# THE DYNAMIC RELATIONSHIP BETWEEN HOUSING MARKET SENTIMENT AND HOME PRICES

LE THI THANH THAO

A0025889M

SUPERVISOR: A/P OOI THIAN LEONG, JOSEPH

A thesis submitted for the degree of

Master of Science (Real Estate and Urban Economics)

DEPARTMENT OF REAL ESTATE
NATIONAL UNIVERSITY OF SINGAPORE

2011/2012

## Executive Summary

*Despite the potential relevance of psychological factors in explaining price movements in the housing markets, no research has directly tested the relationship between sentiment and house prices. This thesis examines the role of sentiment in setting house prices, and thus their well-documented high volatility and susceptibility to bubbles. The two research questions guiding the empirical tests in this thesis are: (1) Does sentiment have a role in explaining house price movements, above and beyond the role of fundamentals, and (2) Do past price movements predict changes in sentiment beyond the predictions of lagged sentiment and fundamentals? On one hand, sentiment is expected to positively drive home prices through its effect on the demand side in the short run. On the other hand, myopic expectations suggest a reverse causality relation from price movements to sentiment.*

*The contribution of this study is twofold. Firstly, it introduces a set of direct proxies that appear to capture the consensus sentiment of three major agents in the housing markets, moving away from the "residual price" measures employed in prior research. Secondly, the use of explicit indicators allows for a dynamic model to directly estimate the timing, magnitude as well as direction of impact between sentiment and house prices. Such insights are not permissible in prior studies as they infer sentiment's effect from house price residuals.*

*To measure and isolate sentiment's effect, this study employs direct indicators that capture the sentiment of three major agents in the U.S. housing markets: homebuyers (demand side), builders (supply side) and lenders (intermediaries). These sentiment measures are derived from surveys conducted on a frequent basis by the University of Michigan, the National Association of Home Builders, and the Federal Reserve Board, respectively. To measure house price appreciation, this study uses the percentage change in the real Case-Shiller U.S. National Home Price Index.*

*The dynamic relationship between housing market sentiment and housing prices is modeled in a VAR framework with two endogenous variables both expressed as linear functions of their own and each other's lagged value. Maximum Likelihood Estimation is employed to simultaneously estimate the VAR system using quarterly data over the 1990:Q2 - 2010:Q3 sample period.*

*Empirical results show that the sentiment of all three market participants positively influences house price appreciations in subsequent quarters. Moreover, this sentiment effect is seen to be highly persistent and takes beyond five years to correct. The sentiment-induced component becomes much larger during episodes of fast escalating and slumping house prices (i.e., the formation and bursting of housing bubbles) than in more moderate periods. This thesis also finds evidence of myopic expectations among homebuyers and lenders; however, homebuilders do not appear to be backward-looking. The dynamic interplay between sentiment and home prices is therefore a self-reinforcing process, which potentially renders housing markets highly prone to bubbles.*

# Acknowledgement

The completion of this thesis has been a very pleasant journey for me thanks to many amazingly great mentors, friends and relatives whom I am lucky to have with me.

My very first and deepest gratitude is owed to Prof Joseph Ooi. I simply cannot thank him enough for his countless, heartfelt care to open the doors leading me to the wonderful destination I am at now. I am indebted to him for each and every accomplishment of mine. He is a great teacher, mentor and friend that anyone would be thankful to have.

I would not have successfully completed this thesis without the valuable advice from Prof David Ling. The physical distance and different time zones between us did not at all deter him from offering his help to me. It is a great pleasure to work with such an experienced and kind-hearted professor.

Prof Tu Yong and Prof Wayne Archer have also kindly offered me useful comments on improving this thesis.

My friends, Cheng Keat, Khairul and Liu Bo, have always been very encouraging and fun to be with. I enjoy and appreciate their company a lot.

Finally, I am grateful for my family who are always a great source of support and motivation for everything I do.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1 - Introduction

## 1. Background and motivation

House prices tend to be quite volatile relative to observable swings in economic fundamentals (Glaeser, Gyourko and Saiz, 2008). In the latest U.S. housing turmoil, a rapid and consistent surge in home prices occurred from 1998 to 2006. Although the boom was fuelled in part by easy credit in its initial years, the later phase saw house prices reaching levels far above what economic fundamentals could reasonably support.

Popular explanations in the housing economics literature for the remarkable house price run-up include cheap credit (Mayer and Sinai, 2009; Glaeser, Gottlieb and Gyourko, 2010), easy credit, and subprime lending (Mayer and Pence, 2009; Lai and Van Order, 2010)[1]. Other explanations for the rapid house price appreciation include supply restrictions (Glaeser, Gyourko and Saiz, 2008) and growth in the demand for second and investment homes (Wheaton and Nechavey, 2008). Regardless of the fundamental models employed, almost all the empirical research suggests that actual house prices are more volatile over the cycle than would be predicted by economic fundamentals. Acknowledging that it is hard to explain the

[1] Mayer and Sinai (2009), in particular, show that the growth in subprime lending was largest in housing markets with the biggest growth in price-rent ratios. Mayer and Pence (2009) further document that the median loan-to-value ratio of a housing loan was an astounding 100% for subprime mortgages originated in 2005, 2006 and the first half of 2007.

rapid rise and fall of housing prices across U.S. markets with a purely rational model, some researchers have suggested that non-rational explanations, such as exuberance and unrealistic expectations of future price appreciation, cannot be ruled out (see Glaeser, Gyourko and Saiz, 2008; Mayer, 2011). Indeed, the very formation and subsequent bursting of such pricing bubbles has been suggested as the most direct evidence of prices being driven in part by misguided beliefs.[2] In their recent study which concludes that easy credit cannot account for the majority of the price changes in the recent housing boom and bust[3], Glaeser, Gottlieb and Gyourko (2010) refer to the work of Case and Shiller (2003) as a compelling alternative explanation.

To gauge the driving forces behind home purchase decisions during the recent housing boom, Case and Shiller (2003) conducted a survey of homebuyers in four U.S. metropolitan areas (MSAs). They put forward several questions to address the role of expectations, emotional charge and the extent of talk about real estate in homebuyers' buying decisions. They conclude that a housing bubble was evident in 2003, for demand was being largely driven by at least three sentiment indicators found from the survey. Firstly, on average respondents expected the price appreciation rate over the following decade to be about 12%-16% per year, a

[2] In fact, it is the very definition of a bubble – that is, a rapid price run-up not justified by economic fundamentals but by unrealistically high expectations of future prices (Stiglitz, 1990).

[3] Although interest rates do influence house prices, GGG (2010) conclude that cost of credit cannot provide anything close to a complete explanation of the great housing market gyrations between 1996 and 2010. Indeed over the long 1996-2006 boom, they observe that falling interest rates only explain about 20% of the rise in house prices. Neither can the approval rates or down payment requirements account for most of the movement in house prices.

strikingly high range[4]. Secondly, at least two thirds of respondents agreed that it was a good time to buy a home because prices might rise further and erode their affordability. Finally, about 30%-46% of the respondents admitted they were being influenced by the "excitement" around them. More recently, Piazzesi & Schneider (2009) use the Michigan Survey of Consumers to find that there was always a small cluster of "momentum" agents who believed prices would continue to rise further. What was unusual about the boom, as noted in their paper, is the increased proportion of these optimists in the market, which peaked precisely at the time when prices rose to their historical high in 2004-2005.

Clearly, it appears from the above work that investor psychology may lead to very large and persistent mispricing in housing markets. When prices start to rise, irrational exuberance may lead buyers and sellers to become carried away by irrational expectations of future price appreciation, causing house prices to spike in booms. As a result, house prices become excessively volatile over the cycle, rising more than fundamentals would suggest in a boom and falling faster than the decline in fundamentals in a bust. Such behavioural biases have long been a popular explanation for various asset-pricing puzzles which are hard to reconcile with a rational framework. The focus, however, has been mainly on the pricing of securitized assets in the finance realm[5]. Much as the inherent characteristics of

[4] Even an appreciation rate of only 12% per year means a tripling of value in ten years (Case and Shiller, 2003).

[5] The possible effect of investor sentiment on asset prices is a long running debate in the financial economics literature. In classical finance models of asset pricing, there is no role for investor sentiment due to the rational actions of informed arbitrageurs. However, in behavioral finance models, investor sentiments can play a role in the determination of asset prices – independent of market fundamentals. Brown and Cliff (2005) note that respected researchers have entered on both sides of the

housing markets imply that home prices are highly prone to sentiment, there is not yet any empirical research aimed at understanding their dynamic relationship thus far. In the housing markets, high transaction costs and limits on short-selling discourage informed traders from taking advantage of profit opportunities in the market to eliminate mispricing. Consequently, housing prices can diverge from fundamentals over a prolonged period, consistent with the argument that housing markets are less efficient than financial markets (Case and Shiller, 1989). Price corrections may also span much longer horizons because sentiment appears to be persistent and reinforced by the positive feedback from home price movements. These characteristics inevitably call for more work on the role that sentiment plays in housing price dynamics.

## 2. Research focus and contribution

Despite the potential relevance of psychological factors in explaining price movements in the housing market, no research has directly tested the relationship between sentiment and house prices due to the difficulty in measuring sentiment. Archer and Smith (2011), in particular, note that "euphoria" cannot be measured directly because it is largely psychological in nature. Therefore, sentiment is often inferred indirectly through price deviations from fundamental values, which are

argument as to whether the stock price run-up and subsequent market collapse of 1929 (as well as the rapid rise and fall of technology stocks) was rational or not. Clayton, Ling and Naranjo, (2009) postulate that the behavioral approach explicitly recognizes that some investors are not rational and that systematic biases in these investor's belief induce them to trade on non-fundamental information (i.e. sentiment).

usually determined based on real construction cost, income multiples, or present value of rents (Abraham and Hendershott, 1996; Clayton, 1996, 1997; Lai & Van Order, 2010; Wheaton & Nechayev, 2008). For instance, Abraham and Hendershott (1996) use the deviation between the actual metropolitan house price level and a "fundamental" price level to explain the large cyclical swings in real house prices.[6] Clayton (1996) also concludes that economic fundamentals can largely explain price movements, but this ability deteriorates in times of large swings in house prices. This suggests his model "misses something in times of rapidly rising or falling prices," which indirectly points to the role of psychology (Clayton, 1996). However, the main limitation of attributing a portion of the perceived mispricing to sentiment is the absence of precise valuation models. This makes it difficult to attribute the deviations to actual mispricing or model misspecification.

To bridge this gap in the literature, this thesis seeks to carry out a direct test on the dynamic relationship between prices and sentiment in the housing markets. The first contribution of this study is the introduction of a set of proxies that appear to capture the consensus sentiment of three major agents in the U.S. housing markets: homebuyers (demand side), builders (supply side) and residential mortgage lenders (intermediaries). These sentiment measures are derived from surveys conducted on a frequent basis by the University of Michigan, the National Association of Home Builders, and the Federal Reserve Board, respectively. Thus, this study moves away

[6] Similarly most authors attempt to determine whether a bubble exists by comparing actual house prices with what house prices should be based on a model of fundamentals. In doing so, researchers sometimes attribute large increases in actual house prices relative to fundamental house prices as representing periods in which a bubble exists. The presumption is that when house prices do not fully reflect fundamentals, a bubble is present (Mayer, 2011).

from the "residual price" measure often employed in prior research – that is, the difference between actual house prices and estimated "fundamental" prices. The use of explicit indicators allows for a dynamic model to directly estimate the timing, magnitude as well as direction of impact between sentiment and house prices, which constitutes the second contribution of this study. Such insights are not permissible in prior studies as they infer sentiment's effect from house price residuals.

The two research questions guiding the empirical tests in this thesis are: (1) Does sentiment have a role in explaining house price movements, above and beyond the role of fundamentals, and (2) Do past price movements predict changes in sentiment beyond the predictions of lagged sentiment and fundamentals? On one hand, sentiment is expected to positively drive home prices through its effect on the demand side in the short run. On the other hand, myopic expectations suggest a reverse causality from prices to sentiment. When prices are increasing (decreasing), market participants (mistakenly) believe that prices will continue to increase (decrease) in the near future, suggesting that sentiment also follows price movements. The dynamic interaction between sentiment and home prices may therefore be "spiral", which presents a compelling explanation for the high volatility and susceptibility of house prices to bubbles.

Examining the link between house prices and sentiment is important for several reasons. Because housing equity is the largest source of wealth for many homeowners, rapid house price increases and declines can have a significant impact on consumption and investment, and thus national economic growth (Riddle, 1999).

The source of house price volatility is therefore of great interest to policy makers. If sentiment is indeed an important driving force for homebuyers, housing prices are then inherently unstable (Case & Shiller, 2003). In the short run, prices can deviate from fundamental values depending on the sentiment agents have about the market, thereby increasing volatility. In the long run, the slow information revelation process means that sentiment-induced mispricing can accelerate more quickly in housing markets than in other more liquid asset markets, causing prices to be highly prone to explosive bubbles. As witnessed in the latest U.S. housing turmoil that subsequently triggered a global financial crisis, house price busts are also closely associated with loan default and foreclosure, leading to high costs in the banking and financial industry. The dynamic relationship between sentiment and house prices therefore has important implications for the economy as a whole.

## 3. Scope of study and research design

This study examines the U.S. housing market between 1990Q2-2010Q3. To measure and isolate sentiment's effect, this thesis employs proxies that appear to capture the consensus sentiment of three major agents in the U.S. housing markets: homebuyers (demand side), builders (supply side) and residential mortgage lenders (intermediaries). These sentiment measures are derived from surveys conducted on a frequent basis by the University of Michigan, the National Association of Home Builders, and the Federal Reserve Board. These surveys purport to gauge the perceptions of respondents on the prevailing housing market conditions and are

available at the national level. The sentiment indicators derived from the surveys are then used to explain movements in the Case-Shiller National Home Price Index over the study period.

Using a vector autoregressive regression (VAR) framework, this study is able to account for a potential bi-directional interaction between sentiment and home prices. To preview the results, it is found that the sentiment of all three housing market agents influences house price appreciations in subsequent quarters. Specifically, high levels of sentiment predict significantly greater price appreciation. A one-standard-deviation shock to household, builder and lender sentiment induces a 60 to 80 basis points increase in house prices over the next two quarters. These figures are economically meaningful relative to the average price change of 0.71% per quarter over the whole sample period. This sentiment-induced component becomes much larger during episodes of fast escalating and slumping house prices (i.e., the formation and bursting of housing bubbles) than in more moderate periods. Moreover, this sentiment effect persists over a sustained period of time. In the long run, however, sentiment is found to be negatively related to the future five-year average house prices as the correction process starts to take place. The results also suggest evidence that increased price appreciation predicts higher levels of sentiment in subsequent quarters among homebuyers and lenders. The dynamic interplay between sentiment and home prices is therefore a self-reinforcing process. However, home builders do not appear to be backward-looking.

The remaining chapters of this thesis are organized as follows:

Chapter 2 reviews some literature on related topics;

Chapter 3 explains the data collection process and empirical models;

Chapter 4 discusses the empirical results;

Chapter 5 presents the robustness tests;

Chapter 6 summarizes the findings and their limitations, as well as discusses some further research directions.

# Chapter 2 - Literature Review

This chapter provides a broad summary of the relevant literature on housing markets that have been established by previous studies. The discussion will start with some major fundamental issues and stylized facts about housing price dynamics. It further goes on to explore the key viewpoints on mispricing in housing markets as well as housing bubbles, where the role of sentiment is acknowledged but not yet directly examined. The third section is a discussion on some prominent work in behavioural housing research, which is still a relatively new field. Finally, this chapter ends with a brief summary and highlight of the contribution this study aims to make for the present literature.

## 1. House price dynamics

### 1.1. Fundamentals of house prices

House price determinants is undoubtedly a well-researched field. Economic theory shows that house prices are essentially determined by supply and demand, which in turns are driven by economic fundamentals. Empirically, the literature that focuses on demand factors often considers population growth, GDP, income, unemployment rate, interest rate, etc., whilst housing supply is influenced primarily

by construction cost and interest rates (see, for example, Mankiw and Weil, 1989; Peek and Wilcox, 1991; Abraham and Hendershott, 1992; DiPasquale and Wheaton, 1994; Meen, 2002; Hwang and Quigley, 2006). On the impact of demographic changes on housing demand, Mankiw and Weil (1989) assess the 1950s baby boom and the 1970s baby bust in the U.S. They demonstrate that the increase in real housing prices in the 1970s was due largely to the entry of the baby boom generation. Following that, they predict housing demand to grow only slowly in the 1990s as the baby bust generation enters their home-buying age, causing real house prices in 2007 to be 47 per cent lower than the 1987 level.

A number of studies focus on the proposition that in a well-functioning market, house prices and economic fundamentals are cointegrated in the long run. For example, Malpezzi (1999) employs data of 133 MSAs in the U.S. to provide evidence for cointegration between house prices and income during 1979-1996. Using transaction-level data from the Parisian market, Meese and Wallace (2003) model housing prices as a function of construction cost, interest rate, employment and real income. They find a long-tem equilibrium relation between economic fundamentals and house price movements over the 1986–1992 period, with an adjustment rate of about 33-40 per cent per month.

When revisiting the issue of cointegration between house prices and fundamentals, Zhou (2010) criticizes that prior studies have always assumed a linear relationship. He argues that, if exists, the cointegration relationship between house prices and economic fundamentals can take any form, depending on the underlying

data generating process. Zhuo (2010) thus examines and finds a non-linear form of cointergation for 10 cities in the U.S. using data from 1978 to 2007.

In contrast, the work of Gallin (2006) fails to find evidence supporting a stable long-run relation between house prices and fundamentals. Arguing that the standard tests applied on times series in previous studies have low power, Gallin (2006) uses a bootstrap approach on panel data, which he claims to be a more powerful test, of 95 MSAs in the U.S. to test the link between house price and income over the period 1975-2002. Contrary to other studies supporting a stable equilibrium condition, his result does not indicate any cointegration relationship between them, at least at levels.

## 1.2. House price anomalies

Notwithstanding the voluminous housing literature, numerous challenges in understanding the housing markets still remain to date. For long, market observers have been puzzled by the high volatility of house prices with not only large time-varying fluctuations but also wide geographical variations. As documented in Glaeser et al. (2010), the three-year real changes of average house prices in their sample of 113 MSAs range from about $6,500 in Sunbelt markets to as high as $30,000 in coastal markets.

On the geographical disparities in house price volatility of different local markets, the common finding among related studies points to housing supply elasticity as the main underlying cause, where supply elasticity is typically measured

by local topography and land use regulation strictness. Miller and Peng (2006), for example, find that housing markets in metropolitan areas with constrained housing supply seem more sensitive and respond more dramatically to shocks, together with longer shock effect. This empirical evidence is in line with the theoretical model of house prices in Glaeser, Gyourko and Saiz (2008) which predicts that places with more elastic housing supply have fewer and shorter bubbles, with smaller price increases.[7] They further employ data of the housing boom and bust cycle 1980 - 2007 to empirically test their model prediction, which show that the price run-ups of the 1980s were almost exclusively experienced in cities where housing supply is more inelastic.

Early studies on housing market efficiency have documented both serial correlation (Case and Shiller, 1989; Cutler, Poterba, and Summers, 1991) and mean reversion in the time-series behaviours of house prices (Abraham and Hendershott, 1996; Malpezzi, 1999; Meen, 2002). The seminal work of Case and Shiller (1989) has been followed by a large stream of literature focusing on the serial correlation and forecastability of housing prices. Their paper establishes that price changes tend to be followed by changes in the same direction in subsequent years; the weak-form efficiency hypothesis for the housing markets is therefore rejected. Following Case and Shiller's work, empirical studies attempting to model house prices usually include lagged values as an important variable. The positive serial correlation, however, persists only in the short-run. As has been observed for other asset classes, over the long horizon house prices also exhibit a negative serial correlation, or a

[7] However, the consequences of bubbles may be higher in more elastic places because those places tend to overbuild more in response to a bubble (Glaeser, Gyourko and Saiz, 2008).

mean-reverting trend. The common approach used for testing this behaviour is an error-correction framework where prices gradually adjust to disequilibrium conditions (see, for example, Malpezzi, 1999; Meen, 2002; Meese and Wallace, 2003).

Although attempts to explain the above price behaviours are voluminous, the consensus in this area is that a fully satisfactory answer is still non-existent to date. Recently, Glaeser, Gyourko, Morales and Nathanson (2010) calibrate a dynamic rational expectation model to reconcile the above puzzling stylized facts about housing prices: high volatility, serial correlation and mean reverting behaviors. They test its prediction power by fitting the model with two sets of data for 1990-2004 and 1980-2003 from the Home Mortgage Disclosure Act and Bureau of Economic Analysis, respectively. Although their model is able to explain mean reversion and price volatility over a one-year horizon, it substantially underpredicts price volatility over time horizons longer than a year; moreover, the model fails to account for serial correlation. The authors offer to two explanations for the inability of their model to predict serial correlation in house prices. The first is that serial correlation might be the result of a learning process during which households and other market participants gradually infer the state of housing demand from prices. The second, less rational, explanation is essentially the extrapolation hypothesis; that is, market participants with myopic expectations infer future price growth from past price changes. Glaeser et al. (2010) thus conclude that the long-term price volatility and serial correlation of house prices remain important unresolved issues.

## 2. Mispricing and housing bubbles

Given their important implications for the overall economy, housing bubbles have long attracted much work from housing economists. At the center of this issue is the identification of the existence of bubbles. In the 1990 symposium on asset bubbles, a widely accepted definition was put forth by Stiglitz (1990): "If the reason that the price is high today is only because investors believe that the selling price is high tomorrow – when 'fundamental' factors do not seem to justify such a price – then a bubble exists". Two main barometers of bubbles emerge from the above definition: the activities of investors, and the extent of deviations from fundamentals. Yet, there are still much debate and contradictory conclusions among studies on this area, owing to the challenges in estimating fundamental prices and quantifying speculation.

The literature has broadly considered three alternative approaches to determine fundamental house prices: (1) finance-based models that relate prices to the discounted values of rents and user cost of ownership (Porteba, 1984, Himmerlberg, Mayer and Sinai, 2005), (2) a cost approach that compares house prices to cost of construction (Glaeser, Gyourko and Saiz, 2008), and finally, (3) assuming an equilibrium state between house prices and fundamental variables such as income or population (Abraham & Hendershott, 1996; Wheaton and Nechayev, 2008). Regardless of the methodology employed, the common inference is that there exists a bubble if actual house prices differ greatly from the estimated fundamental prices.

Lacking an instrument to measure speculation, most researchers adopt an indirect approach which consists of decomposing actual price appreciation into two components: the portion driven by fundamentals and the residual amount assumed to be driven by speculative activities. Riddle (1999) defines speculators or feedback traders as those who trade on price expectations formed by past price movements rather than expected fundamentals.[8] The trading activity of speculators can lead to price undershooting or overshooting, thereby increasing price volatility. Riddle (1999) points to the speculative house price bubble experienced in Santa Barbara South Coast in late 1987 as evidence for his hypothesis. Similarly, Levin and Wright (1997) model speculators as forming their expectations about future capital gains from historical price growth. Based on a UK sample, they find that speculation is a determinant of house prices. More recently, Piazzesi & Schneider (2009) presents a theoretical search model to illustrate the effects of optimistic, or momentum, traders on house prices. In their model, optimists can drive up prices significantly even when they constitute only a small proportion of total market shares in a search market. Empirically, Wheaton and Nechayev (2008) show that forecast errors of house price appreciation between 1998 and 2005 are positively correlated with the proportion of homes sales attributed to investors or second-home buyers. Working on the role of investors in the recent housing crisis, Haughwout et al. (2011) also contend that in states that experienced the largest housing booms and busts, at the peak of the

[8] Riddle (1999) further distinguishes these "feedback traders" from "fundamental traders". Both can be categorized as speculators (as opposed to conventional homebuyers who consume housing services), but the later base their purchase and sale on forecasts of future economic conditions. Hence, fundamental investors are more likely to buy when prices are low relative to expected economic fundamentals, and vice versa. These "fundamental traders" therefore do not cause distortions in the market.

housing market almost half of purchase mortgage originations were associated with investors.

Popular explanations in the housing economics literature for the remarkable house price run-up between 1998 and 2006 include cheap credit (Himmelberg, Mayer and Sinai, 2005; Glaeser, Gottlieb and Gyourko, 2010), easy credit and subprime lending (Wheaton and Nechayev, 2008; Mayer and Sinai, 2009), Lai and Order, 2010;). Mayer and Sinai (2009), in particular, show that the growth in subprime lending was largest in housing markets with the biggest growth in price-rent ratios. Mayer and Pence (2009) further document that the median loan-to-value ratio of a housing loan was an astounding 100% for subprime mortgages originated in 2005, 2006 and the first half of 2007. Another explanation for house price movements considered in the literature concerns the effect of supply restrictions, whereby places with more inelastic supply experience more and longer bubbles, with larger price increases (Glaeser, Gyourko and Saiz, 2008).

In addition to the above fundamental factors, with the rising popularity of behavioural research, there has been growing interest in the role of non-rational factors in explaining bubbles. In a recent working paper, Glaeser, Gottlieb & Gyourko (2010) find that easy credit cannot be fully blamed for the drastic price gyrations in the recent housing boom[9]. They conclude their paper citing the unrealistic price expectations found from the survey in Case & Shiller (2003) as a plausible alternative explanation. Several questions about the role and cause of

[9] Indeed, over the long 1996-2006 boom, they observe that falling interest rates only explain about 20% of the rise in house prices. Neither can the approval rates or down payment requirements account for most of the movement in house prices.

irrational exuberance are then raised as pressing topics for future research to resolve this missing piece of the house price puzzle. The next section will discuss some major work in the field of behavioural housing research.

## 3. Behavioral research in housing economics

The first prominent study to directly employ behavioural theories in housing markets is perhaps that by Genesove and Mayer (2001). Their paper attributes the long-standing puzzle of a positive correlation between housing prices and trading volume to the loss aversion behaviors of sellers in a down market. As loss-averse sellers face potential loss due to house price declines, they have an incentive to set higher reservation prices, resulting in longer time-on-market and lower sale hazard. A number of studies have also provided some insights into how homebuyers form their perceptions about housing markets, albeit limited. Housing market players are found to overreact to income growth in Capozza & Seguin (1996). The initial overreaction of prices to income growth, or what they call “euphoria”, results in lower prices in the next decade. Hamilton & Schwab (1985) examine if homebuyers’ expectations of price appreciation are rational, in the sense that they can efficiently incorporate all available information in their expectations. Measuring expected appreciation rate as embedded in rent-to-value ratios, they find that expectations were systematically wrong during 1974-1976. In particular, the market misinterpreted past price growth as signalling lower capital gains in the future whilst indeed the reverse was true.

Apparently, misusing this important piece of information raises doubt to the rational expectation hypothesis in housing markets[10].

In a more recent paper, Dua (2008) uses the Michigan Survey of Consumers to identify the factors influencing homebuyers' perception about buying houses. Interest rates, income, wealth, financial status and current house prices are found to Granger-cause home buying attitudes, among which interest rate plays the most influential role. Interestingly, the impulse response functions show that a rise in current house prices induces a positive buying sentiment in the short run but becomes a negative impact over longer horizons. Nevertheless, Dua (2008) does not focus on the possibility of a reverse causality from sentiment to home prices.

In fact, no research in the current literature has directly tested the role of sentiment in setting house prices. Instead, it is only indirectly inferred from price deviations from their fundamental values. Clayton (1996, 1997), for example, documents a short run pricing error in house prices which he attributes to psychological factors. Similarly, Meese & Wallace (1994) find that although house prices and market fundamentals are cointegrated in the long run, the relationship does not hold in the short run. Prior studies such as Abraham & Hendershott (1996), Hui & Yue, (2006), and Wheaton & Nechayev (2008) have also adopted this approach to identify the presence of housing bubbles. A common inference drawn from these studies is that housing markets are characterized by irrational price expectations and house prices, at times, deviate from their fundamental values. However, as most

[10] Their result should, however, be interpreted with caution due to the short study period and data limitation (Hamilton & Schwab, 1985).

authors point out themselves, such an indirect approach often suffers from the criticism that their findings of mispricing might be the result of model misspecifications rather than of any meaningful effect.

During the most recent housing boom, Piazzesi and Schneider (2009), using the Michigan Survey of Consumers, conclude there was an increasing cluster of "momentum" agents who believed prices would rise further. Prior to the bursting of the bubble, Case and Shiller (2003) conducted a survey of homebuyers in four U.S. MSAs. They find that home purchase decisions during this period were driven to a large extent by optimistic expectations and "excitement" – an evident sign of a bubble. In addition, euphoria on the part of banks also played a crucial role in the recent bubble (Archer and Smith, 2011). During periods of rising prices, euphoria induces borrowers to willingly take on more risk; in addition, underwriters rationalize their risky lending by pointing to anticipated house price increases and the associated reduction in perceived default risk. Shleifer and Vishny (2010) also propose a theoretical model in which banks cater to investor sentiment through their participation in financial markets. Overall, the literature suggests that although market agents make rational use of much of the information available to them, misbeliefs are still rather pervasive in housing markets.

## 4. Summary

In general, most economists agree that house prices are dictated in the long-run by economic fundamentals such as income, population, employment and rents. However, in the short-run market imperfections might lead to house prices overshooting or undershooting the fundamentals. Despite significant variation in hypotheses, modeling techniques, and data, much of the existing empirical literature acknowledges that actual house prices are seen rather volatile and susceptible to bubbles, which are hard to perfectly reconcile with rational models. Some researchers have therefore suggests non-rational explanations such as exuberance and unrealistic expectations of future price appreciation.

Thus far, no research has directly examined the dynamic relation between sentiment and house prices. Rather, sentiment is often inferred indirectly through price deviations from fundamental values based on real construction cost inflation, income multiples, or present value of rents. However, the main limitation of attributing a portion of the perceived mispricing to sentiment is the absence of precise valuation models. This makes it difficult to attribute the deviations convincingly to either actual mispricing or model misspecification. Unlike previous work, this study employs a set of proxies that appear to capture the consensus sentiment of three major agents in the U.S. housing markets: homebuyers (demand side), builders (supply side) and residential mortgage lenders (intermediaries). This allows for a direct examination of the dynamic relationship between sentiment and house prices,

offering explicit evidence of the timing, magnitude and direction of their impact on each other.

# Chapter 3 - Data and Methodology

## 1. Empirical model – Vector Autoregression

A Vector Autogression (VAR) model is an *n*-equation, *n*-variable linear model in which each variable is in turn explained by its own lagged values, together with past values of the remaining *n - 1* variables. An unrestricted $p^{th}$-order Gaussian VAR model can be represented as:

$$Y_t = \mu + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \cdots + \Pi_p Y_{t-p} + \varepsilon_t \qquad (1)$$

where $Y_t$ is a vector of variables; $\mu$ is a $(p \times 1)$ vector of intercepts; $\Pi_1, \Pi_2 \ldots, \Pi_{t-p}$ are $(p \times p)$ matrices of parameters with all eigenvalues of $\Pi$ having moduli less than one so that the VAR is stationary; and $\varepsilon_t$ is a vector of uncorrelated structural shocks [$\sim$ NID(0,Ω)]. Since the VAR(p) is in the form of a Seemingly Unrelated Regression model where each equation has the same explanatory variables, each equation may be estimated separately by Ordinary Least Squares without loss of efficiency.

In this research, the dynamic relationship between housing market sentiment and housing prices is modeled in a VAR framework with two endogenous variables both expressed as linear functions of their own and each other's lagged values:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i} SENTIMENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i} RETURN_{t-i} + \delta z_t + \nu_{1t} \quad (2)$$

$$SENTIMENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i} SENTIMENT_{t-i} + \sum_{i=1}^{q} \gamma_{2i} RETURN_{t-i} + \nu_{2t} \quad (3)$$

where $RETURN_t$ is the percentage change in real house prices and $SENTIMENT_t$ represents one of the three proxies for homebuyers', builders' and lenders' sentiment. It is essentially a Granger-causality test between market sentiment and house price appreciation. The model includes constant terms to allow for the possibility that the series have non-zero means[11].

The relation between sentiment and house prices is modeled using a lag structure to avoid endogeneity issues and to recognize the nature of housing data. Since a sale transaction typically takes at least three months to be completed, any price movement in a quarter could only be reflected in the next quarter's statistics. As a result, in practice the impact of sentiment on current prices can only be observed with lags. On the other hand, data on house prices (measured by the Case-Shiller U.S. Home Price Index) are released by Standard and Poors only with a two-month lag. Its effect on sentiment, if any, should therefore be felt in the next quarter rather than the concurrent quarter.

In addition to a lag structure of the endogenous variables, Equation (2) controls for a set of economic fundamentals ($z_t$) that have been shown in the literature to be important determinants of housing prices. These control variables include the growth in the population of individuals between 20 and 30 years of age

[11] The mean of RETURN (reported in Table 1 below) is significantly different from zero, while the means of the SENTIMENT variables are zero. For robustness, a model with suppressed constant terms has also been estimated (results not reported), which yielded similar results to the reported model.

*(POP)*, real GDP growth *(GDP)*, real income growth *(INC)*, the change in unemployment rate *(UNEMP)*, the change in the real interest rate *(INT)*, and the change in housing supply *(SUPPLY)*. Changes in the control variables are measured from quarter *t-1* to *t*. Maximum Likelihood Estimation (MLE) is employed to simultaneously estimate Equations (2) and (3) using quarterly data over the 1990:Q2 - 2010:Q3 sample period.

The VAR system specified above allows for a feedback relationship between sentiment and house prices. Of primary interest in this study is the estimated coefficients on sentiment ($\beta_{1i}$). The impact of market sentiment on housing prices is arguably through the demand side. When potential buyers believe that prices are very unlikely to fall, they are more willing to make housing investments because of either less perceived risk or high potential price appreciation.[12] Riddle (1999) posits that once price begins to rise, positive feedback traders enter the market in search of momentum profits and accentuate the rise; in contrast, during periods of price declines feedback traders exacerbate price movements through their selling activities. In the recent housing boom, Case and Shiller (2003) have also demonstrated that many homebuyers made their purchase decisions for fear of being unable to afford one subsequently. Moreover, financial intermediaries were believed to be the key to translate this overoptimism into real actions through their risky lending activities (Archer and Smith, 2011).

[12] The two major types of buyers that will be highly motivated include (1) first-time homebuyers who act swiftly for fear of being priced out of the market, and (2) speculators who search for excess returns from the price momentum. These optimistic buyers will further be encouraged with easier and cheaper access to credit if developers and creditors share the same optimism. As a result, the market will experience an increase in speculative demand, driving prices away from fundamental values. The reverse can occur in "gloomy" periods.

Evidence of the reverse causality from prices to sentiment suggests myopic expectations. When prices are increasing (decreasing), market participants (mistakenly) believe that prices will continue to increase (decrease). The estimated coefficients on $\gamma_{2i}$ is thus expected to be positive. The dynamic relationship between sentiment and home prices may therefore "spiral", with one reinforcing the other. The illiquidity, low transaction volumes, limits to arbitrage, and information asymmetry that characterize housing markets may reinforce such a "spiral" effect and extend the length and magnitude of the cycle, as seen in the recent bubble.

## 2. Measure of housing prices

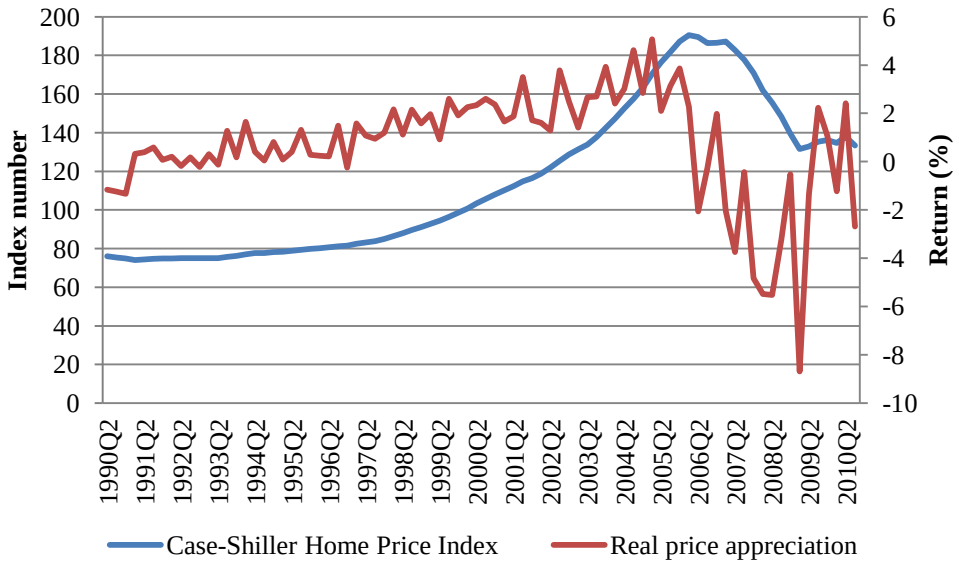
To measure house price appreciation ($RETURN_t$), this study uses the percentage change in the real Case-Shiller U.S. National Home Price Index (nominal index deflated by the CPI-less-shelter). The Case-Shiller U.S. National Home Price Index is a composite price index for single-family homes from all nine U.S. Census divisions, released quarterly by Standard and Poors (see www.standardandpoors.com). It is constructed using the repeat sale method developed by Case and Shiller in the 1980s based on matched sale pairs of existing single-family homes.

Over the past 20 years, U.S. home prices as measured by the Case-Shiller index in ***Figure 1*** have experienced three episodes: stagnation, escalation and collapse. The index stayed relatively flat over a long period between 1990-1996 following the housing burst in 1989. Since 1997, the subsequent ten years saw prices

steadily and rapidly increase to a historical high level in an infamous housing bubble that has spurred abundant debates on its sustainability. Research have largely focused on the role of cheap and easy credit (Glaeser, Gottlieb and Gyourko, 2010), subprime lending (Mayer and Pence, 2009; Lai and Van Order, 2010), and speculation (Wheaton and Nechavey, 2008) in driving the bubble but acknowledge their inability to fully account for the excessive price run-up. When the bubble burst, over just twelve quarters prices plunged by 31% from its peak in 2006Q1, bottoming out in 2009 and remaining stagnant in the last two years.

**Figure 1. U.S. house prices 1990Q2-2010Q3**

This figure plots the Case-Shiller National U.S. Home Price Index and real price appreciation over the 1990Q2-2010Q3 period. Real price appreciation is measured by the quarter-on-quarter change (in percentage) in the Case-Shiller index deflated by the CPI-less-shelter.




In terms of real price appreciation, measured by the quarter-on-quarter change in the nominal Case-Shiller index deflated by the CPI-less-shelter, ***Figure 1*** shows that, notwithstanding some periodic fluctuations, growth rates were also evidently

trending upwards until 2005. This confirms that in parallel with nominal prices, real prices were also escalating during the bubble. Similarly, the rate of real price increase started to moderate before prices fell and became negative in 2006 as the bubble burst. After three years of rapid depreciation, real prices seemed to show minor signs of recovery with some positive growth rates in 2009-2010 but a solid trend is still yet to emerge.

## 3. Measure of housing market sentiment

Sentiment, broadly defined in behavioral finance research, is the misguided belief about future cash flows and investment risks that is not justified by current information (Baker and Wurgler, 2007). Several proxies of investor sentiment have been developed for stock and commercial markets.[13] These sentiment indicators can be broadly categorized as either direct or indirect measures. Indirect proxies for investor sentiment are abstracted from a broad range of quantifiable market indicators, such as closed-end fund discounts, trading volume, mutual fund flows, IPO and SEO volume. In contrast, direct sentiment measures are often derived from surveys of market participants intended to capture their outlook about the markets, such as the UBS/Gallup surveys of randomly-selected investor households, Investor's

[13] Baker and Wurgler (2007) provide a comprehensive discussion of the various common proxies for investor sentiment in the stock market. Note that investor sentiment is by no means a new area in the securitized real estate literature. For example, see Barkham and Ward (1999), Chiang and Lee (2010), and Lin, Rahman and Yung (2009) for the impact of investor sentiment on REIT prices as well as discounts to net-asset-value. In recent years, more attention has also been paid to the noise trader approach in private real estate markets for commercial properties (Clayton, Ling and Naranjo, 2009; Ling, 2005; Ling, Naranjo and Scheick, 2010).

Intelligence surveys of financial newsletter writers, and the University of Michigan Survey of Consumers.[14]

This thesis focuses on the direct measures, or survey-based measures, of market sentiment in its analysis. Baker & Wurgler (2007) and Case and Shiller (2003) have argued that simply by asking people what they are thinking, researchers can gain insights into their decision making process or emotional charge. For example, Carroll, Fuhrer, & Wilcox (1994) use the University of Michigan Index of Consumer Sentiment to forecast homebuyer spending and study if sentiment indeed causes changes in spending. For investor sentiment in the stock market, Brown & Cliff (2004, 2005) use Investors' Intelligence surveys to calculate the "bull-bear" spread as the percentage of bullish stock investment newsletters minus the percentage deemed bearish, and relate it to stock price deviations from fundamental values. In the real estate literature, Ling (2005), Clayton, Ling, and Naranjo (2009) and Ling, Naranjo, and Scheick (2010) employ survey data published by the Real Estate Research Corporation (RERC) as a measure of investor sentiment to examine its impact on commercial property returns. RERC surveys institutional real estate investors, appraisers, lenders, and managers throughout the US to gain information about their expected rates of return, property selection criteria, and investment outlook.

Along similar lines, the sentiment indicators used in this paper are taken from surveys of homebuyers, builders, and mortgage lenders by various institutions in the

[14] Qiu and Welch (2005) provide a comparison of several direct survey-based measures of investor sentiment.

U.S. These are the three major types of agents representing the demand, supply and intermediary side of the housing markets.

## 3.1. Measure of homebuyers' sentiment

The sentiment of homebuyers is addressed in the *Survey of Consumers* conducted monthly by the Survey Research Center at the University of Michigan (see http://www.sca.isr.umich.edu/). Founded in 1946, the center was developed on behavioral economics theories that focus on the "human factor" in economic affairs (Curtin, 1982)[15]. The survey has become a familiar source of information on consumer sentiment that is widely used by not only researchers but also practitioners and institutions[16].

For each monthly sample of respondents, approximately 500 households from all states in the U.S. are chosen using a rotating panel sample design. Each survey contains approximately 50 core questions to track consumers' attitudes and expectations about three main areas: personal finances, business conditions, and buying conditions[17]. For housing sentiment, this thesis focuses on the survey question

[15] Since change in attitudes and expectations occurs in advance of action, measures of consumer attitudes and expectations can act as leading indicators of aggregate economic activities (Curtin, 1982).

[16] The Index of Consumer Expectations, produced from the survey, is included in the Leading Indicator Composite Index published by the U.S. Department of Commerce, Bureau of Economic Analysis.

[17] Personal finances are addressed in questions about expected change in nominal as well as real family income. Attitudes towards business conditions in the economy are measured using questionnaire concerning expected changes in inflation, unemployment, interest rates, and confidence in government economic policies. Finally, there are several questions probing for the respondent's

addressing respondents' attitudes about home buying conditions: "Generally speaking, do you think now is a good time or a bad time to buy a house?" The responses fall into three broad categories: "good", "bad" and "uncertain". The follow-up question asks respondents to provide reasons for their previous answers, which are then classified into six groups for the "good" response and four groups for the "bad" response. The six reasons for the "good" response include: "prices will increase", "prices low", "interest rates low", "rising interest rates", "good investment", and "time's good". From these responses, the percentage of respondents who think it is a "good" time to buy "because price will increase" is chosen as the sentiment proxy for homebuyers[18] *(BUYER)*. This percentage captures the proportion of homebuyers who are bullish about the market due to their optimistic expectations for rising prices. The survey data of the ending month of a given quarter (March for quarter 1, June for quarter 2, September for quarter 3, and December for quarter 4) is used as an indicator for homebuyers' sentiment in that quarter.

***Figure 2*** plots the proportion of "optimistic" respondents against the real price appreciation series (*RETURN*) over the study period. Not unexpectedly, households' sentiment exhibits high volatility with no apparent pattern throughout the years, except in the recent bubble period when it moved in close tandem with price changes. Prior to the bubble, periods of remarkable low sentiment include 1991-1992, 1998Q4 and 2001Q4-2003Q1. Since 2003, parallel with the steady rising trend of

---

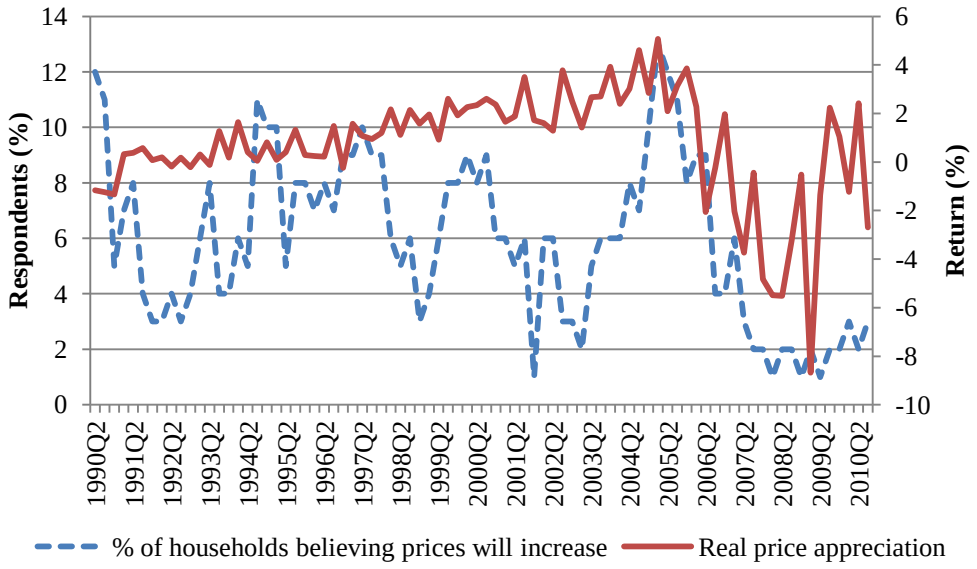
appraisal of buying conditions for large household durables, vehicles, and houses (see http://www.sca.isr.umich.edu/).

[18] The remaining reasons are not as clearly related to sentiment. Specifically, "prices low", "interest rates low", and "rising interest rates" are more likely to reflect prevailing market conditions than the subjective belief of the respondents, whilst "good investment" and "time's good" are too ambiguous for interpretation.

house prices there was an increasing number of optimistic households expecting prices to continue upwards, which peaked precisely in the same quarter price appreciation reached its historical high level (2005Q1). Both series then fell drastically following the burst of the bubble to reach their bottom at about the same time (2008). The last two years (2009-2010) saw some slight improvement in both price growth rate and homebuyers' sentiment.

**Figure 2. Homebuyers' sentiment versus real price appreciation, 1990Q2-2010Q3**

This figure plots the percentage of households who think it is a "good" time to buy a house "because price will increase" from the Michigan Survey of Consumers against real price appreciation rates over 1990Q2-2010Q3. Real price appreciation rate is measured by the quarter-on-quarter change (in percentage) in the Case-Shiller National Home Price Index less the change in the CPI-less-shelter.




## 3.2. Measure of homebuilders' sentiment

To measure the perceptions of homebuilders about housing market conditions *(BUILDER)*, this study employs the *NAHB/Wells Fargo Housing Market Index*

published by the National Association of Home Builders (NAHB). Founded in 1942, NAHB is a federation of more than 800 local associations of home builders and remodelers in the U.S. with the aim of being "the voice of America's housing industry" (see www.nahb.org). An important service by NAHB is providing housing data and in-depth analyses of the industry. Various indices of the housing market constructed by NAHB, such as the NAHB/Wells Fargo Housing Market Index, Housing Opportunity Index or Improving Market Index, are closely followed by various Wall Street firms, government officials, economic analysts, as well as news media.

NAHB reports the *NAHB/Wells Fargo Housing Market Index* monthly as an indicator of "builder confidence" about the housing market. Derived from a survey conducted monthly by NAHB since 1985, this index gauges the opinions of about 400 builders regarding three aspects of housing market conditions: (1) current sales of single-family new homes, (2) expected sales of single-family new homes over the next six months, and (3) traffic of prospective buyers in new homes. The respondents are asked to rate their perceptions of the current and expected sales as "good," "fair" or "poor", and the traffic of prospective buyers as "high to very high," "average" or "low to very low". The final Housing Market Index *(BUILDER)* is a weighted average of the three component indices[19], where their weights are based on their correlations with single-family housing starts. The index can range between 0 and 100, with an index number over 50 indicating that more builders view sales

[19] The three component indices are calculated by applying the formula [(Good - Poor + 100)/2] to the present and future sales series, and [(High/Very High - Low/Very Low + 100)/2] to the traffic series. This formula puts each index on a scale ranging from 0 to 100. The indices are then seasonally adjusted (see http://www.nahb.org/generic.aspx?sectionID=134&genericContentID=532).

conditions as good than poor. Similar to the homebuyer survey series, this research uses the index value of the ending month of a given quarter as an indicator of builders' sentiment for that quarter.

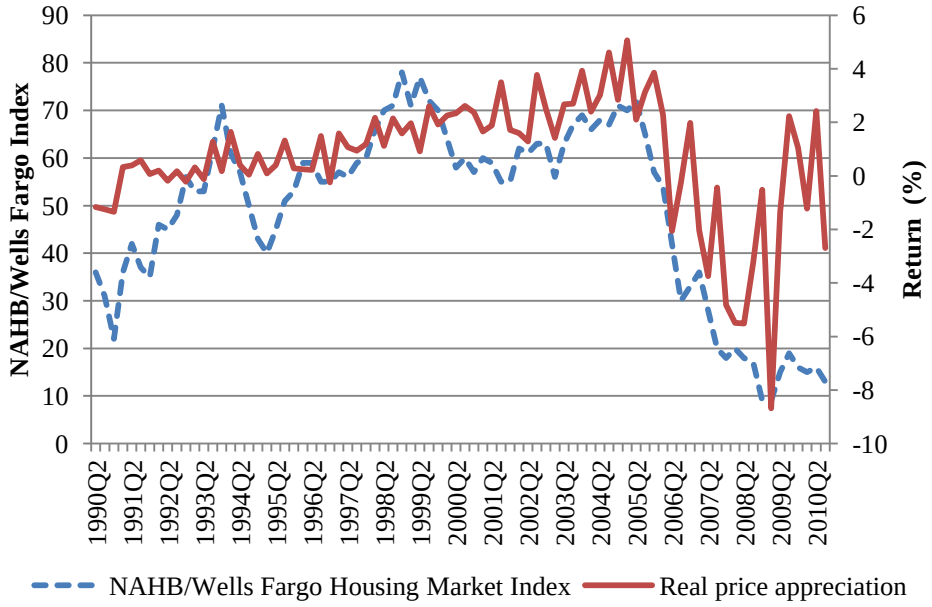
As shown in ***Figure 3***, builders' confidence level was generally increasing from 1990 to 2006 with the exceptions of some major slippage in 1994 and 1999-2001. Both of these downward movements in the confidence index seemed to coincide with two periods of rising interest rates that have a pronounced dampening effect on the mortgage market, which in turn affected home sale prospects. The U.S. economy also experienced a 9-month long recession in 2001 which partly contributed to the low sentiment during 1999-2001. However, overall the index remained high above 50 for more than ten years since 1996, corresponding with increasing price growth over this period. It can be observed that since 2002 throughout the recent bubble, there is a remarkably high correlation between movements in builders' sentiment and real price appreciation rate. At the worst of the crisis the index fell sharply to a historical low of 9 from its previous peak of 72 in 2005Q2. Builders have not shown much improvement in their sentiment even until 2010Q3.

### 3.3. Measure of mortgage lenders' sentiment

The measure of credit suppliers' opinion about housing markets is obtained from the Federal Reserve Board (FED)'s *Senior Loan Officer Opinion Survey on Bank Lending Practices* (see www.federalreserve.gov). This is a quarterly survey of approximately sixty large domestic banks and twenty-four US branches or agencies

**Figure 3. Homebuilders' sentiment versus real price appreciation, 1990Q2-2010Q3**

This figure plots the NAHB/Wells Fargo Housing Market Index against real price appreciation rates over1990Q2-2010Q3. Real price appreciation rate is measured by the quarter-on-quarter change (in percentage) in the Case-Shiller National Home Price Index less the change in the CPI-less-shelter.




of foreign banks. The purpose of the survey is to provide qualitative and limited quantitative information on credit availability and demand, as well as evolving developments and lending practices in the U.S. loan markets[20]. The survey results are reported regularly to the Board of Governors and to the Federal Open Market Committee. Information from the survey also attracts much attention from both the business press and academic research on banking and macroeconomic activities.

[20] Survey questions focus on two broad areas: changes in demand for credit, and changes in bank lending policies. Various types of loans are covered in the survey, including commercial and industrial loans, commercial real estate loans, residential real estate loans, and consumer lending. In addition, a portion of the questions in each survey also covers special topics of timely interest, such as the securitization of mortgage loans, the financial crisis, etc.

This study focuses on the responses to the following survey question: "Over the past three months, how have your bank's credit standards for approving applications from individuals for mortgage loans to purchase homes changed?" Respondents must select one of the following options: "tightened considerably", "tightened somewhat", "remained essentially unchanged", "eased somewhat", or "eased considerably". Since banks are concerned with the ability to recover their loans, it follows that that changes in lending standards for home mortgages reflect their changing perspectives about the riskiness of the housing markets[21]. Thus, in this thesis lenders' sentiment *(LENDER*) is indicated by the net percentage of banks easing their standards (percentage of banks easing their standards minus percentage of banks tightening their standards).

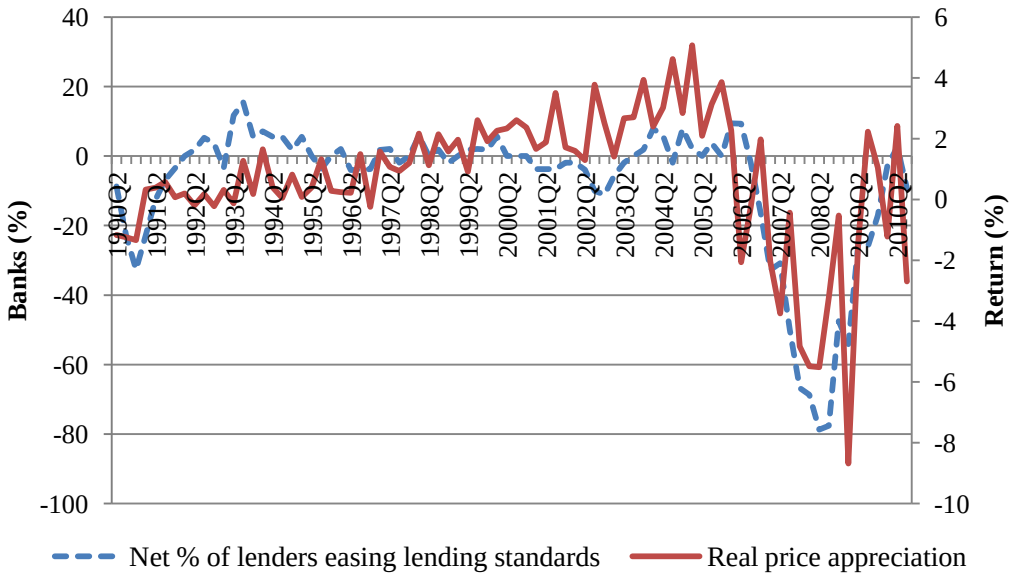
Unlike the homebuyers' and builders' series, lenders' sentiment as shown in ***Figure 4*** stayed relatively stable over the study period prior to the turmoil in 2006. Their sentiment saw significant improvement in only one year after the 1989 housing bust as more and more banks loosened their credit standards. The series peaked in 1993Q2 and stabilized thereafter. The year 2006 marked the start of an exceptionally low sentiment period as the housing market burst. At the worst of the crisis as many as 78% of the surveyed lenders tightened their credit lines for the home mortgage

[21] Take, for example, the 2010Q4 survey. The top three reasons cited by the respondents for tightening their credit standards for commercial and industrial loans include: "less favorable or more uncertain economic outlook", "reduced tolerance of risk", and "increased concerns about the effects of legislative changes, supervisory actions, or changes in accounting standards". Conversely, banks easing their credit standards for commercial and industrial loans cited "more favorable or less uncertain economic outlook", "improvement in industry-specific problems" and "more aggressive competition from other banks or nonbank lenders" as their motivation. Apparently when making loan initiation, banks take into account their outlook for the riskiness of the specific market they are lending to. Although there is no question to address the reasons for changing credit standards for home mortgage loans, it is reasonable to believe banks are motivated by similar factors when making loans to the residential market.

market. Similar to other market agents, banks' sentiment seemed to show signs of recovery in the last two years along with real price growth.

**Figure 4. Lenders' sentiment versus real price appreciation, 1990Q2-2010Q3**

This figure plots the net percentage of banks easing their lending standard (percentage of banks easing their standards minus percentage of banks tightening their standards) against real price appreciation rates over1990Q2-2010Q3. Real price appreciation rate is measured by the quarter-on-quarter change (in percentage) in the Case-Shiller National Home Price Index less the change in the CPI-less-shelter.




## 3.4. Orthogonalization of the sentiment series

*BUYER*, *BUILDER* and *LENDER* are likely to be correlated with economic factors such as GDP or income per capita.[22] As highlighted by Brown and Cliff (2005), when market participants say they are bullish on the market, this can be a rational reflection of prosperous times to come, an irrational hope for the future, or

[22] For example, the contemporaneous correlations of *BUYER*, *BUILDER* and *LENDER* with real GDP growth are 0.28, 0.54, and 0.55 respectively.

some combination of the two. Following Clayton, Ling and Naranjo (2009) and Baker and Wurgler (2007), this study orthogonalizes the three sentiment series against a set of macroeconomic variables to remove the influence of fundamentals. Their residuals are then used in the subsequent VAR analysis. The macroeconomic factors used in the orthogonalization regression include population growth, GDP growth, income growth, change in unemployment rate, and change in interest rate. As a robustness test, this thesis further constructs a composite sentiment index from the three indices which will be discussed in detail in a later section.

The descriptive statistics of the sentiment indices and house price appreciation variable are reported in ***Table 1***. Over the study period, real house prices (*RETURN*) registered an average quarterly appreciation rate of 0.71 per cent, with both the maximum (5.07 per cent in 2005Q1) and minimum (-8.68 per cent in 2009Q1) appreciation occurring during the recent bubble period. The three sentiment indices *BUYER, BUILDER* and *LENDER* are well-behaved mean-zero series, as they are residuals from regressing the original series against a set of fundamental factors.[23]

***Figure 5*** plots the three sentiment indices against contemporaneous real house price changes over the 1990:Q2-2010:Q3 study period, in place of ***Figure 2*** - ***Figure 4*** above. Note that after removing the information on the current economic conditions, lenders' sentiment shows more variations when compared with ***Figure 4*** whilst the homebuyers' and builders' indices stay close to their original series. Among the three indices, home buyer sentiment is the least correlated with house

[23] Tests with sentiment indices in changes have also been carried out, but sentiment in levels work better for the purpose of this thesis.

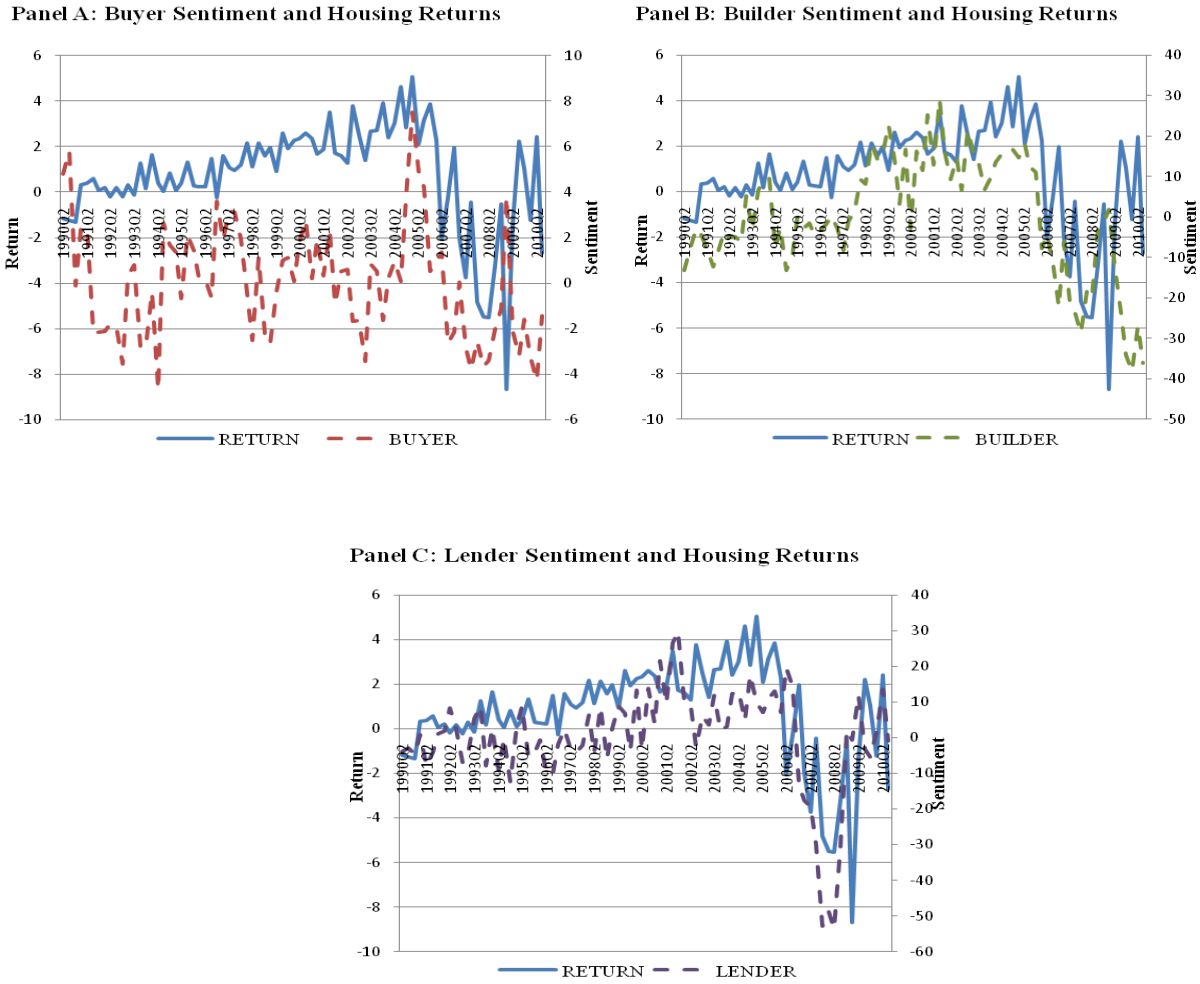
## Table 1. Definition and descriptive statistics of dependent and sentiment variables

This table reports the descriptive statistics for the sentiment proxies and house price appreciation variables. The study period covers 1990Q2-2010Q3. *, ** and *** denote 10%, 5% and 1% significant levels, respectively.

| Definition | Variable | Mean | St.D. | Min | Max | Serial correlation |
|---|---|---|---|---|---|---|
| ***Dependent variable*** | | | | | | |
| Change in Case-Shiller real price index (percentage) | RETURN | 0.71 | 2.31 | -8.68 | 5.07 | 0.60*** |
| ***Sentiment indices*** | | | | | | |
| Percentage of respondents indicating it is good time to buy a home because price will increase (residuals from orthogonalization) | BUYER | 0.00 | 2.52 | -4.67 | 7.53 | 0.52*** |
| NAHB/Wells Fargo Housing Market Index. An index number over 50 indicates that more builders view sales conditions as good than poor (residuals from orthogonalization) | BUILDER | 0.00 | 14.65 | -38.06 | 28.98 | 0.82*** |
| Percentage of banks easing home mortgage lending standards less percentage of banks tightening standards (residuals from orthogonalization) | LENDER | 0.00 | 14.55 | -53.67 | 29.75 | 0.72*** |
| First principal component derived from BUYER, BUILDER and LENDER | DSENT | 0.00 | 1.36 | -3.92 | 2.78 | 0.75*** |
| First principal component derived from home sales transaction volume, residential mortgage flows, and housing starts | INSENT | 0.00 | 1.61 | -4.74 | 2.79 | 0.86*** |

## Figure 5. Sentiment and contemporaneous housing returns

This figure plots the sentiment measures against contemporaneous housing returns for the time period 1990:Q2-2010:Q3. The buyer sentiment measure *(BUYER)* is the residuals from regressing the percentage of homebuyers who think it is a good time to buy a home because price will increase on a set of macroeconomic factors. The builder sentiment measure *(BUILDER)* is the residuals from regressing the NAHB/Wells Fargo Housing Market Index on a set of macroeconomic factors. The lender sentiment measure *(LENDER)* is the residuals from regressing the net percentage of banks easing their standards (percentage of banks easing their standards minus percentage of banks tightening their standards) on a set of macroeconomic factors. The set of macroeconomic factors used for orthogonalization include population growth, GDP growth, income growth, change in unemployment rate, and change in interest rate. The measure of housing returns is the quarterly changes in the real Case-Shiller US National Home Price Index.



price movements, especially during the 1995-2005 period when housing returns and the other two sentiment indices exhibit a consistent upward trend. Homebuyer sentiment, on the other hand, seems to move rather erratically during that period. This observation is confirmed through the contemporaneous correlation between house price changes and the sentiment indices, which are 0.21, 0.55 and 0.61 for homebuyers, builders and lenders respectively.

## 4. Control variables

In addition to the sentiment of the major market participants, housing prices are also modeled as depending on various macroeconomic factors in Equation (2). The control variables (reported in ***Table 2)*** are specified in changes to reduce multicollinearity and ensure stationarity for VAR estimations. Data for these control variables are obtained from various sources, including the Bureau of Economic Analysis, Bureau of Census, and Bureau of Labor Statistics. On average, the U.S population between 20 and 30 years old, who are in their home-buying age (Mankiw and Weil, 1989), grew by 0.09% every quarter over the past 20 years. Meanwhile, the U.S economy expanded by 0.62% whereas real income increased by 0.41% per quarter. Notwithstanding the booming economy, the unemployment rate showed an average quarterly growth rate of 0.84% with the sharpest increase, not unexpectedly, due to the latest crisis. Housing supply as measured by the number of new housing units built in each quarter also increased by 0.55% every quarter. Finally, interest rate is the only control factor experiencing a reduction over the study period with an

average quarter-on-quarter change of -0.08%, consistent with the common accusation that the bubble was partly driven by abundant cheap credit.

**Table 2. Definition and descriptive statistics of control variables**

This table reports the descriptive statistics for our control variables. The study period covers 1990Q2-2010Q3. *, ** and *** denote 10%, 5% and 1% significant levels, respectively.

| Definition | Variable | Mean | St.D. | Min | Max | Serial Correlation |
|---|---|---|---|---|---|---|
| Change in the population aged between 20-30 (percentage) | POP | 0.09 | 0.27 | -0.45 | 0.45 | 0.96*** |
| Real GDP growth (percentage) | GDP | 0.62 | 0.64 | -1.74 | 1.95 | 0.49*** |
| Real income growth (percentage) | INCOME | 0.41 | 0.84 | -2.41 | 2.47 | -0.27 |
| Change in unemployment rate (percentage) | UNEMP | 0.84 | 4.97 | -6.11 | 18.84 | 0.69*** |
| Change in real mortgage interest rate (percentage) | MGTRATE | -0.08 | 12.41 | -31.62 | 74.66 | -0.27 |
| Change in the quarterly number of units completed (percentage) | SUPPLY | 0.55 | 15.72 | -39.08 | 38.57 | -0.34*** |

Overall, ***Table 3*** shows that the sentiment series exhibit high positive correlations with each other but low correlations with other macroeconomic variables. It can therefore be concluded that multicolinearity is not a problem in the model, and more importantly, that the sentiment indicators reflect the sentiment component independent of any economic forces. This allows the estimated sentiment coefficients to be interpreted as the marginal effects of sentiment on house prices beyond fundamentals.

## Table 3. Correlation matrix between explanatory variables

This table reports the pair-wise correlation matrix between the explanatory variables. The sentiment indices include: *BUYER* is the residuals from regressing the percentage of homebuyers who think it is a good time to buy a home because price will increase on a set of macroeconomic factors; *BUILDER* is the residuals from regressing the NAHB/Wells Fargo Housing Market Index on a set of macroeconomic factors; *LENDER* is the residuals from regressing the net percentage of banks easing their standards (percentage of banks easing their standards minus percentage of banks tightening their standards) on a set of macroeconomic factors; *DSENT* is the first principal component extracted from the buyer sentiment measure *(BUYER)*, builder sentiment measure *(BUILDER)* and lender sentiment measure *(LENDER); INSENT* is the first principal component extracted from housing transaction volume, residential mortgage flows and housing starts. The control variables include: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter.

| **Variable** | | **BUYER** | **BUILDER** | **LENDER** | **DSENT** | **INSENT** | **POP** | **GDP** | **INCOME** | **UNEMP** | **MGTRATE** | **SUPPLY** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***Sentiment indices*** | | | | | | | | | | | | |
| Buyer sentiment | **BUYER** | 1.00 | | | | | | | | | | |
| Builder sentiment | **BUILDER** | 0.36 | 1.00 | | | | | | | | | |
| Lender sentiment | **LENDER** | 0.33 | 0.56 | 1.00 | | | | | | | | |
| Direct index | **DSENT** | 0.68 | 0.84 | 0.82 | 1.00 | | | | | | | |
| Indirect index | **INSENT** | 0.35 | 0.74 | 0.42 | 0.65 | 1.00 | | | | | | |
| ***Other control variables*** | | | | | | | | | | | | |
| Population | **POP** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | | | | | |
| GDP | **GDP** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.13 | 1.00 | | | | |
| Income | **INCOME** | -0.01 | 0.07 | 0.09 | 0.07 | 0.05 | -0.04 | 0.31 | 1.00 | | | |
| Unemployment | **UNEMP** | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.13 | -0.70 | -0.22 | 1.00 | | |
| Mortgage rate | **MGTRATE** | -0.06 | -0.03 | 0.05 | -0.01 | -0.04 | 0.05 | -0.10 | 0.03 | -0.04 | 1.00 | |
| Supply | **SUPPLY** | -0.06 | 0.06 | 0.11 | 0.06 | 0.10 | -0.04 | 0.25 | 0.12 | -0.14 | 0.14 | 1.00 |

# Chapter 4 - Empirical results

## 1. VAR estimation results of base models

The Dickey-Fuller tests of all variables shown in ***Table 4*** confirm that they are stationary. ***Table 5*** presents the estimation results of the base VAR model specified in Equation (2) and (3) using MLE. Conventional lag-length selection criteria (Akaike information criterion, Schwarz information criterion, Hannan-Quinn criterion) indicate that two lags of all endogenous variables are most appropriate for the model specified. The first pair of equations in ***Table 5*** corresponds to the homebuyer sentiment index *(BUYER)*, the second pair the builder sentiment index *(BUILDER)* and the last pair the lender sentiment index *(LENDER*). Panel A reports the estimated coefficients and their individual significance levels while Panel B presents the sum of the lagged coefficients and tests of joint significance.

Focusing first on the *RETURN* equations, Panel B shows that the estimated coefficients on all three lagged sentiment measures are positive and significant at 1% level. That is, high levels of sentiment Granger-cause increased price appreciation in subsequent quarters. Examining the individual coefficients on lagged sentiment in Panel A, the two-quarter lagged terms appear more significant than the one-period lags, both statistically and economically. The impulse response functions (IRFs)

## Table 4. Augmented Dicker-Fuller tests of stationarity

This table reports the t-statistics and p-values from the Augmented Dicker-Fuller tests of stationarity of all variables used. The null hypothesis is that the data has a unit root.

| Variable | | t-statistics | p-value |
|---|---|---|---|
| Real price appreciation | RETURN | -4.373 | 0.0000 |
| *Sentiment indices* | | | |
| Buyers' sentiment | BUYER | -5.215 | 0.0000 |
| Builders' sentiment | BUILDER | -2.226 | 0.0144 |
| Lenders' sentiment | LENDER | -3.605 | 0.0003 |
| Direct sentiment index | DSENT | -3.188 | 0.0010 |
| Indirect sentiment index | INSENT | -1.66 | 0.0505 |
| *Other control variables* | | | |
| Population | POP | -1.387 | 0.0846 |
| GDP | GDP | -5.209 | 0.0000 |
| Income | INCOME | -11.779 | 0.0000 |
| Unemployment | UNEMP | -3.803 | 0.0001 |
| Mortgage rate | MGTRATE | -11.684 | 0.0000 |
| Supply | SUPPLY | -12.738 | 0.0000 |
| **Null hypothesis: The data has a unit root.** | | | |

associated with these VAR estimations are displayed in ***Figure 6***. They trace the impact of a one-standard-deviation shock to the residuals of the sentiment variables on subsequent house price changes. A positive shock to household, builder and lender sentiment produces a 13.6, 10.7 and 21.6 basis points increase in house prices in the next quarter, respectively. The price responses rise further by 49.9, 48.0 and 53.9 basis points, respectively, in the second quarter. In total, the effect of a one-standard-deviation sentiment shock amounts to a cumulative price increase of about 60 to 80 basis points within two quarters. These figures are not insignificant considering that the average price change over the whole sample period is 0.71 percent per quarter ***(Table 1).*** Taken together, these findings strongly support the

## Table 5. VAR estimation results of three sentiment indices

This table reports the VAR estimation results with sentiment and housing returns as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, SENTIMENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t}$$

$$SENTIMENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, SENTIMENT_{It-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \nu_{2t}$$

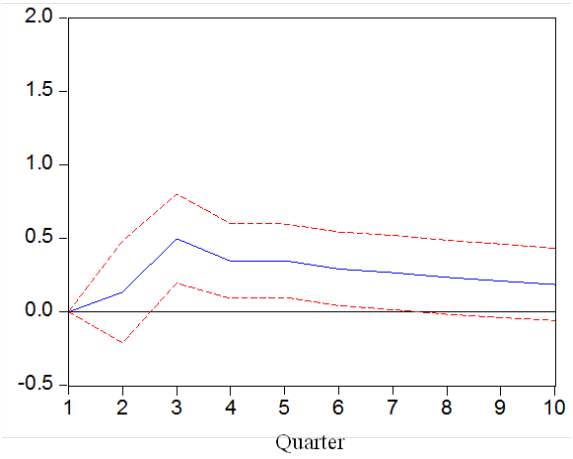
The study period covers 1990Q2-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *BUYER* is the homebuyer sentiment index; *BUILDER* is the builder sentiment index*; LENDER* is the lender sentiment index. All *RETURN* equations include a set of control variables $z_t$ but their coefficients are not reported for brevity: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter, and is used at the one-quarter lag to account for simultaneity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | ***Homebuyer sentiment*** | | ***Builder sentiment*** | | ***Lender sentiment*** | |
|---|---|---|---|---|---|---|
| | **RETURN** | **BUYER** | **RETURN** | **BUILDER** | **RETURN** | **LENDER** |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| **Panel A. Coefficient estimation results** | | | | | | |
| Constant | -0.260 | -0.376 | 0.323 | -0.676 | 0.314 | -1.200 |
| | (0.300) | (0.234) | (0.301) | (0.944) | (0.296) | (1.208) |
| $SENTIMENT_{t-1}$ | 0.0671 | 0.374*** | 0.0157 | 0.545*** | 0.0285* | 0.565*** |
| | (0.0792) | (0.109) | (0.0195) | (0.108) | (0.0157) | (0.110) |
| $SENTIMENT_{t-2}$ | 0.205*** | 0.0573 | 0.0600*** | 0.338*** | 0.0534*** | 0.0206 |
| | (0.0768) | (0.105) | (0.0199) | (0.114) | (0.0169) | (0.118) |
| $RETURN_{t-1}$ | 0.242** | 0.312** | 0.125 | -0.0188 | 0.102 | -0.167 |
| | (0.0976) | (0.124) | (0.0977) | (0.516) | (0.0984) | (0.675) |
| $RETURN_{t-2}$ | 0.301*** | 0.0171 | 0.153 | 0.478 | 0.186** | 1.890*** |
| | (0.0975) | (0.128) | (0.0981) | (0.515) | (0.0922) | (0.636) |
| POP | 0.0658 | | 0.105 | | 0.135 | |
| | (0.601) | | (0.572) | | (0.561) | |
| GDP | 1.025*** | | 0.474 | | 0.492 | |
| | (0.366) | | (0.347) | | (0.342) | |
| INCOME | -0.186 | | -0.215 | | -0.190 | |
| | (0.199) | | (0.190) | | (0.185) | |
| UNEMP | 0.0283 | | -0.0663 | | -0.00892 | |
| | (0.0473) | | (0.0483) | | (0.0455) | |
| MGTRATE | 0.0728*** | | 0.0636*** | | 0.0803*** | |
| | (0.0129) | | (0.0125) | | (0.0124) | |
| SUPPLY | 0.0197* | | 0.0162 | | 0.0151 | |
| | (0.0107) | | (0.0101) | | (0.00997) | |
| Observations | 80 | 80 | 80 | 80 | 80 | 80 |
| R-squared | 0.685 | 0.340 | 0.676 | 0.722 | 0.638 | 0.578 |
| **Panel B. Joint significance tests of estimated coefficients** | | | | | | |
| $SENTIMENT_{t-2 \text{ to } t-1}$ | 0.272*** | 0.431*** | 0.076*** | 0.883*** | 0.082*** | 0.586*** |
| *(F-stat)* | *(13.22)* | *(18.39)* | *(24.16)* | *(109.11)* | *(27.07)* | *(36.91)* |
| $RETURN_{t-2 \text{ to } t-1}$ | 0.543*** | 0.329*** | 0.278* | 0.4592 | 0.288* | 1.723*** |
| *(F-stat)* | *(30.26)* | *(10.44)* | *(5.57)* | *(1.01)* | *(7.16)* | *(9.56)* |

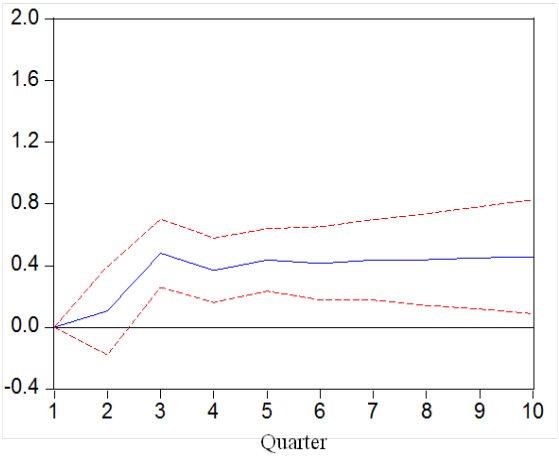
## Figure 6. Impulse response functions of housing returns to three sentiment measures

This figure plots the impulse response of housing returns to a one standard deviation shock in the three sentiment measures, corresponding to the estimated VAR results in *Table 5*. The red lines are the +/- Standard Error responses. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *BUYER* is the homebuyer sentiment index; *BUILDER* is the builder sentiment index; *LENDER* is the lender sentiment index.

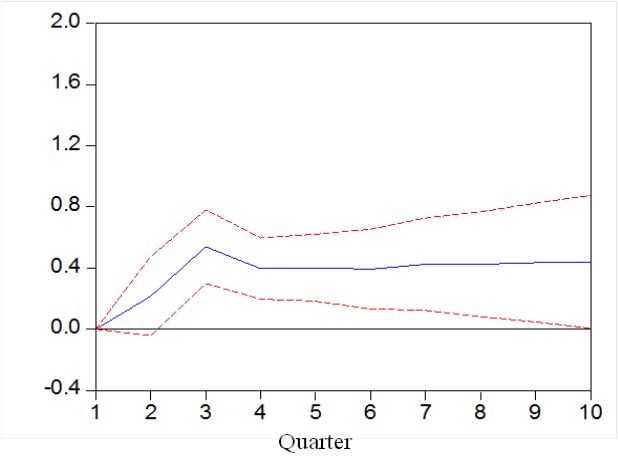
Panel A: Response of *RETURN* to *BUYER*




Panel B: Response of *RETURN* to *BUILDER*




Panel C: Response of *RETURN* to *LENDER*



main proposition in this study that sentiment is a crucial determinant of house prices. To put it another way, the well-documented high volatility of house prices compared with observable changes in fundamentals can be ascribed, at least in part, to housing market sentiment.

Interestingly, the impulse response functions plotted in ***Figure 6*** further show that the sentiment effect is highly persistent, which helps shed light on why housing markets are so susceptible to bubbles[24]. In particular, a shock in the sentiment index of any market-player appears to influence price changes for as long as 10 quarters. This protracted effect reflects the highly illiquid, segmented and informationally inefficient characteristics of housing markets. With some initial stimulating condition, such as an abundance of easy credit, the sentiment-induced mispricing inherent in house prices can spiral into a bubble. The findings in this research provide at least a partial answer to the question raised by Glaeser, Gottlieb and Gyourko (2010) about the interplay between bubble, beliefs and credit condition. Regarding the serial correlation of house prices, this study finds that the estimated coefficient on lagged returns are still positive and jointly significant after controlling for sentiment, even though the significance of lagged returns seems to weaken with the inclusion of builder sentiment in the model (Column 3 of ***Table 5***).

The following discussion turns to the three *SENTIMENT* equations for evidence on the reverse causality from house price movements to sentiment. As expected, homebuyers exhibit myopic expectations as evidenced by the positive and

[24] The impulse responses plotted using Cholesky factorization are essentially similar to the reported graphs.

significant coefficients on lagged returns in the *BUYER* equation, both jointly and individually. Homebuyers, presumably the least sophisticated and most informationally constrained agents in the market, have to rely on past price movements to form their expectations of future house price movements. In contrast, builders appear not to be backward looking in forming their expectations. These findings are consistent with the prior studies of Glaeser, Gyourko, and Saiz (2008) and Ooi and Le (2011) who show that developers lead rather than follow house price movements.

Finally, similar to home buyers, changes in lenders' underwriting standards are associated with past house price changes, but only at the two-quarter lag. Coupled with the persistent impact of sentiment on house prices established above, this finding of a positive feedback loop from prices to sentiment further indicates the potential for housing prices to spiral, which potentially renders housing markets highly prone to bubbles. As a final note, all homebuyer, builder and lender sentiment exhibit serial correlation as do home prices.

## 2. VAR estimation results with a composite direct sentiment index

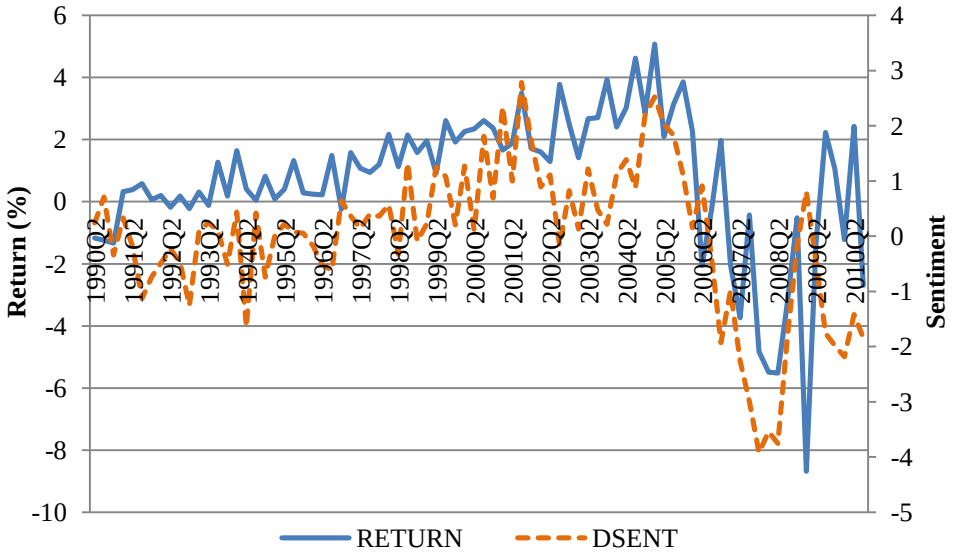
The three individual sentiment indices used in this research are highly correlated (see ***Table 3***). Thus, it is likely that the three indices have a common sentiment element which can be extracted using principal component analysis. This study therefore constructs a composite direct sentiment index *(DSENT)* as the first principal component of the three series, similar to Baker and Wurgler (2007) and

Ling, Naranjo, and Scheick (2010). The descriptive statistics of *DSENT* is reported in ***Table 1***.

Generally, *DSENT* tracks the overall trend of house price changes relatively well in ***Figure 7***, and the two time series have a high contemporaneous correlation of 0.60. ***Table 6*** reports the estimation results of Equation (2) and (3) again using *DSENT* in place of the three individual sentiment indices. The appropriate lag-length for the model is again two lags of endogenous variables.

**Figure 7**. **Composite sentiment index and contemporaneous housing returns**

This figure plots the composite direct sentiment index against contemporaneous housing returns for the time period 1990:Q2-2010:Q3. The composite direct sentiment index *(DSENT)* is the first principal component extracted from the buyer sentiment index *(BUYER)*, builder sentiment index *(BUILDER)* and lender sentiment index *(LENDER)*. Housing returns are measured by the quarterly changes in the real Case-Shiller US National Home Price Index.




The previous central finding from the individual surveys that sentiment strongly affects house prices appears slightly stronger with the composite direct

## Table 6. VAR estimation results of the composite direct sentiment index

This table reports the VAR estimation results with the composite direct sentiment index and housing returns as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t}$$
$$DSENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, DSENT_{It-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \nu_{2t}$$

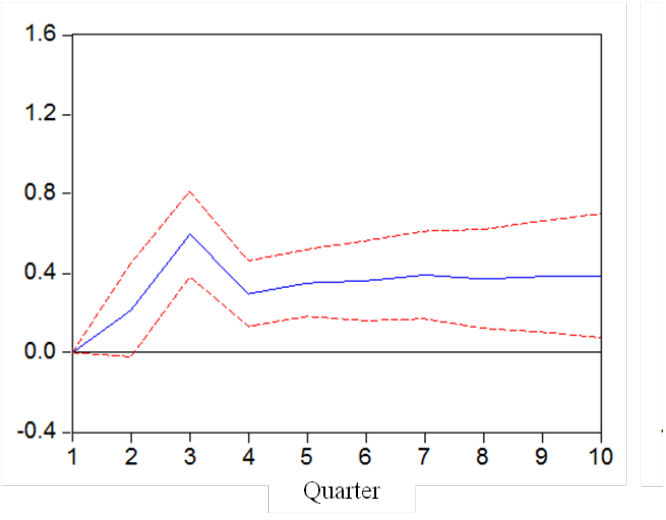
The study period covers 1990Q2-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *DSENT* is the first principal component extracted from the buyer sentiment measure *(BUYER)*, builder sentiment measure *(BUILDER)* and lender sentiment measure *(LENDER)*. The *RETURN* equation includes a set of control variables $z_t$ but their coefficients are not reported for brevity: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter, and is used at the one-quarter lag to account for simultaneity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | **RETURN** | **DSENT** |
|---|---|---|
| | **(1)** | **(2)** |
| **Panel A. Coefficient estimation results** | | |
| Constant | 0.312 | -0.157 |
| | (0.262) | (0.102) |
| $DSENT_{t-1}$ | 0.323** | 0.420*** |
| | (0.163) | (0.107) |
| $DSENT_{t-2}$ | 0.775*** | 0.222** |
| | (0.166) | (0.111) |
| $RETURN_{t-1}$ | 0.0337 | 0.0634 |
| | (0.0889) | (0.0565) |
| $RETURN_{t-2}$ | 0.0938 | 0.107* |
| | (0.0873) | (0.0557) |
| POP | 0.135 | |
| | (0.507) | |
| GDP | 0.678** | |
| | (0.305) | |
| INCOME | -0.239 | |
| | (0.169) | |
| UNEMP | -0.0330 | |
| | (0.0410) | |
| MGTRATE | 0.0713*** | |
| | (0.0110) | |
| SUPPLY | 0.0146 | |
| | (0.00908) | |
| Observations | 80 | 80 |
| R-squared | 0.744 | 0.645 |
| **Panel B. Joint significance tests of estimated coefficients** | | |
| $DSENT_{t-2 \text{ to } t-1}$ | 1.098*** | 0.642*** |
| *(F-stat)* | *(51.81)* | *(38.35)* |
| $RETURN_{t-2 \text{ to } t-1}$ | 0.1275 | 0.1704** |
| *(F-stat)* | *(1.58)* | *(7.17)* |

index. Additionally, the impulse response function in ***Figure 8*** confirms that the impact of sentiment on house prices is persistent over at least 10 quarters. However, there is only weak evidence of a relation between lagged house price changes and sentiment, probably due to the mixed results regarding builders versus homebuyers and lenders when using the three surveys separately. Specifically, only the two-period lag appears weakly significant at 10%, though its joint significance with the one-quarter lagged return slightly improves to 5%. Finally, while sentiment is still positively related to its past values, it seems that the serial correlation of house prices disappear when *DSENT* is controlled for.

**Figure 8. Impulse response function of housing returns to the composite direct sentiment index**

This figure plots the impulse response of housing returns to a one standard deviation shock in the composite sentiment index *DSENT* corresponding to the estimated VAR results in *Table 6*. The red lines are the +/- Standard Error responses. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *DSENT* is the first principal component extracted from the buyer sentiment index *(BUYER),* builder sentiment index *(BUILDER)* and lender sentiment index *(LENDER).*




To summarize, the above analysis suggests that the sentiment of important agents in housing markets is associated with mispricing in subsequent quarters.

Moreover, this mispricing persists over a long period. There is also evidence of myopic expectations in the sense that market agents form their beliefs from past price movements, with builders being a notable exception. The dynamic interplay between sentiment and home prices is therefore a self-reinforcing process, which presents a compelling explanation for the high susceptibility of housing markets to bubbles.

## 3. Sub-period analysis

To further compare the sentiment's effect in different market conditions, this section re-estimates the VAR models above with two sub-periods corresponding with a non-bubble (or normal) and a bubble market. One would intuitively expect sentiment to play a more significant role during the formation and bursting of a housing bubble than in a more normal period. An important step in this test is identifying the start of the recent housing boom. Since there is not yet any clear theoretical guide to define the beginning of a boom, this study follows the empirical approach in Ferreira & Gyourko (2011). They describe the structural breakpoint, or in other words, a significant discrete jump, in the growth rate of house prices as signaling an episode of booming. Essentially, the approach involves finding the quarter during which there is a global structural break by estimating the following equation:

$$RETURN_t = \alpha + \delta I(q_t \geq q_t^*) + \varepsilon_t \qquad \text{for } 1 < t < 61 \qquad (4)$$

where $RETURN_t$ is the real price change in quarter *t;* $\alpha$ is the intercept term; $\delta$ measures the importance of the potential breakpoint; $I(\cdot)$ is an indicator function

which equals 1 if its condition is true and 0 otherwise; $q_t$ is quarter *t;* $q_t^*$ is the location of the potential structural break, and $\varepsilon_t$ is the error term. Estimated structural breakpoint ($\hat{q}_t^*$) is the quarter that maximizes the $R^2$ of the equation; that is, the quarter in which the change in price growth has the highest power in explaining the price growth series itself (Ferreira & Gyourko, 2011). The regression is estimated using data only from 1990Q2 (t=1) to 2005Q1 (t=61) to avoid any influence of the burst period in locating the structural breakpoint[25].

The highest $R^2$ achieved by running Equation (4) is 59.43% at 1998Q1. Hence, this result suggests that the boom began in 1998Q1 when the structural break occurred. Graphically, it can also be observed from ***Figure 9*** that prior to 1998Q1, growth rates fluctuated around a value close to 0.5%, but since 1998Q1 they centered around about 2.5%.

The original study period is then split into two sub-periods based on this reference quarter (1990Q2-1997Q4 and 1998Q1-2010Q3) for re-estimating the VAR base models. The results are reported in ***Table 7***. Prior to the start of the boom (1990Q2-1997Q4), the effect of sentiment on future prices was still present but statistically weak (Column 1). Meanwhile, during the formation and bursting of the bubble the sentiment coefficients became much stronger in both magnitude and significance level as expected. In particular, the change in price growth induced by a change in market sentiment during the bubble was 2.5 times larger than that in a normal market. These findings suggest that market sentiment had a crucial role to

[25] 2005Q1 is chosen because it is the quarter with the highest real price changes *(RETURN).* Thereafter, price growth started to slow down.

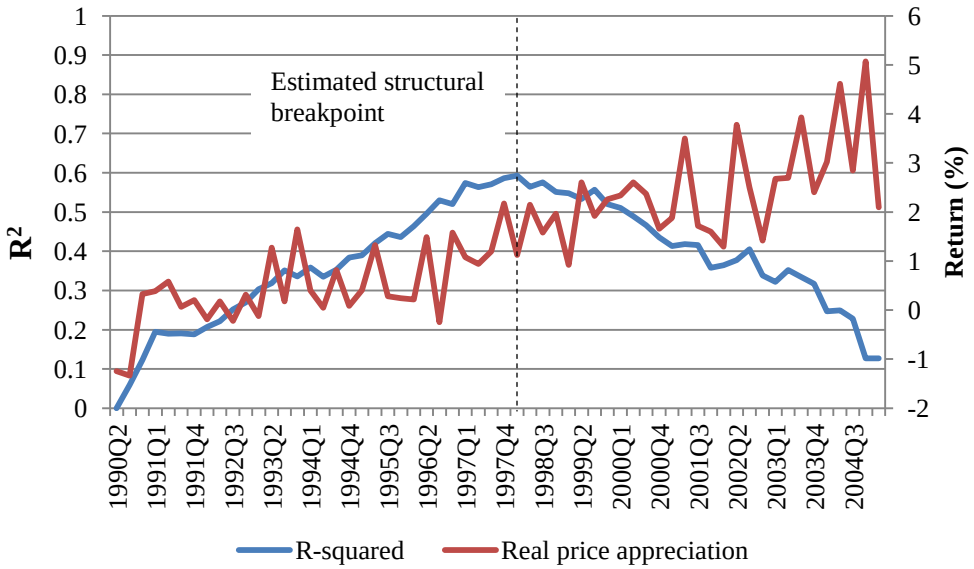
play in the recent housing turmoil. It also provides some direct explanation for the observation from Clayton (1996), who finds that his fundamental equation can largely explain house price movements but fails to perform in periods of drastic price gyrations in the 1980 and 1989 housing bubbles.

**Figure 9. Estimated $R^2$ of Equation (4) and real house changes 1990Q2-2005Q1**

This figure plots the $R^2$ from estimating the following equation:

$$RETURN_t = \alpha + \delta I(q_t \geq q_t^*) + \varepsilon_t \qquad \text{for } 1 < t < 61$$

where $RETURN_t$ is the real price changes in quarter $t$; $\alpha$ is the intercept term; $\delta$ measures the importance of the potential breakpoint; $I(\cdot)$ is an indicator function which equals 1 if its condition is true and 0 otherwise; $q_t$ is quarter $t$; $q_t^*$ is the location of the potential structural break, and $\varepsilon_t$ is the error term. The equation is estimated over 1990Q2-2005Q1. Real house price appreciations $RETURN_t$, measured by the real quarter-on-quarter change (in percentage) in the Case-Shiller index, over the same period are also plotted.




Furthermore, in the second sub-period sentiment seemed to gather momentum such that it could be largely predicted by its previous values (Column 4, $R^2 = 0.696$), whereas the first sub-period saw much more stochastic sentiment levels that could not be explained by either past prices or its past levels (Column 2, $R^2 = 0.133$). At first glance, it seems confusing that the coefficients of lagged returns in the DSENT

## Table 7. VAR estimation results for two sub-periods

This table reports the VAR estimation results with the composite direct sentiment index and housing returns as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t}$$
$$DSENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, DSENT_{It-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \nu_{2t}$$

The models are estimated over two sub-periods, 1990Q2-1997Q4 and 1998Q1-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *DSENT* is the first principal component extracted from the buyer sentiment measure *(BUYER)*, builder sentiment measure *(BUILDER)* and lender sentiment measure *(LENDER)*. The *RETURN* equation includes a set of control variables $z_t$ but their coefficients are not reported for brevity: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter, and is used at the one-quarter lag to account for simultaneity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | **1990Q2-1997Q4** | | **1998Q1-2010Q3** | |
|---|---|---|---|---|
| | **RETURN** | **DSENT** | **RETURN** | **DSENT** |
| | **(1)** | **(2)** | **(3)** | **(4)** |
| **Panel A. Coefficient estimation results** | | | | |
| Constant | 0.210 | -0.155 | 1.410* | -0.164 |
| | (0.205) | (0.118) | (0.802) | (0.143) |
| $DSENT_{t-1}$ | 0.366** | 0.0124 | 0.197 | 0.519*** |
| | (0.172) | (0.173) | (0.224) | (0.138) |
| $DSENT_{t-2}$ | 0.0991 | 0.227 | 0.977*** | 0.159 |
| | (0.183) | (0.172) | (0.224) | (0.143) |
| $RETURN_{t-1}$ | -0.00384 | -0.132 | -0.0369 | 0.0763 |
| | (0.205) | (0.142) | (0.112) | (0.0658) |
| $RETURN_{t-2}$ | 0.386*** | 0.228* | 0.0577 | 0.0806 |
| | (0.137) | (0.136) | (0.111) | (0.0660) |
| POP | -0.267 | | -2.600 | |
| | (0.583) | | (2.018) | |
| GDP | 0.274 | | 0.600 | |
| | (0.253) | | (0.432) | |
| INCOME | -0.138 | | -0.432* | |
| | (0.157) | | (0.239) | |
| UNEMP | -0.0169 | | -0.0854 | |
| | (0.0390) | | (0.0643) | |
| MGTRATE | 0.0521** | | 0.0745*** | |
| | (0.0208) | | (0.0131) | |
| SUPPLY | 0.00581 | | 0.0171 | |
| | (0.00807) | | (0.0128) | |
| Observations | 29 | 29 | 51 | 51 |
| R-squared | 0.561 | 0.133 | 0.779 | 0.696 |
| **Panel B. Joint significance tests of estimated coefficients** | | | | |
| $DSENT_{t-2 \text{ to } t-1}$ | 0.465* | 0.239 | 1.174*** | 0.678*** |
| *(F-stat)* | *(4.77)* | *(1.75)* | *(41.42)* | *(31.61)* |
| $RETURN_{t-2 \text{ to } t-1}$ | 0.382** | 0.096 | 0.021 | 0.157 |
| *(F-stat)* | *(8.04)* | *(3.04)* | *(0.32)* | *(4.12)* |

equation are insignificant even during the bubble (Column 4); or in other words, sentiment was not motivated by past house prices. However, running the sub-period analysis using the three separate sentiment indices (*Appendix A*) reveals that in the bubble, homebuyers' and lenders' sentiment depended largely on past prices whilst builders' did not. This mixed result is likely the reason behind the lack of significance for the coefficients of lagged returns in the composite index *DSENT* equation.

In sum, an important inference from these results is that there always exists a sentiment-induced mispricing component in house prices regardless of market conditions. However, house prices will become much more sensitive to sentiment movements in episodes of market bubbles, and sentiment will also appear more predictable in such periods.

## 4. Summary

In conclusion, the above evidence consistently demonstrates a persistent effect of housing market sentiment on real house prices. All three major participants in the housing market – homebuyers, builders and lenders – contribute to drive house prices away from fundamentals through their psychological impact. This sentiment-induced component becomes much larger during episodes of rapidly rising or falling house prices than in moderate periods.

The VAR model has also revealed a reverse causality from house prices to sentiment in the case of homebuyers and lenders. Builders, however, are not affected

by past prices, due possibly to their leading role in the market. Finally, sentiment is also sticky as it exhibits strong autocorrelation.

# Chapter 5 - Robustness Tests

## 1. VAR estimation with Instrument Variable method

Since the dependent variable *RETURN* is measured as the first difference of the Case-Shiller price index, if the error term $\nu_{1t}$ in Equation (2) is related to the index rather than its change, $\nu_{1t}$ may then be correlated with $RETURN_{t-i}$ in Equation (2). Hence, this study also employs an instrumental variable (IV) for its estimation. Specifically, the IV estimates are calculated in two stages: In the first stage, $RETURN_t$ is regressed on $RETURN_{t-1}$ to obtain the predicted values. In the second stage, Equation (2) is estimated using the predicted values from the first stage regression instead of their actual values. The VAR results using IV for Equation (2) are presented in ***Table 8*** . There is virtually no change compared with previous results: the coefficients in *SENTIMENT* equations remain the same, whilst the coefficients in *RETURN* equations only slightly change in their magnitude.

## 2. VAR estimation with a composite indirect sentiment index

The second robustness check is concerned with the measure of sentiment. In addition to the direct survey-based measures of sentiment used above, many indirect proxies have been widely used in the behavioural finance literature. Baker and

## Table 8. VAR estimation results (by IV method) of three sentiment indices

This table reports the VAR estimation results using instrumental variable with sentiment and housing returns as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, SENTIMENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t}$$
$$SENTIMENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, SENTIMENT_{t-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \nu_{2t}$$

The study period covers 1990Q2-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *BUYER* is the homebuyer sentiment index; *BUILDER* is the builder sentiment index*; LENDER* is the lender sentiment index. All *RETURN* equations include a set of control variables $z_t$ but their coefficients are not reported for brevity: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter, and is used at the one-quarter lag to account for simultaneity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | ***Homebuyers' sentiment*** | | ***Builders' sentiment*** | | ***Lenders' sentiment*** | |
|---|---|---|---|---|---|---|
| | **RETURN** | **BUYER** | **RETURN** | **BUILDER** | **RETURN** | **LENDER** |
| **Panel A. Coefficient estimation results** | | | | | | |
| Constant | -0.462 | -0.376 | 0.102 | -0.676 | 0.130 | -1.200 |
| | (0.337) | (0.234) | (0.356) | (0.944) | (0.344) | (1.208) |
| $SENTIMENT_{t-1}$ | 0.0781 | 0.374*** | 0.00997 | 0.545*** | 0.0205 | 0.565*** |
| | (0.0866) | (0.109) | (0.0212) | (0.108) | (0.0171) | (0.110) |
| $SENTIMENT_{t-2}$ | 0.199** | 0.0573 | 0.0607*** | 0.338*** | 0.0570*** | 0.0206 |
| | (0.0830) | (0.105) | (0.0211) | (0.114) | (0.0176) | (0.118) |
| $RETURN_{t-1}$ | 0.481*** | 0.312** | 0.233 | -0.0188 | 0.242 | -0.167 |
| | (0.170) | (0.124) | (0.173) | (0.516) | (0.163) | (0.675) |
| $RETURN_{t-2}$ | 0.449** | 0.0171 | 0.327* | 0.478 | 0.360** | 1.890*** |
| | (0.175) | (0.128) | (0.174) | (0.515) | (0.171) | (0.636) |
| POP | -0.150 | | -0.0512 | | -0.0833 | |
| | (0.649) | | (0.622) | | (0.602) | |
| GDP | 1.050*** | | 0.556 | | 0.518 | |
| | (0.386) | | (0.375) | | (0.368) | |
| INCOME | -0.304 | | -0.280 | | -0.275 | |
| | (0.213) | | (0.205) | | (0.198) | |
| UNEMP | 0.0435 | | -0.0458 | | 0.0124 | |
| | (0.0509) | | (0.0535) | | (0.0495) | |
| MGTRATE | 0.0678*** | | 0.0612*** | | 0.0772*** | |
| | (0.0139) | | (0.0134) | | (0.0132) | |
| SUPPLY | 0.0274** | | 0.0198* | | 0.0183* | |
| | (0.0111) | | (0.0105) | | (0.0102) | |
| Observations | 79 | 79 | 79 | 79 | 79 | 79 |
| R-squared | 0.647 | 0.340 | 0.682 | 0.722 | 0.699 | 0.578 |
| **Panel B. Joint significance tests of estimated coefficients** | | | | | | |
| $SENTIMENT_{t-2\ to\ t-1}$ | 0.277*** | 0.431*** | 0.0707*** | 0.883*** | 0.078*** | 0.586*** |
| *(F-stat)* | *(5.19)* | *(18.39)* | *(9.49)* | *(109.11)* | *(11.86)* | *(36.91)* |
| $RETURN_{t-2\ to\ t-1}$ | 0.93*** | 0.3291*** | 0.56** | 0.4592 | 0.602** | 1.723*** |
| *(F-stat)* | *(13.27)* | *(10.44)* | *(3.47)* | *(1.01)* | *(4.90)* | *(9.56)* |

Wurgler (2007) provide a comprehensive discussion of the various common proxies for investor sentiment in the stock market, from which this thesis derives many equivalent measures for the housing markets. This thesis constructs an indirect sentiment index *(INSENT)* from three proxies – namely home sale transaction volumes, residential mortgage flows, and housing starts. Data on the quarterly new and used home sales are obtained from the Bureau of Census and National Association of Realtors, residential mortgage flows from the Federal Reserve Flow of Funds Accounts, and housing starts from the Bureau of Census. Even though they do not perfectly represent sentiment, each of the above three series likely contains an element of sentiment. Arguably, the market will observe higher transaction volumes[26], mortgage flows[27], and housing starts[28] during periods of investor overconfidence. Similar to the survey-based data series, the three indirect sentiment proxies are first regressed against a set of macroeconomic factors to remove their influences; the residuals are then used in the principal component analysis to obtain *INSENT* (See ***Table 1*** for descriptive statistics). Estimating the base VAR model specified in Equation (2) and (3) with this indirect sentiment index gives a somewhat

[26] Baker and Stein (2004) posit that *market liquidity* can serve as a sentiment proxy. In the presence of short-sale constraints, irrational investors are more inclined to trade when they are optimistic, hence adding liquidity, than when they are pessimistic. Likewise, housing speculators are more likely to enter the market when their expectations for future price increases are high, coupled with homebuyers worrying about being priced out of the market.

[27] Clayton, Ling and Naranjo (2009) argue that commercial mortgage flows are widely viewed as a barometer of investment sentiment, in part because of the association between past real estate cycles and excessive mortgage flows during periods of low perceived risks. An equivalent measure for the housing market is *residential mortgage flows.*

[28] *Housing starts* are in many ways similar to IPOs and secondary equity offering events in the stock market. In finance, it is well established that the market timing of IPOs and secondary equity offerings is influenced by positive sentiment (Baker and Wurgler, 2000). In the same way, housing developers typically attempt to time their projects to take advantage of the rising prices during periods of optimism.

## Table 9. VAR estimation results of the composite indirect sentiment index

This table reports the VAR estimation results with the composite indirect sentiment index and housing returns as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, INSENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t}$$
$$INSENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, INSENT_{t-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \nu_{2t}$$

The study period covers 1990Q2-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *INSENT* is the first principal component extracted from housing transaction volume, residential mortgage flows and housing starts. The *RETURN* equation includes a set of control variables $z_t$ but their coefficients are not reported for brevity: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter, and is used at the one-quarter lag to account for simultaneity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | **RETURN (1)** | **INSENT (2)** |
|---|---|---|
| **Panel A. Coefficient estimation results** | | |
| Constant | -0.169 | -0.141 |
| | (0.321) | (0.0964) |
| $INSENT_{t-1}$ | -0.0172 | 0.670*** |
| | (0.216) | (0.114) |
| $INSENT_{t-2}$ | 0.528** | 0.217 |
| | (0.254) | (0.138) |
| $INSENT_{t-3}$ | -0.306 | 0.0458 |
| | (0.283) | (0.147) |
| $INSENT_{t-4}$ | -0.347 | -0.149 |
| | (0.249) | (0.120) |
| $RETURN_{t-1}$ | 0.153 | 0.0522 |
| | (0.101) | (0.0552) |
| $RETURN_{t-2}$ | 0.276*** | 0.0755 |
| | (0.106) | (0.0582) |
| $RETURN_{t-3}$ | 0.181 | -0.0408 |
| | (0.116) | (0.0600) |
| $RETURN_{t-4}$ | 0.159 | 0.0678 |
| | (0.114) | (0.0603) |
| Control $z_t$ | Yes | No |
| Observations | 78 | 78 |
| R-squared | 0.668 | 0.777 |
| **Panel B. Joint significance tests of estimated coefficients** | | |
| $INSENT_{t-4 \text{ to } t-1}$ | -0.142 | 0.784*** |
| *(F-stat)* | *(7.32)* | *(121.00)* |
| $RETURN_{t-4 \text{ to } t-1}$ | 0.769*** | 0.155* |
| *(F-stat)* | *(34.78)* | *(9.29)* |

similar but weaker result **(*Table 9*).** Though the indirect sentiment indicator is still found to Granger-cause subsequent price appreciations as in the central hypothesis, only its two-period lagged coefficient appears statistically significant, while there is no evidence for the reverse causal relation. An implication from this exercise is that sentiment measured by direct surveys is more useful in predicting house prices than indirect proxies extracted from market indicators.

## 3. VAR forecast accuracy

This study also computes out-of-sample forecast of house prices to evaluate the model with *DSENT* (Equation 2) versus a conventional model which comprises all the exogenous variables with the exception of the sentiment proxy (Equation 5):

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t} \qquad (2)$$

$$RETURN_t = \alpha + \delta z_t + \gamma RETURN_{t-1} + \nu_t \qquad (5)$$

The forecast process involves randomly selecting a window between 1990Q1-2010Q1 for estimating Equation (2) and Equation (5); the estimated coefficients are then used to forecast house prices in 1-year, 3-year and 5-year ahead. This process is repeated 100 times. Three common measures to evaluate forecast accuracy are adopted, namely mean error (ME), mean percentage error (MPE) and mean squared error (MSE). The results are presented in ***Table 10***. Overall, the VAR model with the sentiment index provides more accurate forecasts of house prices over the short and medium term (1-year and 3-year), i.e. it has lower forecast errors, than the

conventional regression model. This confirms that sentiment is an important determinant of house prices in the short and medium term. However, over a 5-year horizon the fundamental model seems to perform better in predicting housing prices, consistent with the common notion that house prices and fundamentals are cointergrated in the long run.

**Table 10. Measures of forecast accuracy**

This table reports the various measures of forecast accuracy of two models of house prices:

With sentiment: $RETURN_t = \alpha + \sum_{i=1}^{p} \beta_i \, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_i \, RETURN_{t-i} + \delta z_t + \nu_t$

Without sentiment: $RETURN_t = \alpha + \delta z_t + \gamma RETURN_{t-1} + \nu_t$

The following forecast measures are employed:

$$\text{Mean Error} = \text{ME} = \frac{1}{n}\sum_{t=1}^{n}(\text{forecast}_t - \text{actual}_t)$$

$$\text{Mean Percentage Error} = \text{MPE} = \frac{1}{n}\sum_{t=1}^{n}(\text{forecast}_t - \text{actual}_t)/\text{actual}_t * 100$$

$$\text{Mean Squared Error} = \text{MSE} = \frac{1}{n}\sum_{t=1}^{n}(\text{forecast}_t - \text{actual}_t)^2$$

| | **With sentiment** | | | **Without sentiment** | | |
|---|---|---|---|---|---|---|
| | **1-year** | **3-year** | **5-year** | **1-year** | **3-year** | **5-year** |
| Mean Error | -0.077 | 0.376 | 16.797 | 0.429 | 0.404 | 0.420 |
| Mean Percentage Error | 24.0% | 30.0% | 185.5% | 47.7% | 53.0% | 66.2% |
| Mean Squared Error | 0.01 | 0.14 | 282.13 | 7.91 | 12.79 | 16.74 |

## 4. Long-run regressions

The discussion to this point has focused mainly on the short-run relationship between house prices and market sentiment through a VAR framework. Arguably, if periods of optimism (pessimism) lead to house prices overshooting (undershooting) in the initial periods, the market should observe a negative relationship between cumulative long-run returns and sentiment as prices revert to their fundamental values

over time. Such a finding will also be in line with the forecasting results in the above section.

In this section, the long-run effect of sentiment is tested using a framework adopted from Brown & Cliff (2005) and Ling, Naranjo & Scheick (2010). It involves regressing future k-period quarterly housing returns on a vector of control variables $z_t$ and the composite sentiment index $DSENT_t$:

$$(RETURN_{t+1} + \cdots + RETURN_{t+k})/k = \alpha(k) + \theta(k)z_t + \beta(k)DSENT_t + \varepsilon_t \quad (6)$$

where $RETURN_{t+1}, \ldots, RETURN_{t+k}$ are quarterly real price changes as before, $k$ is the number of quarters in the holding period, $\alpha(k)$ is the intercept term, $z_t$ is the same set of control variables employed in the previous VAR models, and $DSENT_t$ is the composite index derived from the homebuyers', builders' and lenders' sentiment measures. The test is carried out over one- to five-year horizons (k = 4 to 20). If sentiment's effect is persistent, as one would expect in the housing markets, this will result in a positive coefficient on $DSENT_t$ over a short and medium horizon that is similar to the results from the VAR models. However, in a well-function market prices should eventually revert to their fundamental values, such that a negative coefficient on $DSENT_t$ in the long run will be observed.

There are two econometric issues in estimating the long-run model specified above. Firstly, the use of overlapping observations in the dependent variable results in a moving average process in the error term, causing the standard errors obtained from an OLS estimation to be biased downwards. As pointed out by Ling, Naranjo & Scheick (2010), the second issue is the potential finite sample bias in the coefficient

estimate of a persistent independent variable (in this case, the coefficient $\beta(k)$ associated with the sentiment index $DSENT_t$). Stambaugh (1999) shows that a persistent explanatory variable, though predetermined, is not strictly exogenous. Although an OLS estimate is consistent and asymtotically normally distributed under the predetermined assumption, it might suffer from biasness with a finite sample setting. To address these econometric issues, this thesis employs a bootstrap simulation procedure similar to Brown & Cliff (2005) and Ling, Naranjo & Scheick (2010). The details of the bootstrap procedure are explained in *Appendix B*.

***Table 11*** reports the bias-adjusted coefficients of the sentiment measure $DSENT_t$ and their standard errors for five different return horizons. The results are revealing. The sentiment coefficients on the one-, two- and three-year average returns are significantly positive, indicating a continuation of price divergence from fundamental values for as long as three years. However, their magnitudes are decreasing as the return horizon increases, which points towards a gradual price correction process at work.

The effect of sentiment then falls sharply in both economic and statistical significance in the four-year return regression, and becomes negative over the five-year horizon, albeit statistically weak. This inevitably suggests that market sentiment at a given time induces very persistent mispricing in future house prices that takes beyond five years to correct. It is consistent with the proposition in the previous sections that housing markets are highly susceptible to prolonged periods of sentiment-induced mispricing. The five-year adjustment period is a marked contrast

## Table 11. Long-run regression results

This table reports the results from estimating the following long-run regression model:

$$(RETURN_{t+1} + \cdots + RETURN_{t+k})/k = \alpha(k) + \theta(k)z_t + \beta(k)DSENT_t + \varepsilon_t$$

*RETURN* is the quarterly percentage change in the real Case-Shiller housing price index and *k* is the number of quarters in the holding period. *DSENT* is the first principal component extracted from the buyer sentiment measure *(BUYER)*, builder sentiment measure *(BUILDER)* and lender sentiment measure *(LENDER)*. The set of control variables $z_t$ includes: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter. The reported coefficients are bias-adjusted OLS estimates derived from the bootstrap simulation procedure documented in the Appendix. *, ** and *** denote 10%, 5% and 1% significant levels, respectively.

| **Return horizon** | **k** | **Adjusted coef. $DSENT_t$** | **Bootstrapped St. Err.** | **$R^2$** |
|---|---|---|---|---|
| One-year | 4 | 1.575*** | 0.14 | 0.709 |
| Two-year | 8 | 1.100*** | 0.12 | 0.475 |
| Three-year | 12 | 0.888*** | 0.11 | 0.301 |
| Four-year | 16 | 0.174* | 0.09 | 0.167 |
| Five-year | 20 | -0.059 | 0.10 | 0.104 |

to the quick correction of stock prices in studies such as Schmeling (2009), Barber, Odean & Zhu (2009), and Brown & Cliff (2005), to name but a few, who find that stock prices revert in only one to twelve months. It is, on the other hand, close to the results from prior studies in the housing literature looking into the price adjustment to supply and demand shocks. For example, Harter-Dreiman (2004) finds that it takes about five years for housing prices to get within 70% of its new equilibrium value in response to an income shock, whilst the adjustment speed in Malpezzi (1999) is approximately 10 years. Nevertheless, a potential limitation of this finding is that the study period (1990Q2-2010Q3) covers mostly the recent housing bubble. Thus, it

cannot be ruled out that such a slow correction found in the long-run regression might be driven partly by the fact that house prices were consistently rising for a long period of time as the bubble built.

## 5. The relationship between sentiment, house prices and trading volume

Finally, this section expands the basic model to a three-equation VAR system to incorporate trading volume as an additional endogenous variable:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \sum_{i=1}^{k} \varphi_{1i}\, VOLUME_{t-i} + \\ + \delta z_t + \nu_{1t} \quad (7)$$

$$DSENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \sum_{i=1}^{k} \varphi_{2i}\, VOLUME_{t-i} + \nu_{2t} \quad (8)$$

$$VOLUME_t = \alpha_3 + \sum_{i=1}^{p} \beta_{3i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{3i}\, RETURN_{t-i} + \sum_{i=1}^{k} \varphi_{3i}\, VOLUME_{t-i} + \\ + \nu_{3t} \quad (9)$$

Along with house prices, trading volume is widely recognized as an essential barometer of the housing markets. The positive correlation between house prices and trading volume has been a familiar question for ongoing research[29]. This section tests a simple proposition that such a positive co-movement is the outcome of both trading activities and house prices being induced by market sentiment. Trading volume is

[29] Stein (1995) posits that during periods when property prices are falling, existing homeowners will find it difficult to purchase a new home due to down-payment constraints. This results in fewer transactions in the market. Genesove and Mayer (2001), on the other hand, argue that sellers who are loss averse tend to raise their reservation prices. This results in higher time-on-market and lower sale hazard.

measured as the total number of new and existing single-family home sales, which are obtained from the National Association of Realtors and U.S. Bureau of Census.

***Table 12*** reports the estimation results of the VAR model with housing returns (*RETURN*), the direct composite sentiment index (*DSENT*) and change in trading volume (*VOLUME*). The model uses four lags of each endogenous variable as indicated by standard selection criteria. Similar to the previous results, market sentiment (*DSENT*) again appears strongly significant in the *RETURN* equation. As expected, trading volume is also driven by sentiment, evidenced by the positive coefficients of *DSENT* in the third column of ***Table 12*****.** Although only the first lag of *DSENT* appears significant, the joint test of all four lags indicates strong significance at 1% level. The evidence thus supports the hypothesis that both trading volume and prices are influenced by sentiment.

**Table 12. VAR estimation results with trading volume**

This table reports the VAR estimation results with the composite direct sentiment index, housing returns and trading volumes as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \sum_{i=1}^{k} \varphi_{1i}\, VOLUME_{t-i} + \delta z_t + \nu_{1t}$$

$$DSENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \sum_{i=1}^{k} \varphi_{2i}\, VOLUME_{t-i} + \nu_{2t}$$

$$VOLUME_t = \alpha_3 + \sum_{i=1}^{p} \beta_{3i}\, DSENT_{t-i} + \sum_{i=1}^{q} \gamma_{3i}\, RETURN_{t-i} + \sum_{i=1}^{k} \varphi_{3i}\, VOLUME_{t-i} + \nu_{3t}$$

The study period covers 1990Q2-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *DSENT* is the first principal component extracted from the buyer sentiment measure *(BUYER)*, builder sentiment measure *(BUILDER)* and lender sentiment measure *(LENDER); VOLUME* is the quarterly percentage change in the number of home sales. The *RETURN* equation includes a set of control variables $z_t$ but their coefficients are not reported for brevity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | RETURN (1) | DSENT (2) | VOLUME (3) |
|---|---|---|---|
| **Panel A. Coefficient estimation results** | | | |
| Constant | 0.441 | -0.295*** | 0.0156** |
| | (0.269) | (0.109) | (0.00682) |
| $DSENT_{t-1}$ | 0.307* | 0.426*** | 0.0138** |
| | (0.169) | (0.110) | (0.00688) |
| $DSENT_{t-2}$ | 0.614*** | 0.415*** | 0.0110 |
| | (0.185) | (0.117) | (0.00735) |
| $DSENT_{t-3}$ | 0.202 | -0.311*** | 0.00772 |
| | (0.183) | (0.118) | (0.00742) |
| $DSENT_{t-4}$ | -0.114 | -0.315*** | 0.000614 |
| | (0.184) | (0.121) | (0.00761) |
| $RETURN_{t-1}$ | -0.145 | 0.158** | -0.00392 |
| | (0.0941) | (0.0637) | (0.00400) |
| $RETURN_{t-2}$ | -0.0481 | 0.243*** | -0.00631 |
| | (0.0985) | (0.0622) | (0.00391) |
| $RETURN_{t-3}$ | 0.0833 | -0.00510 | -0.00767* |
| | (0.0951) | (0.0630) | (0.00395) |
| $RETURN_{t-4}$ | 0.219** | 0.0223 | 0.000730 |
| | (0.0912) | (0.0596) | (0.00374) |
| $VOLUME_{t-1}$ | 7.620** | -0.422 | -0.124 |
| | (3.530) | (2.205) | (0.138) |
| $VOLUME_{t-2}$ | 11.23*** | 0.615 | 0.161 |
| | (3.272) | (2.111) | (0.133) |
| $VOLUME_{t-3}$ | 7.972** | -0.835 | 0.268** |
| | (3.304) | (2.123) | (0.133) |
| $VOLUME_{t-4}$ | -0.632 | -2.936 | -0.176 |
| | (3.549) | (2.369) | (0.149) |
| *Control* $z_t$ | *Yes* | *No* | *No* |
| Observations | 78 | 78 | 78 |
| R-squared | 0.799 | 0.727 | 0.271 |
| **Panel B. Joint significance tests of estimated coefficients** | | | |
| $DSENT_{t-4 \text{ to } t-1}$ | 1.009*** | 0.215*** | 0.033*** |
| *(F-stat)* | *(37.42)* | *(64.64)* | *(18.07)* |
| $RETURN_{t-4 \text{ to } t-1}$ | 0.109* | 0.418*** | -0.107 |
| *(F-stat)* | *(8.89)* | *(23.33)* | *(7.35)* |
| $VOLUME_{t-4 \text{ to } t-1}$ | 26.19*** | -3.578 | 0.129** |
| *(F-stat)* | *(16.26)* | *(2.42)* | *(10.31)* |

Meanwhile, when sentiment is controlled for, there is somewhat weak evidence of trading volume being influenced adversely by past house prices. The negative coefficients of *RETURN*, although marginally significant, indicate that higher prices discourage aspiring buyers from entering the market. The reverse causality is stronger in the direction of past trading volume influencing current housing returns, consistent with the notion that higher trading volume signals increasing demand that pushes up prices and vice versa. Trading volume is also fairly predictable from its past values. Interestingly, sentiment is not driven by trading volume but by house price changes only.

# Chapter 6 - Conclusion

## 1. Summary of findings

The source of house price volatility is of great interest to policy makers because housing plays a crucial part in the economy. Yet, the inherent characteristics of housing markets often render home prices susceptible to sentiment-induced mispricing. Despite the potential relevance of psychological factors in explaining price movements in the housing market, no research has directly tested the relationship between sentiment and house prices due to the difficulty in measuring sentiment. This thesis seeks to provide new evidence on the role of psychology in driving house prices through various direct proxies to quantify sentiment. These measures are derived from surveys on the perceptions of homebuyers, builders and lenders about housing markets that are conducted on a frequent basis by various institutions in the U.S. The two primary research questions guiding the empirical tests are: (1) Does sentiment have a role in explaining house price movements, above and beyond the role of fundamentals and, (2) Do past price movements predict changes in sentiment beyond the predictions of lagged sentiment and fundamentals?

The estimation results from VAR models provide affirmative results for both research questions. The central finding from this thesis is that sentiment, regardless of homebuyers, builders or lenders, indeed has a significant impact on house prices. In

particular, overoptimism appears to drive prices above fundamental values and vice versa. Moreover, this effect is highly persistent in that a sentiment shock will be channelled into housing returns over at least five years. This sentiment-induced component becomes much larger during episodes of fast escalating and slumping house prices than in moderate periods. When forecasting future house prices in the short and medium terms, forecast errors can be substantially reduced by adding sentiment indices into the conventional (fundamental) equation. In the long run, however, sentiment is found to be negatively related to the future five-year average house prices as the correction process starts to take place. Hence, over a five-year horizon a fundamental price model with only supply and demand factors can better predict house prices.

The results also provide some evidence of a feedback relationship between house prices and sentiment. Homebuyers and lenders are affected by past prices in forming their beliefs, whilst builders do not seem to be backward looking. Coupled with the persistent impact of sentiment on house prices established above, this finding of a positive feedback loop from prices to household and lender sentiment further supports the potential for housing prices to spiral, which potentially renders housing markets highly prone to bubbles.

In addition, adding trading volume as another endogenous variable into the VAR system reveals that both trading volume and prices are influenced by sentiment. When sentiment is controlled for, there is somewhat weak evidence of trading volume being influenced adversely by past house prices, while past trading volume strongly

influence current housing returns in a positive manner. This is consistent with the notion that higher prices discourage buying activities whereas higher trading volume signals increasing demand that pushes up prices. Finally, sentiment is not found to be affected by trading volume.

## 2. Implications and limitations

Discovering the link between house prices and sentiment has many important implications for the general understanding of the housing markets. Firstly, mispricing is an inherent component of house prices at least in the short run. Secondly, house price volatility can be better predicted when taking into account the sentiment of market participants. Lastly, in a recent working paper that focuses on the latest housing bubble, Glaeser, Gottlieb and Gyourko (2010) highlight a pressing call for future research on the interplay between bubbles, beliefs and credit condition. This study's finding of a "spiral" sentiment effect thus provides some first insights to the question raised on the part of misbeliefs and bubbles. For a more complete answer to the issue, further research is needed on the long-run effect of sentiment, the price correction process and the interaction between sentiment and other fundamental price determinants.

As a final note, this research is carried out at the aggregate level due to the fact that the sentiment proxies are derived from survey data which are only available at the national level. Much as the study has revealed many important discoveries into

the aggregate price dynamics, it is important to recognize that there exists much heterogeneity in both market sentiment and house prices among local regions. Thus, where available data permit, similar study carried out on local markets will greatly improve the current understanding of this relatively unexplored area.

# Appendix A

## VAR estimation results of individual sentiment indices, 1998Q1-2010Q3

This table reports the VAR estimation results with three sentiment indices and housing returns as endogenous variables:

$$RETURN_t = \alpha_1 + \sum_{i=1}^{p} \beta_{1i}\, SENTIMENT_{t-i} + \sum_{i=1}^{q} \gamma_{1i}\, RETURN_{t-i} + \delta z_t + \nu_{1t}$$

$$SENTIMENT_t = \alpha_2 + \sum_{i=1}^{p} \beta_{2i}\, SENTIMENT_{It-i} + \sum_{i=1}^{q} \gamma_{2i}\, RETURN_{t-i} + \nu_{2t}$$

The models are estimated over the bubble period 1998Q1-2010Q3. Lag-length is chosen based on three selection criteria AIC, HQIC and SBIC. *RETURN* is the quarterly percentage change in the real Case-Shiller housing price index; *BUYER* is the homebuyer sentiment index; *BUILDER* is the builder sentiment index*; LENDER* is the lender sentiment index. All *RETURN* equations include a set of control variables $z_t$ but their coefficients are not reported for brevity: *POP* is the percentage change in the population aged between 20-30; *GDP* is real GDP growth (percentage); *INCOME* is real per capita income growth (percentage); *UNEMP* is the percentage change in unemployment rate; *MGTRATE* is the percentage change in real mortgage rate; *SUPPLY* is the percentage change in the number of residential units completed in each quarter, and is used at the one-quarter lag to account for simultaneity. ADF tests show that all variables are stationary. *, ** and *** denote 10%, 5% and 1% significant levels, respectively. Standard errors are in parentheses, unless otherwise stated.

| | ***Homebuyer sentiment*** | | ***Builder sentiment*** | | ***Lender sentiment*** | |
|---|---|---|---|---|---|---|
| | **RETURN** | **BUYER** | **RETURN** | **BUILDER** | **RETURN** | **LENDER** |
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** |
| Constant | 0.731 | -0.489* | 1.018 | -0.964 | 1.465* | -1.099 |
| | (0.893) | (0.280) | (0.851) | (1.295) | (0.859) | (1.703) |
| $SENTIMENT_{t-1}$ | 0.0618 | 0.500*** | 0.00423 | 0.548*** | 0.0197 | 0.623*** |
| | (0.126) | (0.139) | (0.0254) | (0.133) | (0.0202) | (0.138) |
| $SENTIMENT_{t-2}$ | 0.335*** | -0.0582 | 0.0818*** | 0.375*** | 0.0685*** | -0.000424 |
| | (0.123) | (0.132) | (0.0254) | (0.140) | (0.0220) | (0.148) |
| $RETURN_{t-1}$ | 0.170 | 0.364*** | 0.0470 | 0.0647 | 0.0267 | -0.129 |
| | (0.124) | (0.122) | (0.122) | (0.594) | (0.126) | (0.798) |
| $RETURN_{t-2}$ | 0.262** | -0.0403 | 0.113 | 0.307 | 0.149 | 1.722** |
| | (0.129) | (0.131) | (0.122) | (0.599) | (0.115) | (0.756) |
| Control $z_t$ | Yes | No | Yes | No | Yes | No |
| Observations | 52 | 52 | 52 | 52 | 52 | 52 |
| R-squared | 0.682 | 0.717 | 0.7165 | 0.755 | 0.715 | 0.623 |

## Appendix B

## Bootstrap simulation procedure

This thesis adopts the bootstrap simulation procedure described in Brown & Cliff (2005) and Ling, Naranjo & Scheick (2010). First, the long-run regressions specified in Equation (5) are estimated using OLS and the beta coefficients of the sentiment measure, $\hat{\beta}_{OLS}$, are saved to calculate bias-adjusted coefficients later.

The second step involves generating a pseudo return series under the null hypothesis that sentiment has no effect on house prices. The following vector autoregressive (VAR) model is used as the underlying data generating process of the pseudo return series:

$$\text{VAR(1) for } y_t = [RETURN_t, DSENT_t, z_t] \qquad (1)$$

where $RETURN_t$ is the real price changes, $DSENT_t$ is our measure of sentiment, and $z_t$ is the set of control variables. The beta coefficient on $DSENT_t$ is set to zero to ensure that the pseudo return series is generated under the null. The predicted y-values and residuals from the VAR estimations are saved.

In the third step, each residual is first multiplied by $\{n/(n-v)\}^{1/2}$ to adjust for a downward bias that results from the use of OLS in estimating the VAR model, where *n* is the number of observations and *v* refers to the degrees of freedom of the VAR (MacKinnon, 2002). The next step involves sampling with replacement from these residuals to generate a new set of bootstrapped residuals. This sampling process is

repeated 10,100 times but the first 100 samples are discarded to avoid any startup effects.

Fourth, the remaining 10,000 sets of bootstrapped residuals are added to the predicted y-value saved from the VAR estimation above to create 10,000 sets of pseudo return series. From these pseudo return series, 10,000 new sets of k-period future returns can be created as the dependent variables in the long-horizon model specified in Equation (5). OLS is used to estimate the regressions with these new dependent variables, yielding 10,000 beta coefficients for each of the return horizons.

Calculating bias-adjusted beta:

The bootstrap simulation procedure discussed above can be utilized to correct for the bias in coefficient estimates within finite sample settings when the bias function is not known analytically (see MacKinnon and Smith, 1998). Assuming the bias is constant for a particular return horizon, the estimated bias is expressed as:

$$Estimated\ bias = \bar{\beta} - \hat{\beta}_{OLS} \qquad (2)$$

where $\hat{\beta}_{OLS}$ is the OLS beta coefficient from running the original long-run regression (Step 1 above), and $\bar{\beta}$ is the sample mean of the 10,000 simulated beta coefficients. The adjusted beta coefficient can therefore be specified as:

$$\hat{\beta}_{adj} = \hat{\beta}_{OLS} - \left(\bar{\beta} - \hat{\beta}_{OLS}\right) = 2\hat{\beta}_{OLS} - \bar{\beta} \qquad (3)$$

This bias-adjusted estimator has been widely shown to provide reliable coefficient estimates in the bootstrap literature.

Calculating adjusted t-statistics and p-values:

Since asymptotic standard errors calculated from a small sample with overlapping observations suffer from significant downward bias, the bootstrap standard error can be used as a more accurate alternative. It is simply the standard deviation of the 10,000 simulated beta coefficients:

$$se(\hat{\beta}) = \left(\frac{1}{B-1}\sum_{i=1}^{B}\left(\hat{\beta}_i - \bar{\beta}\right)^2\right)^{1/2} \qquad (4)$$

where $\hat{\beta}_i$ is the estimated beta coefficient from iteration *i* of the simulation process, $\bar{\beta}$ is the sample mean of the 10,000 simulated coefficients, and B is the number of simulations. The new bias-adjusted t-statistic can be calculated as follows:

$$\hat{t}_{adj} = \frac{\hat{\beta}_{adj}}{se(\hat{\beta})} \qquad (5)$$

Because the t-statistics constructed in this manner do not always follow the standard t-distribution, this thesis develops a bootstrap distribution of t-statistics to compute more accurate p-values for testing of the null hypothesis. The bootstrap t-distribution is derived from 10,000 adjusted t-statistics of the simulation process:

$$\hat{t}_i^* = \frac{\hat{\beta}_i - \bar{\beta}}{se(\hat{\beta})} \qquad (6)$$

where $\hat{\beta}_i$ is the estimated beta coefficient from iteration *i* of the simulation process, $\bar{\beta}$ is the sample mean of the 10,000 simulated betas, and $se(\hat{\beta})$ is the bootstrap standard error. This technique generates an empirical distribution of the t-statistics that provides new critical values for hypothesis testing.

Finally, there are two approaches for computing p-values. The first approach assumes the distribution is symmetric around zero; thus, p-values can be calculated as:

$$\hat{p}_s = \frac{1}{B}\sum_{i=1}^{B} I\left(|\hat{t}_i^*| > |\hat{t}_{adj}|\right) \qquad (7)$$

where $I(\cdot)$ denotes the indicator function, which equals 1 if its argument is true and 0 otherwise, $|\hat{t}_i^*|$ is the absolute value of the t-statistic of iteration *i* as in (6), $|\hat{t}_{adj}|$ is the absolute value of the bias-adjusted t-statistic specified in (5), and B is the number of simulations.

On the other hand, the second approach assumes an asymmetric distribution:

$$\hat{p}_{ns} = 2\ min\left[\frac{1}{B}\sum_{i=1}^{B} I\left(\hat{t}_i^* > \hat{t}_{adj}\right), \frac{1}{B}\sum_{i=1}^{B} I\left(\hat{t}_i^* < \hat{t}_{adj}\right)\right] \qquad (8)$$

This thesis employs both approaches in calculating the p-values for the bias-adjusted sentiment coefficients and the results are similar (except in the four-year horizon regression, where (7) indicates a significant sentiment coefficient at 10% level whilst (8) gives an insignificant beta). Hence, only the results from the first specification in (7) are reports in *Table 11*.

## Bibliography


Abraham, J. M. and Hendershott, P. H. (1996). Bubbles in metropolitan housing market. *Journal of Housing Research, 7*(2), 191-207.

Archer, W. R. and Smith, B. C. (2011). Residential Mortgage Default: The Roles of House Price Volatility, Euphoria and the Borrower's Put Option. *Working paper*.

Barber, B. M., Odean, T., and Zhu, N. (2009). Do Retail Trades Move Markets? *Review of Financial Studies, 22*(1), 151-186.

Baker, M. and Stein, J. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets, 7*, 271-299.

Baker, M. and Wurgler, J. (2000). The equity share in new issues and aggregate stock returns. *Journal of Finance, 55*(5), 2219-2257.

Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives, 21*(2), 129-151.

Barkham, R. J. and Ward, C. W. R. (1999). Investor sentiment and noise traders: Discount to net asset value in listed property companies in the U.K. *Journal of Real Estate Research, 18*(2), 291-312.

Brown, G. W. and Cliff, M. T. (2004). Investor sentiment and the near term stock market. *Journal of Empirical Finance, 11*, 1-27.

Brown, G. W. and Cliff, M. T. (2005). Investor Sentiment and Asset Valuation. *Journal of Business, 78*(2), 405-440.

Brown, G. W. and Cliff, M. T. (2005). Investor sentiment and asset valuation. *Journal of Business, 78*(2), 405-440.

Capozza, D. R. and Seguin, P. J. (1996). Expectations, efficiency, and euphoria in the housing market. *Regional Science and Urban Economics, 26*, 369-386.

Carroll, C. D., Fuhrer, J. C. and Wilcox, D. W. (1994). Does Consumer Sentiment Forecast Household Spending? If So, Why? *American Economic Review, 84*(5), 1397-1408.

Case, K. E. and Shiller, R. J. (1989). The efficiency of the market for single-family homes. *American Economic Review, 79*(1), 125-137.

Case, K. E. and Shiller, R. J. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity, 2*, 300-361.

Chiang, K. C. H. and Lee, M. L. (2010). The role of correlated trading in setting REIT prices. *Journal of Real Estate Finance and Economics, 41*, 320-338.

Clayton, J. (1996). Rational expectations, market fundamentals and housing price volatility. *Real Estate Economics, 24*, 441-470.

Clayton, J. (1997). Are housing price cycles driven by irrational expectations? *Journal of Real Estate Finance and Economics, 14*(3), 341-363.

Clayton, J., Ling, D. C. and Naranjo, A. (2009). Commercial real estate valuation: Fundamentals versus investor sentiment. *Journal of Real Estate Finance and Economics, 38*, 5-37.

Curtin, R.T. (1982). Indicators of Consumer Behavior: The University of Michigan Surveys of Consumers. *The Public Opinion Quarterly, 46*(3), 340-352.

Cutler, D., J. Poterba, and L. Summers (1991): "Speculative Dynamics," *Review of Economic Studies*,58(3), 529-46.

Dua, P. (2008). Analysis of consumers' perceptions of buying conditions for houses. *Journal of Real Estate Finance and Economics, 37*, 335-350.

Ferreira, F. and Gyourko, J. (2011). Anatomy of the beginning of the housing boom: U.S. neighborhoods and metropolitan areas, 1993-2009. *Working paper*.

Gallin, J. 2006. The Long-Run Relationship between House Prices and Income: Evidence from Local Housing Markets. *Real Estate Economics* 34: 417–438.

Genesove, D. and Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics, 116*(4), 1233-1260.

Glaeser, E. L., Gottlieb, J. D. and Gyourko, J. (2010). Can cheap credit explain the housing boom? *Working paper*.

Glaeser, E. L., Gyourko, J., Morales, E. and Nathanson, C. G. (2010). Housing dynamics. *Working paper*.

Glaeser, E. L., Gyourko, J. and Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics, 64*, 198-217.

Goodman, J. L. Jr. (1994). Using attitude data to forecast housing activity. *Journal of Real Estate Research*, 445-453.

Hamilton, B. W. and Schwab, R. M. (1985). Expected appreciation in urban housing markets. *Journal of Urban Economics, 18*, 103-118.

Harter-Dreimanm, M. (2004). Drawing inferences about housing supply elasticity from house price responses to income shocks. *Journal of Urban Economics, 55*, 316-337.

Haughwout, A., Lee, D., Tracy, J. and Klaauw, W.V.D. (2011). Real estate investors, the leverage cycle and the housing market crisis. Federal Reserve Bank of New York Staff Reports No. 514.

Himmelberg, C., Mayer, C., and Sinai, T. (2005). Assessing high house prices: bubbles, fundamentals, and misperceptions, *19*(4), 67-92.

Hui, E. C. M. and Yue, S. (2006). Housing price bubbles in Hong Kong, Beijing and Shanghai: A comparative study. *Journal of Real Estate Finance and Economics, 33*, 299-327.

Lai, R. N. and Order, R. A. V. (2010). Momentum and house price growth in the United States: Anatomy of a bubble. *Real Estate Economics, 38*, 753-773.

Levin, E. J. and Wright, R. E. (1997). Speculation in the housing market. *Urban Studies, 34*(9), 1419-1437.

Lin, C. Y., Rahman, H. and Yung, K. (2009). Investor sentiment and REIT returns. *Journal of Real Estate Finance and Economics, 39*, 450-471.

Ling, D. C. (2005). A random walk down main street: Can experts predict returns on commercial real estate? *Journal of Real Estate Research, 27*(2), 137-154.

Ling, D. C., Naranjo, A. and Scheick, B. (2010). Investor sentiment and asset pricing in public and private markets. *Working paper*.

Mankiw, N.G. and Weil, D.N. (1989). The baby boom, the baby bust, and the housing market. *Regional Science and Urban Economics, 19*, 2235-258.

Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics, 8*, 27-62.

Mayer, C., and Pence, K. (2009). Subprime mortgages: what, where, and to whom? In E. Glaeser and J. Quigley (Eds.), *Housing Markets and the Economy: Risk, Regulation, and Policy* (pp. 149–96). Cambridge, MA: Lincoln Land Inst.

Mayer, C., and Sinai, T. (2009). U.S. house price dynamics and behavioral economics. In C. L. Foote, L. Goette, and S. Meier (Eds.), *Policy Making Insights on Behavioral Economics* (pp. 261–95). Boston: Fed. Reserve Bank Boston.

MacKinnon, James G. (2006). Bootstrap Methods in Econometrics. *The Economic Record, 82*, S2-S18.

Meen, G. (2002). The time-series behaviors of house prices: A Transatlantic divide? *Journal of Housing Economics, 11*, 1-23.

Meese, R. and Wallace, N. (1994). Testing the present value relation for housing prices: Should I leave my house in San Francisco? *Journal of Urban Economics, 35*, 245-266.

Meese, R. and Wallace, N. (2003). House price dynamics and market fundamentals: The Parisian housing market. *Urban Studies, 40*(5-6), 1027-1045.

Miller, N. & Peng, L. (2006) Exploring Metropolitan House Price Volatility. *Journal of Real Estate Finance and Economics,* 33 (5), 5-18.

Ooi, T. L. J. and Le, T. T. T. (2011). New supply and price dynamics in the housing market. *Urban Studies, forthcoming*.

Piazzesi, M. and Schneider, M. (2009). Momentum traders in the housing market: Survey evidence and a search model. *American Economic Review: Papers and Proceedings, 99*(2), 406-411.

Poterba, J. (1984). Tax subsidies to owner-occupied housing: an asset market approach. *Quarterly Journal of Economics, 99*, 729-52.

Qiu, L. and Welch, I. (2005). Investor Sentiment Measures. *NBER Working Paper*.

Riddle, M. (1999). Fundamentals, Feedback trading, and housing market speculation: Evidence from California. *Journal of Housing Economics, 8*, 272-284.

Schmeling, M. (2009). Investor sentiment and stock returns: Some international evidence. *Journal of Empirical Finance, 16*(3), 394-408.

Shleifer, A. and Vishny, R. W. (2010). Unstable banking. *Journal of Financial Economics, 97*, 306-318.

Shilling, J. D. and Sing, T. F. (2007). Do institutional real estate investors have rational expectations? *Working paper*.

Stambaugh, Robert F. (1999). Predictive Regressions. *Journal of Financial Economics, 54*, 375-421.

Stiglitz, J. (1990). Symposium on bubbles. *Journal of Economic Perspectives, 4*(2),13–18.

Wheaton, Q. C. and Nechayev, G. (2008). The 1998-2005 Housing "bubble" and the current "correction": What's different this time? *Journal of Real Estate Research, 30*, 1-26.

Zhou, J. (2010). Testing for cointegration between house prices and economic fundamentals. *Real Estate Economics, 38*(4), 599-632.