureau of the Census index only samples new homes sold, there is a danger of underestimating overall home price inflation with such an index (Hulten 2003).

Each series has its strong and weak points. From the above analysis, I believe that the OFHEO index is most appropriate for my assessment of the present value relation in the housing market. While the U.S. Census Bureau index might come the closest to controlling for quality, it is still biased towards new homes, which presents a significant sampling problem when you take into account the inelasticity of supply critique noted above. In addition, the index does not provide metropolitan area specific price data and instead focuses on the four major census areas – Northeast, South, Midwest, and West. The lack of metropolitan city specific data makes the dataset virtually unusable, since I need very local data to make predictions about the San Francisco housing market. Thus, I use the OFHEO index as my measure of housing price inflation.

Estimating the rental inflation in San Francisco is relatively straightforward compared with housing prices. For my model, I will need a measure of a rental series that captures the rent or monthly return to owning a home. Luckily, the Consumer Price Index publishes such a series called the “owners equivalent rental series”, and as noted above, it is available specifically for San Francisco. The series is adjusted for seasonality and is updated quarterly, which is ideal for my analysis. The rental series is obtained
from the “owners-equivalent rents” series of the Consumer Price Index and the value is determined by a survey of homeowners asking the following question:

“If someone were to rent your home today, how much do you think it would rent for monthly, unfurnished and without utilities?” (Bureau of Labor Statistics).

I use the same data as McCarthy and Peach (OFHEO for sales prices and the owners-equivalent rent index of the CPI for rental prices) but like Meese and Wallace, I will concentrate on the San Francisco Bay Area to assess the existence of a bubble in a local market. I define my market specifically, as the City of San Francisco, since the housing price and rental data is readily available for such a market definition.

There is one more data series that is imperative to the analysis of the present value relation: cost of capital. Cost of capital, simply put, is the cost of borrowing money. Since many homeowners take out loans to finance the purchase of a home, the cost at which they can do this should be a factor in their decision-making process. The cost of capital series will take into account various factors that affect the cost of borrowing money. It is standard in the literature on housing price studies to incorporate income taxes, property taxes, depreciation, and interest rates into the cost of capital series. I follow this example and define my cost of capital in the following way:

\[ K(t) = (I(t) (1-T^Y(t)))) + d(t) + T^P(t)( 1 + T^Y(t)), \]  
where

\[ I(t) = \text{interest rate, defined as 3mo. Treasury bill rate} \]

\[ T^Y(t) = \text{income tax rate (federal + state)} \]
\[ d(t) = \text{depreciation rate (constant 5\%)} \]

\[ T^p(t) = \text{property tax rate} \]

It is conventional to use a constant depreciation rate of 5\% and the three-month Treasury bill rate as a proxy for interest rates.

**Descriptive Overview**

I now discuss some descriptive data that will shed some light on the current situation in the housing market in San Francisco. I start with the rent-to-price ratio. As noted earlier, housing prices and implicit rents, as defined by the rent a homeowner would have to pay were he to rent a similar home, should share a common trend. This is because if the rent-to-price ratio is high, this suggests that the return on owning the home is high relative to other assets. This relatively high return is unlikely to persist and in order for it to reach the level of competing assets, housing prices must rise or rents must fall.

I analyze the rent-to-price ratio for the City of San Francisco over the entire period data is available (from the first quarter of 1983 to the second quarter of 2005). I use the owners’ equivalent rent series for the City of San Francisco from the Consumer Price Index for the rents and the quality-adjusted measure of prices from the OFHEO for the housing prices. I normalize the results to have a ratio of unity in the first quarter of 1983. A graph of the ratio over the data period is shown in Figure 3. Over the last sixteen quarters the ratio has fallen significantly. The last time the ratio fell below its long run average was in the late 1980s and prices fell significantly a few years later.
Figure 4 displays the lagged price appreciation with the rent-to-price ratio to show the relationship between changes in the rent-to-price ratio and changes in price. The lagged price appreciation is obtained by taking the quarter-on-quarter percentage growth in prices and shifting it two quarters to effectively create a two quarter lag of the price appreciation. As can be seen in the graph, the quarters where the rent-to-price ratio drops significantly are followed by a string of quarters with depreciation in price, or at least a slowed rate of price appreciation. This provides more evidence that there exists a common trend between implicit rents and prices. Specifically, when the rent-to-price ratio is low, housing prices must decrease in order for the return on owning a home to reach the level of competing assets.

However, the picture is not that simple. One problem with this analysis is that the rent-to-price ratio does not take into account any cost of capital variables – interest rates, income tax rate, property tax rate, or depreciation. Certainly, interest rates have a significant effect on demand for housing. Just like an equity position in a company, a house is an income-producing asset. The value of such an asset “is the discounted present value of the net income it provides, with the discount rate being the current yield on a comparable asset with comparable risk characteristics,” (McCarthy and Peach, 2004). Thus, a fall in the discount rate, even with net income constant, would increase the equilibrium value of the asset. If interest rates are low, the cost of financing for a house is relatively low as well and all else equal, one should observe an increased demand for housing leading to price appreciation. In addition, interest rates represent the yield on a competing asset. This means that low interest rates imply low returns on competing assets, and all else equal, a demand based growth in housing prices. A major
feature of the housing market over the last few years has been the steady decline in
nominal mortgage interest rates. It is quite possible that a decline in this rate led to a
higher affordability of homes which in turn led to a steady price increase in homes that
had nothing to do with irrational expectations – merely increased affordability.

I modify my rent-to-price ratio to include interest rates. As suggested in Poterba (1984), in equilibrium, “homeowners equalize the marginal cost and benefit of the
services derived from the housing assets they own.” I take the marginal benefit to be the
implicit rental price from owning the home. Jorgensen (1963) defines the marginal cost
as “the sum of the after-tax opportunity cost of holding the capital asset, after-tax
property taxes, and depreciation and repair, minus the expected capital gain of the asset.”
In other words the equation below is a quarterly cash flow representation of buying or
selling a home. Rent is the marginal benefit while the marginal cost is the price weighted
by the cost of capital. Equating marginal benefit and marginal cost,

$$r(t) = P(t) \left[ (1-TY(t))(I(t)+TP(t)) + d(t) - E(\pi(t)) \right]$$

and rearranging the terms, gives me the adjust rent-to-price ratio:

$$r(t)/P(t) - \left[ (1-T^Y(t))(I(t)+TP(t)) + d(t) \right] = -E(\pi(t))$$

where $E(\pi(t))$ is the expected capital gain, $r(t)$ is the rental series, $P(t)$ is the price
series, and the other variables are the same as the inputs into the cost of capital series.
This last equation shows that the rent-to-price ratio can be adjusted by subtracting the
depreciation rate, taken to be 5%, and after-tax basis of the interest rate and property tax
rate. Following from the first equation, the adjusted rent-to-price ratio is inversely
related to the expected home price appreciation. This can be interpreted the following
way: in equilibrium, relatively low levels of the adjusted rent-to-price ratio imply that
potential home buyers expect relatively high inflation in housing prices, which would be evidence for the existence a bubble. Figure 5 displays the graph of the adjusted rent-to-price ratio.

As can be seen in the graph, the adjusted rent-to-price ratio has not been falling over the last few years. In fact, after taking into account interest rates, property taxes, income taxes, and depreciation, it has been rising slightly. This is in contrast to the unadjusted rent-to-price ratio falling over the same time period. Notice that the adjusted ratio is at its low point during the four quarters of 2000. McCarthy and Peach (2004) identify the strong income growth in the United States during that time as a primary reason that prevented prices from falling. The subsequent decline in interest rates and other key fundamental factors raised the adjusted ratio over the next few years to put it in the high part of its historical range (McCarthy and Peach, 2004). I take this as descriptive evidence that there is no bubble in the housing market in San Francisco over the last few years.

This analysis isn’t without its flaws. It might be easier to control for quality with a price and rental index that is self-constructed from observed transactions documented by state-specific Real Estate Agencies and classifieds from newspapers. In addition, the analysis would be more meaningful when other cities are included in the data. Comparisons across geographic areas have the advantage of controlling for various other factors like inelasticity of supply, construction costs, demographic effects, etc.

However, obtaining data on multiple geographic areas is quite difficult. The CPI “owners equivalent” rental series is only available for a select few metropolitan areas, and it is not available for much of the Midwest, an area many would consider a prime
testing location for the effects of changing elasticity of supply. In addition, while indices are sometimes flawed, obtaining data through newspaper classifieds and transaction databases also has its limitations. Many households do not advertise in newspaper classifieds and so selection bias would be an enormous problem. In addition, the “owners equivalent” rental series fits the exact description of the data one wants for an index of rental prices, “how much would an owner obtain in rent where he to rent out his current home,” (U.S. Bureau of Labor Statistics). It is doubtful that constructing one through transaction databases would be better than this index.

**Constructing the PDV model**

The above analysis provides a descriptive overview of specifics of the San Francisco housing market. However, in order to assess the existence of a bubble, I need something more rigorous than the relative levels of the rent-to-price ratio. It would be interesting to assess the short-run and long-run implications of a bubble, if it exists. Do short-run inefficiencies persist over time, or are they arbitraged away? In order to present a more rigorous analysis and to looking at longer horizons, I construct a present value model similar to Meese and Wallace (1994). First, I will have to determine the degree of persistence in the three data series, since this will determine the testable economic content in each series. If the series are stationary with a high degree of persistence, then the testable economic content of the present value model is that deviations from the levels of house price and capitalized rents are stationary. If the deviations are not stationary, then there is no ground to the efficient markets theory since any deviations from the present value relation are permanent and never dissipate in the long run. If the deviations
are stationary, then the integrated price, rent, and capital series are cointegrated (Granger, 1981). If instead, the levels of prices, rents, and the cost of capital are well approximated by a stationary series, then any linear combination of these series is stationary. This means that there is no long-run testable implication, since all deviations are guaranteed to be stationary. Thus, testing for degree of persistence is important to further analysis, (Meese and Wallace 1994).

I use a Dickey-Fuller unit root test to determine whether the three series are best approximated by a stationary series in levels or by stationary series in first differences. I use a time trend for the housing price index to account for the positive drift in the growth rate of that series. Since the rental and housing series are already seasonally adjusted, I do not need to include any lagged changes to test for seasonality. The results of the test are shown in Figure 6. The test statistics are consistent with a unit root null at all the usual significance levels. On the basis of the tests, all three series can be approximated with a unit root null, and I can continue with my short-run and long-run tests of the present value relation.

I start my present value model by assuming the base case, that the housing market has no transaction costs, common expectations, and a competitive market for a homogenous asset. Under this scenario, investors will equate the real return to owning a home with the cost of owning a home, or the cost of capital. I define my return in the following way,

\[ R(t+1) = \left[ \frac{P(t+1) + r(t+1) - P(t)}{P(t)} \right] \]

where \( R(t) \) represents the real return to owning a home, \( P(t) \) represents the home price, and \( r(t) \) represents the rental price at time \( t \).
Setting the real return equal to the cost of capital and solving for $P(t)$ gives us the classic asset pricing relationship,

$$P(t) = \left[ P(t+1) + r(t+1) \right] / (1+K(t))$$

**Testing the Present Value Model**

I conduct two tests for this present value relation: one for the short-run and one for the long-run. Since predictable serial correlation in returns violates the present value model for the short-run, I will test the present value relation by checking for serial correlation in returns. When testing for serial correlation in returns, I must be careful to choose the correct number of lags of the dependent variable. A model with many parameters (many lags) will provide a very good fit to the data, but will have few degrees of freedom and be of limited utility. On the other hand, including too few lags can reduce the predictive power of the regression. To decide the appropriate number of lags, I use the Akaike’s information criterion (AIC), which tries to balance the complexity of the model with the goodness of fit to the sample data. The definition for AIC is the following:

$$AIC = 2k + N* \ln (RSS / N)$$

where $N$ is the number of observations and $RSS$ is the residual sum of squares. Here, the preferred model is the one with the lowest AIC value. Using this metric, three lags of $R(t)$ minimizes AIC, and so my regression to test for serial correlation in returns looks like:

$$R_t = \beta_0 + \beta_1 R_{t-1} + \beta_2 R_{t-2} + \beta_3 R_{t-3} + \varepsilon_i$$
I employ an F-test to test for the joint significance of the $\beta$ coefficients on the lags of the dependent variable. The F-statistic, along with the coefficient values are reported in Figure 7. The F-test is resoundingly clear: I easily reject the null that there is no serial correlation, as there is joint significance of the three lags.

Notice, however, that the previous test does not allow the cost of capital to change over time. The cost of capital series does not enter the formula for return, and so the test for serial correlation in returns does not take into account any changes to the cost of capital series. Thus, one possible explanation for the serial correlation in the returns is that changes in the mortgage rate, property tax rate, or income tax rate have a great deal of influence on the yearly return to owning a home. To account for this, I introduce the forecast error, $E(t)$, as defined by:

$$E(t) = P(t) + r(t) - P(t-1)(1+(k(t-1)))$$

By incorporating the cost of capital series in the forecast error, I allow any changes to the cost variables to change over time. This leads to my second test for the short-run analysis of the present value model. I test for serial correlation in the forecast error by linear regression of $E(t)$ on a constant and three lags:

$$E_t = \beta_0 + \beta_1 E_{t-1} + \beta_2 E_{t-2} + \beta_3 E_{t-3} + \varepsilon_i$$

Again, the appropriate number of lags was chosen by calculating the Akaike’s information criterion (AIC) for a regression of $E(t)$ on a various number of lags and selecting the number of lags that minimized the AIC. Just like in the test for serial correlation in returns, I employ an F-test to test for joint-significance in the coefficients. The results of this regression are reported in Figure 8.
With an F-statistic of 281.45 the coefficients are jointly significant and so I again easily reject the null that the forecast error is not serially correlated. Thus, the short-run present value relation does not hold in the San Francisco housing market over the 1983 to 2005 time period. However, this does not imply a bubble. As long as any short-run deviations from the present value relation are temporary, the long-run relation will hold.

**Long-run test of the PVR**

In the long-run, the implication is that a linear combination of prices and capitalized rents are covariance stationary. In other words, the series for price, rentals, and cost of capital are cointegrated. Series with a unit root are cointegrated if some linear combination of the series is stationary. Since my previous tests showed that $P(t)$, $R(t)$, and $k(t)$ are all unit root, and my forecast error is defined as a linear combination of these three series, the long-run implication holds if the forecast error is stationary (Meese and Wallace 1994). Thus, I will test for this using a Dickey-Fuller unit root test on the forecast error to examine if the prices and capitalized rents are cointegrated.

The results of this test are reported in Figure 9. At all conventional significance levels, I can reject the null of a unit root, and thus $P(t)$, $R(t)$, and $P(t-1)(1+k(t-1))$ appear to be cointegrated. The interpretation here is that any deviations from the present value relation in the San Francisco housing market are mean reverting in the long-run. Since deviations from the forecast error term are not permanent (not unit root), I can interpret this as long-run support of the present value model in the housing market.

**Comparisons of Actual and Present Value Prices**
The present value relation is a formula for the current price based on expectations of future price and rents. In order to generate a graph that shows the present value home prices over time, I must solve for the forward solution of the present value relation. I do this by recursive substitution of the expected future price terms in the present value relation. I must now make an assumption about how the rents and cost of capital is expected to change from year to year. I assume that each year, the representative agent recalculates the present value price by observing the current cost of capital, \( k(t) \). In addition, the expected growth rate of the rental series is \( Q(t) \) given by:

\[
Q(t) = r(t) - r(t-1)
\]

Under these assumptions, I can generate a present value home prices series that is given by:

\[
P^*(t) = \frac{r(t)}{k(t)} + \frac{Q(t)(1+k(t))}{k(t)}.
\]

Figure 10 plots the present value home price series, \( P^*(t) \), and the observed home price series, \( P(t) \), over the sample period from 1983Q1 to 2005Q2. The graph supports the conclusions of the short-run and long-run tests of the present value relation. The graph clearly shows that there are many short-term deviations between the actual and present value price. However, in the long-run, the short-run deviations are erased and the actual and present value price series track each other pretty closely over the sample period.

For example, one can interpret the graph by assessing the forecast error, or the difference between \( P(t) \) and \( P^*(t) \). From the period 1987Q3 to about 1992Q1, the forecast error is positive since \( P(t) \) is greater than \( P^*(t) \) during that time. This means that housing prices were overvalued during this time period with respect to the forecasted
prices based on fundamental factors. However, this error turns from positive to negative over the next seven years. From 1992Q1 to 1999Q3, \( P(t) \) is consistently less than \( P^*(t) \) which means that housing prices were undervalued with respect to the forecast. In the last five years, \( P^*(t) \) tracks \( P(t) \) pretty closely and so the forecast error is close to zero for this period. This means that currently, housing prices are on par with forecast levels, which lends credence to the conclusion that there is currently no bubble in the market.

**Discussion**

I find that the present value model holds in the long-run, but not in the short-run. I discuss a few reasons for the short-run violation. The possibility of measurement errors in the price, rental, or cost of capital series that are serially correlated would in turn cause serial correlation in returns and forecast errors. However I have reason to believe this is not a significant factor. First, any serially correlated measurement error in the price series is likely to be caused by serially correlated quality adjustments to homes. However, recall that the price series is a “constant-quality” index since it is from the Office of Federal Housing Enterprise Oversight (OFHEO) and tracks repeat home sales based on mortgages bought by Fannie Mae and Freddie Mac more than once. So serially correlated measurement errors in price are probably not a significant factor. Second, the cost of capital series is constructed to take into account any changes in the property tax rate, income tax rate, and interest rates. It is unlikely that there is any serial correlation in measurement errors here, since all three data sources are not very measurement intensive. That is, the income tax rate and the property tax rate are values declared by the government and these are unlikely to be cause for measurement error. In addition, the
three-month Treasury bill rate, serving as a proxy for the interest rate, is unlikely to be in error since it is a very liquid estimate of the actual rate. Thus, any serial correlation in returns cannot be due to measurement error in the cost of capital series. Finally, serially correlated measurement errors in the rental series could be a factor causing serially correlated returns. However, there is no indication that there are serially correlated measurement errors for the rental series. The rental series is obtained from the “owners-equivalent rents” series of the Consumer Price Index and the value is determined by a survey of homeowners asking the following question:

“If someone were to rent your home today, how much do you think it would rent for monthly, unfurnished and without utilities?” (Bureau of Labor Statistics).

The only way to observe serially correlated measurement errors in rents is if homeowners, on average, are making an assessment error that is serially correlated. There is no reason, however, to believe that this is true. In addition, there is another component of the Consumer Price Index that reveals a measure of rental prices without surveying homeowners, the rental sub-series. This sub-series is obtained by sampling current units for rent and observing the rental price. The serially correlated returns are robust to the use of rental series. Thus, I have no reason to believe that both components of the CPI – one obtained by homeowner surveys, the other from sampling rental prices – have serially correlated measurement errors.

One very plausible reason for the short-run violation is transaction costs. It is well known that large transaction costs are associated with buying and selling a home. The search process, real estate agent commissions, and zoning restrictions are only a few
of the costs that add friction in the market. These transaction costs cause a separation between the present value of a home and its actual price, as any arbitrage opportunity would not be able to be seized immediately. Imagine a potential homeowner solving the first order condition given by the present value relation:

\[ P(t) = \frac{\left[ P(t+1) + r(t+1) \right]}{(1+K(t))} \]

One reason for this separation is that before the potential buyer purchases the home, the transaction costs outlined above would have to be outweighed by the utility gain in owning the home. The fact that the present value relation holds in the long-run, but not the short-run is reason to believe that transaction costs play a significant role in the decision making process (Meese and Wallace 1994).

The definition for the cost of capital series does not factor in a risk premium to owning a home. One might think that the lack of a risk premium might cause the forecast error to be serially correlated. However if this were true, the risk premium could not be constant over time, as a constant risk premium would merely translate the forecast error by a constant and not affect the change in forecast over time. Even though it is generally assumed that a risk premium would be relatively constant over time, for the sake of argument, I assume it varies significantly. Even then, it should not cause any serial correlation in the forecast error, since \( K(t) \) does not enter the equation for return linearly. Rather, it is scaled by \( P(t) \). Thus, I cannot conclude that any serial correlation in a non-constant risk premium is to blame for serial correlation in the forecast error.

Conclusion
The findings of this paper could be greatly improved with more precise data. Despite obtaining specific data sets for the city of San Francisco, the price and rental indices are inexact and have their respective biases. The ideal price and rental index would be self-constructed from observed transactions documented by state-specific Real Estate Agencies and classifieds from newspapers. In addition, the analysis would be more meaningful when other cities are included in the data. Comparisons across metropolitan areas have the advantage of controlling for various other factors like inelasticity of supply, construction costs, demographic effects, etc.

Given the obtained data, I conclude that for the city of San Francisco, the housing market is marked with inefficiencies in the short-run but is ultimately efficient in long-run. The short-run inefficiencies are most likely due to the high prevalence of transaction costs in the market, and do not stem from an irrational bubble.
Figure 3
Rent to Price Ratio

Figure 4
Rent to Price Ratio with Lagged Price Appreciation
Figure 5

Adjusted Rent-to-Price Ratio

adjusted R(t)/P(t) (1983Q3=1)
**Figure 6**
Augmented Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th></th>
<th>TEST STATISTIC</th>
<th>1% CRITICAL VALUE</th>
<th>5% CRITICAL VALUE</th>
<th>10% CRITICAL VALUE</th>
<th>NUMBER OF OBS.</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(t)</td>
<td>0.184</td>
<td>-4.073</td>
<td>-3.465</td>
<td>-3.159</td>
<td>85</td>
</tr>
<tr>
<td>R(t)</td>
<td>-2.163</td>
<td>-4.073</td>
<td>-3.465</td>
<td>-3.159</td>
<td>85</td>
</tr>
<tr>
<td>K(t)</td>
<td>-1.239</td>
<td>-3.525</td>
<td>-2.899</td>
<td>-2.584</td>
<td>89</td>
</tr>
</tbody>
</table>

**Figure 7**
Results from linear regression of R(t) on three lags of R(t)

|       | COEF. | STD. ERR. | T   | P>|T| |
|-------|-------|-----------|-----|-----|
| lag1  | .0865125 | 13.69    | 0.000 | 1.012114|
| lag2  | .1448213 | 0.56   | 0.574 | -.2064527|
| lag3  | .1012424 | -2.77   | 0.007 | -.4816013|
| _cons | 1.468498 | 0.59    | 0.555 | -2.051356|
### Figure 8
Results from linear regression of E(t) on three lags of E(t)

| E(T) | COEF.  | STD. ERR. | T     | P>|T| |
|------|--------|-----------|-------|-----|
| lag1 | 0.7592047 | 0.1422589 | 5.34   | 0.000 |
| lag2 | 0.0720354 | 0.1361674 | 0.53   | 0.598 |
| lag3 | 0.1696855 | 0.1532655 | 1.11   | 0.271 |
| _cons | 2.622913 | 2.273486 | 1.15   | 0.252 |

### Figure 9
Augmented Dickey-Fuller test for unit root on E(t)

<table>
<thead>
<tr>
<th>TEST STATISTIC</th>
<th>1% CRITICAL VALUE</th>
<th>5% CRITICAL VALUE</th>
<th>10% CRITICAL VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>E(t)</td>
<td>-3.044</td>
<td>-3.527</td>
<td>-2.900</td>
</tr>
</tbody>
</table>
Figure 10

PDV vs Actual P(t)

Price index (1983Q1=100)

- PDV
- Actual
Reference List


