ESSAYS ON HOUSING BUBBLES: TESTING, ESTIMATION, AND FUNDAMENTAL ANALYSIS

By

KELI XIAO

A dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

written under the direction of

Dr. Yangru Wu

and approved by

____

____

____

____

Newark, New Jersey

October, 2013

ABSTRACT OF THE DISSERTATION

Essays on Housing Bubbles: Testing, Estimation, and Fundamental Analysis

By KELI XIAO

Dissertation Director: Dr. Yangru Wu

Housing bubbles often cause serious problems to local economies and sometimes even cause global financial crisis. This dissertation addresses the following important questions about housing bubbles: What are the origins of housing bubbles? How do we estimate housing bubbles? What are the possible economic variables that relate to housing bubbles? I focus on the estimation of housing bubbles and the analysis of housing fundamentals and bubbles.

The first essay studies housing bubbles in the U.S. during the subprime crisis. Using the Kalman filter, I find significant evidence of a housing bubble in the U.S. city-average level and in six metropolitan areas in the U.S., including Los Angeles, San Francisco, Chicago, New York, Boston, and Miami, between 2004 and 2007. Population and income are found to be the main factors that affect housing fundamentals. Expected capital gains are found to be significantly associated with housing bubbles. Furthermore, I find a high correlation between housing bubbles and housing vacancy rates.

The second essay locates the time stamp of the origin and collapse of housing bubbles using augmented Dickey-Fuller (ADF) tests. An accumulated recursive regression and a rolling recursive regression are applied to test for U.S. housing bubble. Test results confirm explosive behavior in the housing price process, which serves as a necessary condition for the existence of housing bubbles, for the U.S. aggregate and some U.S. cities during the subprime crisis. I find significant evidence of explosiveness in the rent data for most of the cities in the sample, but not in the U.S. aggregate data. Furthermore, the results show a trend of the U.S. housing bubble moving from the east coast to the west coast.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to Professor Yangru Wu, the dissertation committee chair and my advisor, for the care, understanding, help, and support that he provided to me during the past six years. I appreciate that he brought me into the finance area, taught me from the beginning, and led me stepping into the career in finance academia. It was my great honor to work closely with him and receive the supervision and advises from him. This dissertation would not have been possible without his guidance.

I would like to extend my thanks to Professor Hui Xiong for his constant help, advises, and care during my entire Ph.D. study. Especially, I appreciate that he brought the sense of “big data” into my studies and life. I would also like to thank Professor Robert H. Patrick for the constructive suggestions to complete my dissertation and the help and encouragement he has provided to me during the job interview process. I also wish to thank Professor Ren-Raw Chen from Fordham University for being the external committee member for this dissertation and providing the valuable comments.

Special thanks go to Professor Glenn Shafer, Dean of the Rutgers Business School, and Professor Ivan Brick, chair of the Finance and Economics Department, for their understanding and care after I entered the Ph.D. program. I would also like to express my gratitude to Professor Abraham Ravid, Professor Darius Palia, Professor Tavy Ronen, Professor Ben Sopranzetti, and Professor Phil Davis for their lectures, advises, and help. In addition, thank Goncalo Filipe, Jane Foss and Luz Kosar for the great support in administration. Thank my classmates, colleagues, and friends in the Ph.D. program of Rutgers Business School for their help and friendship.

Last but not the least, I own a debt of gratitude to my family. I want to express my great appreciation to my wife and my parents for the warmest care and understanding they have been providing to me, especially in the past several years when I was completing this dissertation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Many times in the history, housing bubbles caused extremely serious financial crisis. In 1980s, the housing price in Japan increased up to 30% each year. After the burst of the bubble in 1991, the increase of bankrupt companies, unemployment rate, bad debt in banks and financial institutes, and the decrease of national wealth and consumption power caused a huge depression of the Japan economy. The housing bubble in Thailand from 1988 to 1997 caused 20-30% annual increasing rates in the housing price. In 1997, the bubble burst, and triggered a serious financial crisis of South-Eastern Asia. In 2006, a global financial crisis was caused by the crash of the U.S. subprime loan market. Housing bubbles generally result over consuming and a high housing vacancy rate. On one hand, the limited resources in a nation cannot support high speed increase in its housing market in long term. On the other hand, when the bubble increases, it reduces the fairness of wealth distribution and enlarges the gap between the rich and the poor.

1.1 The Definitions of Bubbles

The literature on bubbles suggests three ways to define bubbles. First of all, the most common way is to define bubbles based on the fundamental value of related assets. Garber (2000) suggests that a bubble is a price movement that is in explicable based on the fundamental. Kindleberger (1992) explains bubble as a status that assets or output prices increase at a rate that is greater than can be explained by the market fundamental. This group of definitions provide straight forward approaches to test and estimate bubbles. However, some researchers argue that the fundamental value is difficult to measure because of the uncertainty of the future. Second, a bubble can be defined based on the price movements. Kindleberger (1996) claims that a bubble as an upward price movement over an extended range that implodes then. Brunnermeier (2007) describes that “bubbles are typically associated with dramatic asset price increases, followed by a collapse”. These definitions avoid the assessment of fundamental values, whereas abnormal price movements are not the only market characteristics during bubble period. Many other features are expected to associate with bubbles, such as the increase of the gap between the rich and the poor, high housing vacancy rate, and imbalanced market structure. Moreover, the fluctuations in price are hard to be simply explained by bubbles alone. The seasonal fluctuations of industries, the seasonal fluctuations of sales, and new applied government policies affect the price as well. Finally, bubbles can be defined by the generating process. Case and Shiller (2003) defines bubble as a status in which excessive public expectations of future price increases cause prices to be temporarily elevated. However, the generation of an asset bubble is an extremely complex process, hence it should not be concluded by a simple causation. Previous research shows the evidences that bubbles can be caused by different factors. As the earliest, Adam Smith (1776) states that "overtrading" moves the market price away from its fundamental. Kindleberger (1996) argues that it is due to "some sudden advice many times unexpected". Shiller (2000) introduces 12 factors that contributed in the Internet bubble. Also, some other factors have been listed as well, such as synchronization risk (Abreu & Brunnermeier, 2002; 2003), heterogeneous beliefs about asset fundamentals (Scheinkman & Xiong, 2003), and delegated investments (Allen & Gorton, 1993). Therefore, I consider the first type of definitions for the bubble estimation and testing in this dissertation, and detailed discussions are provided in Chapter 2 and Chapter 3.

1.2 Asset Fundamental Values

According to the first type of definition, a bubble is defined as an unexpected increase of the asset price away from its fundamental value. However, another important question arisen after that: what is the fundamental value? In the commodity market, the fundamental value of a product is the value of utility for its consumption. It can be measured by the price which is determined by the consumption based supply and demand. Speculators can push the price away from the fundamental by hoarding related commodities. The famous commodity bubbles happened in the history are Netherland's Tulip mania (1630s), Great Britain's South See Bubble (1720s), the Mississippi Company Bubble (1720s) and the Clivia Bubble in China in 1980s. In the security market, the fundamental value is determined by the constant future cash flows and the expected capital gain, so, the fundamental value is measured by related risk and return. In the real estate market, the same method in assessing the fundamental value of securities can be applied if considering houses as investment products. However, houses can be considered as not just investment opportunities but also consumption products. Using residential real estate as an example, houses are initially built to satisfy the consumption demands. The fundamental value should be the equilibrium price determined by the owner occupations by assuming zero capital gains (Levin & Wright, 1997). So, the investment transactions affect the market supply and demand, hence the housing prices move away from the original equilibrium price. However, one problem is that it is hard to find this equilibrium in the real world. Considering the advantages and drawbacks, on one hand, this dissertation agrees that the housing fundamental is the equilibrium price determined by the housing market without investment activities. The idea is used in the proposing hypothesis for the fundamental analysis and bubble analysis. On the other hand, the estimation of housing bubbles is based on the traditional pricing model to fit the data.

In the following chapters, I present two essays on the study of housing bubbles. The first essay, in Chapter 2, focuses on the estimation of housing bubbles and the fundamental analysis. In Chapter 3, the second essay attempts to locate the timeline of the U.S. housing bubbles and study the migration of housing bubbles during the subprime crisis. Finally, in Chapter 4, I conclude this dissertation and discuss the potential future work.

CHAPTER 2

MEASURING HOUSING BUBBLES: AN EMPIRICAL STUDY

2.1 Introduction

As suggested by the title of this chapter, the main objective of this essay is to investigate the generating process of housing bubbles and study the experiences of the price bubble of the U.S. housing market. The related work on this topic can be summarized in two parts: the fundamental analysis and the bubble analysis. For the fundamental analysis, many studies use the market price in an non-bubble period as the asset fundamental value to apply the empirical investigation and identify the major factors which highly contribute to the housing fundamental (Levin & Wright, 1997; Muellbauer & Murphy, 1997; Stevenson, 2008); however, lack of studies provide evidence to show these factors have the same effects on housing fundamental in the bubble period. Regarding the bubble analysis, a variety of models are introduced for the bubble estimation (Blanchard & Watson, 1982; Wu, 1995; Campbell & Shiller, 1988), and some studies focus on the analysis of the major factors of housing bubbles (Case & Shiller, 2003; Cambell, Davis, Gallin, & Martin, 2009) and the behavior of traders (Case & Shiller, 1988). However, lack of the studies apply the factor analysis based on the bubble estimation. In this essay, I estimate the housing bubbles using the Kalman filter. Based on the estimation, I jointly study the housing fundamental for both the non-bubble period and bubble period. And I empirically investigate the major factors which are highly correlated to the housing bubbles.

Based on previous research on asset bubbles, a bubble can be defined as an unexpected deviation of the market price from its fundamental value. In the real estate market, the assessment of market fundamental can be different. A common approach is to use the rental flows and the expectation of capital gain to value a property. However, some studies state that the expectations are based on subjective, hence the valuation is inherently subjective but not necessary to reflect the true fundamental value (Shiller, 2000). Some studies define the fundamental value of a property by assuming no expected capital gains. For example, Levin and Wright (1997) believe that the fundamental valuation of a property can be defined as the sum of the price determined by owner occupation if the capital gain and the present value of expected capital gain are both zero. Stevenson (2008) verifies this argument by investigating the Irish housing market. Case and Shiller (2003) also supports this argument and links it to the definition of housing bubbles. They believe that "the tendency to view housing as an investment is a defining characteristic of a housing bubble."

Brunnermeier (2007) summarizes the models for asset bubbles as four categories: rational bubbles with symmetric information, bubbles with asymmetric information, bubbles caused by limited arbitrages, and bubbles caused by heterogeneous beliefs. I think the last three categories can be combined into one. Therefore, the models in estimating bubbles can be generally classified into two types. The first type of models is the rational bubbles under symmetric information which assumes all agents have rational expectations and identical information. In these models, bubbles are considered as the deviation the equilibrium price determined by the efficient market. For the market price an asset, p_t, it can be indicated by the sum of its fundamental value, p_t^f, and the bubble component b_t which follows an explosive process, such that $b_t = E_t[\frac{1}{(1+i)} b_{t+1}]$ (Blanchard & Watson, 1982). The empirical tests under this class of models initially presented by Flood and Garber (1980) who stated that "at any point in time the price must have a positive part that grows at an expected rate of *r*" then bubbles are able to start. Diba and Grossman (1988) test whether the explosive process of stock price is larger than that of dividend process to check whether bubble exists. In addition, Wu (1995; 1997) provides the evidences of the existence of bubbles both in foreign exchange market and the US stock market by estimating bubbles using Kalman filter. Theoretically, rational bubble does not necessary to end because the bubble component is expected to grow at a constant rate. Traders will keep buying bubble assets even if they know the prices are already above the fundamentals. However, one significant advantage of this form of models is that the estimating process is straight forward under its assumptions. In the second type of models, studies challenge symmetric information framework from generally three aspects. First, they believe symmetric information does not exist in real-world. An evidence provided by Allen and Gorton (1993) is that fund managers can get benefit from buying bubble assets and resell them to their customers. Second, some studies challenge the efficient market hypothesis, which states that well-informed rational investors will pull the price back to its fundamental by claiming the cost and risk to trade against bubbles. Shleifer and Vishny (1997) find that fund managers consider more about short-term price variability because they do not want to see any temporary losses. It has also been pointed out by Abreu and Brunnermeier (2003) that rational traders generally like to ride bubbles but not attack it. Many empirical studies provide evidence for this statement. For example, Brunnermeier and Nagel (2004) find that between 1998 and 2000, many hedge funds were continually holding and buying overpriced technology stocks. They, as the informed investors in the market, did not correct the market price as expected. Third, heterogeneous beliefs would also be a reason for the emergence of bubbles. The idea is that even if a trader knows an asset is overpriced in the market, (s)he would still buy it, because it is possible that somebody else would have higher evaluation on the same asset in the future. In such case, the price would be able to continually increase until all people have the same valuation of the asset.

Regarding the real estate market, Case and Shiller (1988) provide the evidences of the behaviors that people keep buying when the price is already too high. Also, rents are generally used to understand the fundamental of houses. Many methods to measure bubble levels are based on the price-to- rent or rent-to-price ratios (Himmelberg, Mayer, & Todd, 2005; Cambell, Davis, Gallin, & Martin, 2009). Stevenson (2008) provides evidence from Irish housing market in modeling housing market fundamental by population, income, housing stock, and interest rates (also see Levin and Wright (1997); Muellbauer and Murphy (1997)). Kim and Suh (1993) examine the speculation and price bubbles in the Korean and Japanese real estate market by testing the existence of rational bubbles and find significant evidence of bubbles during the 1974-1989. In addition, Downs and Guner

(2013) investigate how different factors of production complement or substitute with each other across commercial real estate markets, can conclude the supports of the substitution hypothesis that appraisal-based information substitutes for transactions-based information.

In this essay, I follow the rational bubble under the symmetric information framework to estimate the housing bubbles for two main reasons. First, rational bubble models provide straightforward methods to model and estimate bubbles, while the main focus of other types of models is on the behaviors of traders. Second, it is reasonable to assume the symmetric information in the housing market. The transactions in the housing market are not processed in high frequency. Most traders can be assumed to have enough research on the assets they want to trade; hence they understand the fundamental value of the assets. Then, I empirically test the major factors that affect the housing fundamentals and bubbles.

I describe the process of the housing bubble generation as three stages. In the first stage, non-bubble stage, only consumption transactions are processed in the housing market. I can assume that traders have symmetric information about the housing fundamental in this stage, because people would usually do enough research on the houses they want to purchase for living. Therefore, the housing price increases stably because of the growth of population and income and the increasing user cost of houses (Muellbauer & Murphy, 1997). In this stage, the housing price is very close to its fundamental. In the second stage, bubble growth stage, investment transactions increase in the housing market, but no investor can affect the market by themselves. I assume that traders still have symmetric information about the housing fundamental; however, it is not necessary to be common knowledge that everyone knows that all the other traders also know that the fact. In such case, even if most of traders understand the housing price exceeds its fundamental value, they would still buy. In other words, in this stage, although people understand that they are buying bubble houses, most traders in this stage would pretend that they still have a very high expectation on the price and keep buying. This expectation on price, in fact, must be considered as the expectation on bubble. Thus, rational bubble models would be well fitted in measuring the level of bubbles. Furthermore, when the price increases with bubbles, the rental housing vacancy rate goes up as well the supply of rental houses increases as rational housing investors would put their houses into the rental market. Second, regarding the homeowner houses, because the houses they rent out would be counted as rental houses, the homeowner housing vacancy rate would be relatively stable in such stage. In the third stage, bubble burst stage, the housing bubbles burst for different reasons. For example, some negative news and reports related to the housing market suddenly reduce traders' expectation, and lead market collapse. Even if there is no such event, the housing bubbles are still going to burst. Let us consider the following scenario of the housing market. In the stage that investment transactions have become the main demand in the housing market, and some big investors have the power to affect the market by themselves, holding a large amount of bubble assets is very risky. Thus, those big bubble asset holders will seek the ways to quit the market; then the market collapses. While the analysis of the reasons of the burst of housing bubbles (in the third stage) is important, I focus on the study of the generation of housing bubbles (the first and second stage) in this essay.

The remainder of the essay is organized as follows. In section 2.2, I introduce the procedure in estimating the housing bubbles and the regression models that I use for the study on the effects of related variables on housing bubbles. In section 2.3, the data for the estimation and the empirical study is introduced. And the testing results and analysis are discussed in section 2.4. Finally, I summarize this essay in section 2.5.

2.2 Models

2.2.1 The Estimation of Housing Bubbles

Housing bubbles can be estimated by using the Kalman filter technology. Wu (1995; 1997) provides its implementation on the estimations of stock market bubbles and foreign exchange bubbles. It starts with a common real estate pricing model as following. First, the housing price is determined by the discounted future price and rent flows, which can be described as follows,

$$P_t = \frac{1}{1+C_t} E_t(P_{t+1} + R_{t+1}), \tag{2.1}$$

where P_t is the housing price at time t, R_t is the rent flow at time t, and C_t is the required return on the housing investment (the annual cost rate of holding houses), which is assumed to be a constant in this model, say *C*. Then, according to Campbell and Shiller (1988), by taking log of the both sides of (2.1), I would get following equation as one of the solutions.

$$p_t = p_t^f + b_{t,} \tag{2.2}$$

where

$$p_t^f = \frac{k-\gamma}{1-\rho} + (1-\rho)\sum_{i=0}^{\infty} \rho^i E_t r_{t+1+i} \tag{2.3}$$

and

$$b_t = \lim_{i\to\infty} \rho^i \, E_t p_{t+i}. \tag{2.4}$$

A different form of (2.4) can be written as

$$E_t(b_{t+1}) = \frac{1}{\rho} b_t = (1+exp(\overline{r-p}))b_t, \tag{2.5}$$

where $p_t = \log(P_t)$, $r_t = \log(R_t)$, $\gamma = \log(C)$, $\rho = \frac{1}{1+exp(\overline{r-p})}$, $\overline{r-p}$ is the average log rent-price ratio, $k = -\log(\rho) - (1-\rho)\log(\frac{1}{\rho} - 1)$, and 0< ρ<1. Thus, the logarithmic housing price, p_t, can be expressed by a fundamental component, p_t^f, and a bubble component, b_t.

As shown in (2.3) - (2.5), p_t^f, the fundamental of log price, is determined by the expected rental payments. And b_t in (2.5) can be rewritten as:

$$b_t = \frac{1}{\rho} b_{t-1} + \varepsilon_{b,t} = (1+q)b_{t-1} + \varepsilon_{b,t}, \tag{2.6}$$

where $\mathrm{E_t}(\varepsilon_{b,t}) = 0$ with variance $\sigma^2_{\varepsilon_{b,t}}$, $q = \frac{1}{\rho} - 1 = exp(\overline{r-p})$. This explosive process for the bubble component, b_t, drive the market price, p_t, away from the its fundamental value, and also drive it to behave following an explosive process. In the housing market, the origin of bubbles can be assumed to be the initial event of investment, and its explosive process is driven by a continually increasing expectation on the property. As an important characteristic of rational financial bubbles, a pattern of stochastically explosive or mildly explosive behavior is considered as the sign of the existence of bubble. In other words, an explosive p_t is the necessary condition to conclude there is a price bubble. In addition, if $\{p_t^f\}$ does not behave as an explosive process, an explosive p_t is the sufficient condition of the presence of price bubble.

Based on the data, since both the price and rent index appear to have unit root. I consider apply the estimation using the difference form of (2.2). And I assume that the difference of the housing fundamental, log rents, follows an ARIMA(h,1,0) process as follows:

$$\Delta r_t = \mu + \sum_{j=1}^{h} \varphi_j \, \Delta r_{t-j} + \delta_t, \quad \delta_t \sim N(0, \sigma_\delta^2) \tag{2.7}$$

where *h* is determined by the data. This equation can be written in the companion form:

$$Y_t = U + GY_{t-1} + \upsilon_t, \tag{2.8}$$

where $\mathrm{Y_t} = (\Delta r_t, \Delta r_{t-1}, \dots, \Delta r_{t-h+1})'$ and $\upsilon_t = (\delta_t, 0, 0, \dots, 0)'$ are both *h-vectors*, and $\mathrm{G} =$

$\begin{bmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_h \\ 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \vdots & \cdots \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is *h* x *h-matrix*.

Taking the first difference of (2.2) and using (2.8), I get:

$$\Delta p_t = \Delta r_t + M\Delta Y_t + \Delta b_t + \varepsilon_{p,t}, \qquad \varepsilon_{p,t} \sim N(0, \sigma^2_{\varepsilon_{p,t}}) \tag{2.9}$$

where $M = (m_1, m_2, \dots, m_h)'$ and $\mathrm{g} = (1,0,0,\dots,0)$ are both *h-vectors*, and M can be obtained by $M = gG(I - G)^{-1}[I - (1 - \rho)(I - \rho G)^{-1}]$.

To estimate the bubble component which is unobserved, I use Kalman Filter by forming (2.6) which expresses the bubble process, (2.7) describes the log rent process, and (2.9) which states the price process into a state-space form as follows.

$$\hat{X}_{t|t-1} = A\hat{X}_{t-1|t-1} + Bu_t, \tag{2.10a}$$

$$\hat{P}_{t|t-1} = A\hat{P}_{t-1|t-1}A' + \Omega, \tag{2.10b}$$

$$\hat{Z}_{t|t-1} = C\hat{X}_{t|t-1} + Du_t, \tag{2.10c}$$

$$K_t = \hat{P}_{t|t-1}C'[C\hat{P}_{t|t-1}C' + R]^{-1}, \tag{2.10d}$$

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(\hat{Z}_{t|t} - \hat{Z}_{t|t-1}), \tag{2.10e}$$

$$P_{t|t} = [I - K_t C]P_{t|t-1}, \tag{2.10f}$$

where

$$X_t \equiv (b_t, b_{t-1})',$$

$$Z_t \equiv (\Delta r_t, \Delta p_t)',$$

$$u_t = (\Delta r_t, \Delta r_{t-1}, \dots, \Delta r_{t-h})',$$

$$A \equiv \begin{bmatrix} \frac{1}{\rho} & 0 \\ 1 & 0 \end{bmatrix}, B \equiv \begin{bmatrix} 0 & \dots & 0 \\ 0 & \dots & 0 \end{bmatrix}, C \equiv \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, D \equiv$$
$$\begin{bmatrix} 0 & \varphi_1 & \varphi_2 & \dots & \dots & \varphi_h \\ (1+m_1) & (m_2 - m_1) & (m_3 - m_2) & \dots & \dots & -m_h \end{bmatrix},$$

$$\Omega \equiv \begin{bmatrix} \sigma^2_{\varepsilon_{b,t}} & 0 \\ 0 & 0 \end{bmatrix}, \text{and } R \equiv \begin{bmatrix} \sigma^2_{\delta} & 0 \\ 0 & \sigma^2_{\varepsilon_{p,t}} \end{bmatrix}.$$

Also, $\hat{P}_{t|t-1} = E[(X_t - \hat{X}_{t|t-1})(X_t - \hat{X}_{t|t-1})']$ and $\hat{P}_{t|t} = E[(X_t - \hat{X}_{t|t})(X_t - \hat{X}_{t|t})']$ are the error covariance matrices. By running the above forward recursive procedure, I can get the estimated bubbles $\hat{X}_{t|t}$ at each time *t*. To get more efficient estimation and error covariance, I apply a recursive backward pass smoother which the uses all information up to time *T* as following:

$$\hat{X}_{t-1|T} = \hat{X}_{t|t} + L_t(\hat{X}_{t|T} - \hat{X}_{t|t-1}), \qquad (2.11a)$$

$$P_{t-1|T} = P_{t|t} + L_t(P_{t|T} - P_{t|t-1})L_t, \qquad (2.11b)$$

where
$$L_t = P_{t|t}A'P^{-1}_{t|t-1}, \ \mathrm{t = T-1, T-2, \dots 1}. \qquad (2.11c)$$

Before I run the process of the Kalman filter with smoother, I estimate the parameter matrices *A, D*, Ω, and R. The unknown parameters can be estimated by maximizing the following log likelihood function:

$$L = -\frac{1}{2}\sum_{t=1}^{T}(ln[det(CP_{t|t-1}C' + R)] + \xi_t(CP_{t|t-1}C' + R)^{-1}\xi_t) \qquad (2.12)$$

where $\xi_t = \hat{Z}_{t|t} - \hat{Z}_{t|t-1}$. Let β be the unknown parameter vector. After I get the optimal β by maximizing (2.12), the state vector $\hat{X}_{t|t}$, and the error covariance matrix $P_{t|t}$ can be obtained by the Kalman filter procedure, (2.10*a*)-(2.10*f*), and the backward smoother, (2.11*a*)-(2.11*c*).

2.2.2 Housing Fundamental Estimation

According to the analysis in before, the housing fundamental value can be estimated by the inverted demand model (See Muellbauer and Murphy 1997) as follows,

$$P^f = f(\frac{D^T}{POP}, INC, V, O), \qquad (2.13)$$

where D^T is the total demand of houses, *POP* is the population, *INC* is the personal income, *V* is the rate of return, such as the after-tax interest rates, *O* are control variables. For empirical testing, I apply following regression in log-level form:

$$lnP_t = a + \beta_1 \ln(POP_t) + \beta_2\, ln(INC_t) + \beta_3 \ln(V_t) + \varepsilon_t. \qquad (2.14)$$

Also, an error correction version of above regression may be tested as well based on the unit tests for each variables and the cointegration test:

$$\Delta lnP_t = a' + \beta_1{}'\Delta \ln(POP_t) + \beta_2{}'\Delta\, ln(INC_t) + \beta_3{}'\Delta \ln(V_t) + \lambda\, \Delta ec_{t-1} + \varepsilon_t{}', \qquad (2.14')$$

where *ec* is the error correction term which is the residual vector, ε_t, which can be obtained by (2.14). I can apply above two empirical testing models for the no-bubble period in the housing market.

I similarly apply above regressions for the bubble period which is identified by the bubble estimation. By adding the expected present value capital gains, *G*, and some variables which is assumed to have significant effects on the housing bubble, I reform the regressions above as

$$lnF_t = c + \gamma_1 \ln(POP_t) + \gamma_2\, ln(INC_t) + \gamma_3 \ln(V_t) + \gamma_4(G_{t-1}) + \sum_{i=1}^{n} \mu_i O_{it} + \varepsilon_{1t} \tag{2.15}$$

and

$$\Delta lnF_t = c' + \gamma_1'\Delta \ln(POP_t) + \gamma_2'\Delta\, ln(INC_t) + \gamma_3'\Delta \ln(V_t) + \gamma_4'\Delta \ln(G_{t-1}) + \textstyle\sum_{i=1}^{n} \mu_i'\Delta O_{it} + \zeta\, \Delta ec'_{t-1} + \varepsilon_{1t}', \tag{2.15'}$$

where O is a set of the control variables including rental housing vacancy rate, homeowner housing vacancy rate, and number of building permits. Three models are applied for each regression by setting F_t to the housing price index, P_t, the estimated bubble component, B_t, and the estimated fundamental component, P_t^f.

I summarize the testable hypothesis which will be tested and discussed in this essay as follows, *Hypothesis 2-1*: In the no bubble period, the housing market price is determined by the fundamental factors such as population, personal income, and interest rates.

Hypothesis 2-2: In the bubble growing period, the housing market follows the rational bubble framework.

Hypothesis 2-3: In the bubble period, the housing market price is additionally determined by the expected present value capital appreciation besides the fundamental factors. Also, associated with the increase of bubble, the rental housing vacancy rate would increase while the homeowner housing vacancy rate would decrease.

2.3 Data

For the estimation of housing bubbles, I collect the monthly housing price index and rent index in eight cities of the U.S. (Los Angeles, San Francisco, Miami, Chicago, Boston, Detroit, New York, and Cleveland) plus the U.S. city-average price and rent data from January 1991 to January 2012 (254 observations). For prices, I use the Case-Shiller S&P 500 Home Price Indices. For rents, I use the Rent of primary residence from CPI monthly reports published by U.S. Bureau of Labor Statistics (BLS). In *Figure 2.1*, I plot the time series trajectories of the real U.S. city-average housing price index and rent index. Both the price and rent are normalized to 100 at the beginning of the samples (January 1991). It can be seen that the price series began to surge in around 2000, and reached the peak in between late 2005 and early 2006, while the rent series increased smoothly in the general time. *Figure 2.3* plots the trajectories of the real housing price index and rent index for the selected eight U.S. cities.

*** *Figure 2.1* and *Figure 2.3* ***

In addition, a variety of data series were collected for the studies on the bubble and no bubble period of the U.S housing market. The data include the variables which were generally used in related studies on real estate fundamental, such as population, personal income, risk free interest rates, and the variables which are assumed to be highly correlated to housing bubbles, such as the housing vacancy rates[1] (quarterly) and the number of approved building permits. For the population, I use the monthly estimated population data provided by the U.S. Census Bureau. In the tests, I choose the age group of 25-44 years as the major buyers in the housing market. For the personal income, I use the real disposable personal income per capita which is provided by the U.S. Bureau of economic analysis. For the risk free interest rates[2], I use the average rate of the U.S. treasury bonds. In addition, the vacancy rates are separately recorded quarterly as the rental housing vacancy rate and homeowner housing vacancy rate, which are plotted in *Figure 2.2* (from 1992 to 2012). I

[1] The vacancy rates are quarterly reported based on the Housing Vacancy Survey (HVS), which is a part of Current Population Survey (CPS). Both rental housing vacancy rates and homeowner housing vacancy rates are provided. About 72,000 housing units are selected as the sample to do the CPS/HVS questionnaire. According to the methodology of CPS, "The rental vacancy rate is calculated as the ratio of vacant year-round units for rent to the sum of renter- occupied units, vacant year-round units rented but awaiting occupancy, and vacant year-round units for rent. The homeowner vacancy rate is calculated as the ratio of vacant year-round units for sale to the sum of owner-occupied units, vacant year-round units sold but awaiting occupancy, and vacant year-round units for sale."

[2] Average mortgage rates are tested as well. The results are very similar.

can see that both of the rates increase until 2004, then the homeowner housing vacancy rate keep increasing until the mid-2008 while the rental housing vacancy rate decreases until the mid-2008.

*** *Figure 2.2* ***

2.4 Testing and Results

2.4.1 The Estimation of Housing Bubbles

I first estimate the bubbles in the U.S. housing market by using the Kalman filter. The housing price (2.9) depends on the rent process (2.6) which is found contains a unit root. Thus the first differenced process (2.6) can be well approximated by an ARIMA (h, 1, 0) process. To determine the optimal autoregressive order, h, I estimate the log rent process by the maximum likelihood method for different selections of h. I set up 2 as the cap of h. Then for each selection of h, the Akaike information criterion (AIC) is calculated to determine the best h. The model with the smallest AIC, which indicates better fitness on the data, is selected to estimate the rent process.

After that, the Kalman filter framework (2.10a) – (2.10f) can be used to estimate the housing bubbles. I set the initial value (at January 1992) of the housing bubble to 0 with a standard deviation, 10, which indicates that I believe there is no bubble in the year around 1992 but not sure about that. Before the estimation can be processed, six unknown parameters, (ρ, $\sigma_{\varepsilon_{b,t}}$, φ_1, φ_2, σ_δ, $\sigma_{\varepsilon_{p,t}}$), must be determined first. By maximizing the log likelihood function, (2.12), and the optimal choice of the unknown variables are identified.

Table **2.1** reports the point estimates of these variables and the t-statistics for the eight selected urban areas and the U.S. average. It is shown that most of the parameters are significantly different from zero. ρ in around 0.9 denotes that the estimated expectation of the bubble appreciation is around 11%. Based on the optimized parameters, I recursively run the process of Kalman filter, (2.10*a*) – (2.10*f*). And the estimated bubbles, $\hat{X}_t$, are updated and recorded with the error covariance, $\hat{P}_t$, every time.

*** *Table 2.1* and *Figure 2.4* ***

In *Figure 2.4*, the estimated bubbles over the housing prices and the t-statistics are plotted for the U.S. city-average. Several observations can be summarized from this figure. First, based on the estimation results, the actual prices might have been caused by speculative bubbles from April 2002 to August 2008. Second, the estimated bubble to price ratio reaches 0.263 as the maximum in November 2005, which indicates that 26.3% of the actually price is contributed by housing bubbles. Third, the estimated bubbles grew from zero in May 2002 to the peak in November 2005 for more than three years, while it shrunk to zero in less than two years. Finally, the negative housing price bubbles are found from1992 to 2000 and after 2008, which indicate that the rent process grew faster than the price process during these periods. Fourth, based on the related t-statistics, the significance of the estimation at the 5% level is identified during the period July 2004 to August 2007.

However, if the significant estimating numbers are reliable, it is reasonable to believe that the housing bubble does not suddenly come to 17.4% in July 2004. Hence, I assume that housing bubbles exist before July 2004 based on the estimation. Furthermore, I plot the estimated bubbles over the actual housing prices and their t-statistics for the eight urban areas in sample in *Figure 2.5*. As plotted in the figure, the highest level of bubbles is found in Miami in May 2006 up to 32.8%. High level bubbles (above 20%) are also discovered in Los Angeles, San Francisco, and New York in between the late 2005 and middle 2006. Also, based on their t-statistics, the estimations of the four urban areas are found significant at the 5% level in the time period from 2004 to 2007. Also, it shows that in Chicago, the significant housing bubbles are identified from December 2005 to April 2007. In Boston, only several months of significant housing bubbles are found in between January 2005 and November 2005. However, in Detroit and Cleveland, no significant housing bubbles are identified in the 5% level.

*** *Figure 2.5****

After the estimation, I investigate the effectiveness of the estimation. Let $\hat{p}_t$ indicates the estimated value of log housing price index at time *t*, based on (2.2) to (2.5), the estimated bubbles and parameters, I compute the in-sample root mean square error (RMSE) and mean absolute error (MAE) of the estimations for the eight tested cities and

U.S. average. The results are reported in *Table 2.2*. I can see that the RMSE and MAE are around 10% or less for most of estimated samples. However, both of the two error estimators are at about 20% for the estimation in Miami.

*** *Table 2.2* and *Table 2.3****

2.4.2 Housing Fundamentals, Bubbles, and Other Variables

As the discussion in section 2.2, I apply the regression for the inverted demand model first. Based on the estimation of housing bubbles, I select the data from January 1994 to April 2000 (76 observations) as the subsample to test the state of non-bubble period. The major concerns in using the original version of the model, (2.14), are the correlations between the independent variables, their stationarity, and also the cointegrating relationships among the variables. With these issues, the error correction version of the model may serve as an improved regression of the original one. As the necessary procedures to apply the error correction estimation, unit root and cointegrating tests are undertaken first. The statistical results of the Phillips-Perron test are reported in *Table 2.3*. As shown in the table that most variables are found not stationary in levels except the expected capital gain; however most of them are stationary if they are first order differenced except the population. Hence, directly using the variables in levels may result in spurious relationships, while the use of first differenced variables would improve the model. The Johansen cointegration results are reported in *Table 2.4*. In the Johansen test, the trace test has the null hypothesis that the number of cointegrating vectors is less than r (here between one and five); while the null hypothesis in max-eigen test is that the number of cointegrating vectors is r, against an alternative of r+1. It shows in the table that both the trace test and the max-eigen test provide significant evidence of one cointegrating vector among the tested variables. Thus, the error correction model should be used to correct the short-run relationships.

*** *Table 2.4* and *Table 2.5* ***

The results of the error correction model, (2.14'), are shown in *Table 2.5*. It shows the coefficients are found anticipated in most of the cases as the hypothesis I summarize in section 2.2. First, both the population and income are significant in expected positive sign, which verify their positive impacts upon the housing fundamental. Second, as expected, the present value of capital gain is found insignificant. It provides the evidence to show that the expected capital gain is not a major factor to affect the housing fundamental. Third, negative sign of the error correction term indicates that it successfully pull the estimations back to the actual data. Finally, the only unexpected result comes to the interest rate, which is found insignificant to the housing fundamental.

After testing the non-bubble period, I also examine the associations between estimated housing bubbles and the same variables. Based on the estimation, for the bubble period, I select the data in the period from May 2002 to August 2008 (76 observations). Similar to the previous test, it has the same issue in directly using the model with variables in levels, (2.15), by considering the stationarity and cointegration. Thus, Phillips-Perron unit root test and Johansen cointegration test are both undertaken before the regression is implemented. Especially, besides the recoded housing price index, I extract the estimated bubble component and compute the fundamental component in order to test the effects of the selected variables on different portions in the price. *Table 2.6* reports the unit root test results. I shows that most of the variables are not stationary in levels expect the income, while most of them are stationary in the first difference expect the interest rate. Lack of stationary in levels indicates that it would cause problems by using (2.15). Also, the Johansen cointegration testing results of the three models are reported in *Table 2.7*. The entire housing price, estimated bubble component, the computed fundamental component are respectively tested in Model (a), Model (b), and Model (c). The Johansen tests show there is significant evidence of one cointegrating vector for each model. Hence, the short-run error correction regressions, (2.15'), should be undertaken.

*** *Table 2.6* and *Table 2.7* ***

The statistical results of the three error correction models are shown in *Table 2.8*. Two major observations can be made as follows. First, significant positive coefficients of expected capital gain are found in Model (a) and Model (b) which respectively test the whole housing price and the bubble component; while its coefficient is insignificant in Model (c) which serve as the test for the housing fundamental. Second, the real disposable income per capita is found has significant impact upon the fundamental component of housing price, while it is found does not contribute in affecting the housing price and its bubble component. These two findings follows the hypothesis of non-bubble period and bubble period which I propose in section 2.2. In addition, the population is found significantly positive for all three models; the coefficients of the error corrections terms in three models are negative as anticipated.

*** *Table 2.8* ***

Finally, I additionally test the effects of the rental housing vacancy rate and the homeowner housing vacancy rate in the three models I apply for the bubble period. Since the vacancy rate data are quarterly reported, I first convert other variables from monthly data into quarterly data. The selected sample after the conversion is from 2002Q1 to 2008Q4 (28 observations). Reasonably, all the variables are assumed to be stationary in the first differenced and have one cointegrating vector as their features in monthly. Thus, the short-run error correction models are undertaken. *Table 2.9* reports the testing results. Several observations can be made as follows, first, the rental housing vacancy rate is significantly positive related to the housing price (Model (a)) and the bubble component (Model (b)), while it is insignificantly associated to the fundamental component (Model (c)). One explanation of this finding is that people are more likely to buy houses in the bubble period for the increasing housing prices. People get their own houses to live, so the rental vacancy rates increase. Second, I find significant evidence of negative association of the homeowner vacancy rate with the housing price (Model (a)) and the bubble component (Model (b)); however, no significant evidence is found for its association with the fundamental component (Model (c)). The explanation is that the raising housing bubbles result the increase of the sales of new houses, and reduce the homeowner vacancy rates. Third, the same as the results from the tests of monthly data, the present value of capital gain is found has a significant positive impact upon the housing price and the bubble component (Model (a) and Model (b)); however, it has insignificant effect on the fundamental component (Model (c)). In addition, negative signs of the error correction terms in the three models are found as anticipated.

*** *Table 2.9* ***

2.5 Summary

In this essay, I show the process to estimate the price bubbles in the U.S. housing market using the Kalman filter under the rational bubble framework. Based on the estimated results, I additionally test the movements of several factors in no bubble and bubble periods. Some interesting findings are eventually reported. First, in the no bubble period, the housing price is positive correlated to population and personal income, while the effect of the expected capital gain on the price is found insignificant. Second, in the bubble period, I find the population and income are significantly positive related to the price as well as the bubble component, while they are insignificantly associated to the fundamental component. Furthermore, a significant positive correlation between the housing bubbles and rental housing vacancy rates are found, and a significant negative correlation between the housing bubbles and homeowner housing vacancy rates are identified. In addition, the testing results are stick with the expectation and hence provide the empirical evidence to support the credibility of the bubble estimation model in fitting the data.

Figure 2.1 Real Price Index and Real Rent Index (U.S City-average)

This figure plots the U.S. city-average price and rent data from January 1991 to January 2012. (254 observations) For prices, I use the Case-Shiller S&P 500 Home Price Indices. For rents, I use the Rent of primary residence from CPI monthly reports published by U.S. Bureau of Labor Statistics (BLS). Both the price and rent are normalized to 100 at the beginning of the samples (January 1991). As shown in this figure, the price series began to surge in around 2000, and reached the peak in between late 2005 and early 2006, while the rent series relatively increased smoothly. After the price bubble burst, price series tend to decrease but rent series still increase smoothly for a long time.

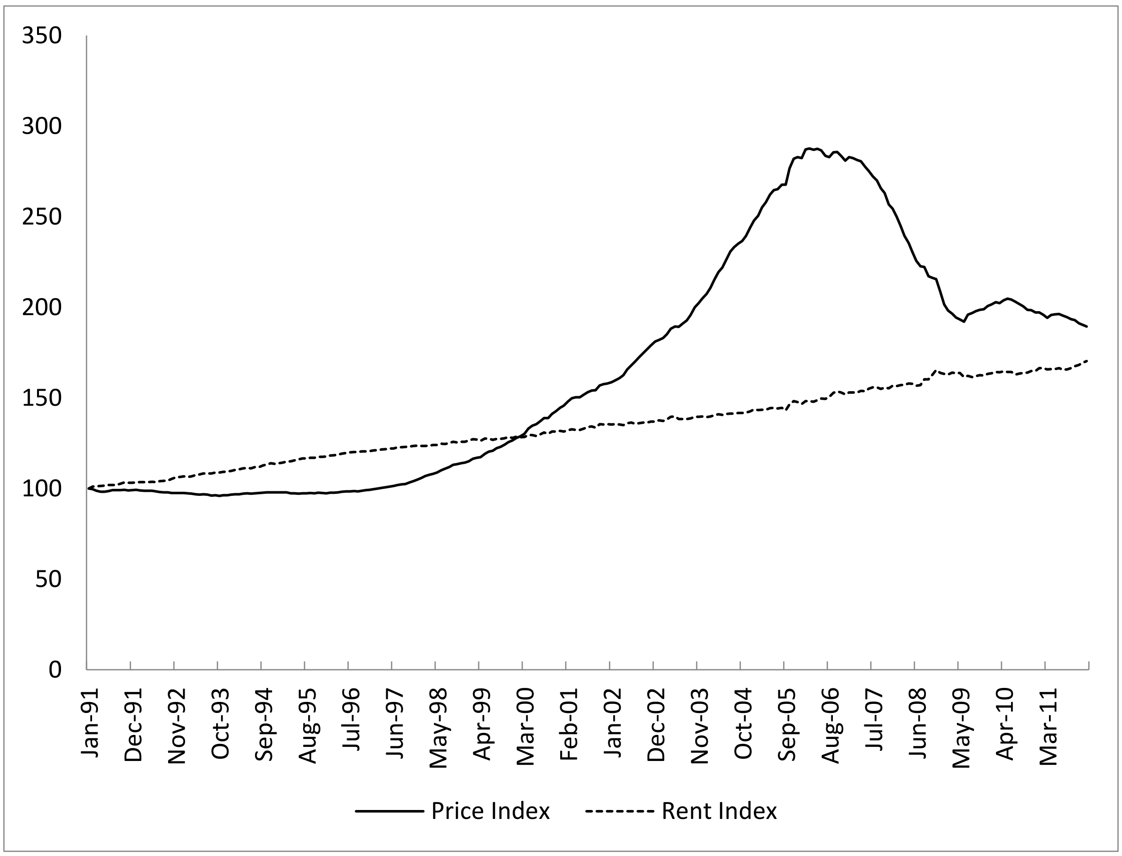


Figure 2.2 The Housing Vacancy Rates (%)

This figure plots the rental housing vacancy rates and the homeowner housing vacancy rates in percentage. The rental housing vacancy rate starts to rapidly increase from early 2000 and become smooth after the second quarter in 2004 when reaching the first peak at 10.4%. The second peak of the rental housing vacancy rates is 11.1% which appears in the third quarter in 2011. An abnormal increase of the homeowner housing vacancy rate appears in late 2004. It reaches the peak at 2.9% in 2008. Based on the Housing Vacancy Survey (HVS), the rental vacancy rate is calculated as the ratio of vacant year-round units for rent to the sum of renter- occupied units, vacant year-round units rented but awaiting occupancy, and vacant year-round units for rent; the homeowner vacancy rate is calculated as the ratio of vacant year-round units for sale to the sum of owner-occupied units, vacant year-round units sold but awaiting occupancy, and vacant year-round units for sale.

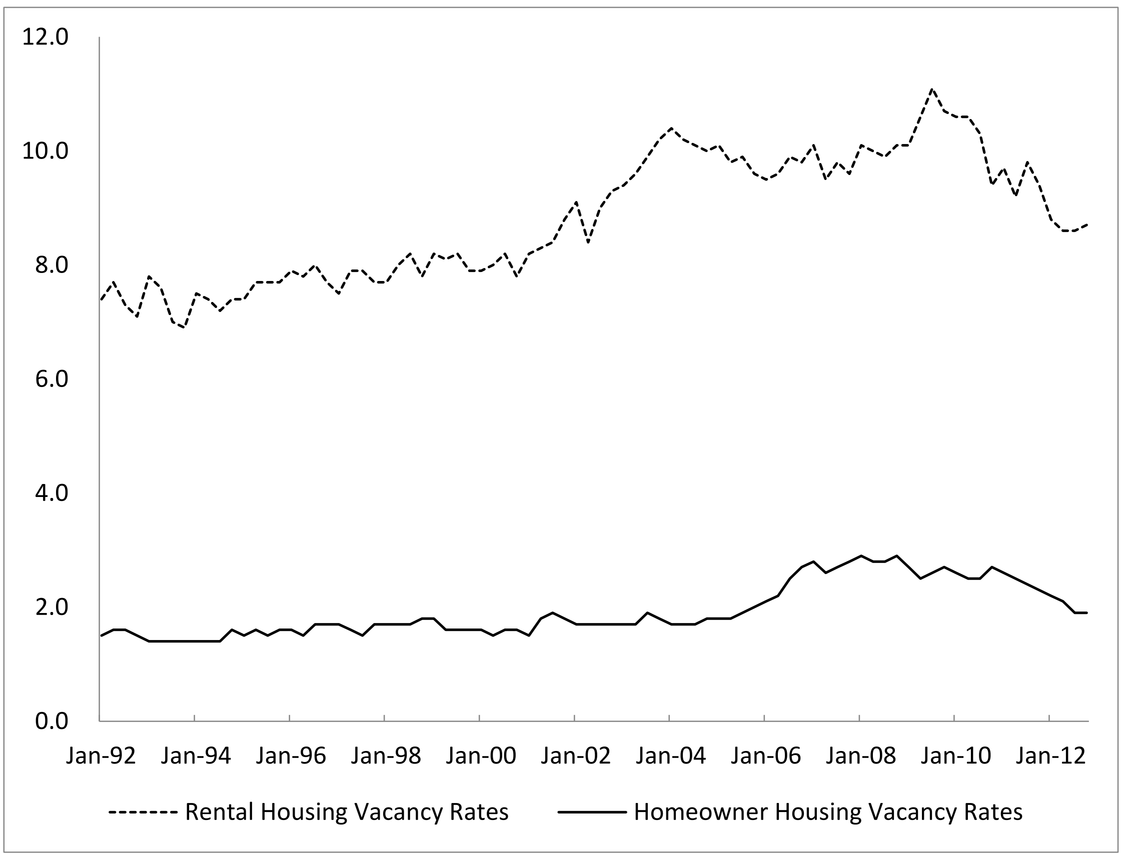


Figure 2.3 Normalized Time Series of Real Price Index and Real Rent Index

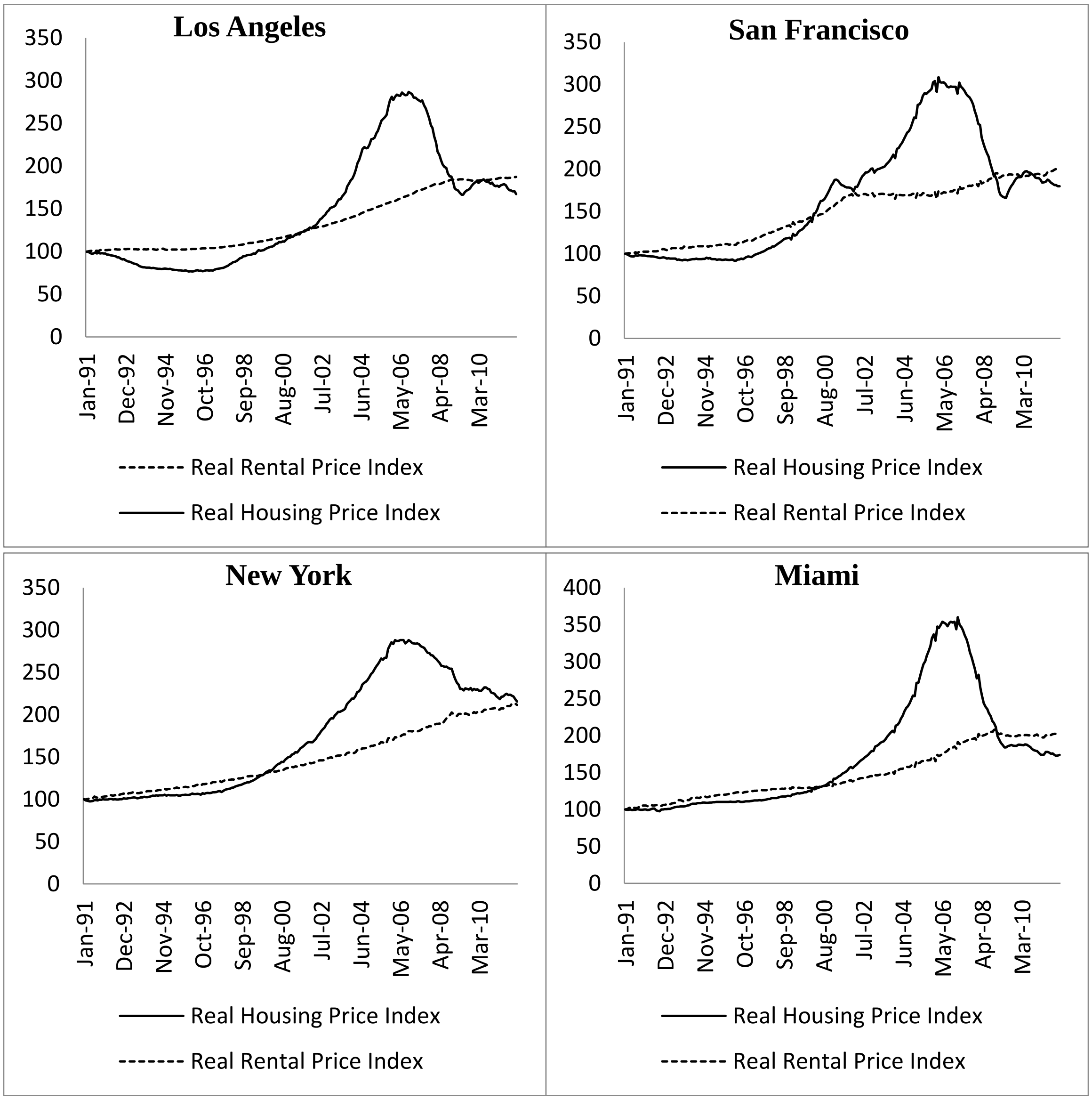


Figure 2.3(Cont.) Normalized Time Series of Real Price Index and Real Rent Index

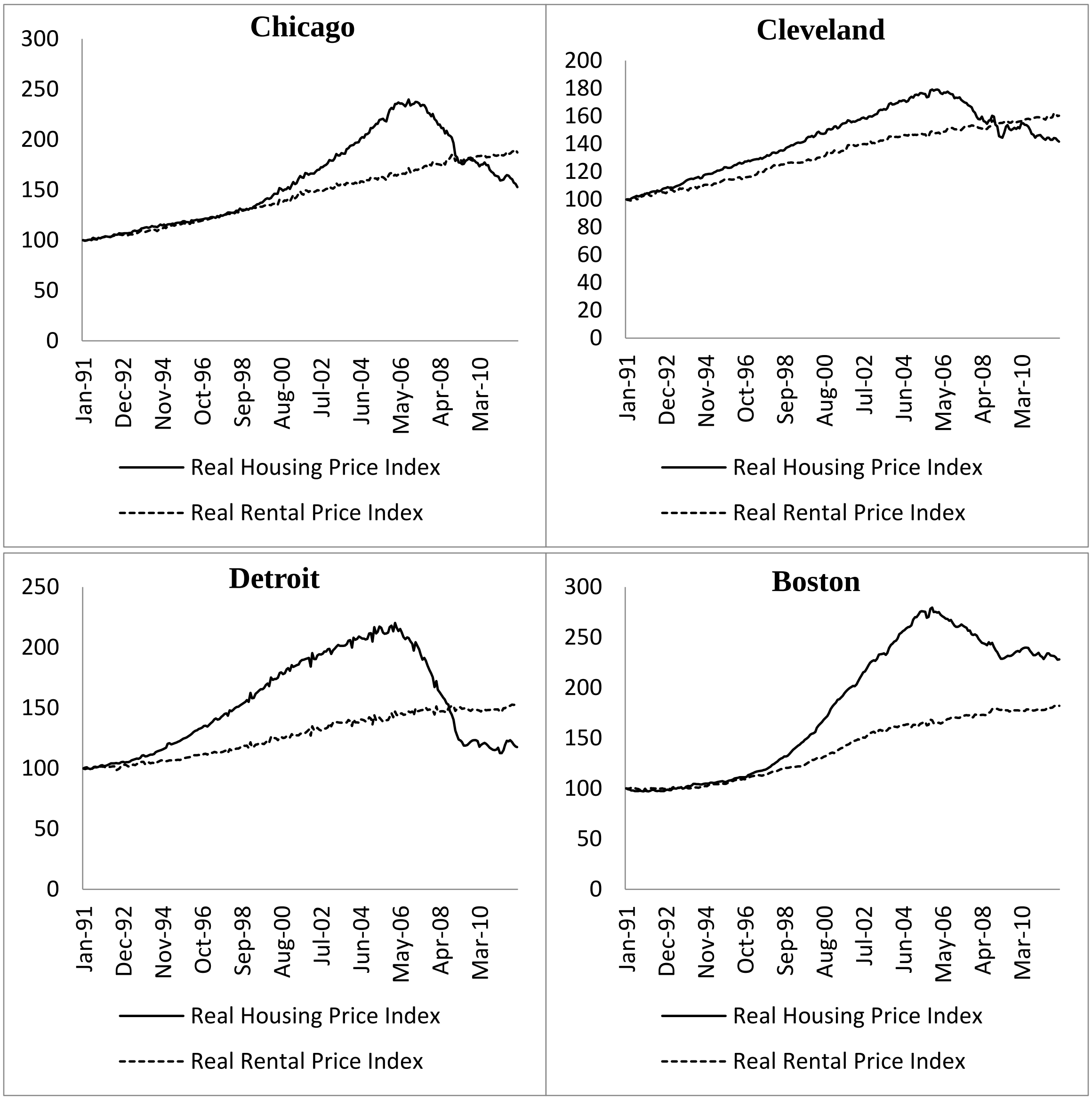


Figure 2.4 Estimated Bubbles-to-Price Ratio and the T-statistic (U.S. Average)

In this figure, the estimated bubbles over the actual housing prices are plotted. Several observations can be summarized from this figure. First, the results show that about the actual prices might have been caused by speculative bubbles from May 2002 to August 2008. Second, the estimated bubble to price ratio reaches 0.263 to the maximum in November 2005, which means that 26.3% of the actually price is contributed by housing bubbles. Third, the estimated bubbles grew from zero in May 2002 to the peak in November 2005 for more than three years, while it shrunk to zero in only one year. Finally, the negative housing price bubbles are found from1992 to 2000 and after 2008.

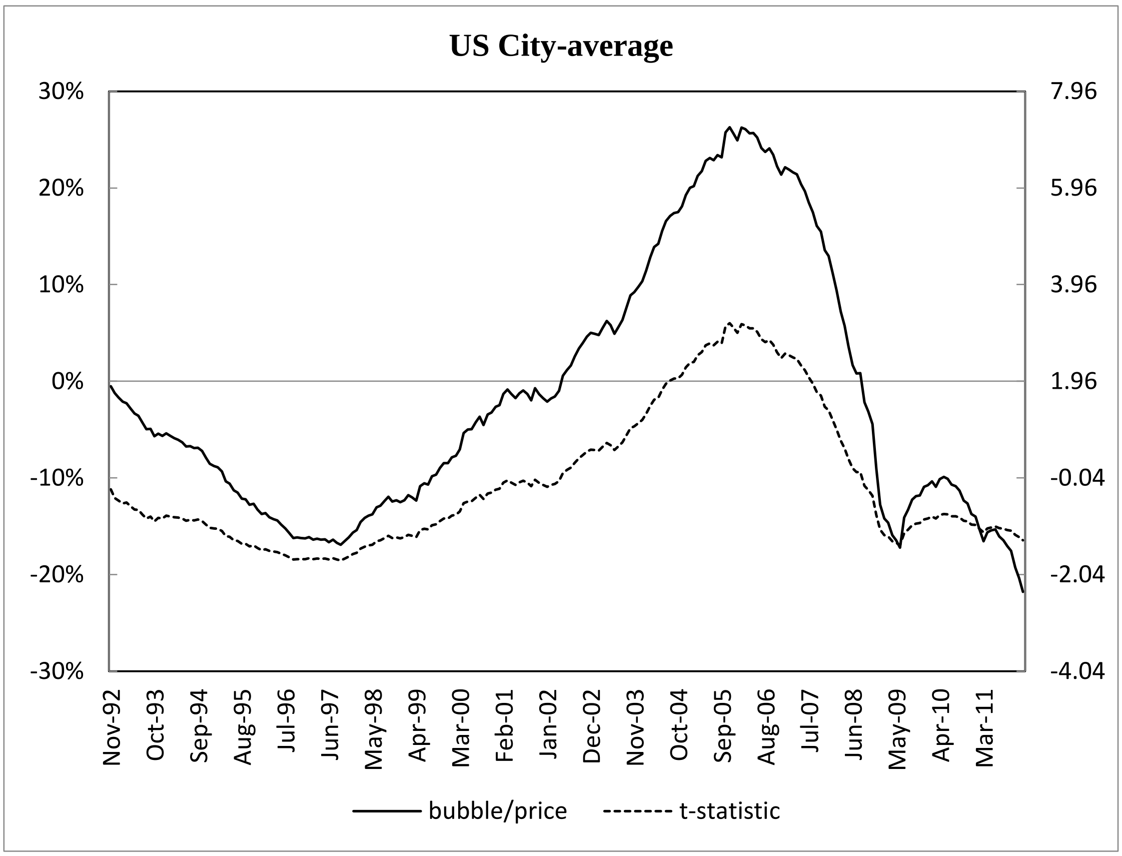


Figure 2.5 Estimated Bubble to Price Ratio and the T-statistic (City-level)

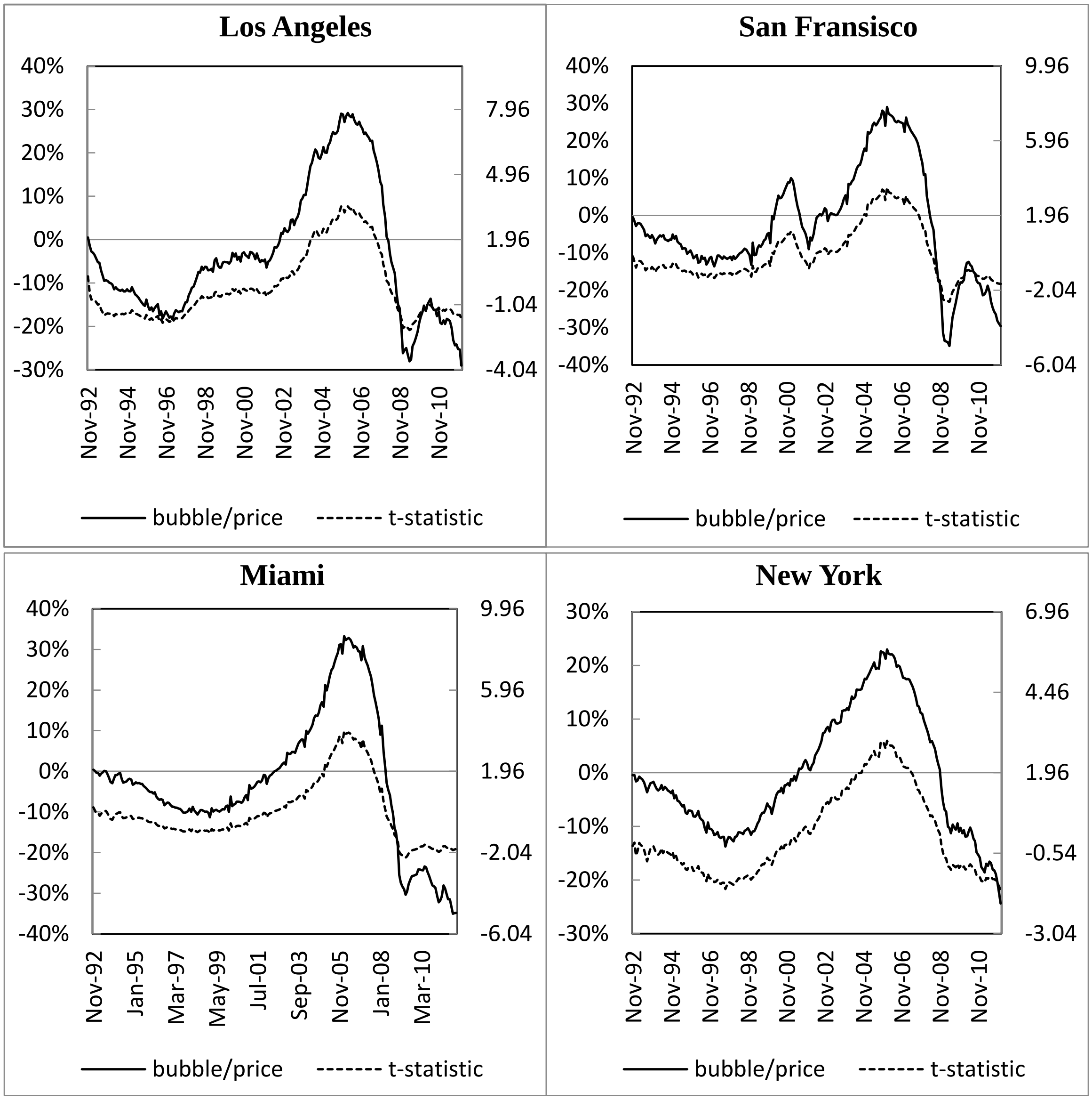


Figure 2.5(Cont.) Estimated Bubble to Price Ratio and the T-statistic (City-level)

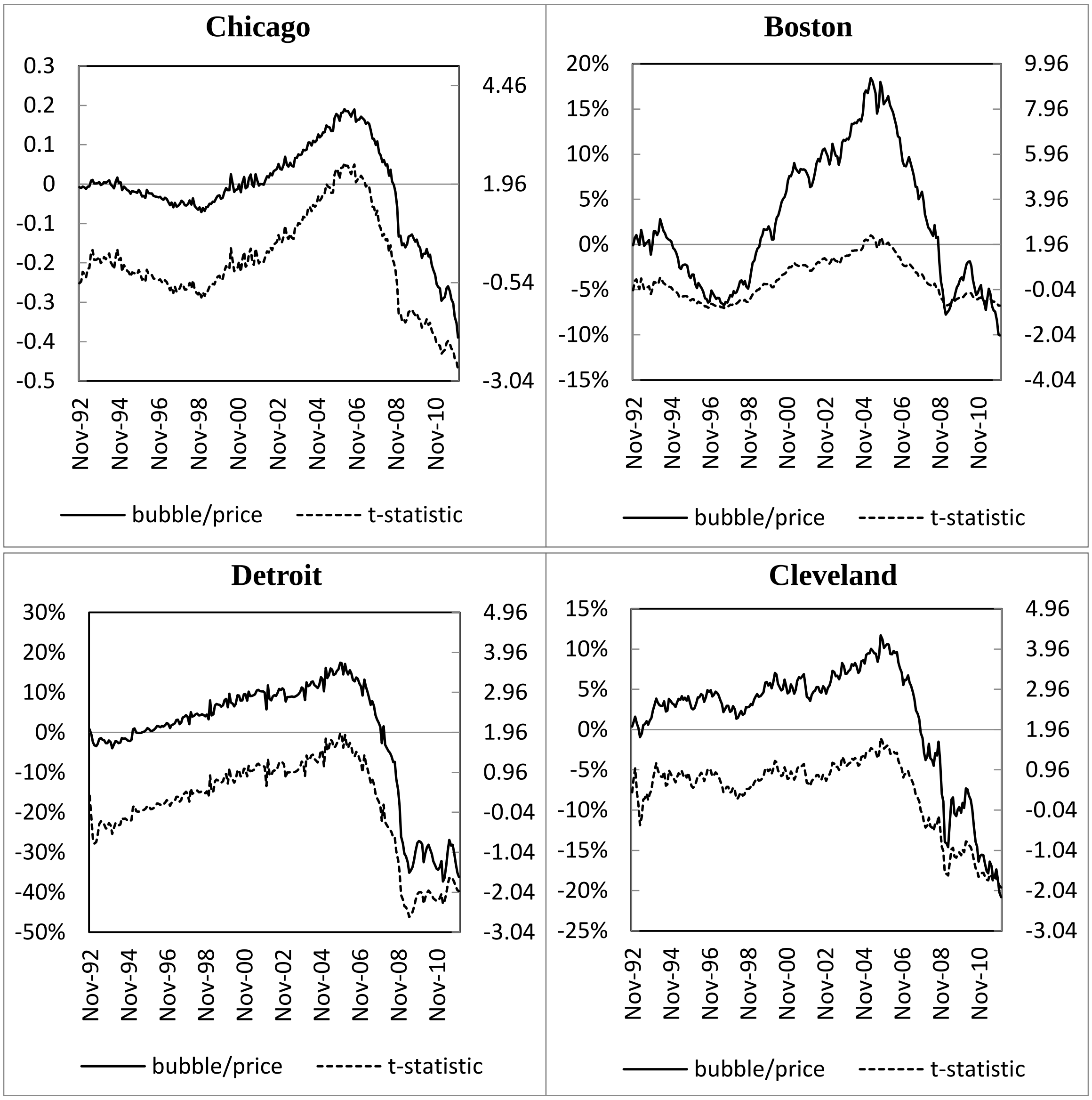


Table 2.1 MLE Estimates of Housing Price and Rent Equations

The optimal estimations of the unobserved parameters and the absolute values of the t-statistics, which would be processed in the Kalman filter, are reported in this table. The six unknown parameters are estimated by the maximizing log likelihood method. As shown in this table, most of the parameters are significant, except φ_1 and φ_2 in some cities. The numbers in brackets are the t-statistics; **statistically significant at the 5% level; *statistically significant at the 10% level

	ρ	$\sigma_{\varepsilon_{b,t}}$	φ_1	φ_2	σ_δ	$\sigma_{\varepsilon_{p,t}}$
Los Angeles	0.901**	0.100**	-0.704	-0.366	0.011**	0.021**
	(21.975)	(6.023)	(1.452)	(0.980)	(22.489)	(16.031)
San Francisco	0.901**	0.100**	-1.045**	-0.573**	0.020**	0.021**
	(43.987)	(3.640)	(3.193)	(2.141)	(20.337)	(9.344)
Miami	0.902**	0.099**	-1.016	-0.539**	0.017**	0.022**
	(28.375)	(5.509)	(2.713)	(1.475)	(22.351)	(14.993)
Chicago	0.900**	0.099**	-0.825**	-0.480**	0.012**	0.014**
	(15.426)	(2.049)	(2.847)	(1.405)	(21.964)	(3.158)
Boston	0.901**	0.101**	-0.620	-0.516	0.009**	0.015**
	(35.467)	(8.153)	(0.892)	(1.474)	(22.238)	(22.333)
Detroit	0.901**	0.100**	-1.046**	-0.560**	0.019**	0.021**
	(38.141)	(4.708)	(3.605)	(2.122)	(22.128)	(13.887)
New York	0.901**	0.100**	-0.711*	-0.346	0.009**	0.016**
	(27.222)	(7.037)	(1.759)	(0.939)	(22.610)	(17.411)
Cleveland	0.900**	0.099**	-0.450*	-0.487**	0.007**	0.010**
	(30.947)	(2.896)	(1.723)	(2.604)	(22.455)	(6.438)
US-average	0.902**	0.099**	-0.680	-0.398	0.007**	0.020**
	(12.889)	(7.464)	(0.822)	(0.703)	(21.788)	(19.737)

Table 2.2 Root Mean Square Error and Mean Absolute Error

The in-sample root mean square error (RMSE) and mean absolute error (MAE) are defined as: $RMSE = \left[(\frac{1}{T})\sum_{t=1}^{T}(\hat{p}_t - p_t)^2\right]^{1/2}$ and $MAE = (\frac{1}{T})\sum_{t=1}^{T}|\hat{p}_t - p_t|$.

	Los Angeles	San Francisco	Miami	Chicago	Boston	Detroit	New York	Cleveland	US Average
RMSE	0.076	0.116	0.200	0.110	0.097	0.069	0.105	0.066	0.113
MAE	0.054	0.093	0.181	0.094	0.082	0.050	0.086	0.052	0.091

Table 2.3 Phillips-Peron Unit Root Tests (January 1994 – April 2000, No Bubble)

This table shows the unit root test results for the variables used in the inverted demand model for the subsample during January 1994 to April 2000 (76 observations). Phillips-Peron tests are applied for each variable by assuming the existence of intercept and trend. The critical value for significance at 90% for the tests is -3.157; * indicates the significance at 10% level. I can see that most of the variables are insignificant in levels except the capital gain, while most of them are found significant in the first difference excluding the resident population. It indicates non-stationary of these variables in levels. Therefore, directly using them in the regression may result spurious relationships, but using them in the first difference is fine. Based on the test of cointegration, I decide whether to use the error correction model or not.

	Levels	First Difference
Housing Price Index	2.362	-8.320*
Resident Population (age 25-44)	-1.647	-10.918*
Interest Rate	-1.784	-11.040*
Real Disposable Income per Capita	-2.592	-13.652*
PV Capital Gain	-9.795*	-46.449*

Table 2.4 Johansen Cointegration Results (January 1994 – April 2000, No Bubble)

This table reports multivariate Johansen cointegration results and the appropriate critical value at the 0.05 level for the data during January 1994 to April 2000 (76 observations). The lag length in the Johansen vector autoregressive model is specified as two. As shown in the table, both the trace test and the max-eigenvalue test result significant statistics with null hypothesis that it has zero cointegrating equation. Thus, I apply the error correction model to the data.

Hypothesized # of cointegrating equations	Trace Statistic	Trace Critical Value (95%)	Max-Eigen Statistic	Max-Eigen Critical Value (95%)
R = 0	117.875*	88.804	58.986*	38.331
R < 2	58.889	63.876	25.991	32.118
R < 3	32.898	42.915	16.838	25.823
R < 4	16.061	25.872	10.635	19.387
R < 5	5.426	12.518	5.426	12.518

Table 2.5 The Statistical Results of the Error Correction Model (January 1994 – April 2000, No Bubble)

This table reports the coefficients and the t-statistics for the error correction model:

$$\Delta lnP_t = a' + \beta_1' \Delta \ln(POP_t) + \beta_2' \Delta\, ln(INC_t) + \beta_3' \Delta \ln(V_t) + \beta_4' \Delta(G_{t-1}) + \lambda\, ec_{t-1} + \varepsilon_t'.$$

All variables are first order differenced. The data used in the test is from January1994 to April 2000, contains 76 observations (75 after adjustments). All the variables are in the logarithmic level. For the capital gain, I use the six-month average price gain with one month lag. As shown in this table, the adjusted R^2 is 0.165. Also, the resident population and real disposable income per capita are significant with positive signs; the interest rate is found insignificantly negative; the present value capital gain is insignificantly positive. In addition, as expected, the error correction term is significantly negative.

Variable	Test Statistics
Intercept	-0.001 (0.784)
Resident population (age 25-44)	12.172** (3.777)
Interest rate	-0.002 (0.274)
Real disposable income per capita	0.448** (3.302)
PV capital gain	0.011 (0.177)
Error correction term	-0.021* (1.659)
Adjusted R^2	0.165

*Notes:*The numbers in brackets are the t-statistics; ** indicates statistical significance at the 5% level; * indicates statistical significance at the 10% level.

Table 2.6 Phillips-Peron Unit Root Tests (May 2002 – August 2008, Bubble Period)

This table shows the unit root test results for the variables used in the inverted demand model for the subsample during May 2002 to August 2008 (76 observations). Phillips-Peron tests are applied for each variable by assuming the existence of intercept and trend. The critical value for significance at 90% for the tests is -3.157; * indicates the significance at 10% level. I can see that most of the variables are insignificant in levels except the capital gain, while most of them are found significant in the first difference excluding the resident population. It indicates non-stationary of these variables in levels.

	Levels	First Difference
Housing Price Index	3.119	-5.480*
Estimated Bubble Component	0.46	-5.264*
Estimated Fundamental Component	-0.11	-4.348*
Resident Population	-2.510	-3.740*
Interest Rate	2.310	-0.965
Real Disposable Income per Capita	-4.461*	-14.068*
PV Capital Gain	-2.123	-7.165*

Table 2.7 Johansen Cointegration Results (May 2002 – August 2008, Bubble Period)

This table reports multivariate Johansen cointegration results and the appropriate critical value at the 0.05 level for the data during May 2002 to August 2008 (76 observations). I apply the tests for three models. Model (a) tests the effects of selected variables on the housing price; Model (b) tests their relationships with bubbles; Model (b) checks how the housing fundamental reacts after these variables. The lag lengths of the three Johansen vector autoregressive model are specified as two. As shown in the table, for Model (a), the trace test indicates 3 cointegrating equations at the 0.05 level, while the max-eigen test indicates one cointegrating equation at the 0.05 level; for Model (b), both the trace test and the max-eigen test denote the existence of one cointegrating equation at the 0.05 level; for Model (c), I find significant trace statistic to reject the null hypothesis that the number of cointegrating equations is less than three, while the max-eigen statistic shows the existence of one cointegrating equation in the 0.05 level.

	Hypothesized # of cointegrating equations	Trace Statistic	Trace Critical Value (95%)	Max-Eigen Statistic	Max-Eigen Critical Value (95%)
Model (a)	R = 0	121.862*	69.819	60.540*	33.877
	R < 2	61.321*	47.856	25.689	27.584
	R < 3	35.632*	29.797	22.214*	21.132
	R < 4	13.417	15.495	12.507	14.265
	R < 5	0.911	3.841	0.911	3.841
Model (b)	R = 0	86.240*	69.819	46.857*	33.877
	R < 2	39.384	47.856	18.248	27.584
	R < 3	21.135	29.797	12.270	21.132
	R < 4	8.865	15.495	8.707	14.265
	R < 5	0.158	3.841	0.158	3.841
Model (c)	R = 0	94.145*	69.819	35.660*	33.877
	R < 2	58.484*	47.856	26.955	27.584
	R < 3	31.530*	29.797	19.022	21.132
	R < 4	12.508	15.495	10.402	14.265
	R < 5	2.106	3.841	2.106	3.841

Table 2.8 The Statistical Results of the Error Correction Model (May 2002 – August 2008, Bubble Period)

This table reports the coefficients and t-statistics for the following regression:

$$\Delta lnF_t = c' + \gamma_1{}'\Delta \ln(PoP_t) + \gamma_2{}'\Delta\, ln(INC_t) + \gamma_3{}'\Delta \ln(V_t) + \gamma_4{}'\Delta(G_{t-1}) + \zeta\, ec'_{t-1} + \varepsilon_{1t}{}'$$

where $F_t = P_t$ in Model (a), $F_t = B_t$ in Model (b), and $F_t = P_t^f$ in Model (c). The three models are applied for the data during May 2002 to August 2008 (76 observations). As shown in this table, the test statistics show that the interest rate is only significant in Model (b), and the real disposable income per capita is only significant in Model (c). The expected capital gain is significant in both Model (a) and (b) but in Model (c).

Variable	Test Statistics		
	Model (a) ΔlnP_t	Model (b) ΔlnB_t	Model (c) ΔlnP_t^f
Intercept	-0.048	-0.863**	-0.004
	(1.480)	(4.853)	(1.300)
Resident population (age 25-44)	68.596**	1161.785**	11.240**
	(4.403)	(4.853)	(2.522)
Interest rate	0.006	-0.257**	0.002
	(0.669)	(1.849)	(0.847)
Real disposable income per capita	0.063	2.696	0.070*
	(0.437)	(1.245)	(1.700)
PV capital gain	2.313**	42.310**	0.936
	(2.710)	(3.404)	(1.591)
Error correction term	-0.014	-0.069	-0.031
	(0.294)	(1.454)	(0.657)
Adjusted R^2	0.213	0.170	0.195

Notes: T-statistics are in brackets; ** indicates statistical significance at the 5% level; * indicates statistical significance at the 10% level.

Table 2.9: The Error Correction Models Using Quarterly Data (2002Q1 –2008Q4, Bubble Period)

This table reports the coefficients and t-statistics for the following regression:

$$\Delta lnF_t = c' + \gamma_1' \Delta \ln(POP_t) + \gamma_2' \Delta\, ln(INC_t) + \gamma_3' \Delta \ln(V_t) + \gamma_4' \Delta(G_{t-1}) + \sum_{i=1}^{n} \mu_i' \Delta O_{it} + \zeta\, \Delta ec'_{t-1} + \varepsilon_{1t}'$$

where $F_t = P_t$ in Model (a), $F_t = B_t$ in Model (b), and $F_t = P_t^f$ in Model (c). The three models are applied for the data during 2002Q1 to 2008Q4 (28 observations). O contains the rental housing vacancy rates, homeowner housing vacancy rates, and the number of building permits.

Variable	Test Statistics Model (a) ΔlnP_t	Model (b) ΔlnB_t	Model (c) ΔlnP_t^f
Intercept	-0.071	-0.778*	-0.004
	(1.048)	(1.718)	(0.243)
Resident population (age 25-44)	37.667	393.784*	6.259
	(1.254)	(1.718)	(0.856)
Interest rate	1.259	-3.408	-0.467**
	(1.202)	(0.534)	(2.806)
Real disposable income per capita	0.113	0.549	0.024
	(1.490)	(1.326)	(1.280)
PV capital gain	6.222**	65.562**	0.862
	(2.596)	(4.172)	(1.604)
Rental housing vacancy rate	0.357*	5.261**	0.159
	(1.720)	(2.882)	(0.818)
Homeowner housing vacancy rate	-0.145	-0.482*	-0.017
	(1.020)	(1.861)	(0.481)
Number of building permits	-0.114	-1.067**	0.009
	(1.631)	(2.103)	(0.506)
Error correction term	0.357**	-1.067**	0.009
	(2.566)	(2.103)	(0.506)
Adjusted R^2	0.210	0.335	0.224

Notes: T-statistics are in brackets; ** indicates statistical significance at the 5% level; * indicates statistical significance at the 10% level.

CHAPTER 3

DATING THE TIMELINE OF HOUSING BUBBLES

3.1 Introduction

The main purpose of this essay is to empirically investigate the experience of U.S. housing market during the subprime crisis. By estimating the timestamp of housing bubbles for both the U.S. aggregate data and the city-level data, I study the housing bubbles, the exuberance in the rental market, and the bubble migration. Based on the literature, a bubble can be defined as an unexpected deviation between the asset market price and its fundamental value. Generally, bubbles begin with an asset which has infinite maturity periods. The uncertainty in the infinite future makes it possible for an asset to have a continually increasing expectation on the price even if its fundamental value does not change. On one hand, according to the theory of rational bubbles, the asset price can be separated into two parts: the fundamental and the bubble. The fundamental part is based on the basic ability of an asset to generate future cash flows, while the bubble part is only contributed by the expectation of present value capital gain. If the maturity period tends to be infinite, the present value capital gain of a non-bubbled asset has to be zero hence the bubble part equals to zero. However, for a bubbled asset, the present value capital gain would tend to be infinite and contribute an infinite future increasing of the price. These arguments can be found in many existing research, such as Tirole (1982), Blanchard and Watson (1982), Campbell and Shiller (1988). On the other hand, irrational bubbles are also claimed to exist by some researchers. For example, an irrational bubble may exist in a double auction market where the traded assets pay random dividend in different time (Smith, Suchankek, & Williams, 1998). Moreover, traders may buy an asset which has already been overpriced even if they are not sure whether they can resell it because of a consistent risk-loving preferences and judgemental errors (Lei, Noussair, & Plott, 2001). In additoin, irrational bubbles may happen in some other markets, such as lottery markets (Ackert, Charupat, Church, & Deaves, 2005), the markets with inexperienced traders as the majority (Dufwenberg, Lindqvist, & Moore, 2005), and the futures market with unlimited short-sell (Ackert, Charupat, Deaves, & Kluger, 2006). Although the rational bubble framework is argued to be failed in modeling the market with asymmetric information, it is still a common agreement that the rational bubble framework works well in modeling mature asset markets with a long-term view.

In this essay, I follow the rational bubble frame work to study the U.S. housing market. Under the rational bubble framework, the related studies can be summarized as four different topics. First of all, regarding to the bubble estimation, unit root test is used in detecting financial bubbles. Generally, non-stationarity or explosiveness is considered as the necessary condition of the existence of bubbles. For example, Diba and Grossman (1988) claim that if the discount rate is constant, testing the explosive behaviors in the data is equivalent to identify stock bubbles. West (1987) identifies bubbles in the U.S. stock market using a Hausman-specification test. Second, the estimation of bubbles can also be implemented under the rational bubble framework. Wu (1995; 1997) studies the rational bubbles in stock market and foreign exchange market by estimating them using the Kalman filter. Third, researchers use a series of unit root tests to timestamp the origin and collapse of bubbles. For example, Phillips, Wu and Yu (2011) test the 1990's Nasdaq bubbles using a ADF test based forward recursive regressions and identify its bubble period between 1995 and 2000. Although approaches based on unit root tests in detecting bubbles are claimed to be insufficiently effective to study the explosive non-linear process, especially in the period of bubble collapse (Evens, 1991), they are still wildly used in the studies on bubbles. Finally, some studies focus on the bubble migration from one market to another one. Caballero, Farhi, and Gourinchas (2008a) study global asset markets and link the bubble in commodities markets to the financial crisis in 2008. Phillips and Yu (2011) test the bubble migration from the real estate market to the oil market and to the bond market.

For the real estate market, above topics are also studied. Kim and Suh (1993) examine the real estate rational bubbles in Korea and Japan and conclude the existence of bubbles in between 1974 and 1989. Stevenson (2008) proposes an approach in modeling housing market fundamental based on population, income, housing stock index and interest rates (Similar studies can be seen in Muellbauer and Murphy (1997) and Levin and Wright (1997)). In addition, Phillips and Yu (2011) use Dickey-Fuller test to locate the start and ending times of bubbles in three different markets, including the real estate market, the commodity market, and the bond market during the subprime crisis. However, compare to the stock market and foreign exchange market, the studies on real estate bubbles are relatively more difficult. First, the lack of rent data makes it difficult to study both of the price behavior and the fundamental behavior in the real estate market. Second, different from stock dividend which is determined by the board of directors of a company, rent is determined by the supply and demand of the rental market. Therefore, the activity in the rental market also needs to be considered when studying the housing market.

This essay differs from previous research on real estate bubbles in three aspects. First, I apply the Phillips, Wu, and Yu (2011)'s approach to the housing price index, rent index, and price-to-rent ratios over the period from 1991 to 2012 for both the U.S. city-average level and city level data. Based on the empirical results, I estimate the timeline of the U.S. housing bubbles during the subprime crisis, and verify the hypothesis that the rental markets in different cities are independent with each other. Second, after discussing the situation of having a time varying discount rate in the housing market and comparing the results from the price adjusted by rent and the separated price and rent series, I conclude that the price-to-rent ratio would be used as an alternative series in locating the times of housing bubbles when the price and rent series are both found to be explosive. Third, based on the studies on the eight metropolitan areas, I date the timeline for the expanding of the U.S. housing bubbles among cities and test the bubble migration by econometric models.

The remained part of this essay is as follows. Section 3.2 introduces the econometric models I use to date the origin and conclusion of housing bubbles, and discusses how the explosive behavior analysis can be used in the time-stamping process. I also discuss the exuberance of rent and the approach for testing bubble migration. In Section 3.3, I describes the data used in the empirical analysis of this essay. After that, I report and analyze the test results in Section 3.4. Finally, I summarize the whole essay in Section 3.5.

3.2 Models

3.2.1 Locating the Timestamp of Exuberance

Calling the pricing model used in Chapter 2, the test for the housing market also begins with the standard no arbitrage condition:

$$P_t = \frac{1}{1+i} E(P_{t+1} + R_{t+1}), \qquad (3.1)$$

where P_t is the housing price at time t, R_t is the rent received by the householder from time $t-1$ to time t, and i is the required return of the house. To simplify the problem, I assume a constant i. Similar to stock market, here the rent can be considered as the dividend of the house, so the housing fundamental value would be assessed based on it.

According to Campbell and Shiller (1988), by taking log of the both sides of (3.1), its solution can be written as

$$p_t = p_t^f + b_t, \qquad (3.2)$$

where

$$p_t^f = \frac{k-\theta}{1-\rho} + (1-\rho) \sum_{i=0}^{\infty} \rho^i E(r_{t+1+i|t}) \qquad (3.3)$$

and

$$b_t = \lim_{i \to \infty} \rho^i E(p_{t+i|t}). \tag{3.4}$$

Thus,

$$E_t(b_{t+1}) = \frac{1}{\rho} b_t = (1 + exp(\overline{r-p}))b_t, \tag{3.5}$$

where $p_t = \log(P_t)$, $\mathrm{r_t} = \log(R_t)$, $\theta = \log(1+I)$ $\rho = \frac{1}{1+exp(\overline{r-p})}$, $\overline{r-p}$ is the average log rent-price ratio, $k = -\log(\rho) - (1-\rho)\log(\frac{1}{\rho} - 1)$, and $0 < \rho < 1$.

As shown in (3.3), p_t^f is determined by the expected rental payments. And if b_t presents the rational bubble component, (3.5) can be rewritten as:

$$\begin{aligned} b_t &= \frac{1}{\rho} b_{t-1} + \varepsilon_{b,t} \\ &= (1+g)b_{t-1} + \varepsilon_{b,t}, \end{aligned} \tag{3.6}$$

where $\mathrm{E_t}(\varepsilon_{b,t}) = 0$, $\mathrm{g} = \frac{1}{\rho} - 1 = exp(\overline{r-p})$.

Following (3.2), if b_t behaves as an explosive process, p_t will be explosive as well. Therefore, identifying the explosive behavior in p_t and nonexplosive behavior in r_t by right-tailed unit root tests is one possible approach to discover the existence of bubbles.

The tests are based on the augmented Dickey-Fuller (ADF) test. For each time series x_t, log housing index or log rent index or the ratio of price-to-rent, I test the following autoregressive specification:

$$\Delta x_t = \mu_x + \delta x_{t-1} + \sum_{j=1}^{J} \varphi_j \Delta x_{t-j} + \varepsilon_{x,t}, \tag{3.7}$$

where $\mathrm{E}(\varepsilon_t) = 0$, and J is the given lag parameter[3]. Then I compare the t-statistics of x_{t-1}

[3] J is decided in the following way. Starting with 12 lags in the model (J = 12), coefficients are sequentially tested for

with the related critical values to test for the hypotheses that x_t is:

$$\mathrm{H_o}\text{: Non-explosive process } (\delta = 0)$$

$$\mathrm{H_a}\text{: Explosive process } (\delta > 0)$$

Two approaches of tests are considered in this essay: forward recursive regressions and rolling regressions. On one hand, to implement the tests by forward recursive regressions, the model (3.7) is repeatedly processed by including one more sample to the tests every time. For each subset used in the tests, I assume it covers $\tau = \lfloor nr \rfloor$ observations, where n is the full sample size, and $0 < r < 1$. If denoting the t-value of the ADF coefficient for this subset by ADF_r, then ADF_{r_0} can be used to describe the ADF coefficient t-value for the first subset and ADF_1 can be used to describe the ADF coefficient t-value for the full sample. Finally, the emergence date of bubbles, $\hat{r}_e$, and the moment of collapse of bubbles, $\hat{r}_f$, are estimated as follows:

$$\hat{r}_e = \inf_{s \geq r_0}\{s: ADF_s > cv_{\beta_n}^{adf}(s)\}, \hat{r}_f = \inf_{s \geq \hat{r}_e}\{s: ADF_s < cv_{\beta_n}^{adf}(s)\}, \tag{3.8}$$

where $cv_{\beta_n}^{adf}(s)$ is the right-side critical value of ADF_s at a significance level of β_n.[4] On the other hand, rolling regressions can also be implemented as additional robustness check. In this approach, the selection of each subsample is based on a moving window. Assume

significance at the 5% level for possible models (J = 11, 10, 9... 1). I select the model if the coefficient of the last included lag is significant at the 5% level. (Notes: For the rolling regressions, the maximum value of J is set to 6.)

[4] The setting applied in this essay is $cv_{\beta_n}^{adf}(s) = \log(\log(ns))/100$, which leads to critical values around the 4% significance level.

the window size is *m*, then all the repeated tests contain *m* samples. The first subsample to be tested is $\{x_1, x_2, ..., x_m\}$, then the window moves to later period until $\{x_{n-m+1}, x_{n-m-2}, ..., x_n\}$. And the total number of tests need to be applied for each series is *n*-*m*, denoted by *u*. Let ADF^i denotes the t-value of the ADF coefficient for the *i-th* subsample, then the date of emergence of bubbles, $\hat{r}_e$, and the moment of collapse of bubbles, $\hat{r}_f$, can be estimated by:

$$\hat{r}_e = \inf_{i>1}\{i : ADF^i > cv_{\beta_n}^{adf}(i)\}, \hat{r}_e = \inf_{i>\hat{r}_e}\{i : ADF^i < cv_{\beta_n}^{adf}(i)\}, \tag{3.9}$$

where $cv_{\beta_n}^{adf}(i)$ is the right tailed critical value of the ADF coefficient at a significance level of β_n.

Based on the theory of rational bubbles, the housing price series has to be explosive during the bubble period. This explosive process is caused by the explosive bubble component b_t, while it can also be affected by the fundamental component p_t^f, if the fundamental series appears to be explosive. PWY (2011) use forward recursive regressions to test the exuberance for 1990s Nasdaq data. As concluded, the price series is found to be explosive in some periods while no significant evidence can be found to support the exuberance for the dividend series. Thus the origination and collapse of the 1990s Nasdaq bubble can be located by solely checking the explosive behavior for the log price series. In the real estate market, similar methods can be applied to test the exuberance for the housing price and rent data. However, compare to stock dividends, rent series has some special characteristics.

3.2.2 The Exuberance in the Rental Market

Rents can be considered as the dividends of holding a house. However, the most important distinction between rents and stock dividends is that rents are determined by the supply and demand in the rental market while stock dividends are decided by the shareholders of a company every year. To model the rental market, I consider following supply equation

$$Q_t^s = \alpha_0 + \alpha_1 R_t - \beta_1 E(R_{t+1|t}) + \gamma_1 X_t + \eta_t \tag{3.10}$$

and the demand equation

$$Q_t^d = \alpha_0^{'} - \alpha_1^{'} R_t + \gamma_2 Y_t + \eta_t^{'}, \tag{3.11}$$

where Q_t is the quantity of houses in the rental market, R_t is the rent at time t, η and η' are the error terms, X_t and Y_t are the sets of exogenous variables for the two equations, and all coefficients ($\alpha_0, \alpha_1, \beta_1, \gamma_1, \alpha_0^{'}, \alpha_1^{'}, \gamma_2$) are assumed to be positive. (3.10) shows that the supply quantity is an increasing function of current rent but a decreasing function of the expected rent of time t+1 based on the information of time t. Also, it is correlated to a set of exogenous variables X_t. For the supply quantity, I assume a standard supply function in the rental market. (3.11) says that the supply quantity is positively correlated to current rent and is affected by a set of exogenous variables Y_t.

Let $Q_t^s = Q_t^d$, the equilibrium function can be written as

$$\omega_0 + \omega_1 R_t - \beta_1 E(R_{t+1|t}) + \gamma \, \mathrm{Z}_t = u_t, \tag{3.12}$$

where $\omega_0 = \alpha_0 - \alpha_0^{'}$, $\omega_1 = \alpha_1 + \alpha_1^{'}$, $\gamma = [\gamma_1, \gamma_2]$, $Z_t = [\mathrm{X}_t, -\mathrm{Y}_t]$, and $u_t = \eta_t - \eta_t^{'}$. Based on the approach provided by Blanchard & Kahn (1980), one solution to (3.12) can be written as follows,

$$\omega_0 + \omega_1 R_t - \beta_1 [\frac{-\omega_0}{\omega_1 - \beta_1} - \frac{1}{\omega_1} \sum_{h=0}^{\infty} (\frac{\beta_1}{\omega_1})^h \gamma Z_{t+h+1|t}] + \gamma \mathrm{Z}_t = u_t \tag{3.13}$$

Rearranging the equation, yields,

$$R_t = \frac{1}{\omega_1} \{-\omega_0 + \beta_1 [\frac{-\omega_0}{\omega_1 - \beta_1} - \frac{1}{\omega_1} \sum_{h=0}^{\infty} (\frac{\beta_1}{\omega_1})^h \gamma Z_{t+h+1|t}] - \gamma \mathrm{Z}_t - u_t\}. \tag{3.14}$$

Based on this equation, it indicates that $\left|\frac{\beta_1}{\omega_1}\right|$ should be less or equal to one, otherwise forward expectations and rents would not converge, and the market would crash. If $\left|\frac{\beta_1}{\omega_1}\right| < 1$, then (3.12) can be rewritten as follows,

$$E(R_{t+1|t}) = c + \lambda R_t + v_t \tag{3.15}$$

where $c = \frac{\omega_0}{\beta_1}$, $v_t = \frac{1}{\beta_1}(\gamma \mathrm{Z}_t - u_t)$, and $\lambda = \frac{\omega_1}{\beta_1} > 1$. Assuming the information set Z_t can be described by a sequence of rent changes in the past, say, $\gamma \mathrm{Z}_t = \sum_{j=0}^{J} \tau_j \Delta R_{t+j}$, an explosive process of the expected rent can be obtained,

$$R_{t+1} = c + \lambda R_t + \sum_{j=0}^{J} \varphi_j \Delta R_{t+j} + \varepsilon_{R,t}, \tag{3.16}$$

where $\lambda > 1$, $\varphi_j = \frac{\tau_j}{\beta_1}$, J is the optimal lag length, and $\varepsilon_{R,t} = \frac{-u_t}{\beta_1} + \xi_t$ is the error term where ξ_t is an adjustment term with the expected value of zero. As we know that $E(u_t) = 0$, so the expected value of ε_t is also zero.

Recalling that $\omega_1 = \alpha_1 + \alpha_1^{'}$, the condition $\left|\frac{\beta_1}{\omega_1}\right| < 1$ can be written as $\beta_1 < \alpha_1 + \alpha_1^{'}$, which says that the sum of the influence of the current rent on the supply and demand is greater than the influence of the expected future rent on the supply. In such case, the rent series would follow an explosive process as shown in (3.16). If $\left|\frac{\beta_1}{\omega_1}\right| = 1$ or $\beta_1 = \alpha_1 + \alpha_1^{'}$, then the rent series would not be explosive. If assuming $J = 0$, it becomes to a random walk process with drift. In other words, explosive rent series would be found in some periods if the condition $\left|\frac{\beta_1}{\omega_1}\right| < 1$ satisfied in tested rental markets. I assume that the explosive behaviors can be found in different time periods for the rent data of different cities, because the rental market in each city is relatively independent. The rent related information updated in one city has limited effect on the market behaviors in another city. Therefore, (3.8) and (3.9) can be applied to detect the explosive behaviors in the rent series for testing following hypothesis.

Hypothesis 3-1: In the rental market, assuming people consider more about the present than the future at some points, rent series would be found to be explosive.

Hypothesis 3-2: The rental market of each city is independent with each other. Therefore, even if the explosive behavior can be easily found to exist in city-level rental markets, the country-level data does not need to be explosive.

3.2.3 The Sufficient Evidence of Bubbles: Explosive Price-to-Rent Ratio

As the above discussion, discovering explosiveness for both the price series and rent series is highly expected in the city-level data. Based on original settings of the housing pricing model, the explosiveness in price is the necessary condition for bubbles. However, this essay follows the discussion of Phillips and Yu (2011) and shows that this condition can be considered as sufficient evidence of housing bubbles under the assumption of a varying discount rate.

Assuming rents constantly grow at a rate of g_R, and $g_R < i$, the fundamental housing price can be written as

$$F_t = \frac{R_t}{i - g_R}. \tag{3.17}$$

And the equation can be rewritten as a continuous time version as follows,

$$F_t = \int_0^\infty \exp(-si_{t+s}) E_t R_{t+s} ds, \tag{3.18}$$

where $E_t D_{t+s} = \exp(g_R s) R_t$.

Rearranging (3.18), yields

$$F_t = \int_0^\infty \exp(-s(i_{t+s} - g_R)) R_t ds, \tag{3.19}$$

where i_{t+s} is assumed to follow the time profile below,

$$i_{t+s} = \begin{cases} g_R + \dfrac{t_b - t - s}{s} C_a + \dfrac{\chi_1}{s}, & \text{for } 0 \le s < t_b - t, \\ g_R + C_a + \dfrac{\chi_2}{s}, & \text{for } s \ge t_b - t, \end{cases} \tag{3.20}$$

where t_b is a fixed time point to separate the bubble and normal period, $c_a > 0$, and $\chi_1 > \chi_2 > 0$. Based on this assumption, the discount rate decreases to $g_R + \dfrac{\chi_1}{t_b - t}$ when time moves from t to t_b; then it jumps to $g_R + C_a + \dfrac{\chi_2}{t_b - t}$ and follows an new decreasing process. Here t_b is the break point in separating the bubble period and normal.

Then, (3.19) can be rewritten as follows,

$$\frac{F_t}{R_t} = \int_0^{t_b - t} \exp(-C_a(t_b - t - s) - \chi_1) ds + \int_{t_b}^{\infty} \exp(-c_a s - \chi_2) ds \tag{3.21}$$

Based on above settings, the time path of F_t / R_t can be proved to be explosive for $0 < t < t_b$, while it is not explosive for $t \ge t_b$. [5] In other words, an explosive price-to-rent ratio can be considered as sufficient evidence to support the existence of housing bubbles. Hence, studying the explosiveness in the series of price-to-rent is an alternative to the traditional method.

3.2.4 The Migration of Housing Bubbles

As mentioned in before, I assume that the housing markets in different cities are independent with each other. Then, after dating the timeline of housing bubble in each city, another interesting study is to test the bubble migration in different housing markets. If

[5] The proof is provided by Phillips and Yu (2011). See the Appendix.

migration can be identified, the results would show us how housing bubbles spread from one place to others, hence it would help in understanding the real estate market activities. Below I present the testing procedure I use for identifying the bubble migration moving from a market with series X_t to another market with series Y_t.

Let $\delta_X(\tau)$ denote the ADF coefficient for series X_t with the sample from $t=1$ to $t=\tau=\lfloor nr \rfloor$, which has been obtained by the accumulated recursive regressions as discussed at the beginning of this section. The basic idea of the testing procedure is to check whether the behavior of another series Y_t can be explained by the changes in the ADF coefficients of X_t during the explosive period. Following Phillips and Yu (2011), I extract the sample start from $t=\tau_P X$ to $t=\tau_p Y$, where $\tau_P X$ is the time when the estimated ADF coefficient of X_t reaches the peak after $t=\tau_e X$ when the explosiveness is first identified in X_t, and $\tau_p Y$ denotes the same for series Y_t. In addition, it is assumed that the effects raised by the changes in X_t become stronger as the time gets closer to present. Therefore, the bubble migration from one market to others can be tested by the model as follows,

$$\delta_Y(\tau) = \beta_{0n} + \beta_{1n}\delta_X(\tau)\frac{\tau - \tau_p X}{m} + error, \tag{3.22}$$

where $m = \tau_p Y - \tau_p X$ and $\tau = \lfloor nr_p X \rfloor + 1, \ldots, \lfloor nr_p Y \rfloor$. The sample contains both the period when explosive series X_t collapses and the period when the explosiveness can be identified in Y_t. The null hypothesis is ($\beta_{1n} = 0$) is no bubble migration from X_t to Y_t, while the alternative hypothesis ($\beta_{1n} < 0$) is bubble migration. The key points of this model can be summarized as follows. First, the peak of ADF coefficient for X_t has to be located earlier than the peak of ADF coefficient for Y_t. So that the sample would just cover the collapse period of X_t and the boom period of Y_t. Second, the sign of β_{1n} is expected to be negative to claim bubble migration because it is assumed that if there is a bubble migration from X_t to Y_t then the increasing of Y_t is caused by the decreasing of X_t. A positive sign would just indicate a positive correlation but not evidence for bubble migration. Third, the exuberance in X_t does not end too early (say, 12 months) before the exuberance in Y_t starts.

3.3 Data

The data includes housing price index and rental price index. For prices, I use the Case-Shiller S&P 500 Home Price Indices. For rents, I use the Rent of primary residence from Consumer Price Index (CPI) monthly reports published by U.S. Bureau of Labor Statistics (BLS). Also, I use the collected CPI to convert nominal series to real series. The sample contains two west coast cities (Los Angeles and San Francisco), three east coast cities (Boston, New York, and Miami), three Midwest cities (Detroit, Cleveland and Chicago), and the U.S. city-average. The full sample covers the period from January 1991 to January 2012 and contains 254 observations in total.

*** *Figure 3.1* ***

As shown in *Figure 3.1*, the time series trajectories of the U.S average real housing price indices, the real rent indices, and the price-rent ratios are plotted. By normalizing the real housing price index and rent index to 100 at the beginning of the sample, it can be seen that the price series was flat until the year 2000. Then it explosively grew to the peak in around June 2006 when the subprime crisis happened. After that, it suddenly started to drop until the end of the sample (January 2012). Regarding the rent, the series grew smoothly from the beginning to the end of the sample. Also, the movements of the price-rent ratio series are similar to the price series.

3.4 Testing and Results

3.4.1 Estimating the Timeline of Housing Bubbles

ADF tests are applied based on two approaches: forward recursive regressions and rolling regressions. *Table 3.1* reports the key statistics for the tests of the null hypothesis of non-stationary against the alternative of stationary in the U.S. city-average level. Two types of tests are implemented, one is the forward recursive regressions, and the other one is the rolling regressions (moving window). The tests are based on the data from January 1991 to January 2012 (N = 252 observations). The optimal lag length for the ADF test is determined by a top-down sequential significance testing with 12 as the maximum lag and 5% as the significant value. The model for the ADF tests contains a constant but no trend. For the forward recursive regressions, the first sample is from January 1991 to May 1992

(17 observations, $r_0 = 0.06$). For the rolling regressions, the moving window size is set to 48 observations, 4-year data. The first sample is from January 1991 to December 1994. As shown in the table, the ADF coefficient's t-statistics of the log price index and log rent index for both the first subsample and the full sample are negative and below the critical values, hence the null hypothesis in favor of non-explosive process cannot be rejected. However, the $\sup_{r\in[r_0,1]} ADF_r$ in the forward recursive regressions and the $\sup_{i\in[n-m+1,n]} ADF^i$ in the rolling regressions both provide significant evidence of exuberance in the price and rent data. This indicates two points: first, applying the ADF tests to the full sample do not provide any significant evidence to support the existence of explosive behaviors for all the tested series hence the existence of housing bubbles cannot be supported by this test alone; second, even if the existence of exuberance in the price and rent data are supported by $\sup_{r\in[r_0,1]} ADF_r$ and $\sup_{i\in[n-m+1,n]} ADF^i$, it is still possible to have no significant evidence to support the existence of housing bubbles because the rent data, which is used to describe the housing fundamental, appears to explosive behavior in both tests. Therefore, the verification of housing bubbles cannot be finished up to this point, and no information reveals the start and ending dates of bubbles.

*** *Figure 3.1* and *Table 3.2* ***

The city-level testing results are reported in *Table 3.2*. The same as the tests for the U.S. city-average data, both the forward recursive regressions and rolling regressions are implemented. The time period of data and all setting of tests are the same as the tests for the U.S. city-average data. According to the results, several conclusions can be made. First, for both tests using the forward recursive regressions and the rolling regressions, the maximum ADF coefficient's t-statistics of the price and rent data are greater than the related critical values, so both of the two tested series are explosive in one or more points during the full period. Second, it can be seen that both the price and rent data appear non-stationary for most of the tests of full sample except for Los Angeles and New York. On one hand, using forward recursive regressions, the log price ADF_1 of these two cities are -1.031 and -1.114 respectively, and the log rent ADF_1 are 1.095 and 0.817. This shows that the rent series, which describes the housing fundamental, is found to be explosive but the price series is not. It may indicate market crash in these two cities. On the other hand, similar results can be found using rolling regressions (The log price ADF^T is -4.795 for Los Angeles and -1.733 for New York; the log rent ADF^T is 2.535 for Los Angeles and 1.737 for New York). Finally, based on both of the two types of tests, Detroit is found to have an explosive price series but non-explosive rent series for the first subsample. However, since the number of observations of the first subsample for each type of test is different, ADF_0 and ADF^0 show the ADF t-statistics in different time period. For the data of Detroit, ADF_0 provides significant evidence of explosive log price series up to May 1992 while ADF^0 provides significant evidence of explosive log price series up to December 1994. Besides that, significant evidence is found to support explosive log price series in Boston, Miami, and Cleveland under the rolling regressions, while no significant evidence can be detected to support the explosive behaviors for the log rent data of these cities. Based on the results posted in this table, it is impossible to identify the dates of the origination and conclusion of the explosive behaviors of housing price and rent until more detailed testing results can be provided.

*** *Table 3.2* ***

To show more detailed information about the tests, *Figure 3.2* plots the ADF t-statistics of all three tested series, log price, log rent and the ratio of price-to-rent, for the U.S. city-average data using forward recursive regressions. The full sample of the data is from January 1991 to January 2012 (N = 252 observations). Every ADF t-statistic is obtained from the forward recursive regression with $r_0 = 0.06$ (the first ADF t-statistic is obtained based on the first 17 samples). The obtained ADF t-statistics are from May 1992 to January 2012. As can be seen, the rent series generally maintains negative ADF t-statistics in most of time, which indicates that the null hypothesis in favor of a non-explosive process cannot be rejected based on this test (except for some points in between May 1994 and September 1995). Also, in between December 1997 and January 2008, significant evidence can be found to support the explosive behavior of the price series. In such case, the result can be considered as significant evidence to support the existence of housing bubbles in the period from December 1997 to January 2008. In other words, the origination of the U.S. housing bubbles is located to December 1997 and the collapse is located to January 2008. However, it seems December 1997 is a little early for the bubbles to start and January 2008 is a litter late for the bubbles to conclude, based on the common knowledge. Meanwhile, looking at the ADF t-statistic for the price-rent ratio, the first occurrence date for exuberance is February 2002, and the evidence of the explosiveness is detected until June 2006. If rent series is found to be explosive, an explosive price series cannot be considered as sufficient evidence of a housing bubble. However, if a varying discount rate is assumed, applying the same test on the price-to-rent series is an alternative way to test housing bubbles. The plots also show that reasonable results can be found by solely testing the price-to-rent series.

*** *Figure 3.3* .***

As additional check, I also plot the ADF t-statistics of the same tested series for the U.S. city-average data based on rolling regressions in *Figure 3.3*. The data is also from January 1991 to January 2012 (N = 252 observations). Every ADF t-statistic is obtained from the rolling regressions with the first subsample start from January 1991 to December 1994 (48 observations, 4-year data). The obtained ADF t-statistics are from December 1994 to January 2012. Based on ADF t-statistic of the price-to-rent ratio, the first occurrence date for exuberance is August 2003, and the evidence of the explosiveness is detected until June 2006, which is very similar to the dates located by forward recursive regressions. However, some different findings can be drawn based on the testing results. First, the log rent series does not always appear non-explosive. For example, besides some points in between 1994 and 1995, explosive process can also be identified in the period from 2001 to 2003, from 2006 to 2008, and in 2009. Second, the log price series does not maintain explosive all the time between December 1997 and January 2008. The ADF t-statistics drop below related critical values for one year since July 2002, and finally the significant evidence of the exuberance disappear in February 2007. To summarize based on these two figures, if only considering the results during the subprime crisis, both the tests using forward recursive regressions and the test using rolling regressions result reasonable dates of origination and conclusion of housing bubbles for the U.S. city-average data. However, for the following city level tests, I would focus on the results from forward recursive regressions. The reason is that the window size needs to be determined before applying the rolling regressions. Too small window size would make the results sensitive to some abnormal observations; too large window size would reduce the number of subsamples. Although I also apply the test based on different sample sizes (36 observations and 70 observations) and get very similar results, forward recursive regressions do not have this problem and would result constant result.

*** *Figure 3.4* ***

3.4.2 Testing the Explosiveness for Rent Data

Figure 3.4 plots the ADF t-statistics for the log price, the log rent, and the ratio of price-to-rent of eight U.S. cities based on forward recursive regressions. First, it shows in the figure that the log rent series in all tested cities are not non-explosive all the time. During the subprime crisis period, most of the log rent series of the tested cities appear significant evidence to support the explosiveness, except for Cleveland. For example, in Los Angeles, significant evidence of exuberance of the log rent series can be found since December 1998; New York rent data is also tested to be explosive since January 2000, and the feature is maintained till the end of the sample; In addition, San Francisco, Miami, Chicago, Boston, and Detroit are all detected to be explosive for the rents during the subprime crisis. Second, the log price data of all tested cities are also found to be exuberance at some time periods. However, similar to the U.S. city-average data, the located dates of exuberance of cities are found too early for the origination and/or too late for the collapse. For example, the log housing price of New York City is detected to be explosive in between September 1997 and May 2008; in San Francisco, the evidence to support the explosiveness of its housing price is found to be maintained in the period from September 1997 to April 2008. Third, based on the price-to-rent data, more reasonable results are detected. Most of the cities are found to be explosive for the price-to-rent series in the period between 2000 and 2007 except Detroit and Cleveland. In Detroit, the evidence of exuberance is found in between 1995 and 2001. In Cleveland, no significant evidence can be found to support the exuberance until 2010.

To verify *Hypothesis 3-1*, I compare the results for the U.S. city-average plotted in *Figure 3.2* and *Figure 3.3* with the results for the city-level data plotted in *Figure 3.4*. On one hand, as can be seen, the exuberance for the U.S. average rent data are detected to exist but in very short periods. For example, based on the forward recursive regressions, the significant evidence to support the exuberance of the U.S. average rent can only be found at some points in between November 1994 and September 1995; based on the rolling regressions, the U.S. average rent is found to be explosive at some points in between December 1994 and July 1997, May 2001 and January 2003, and December 2006 and May 2008. On the other hand, for the city-level rent data, the results show that the explosive period for each city is different from each other. For example, the explosive behavior for the rent data is identified to start from January 1999 to the end of the testing sample in Los Angeles; in New York, the rent series is found to be explosive since May 2000; in San Francisco and Chicago, the period is from 1997 to 2004; in Miami, it is from September 2005 to September 2009; in Boston, it is from February 2006 to November 2009; in Boston, it is from February 1996 to November 2009; in Detroit, it is from May 1996 to December 2007; in Cleveland, only some points in between June 1997 and June 2002 are found to be explosive. To sum up, *Hypothesis 3-1* and *Hypothesis 3-2*, which state that explosive behaviors can be easily identified for city-level data but country-level data, are verified based on the U.S. rent data.

****** *Figure 3.2* and *Figure 3.3****

Based on above analysis, the dates of the origination and collapse of housing bubbles can be estimated by testing the unit root for the price-to-rent series. *Table 3.3* reports the estimated dates of the origination and collapse of housing bubbles in the tested U.S. cities. The data used in the estimations are the housing price-to-rent ratios for the period from January 1991 to January 2012. This estimations are obtained by ADF tests using forward recursive regressions with the first sample from January 1991 to May 1992 which includes 17 samples ($r_0 = 0.06$). On one hand, the tested cities with exuberance detected for their price-to-rent data can be separated into three groups by the stating time. First, the evidence of the earliest occurrence of housing bubble is detected to appear in Detroit in February 1995, and the collapse of this bubble is November 2001 as detected. Second, the next group of cities, where the evidence of housing bubbles are detected to start in between 1998 and 2000, contains all the tested cities in the east coast, including Boston, New York, and Miami, plus one mid-west city, Chicago. Third, the two tested west coast cities, Los Angeles and San Francisco as the third group of cities, are found to have explosive price-to-rent ratios start in 2003. Finally, Cleveland is the only city where the evidence of housing bubbles is identified after the subprime crisis, in 2009. On the other hand, two important findings on the collapse dates can be drawn in this table as well. First, even if the start points of housing bubbles are distinct in different cities, the collapse points of most tested cities are in 2006 or 2007, expect for Detroit and Cleveland. The evidence of the housing bubble in Detroit disappears after November 2001. And there is no evidence to show a collapse of the Cleveland housing bubble after January 2009. Second, it is noticeable that the lengths of the existence of exuberance are relatively longer in the east coast cities than in the west coast cities during the subprime crisis. The east coast cities are detected to be explosive for 92.6 months in average, while the west coast cities maintain the explosive behaviors for 36 months in average based on the tests during the subprime crisis. *Figure 3.5* plots the ADF t-statistics for the tested cities. It can be clearly seen the three groups of cities ordered by the existing time of exuberance.

*** *Table 3.3* ***

*** *Figure 3.5* and *Figure 3.6* ***

For robustness checking, I also apply the tests based on quarterly data for all series. *Figure 3.6* reports the ADF t-statistics for the U.S. city-average data. After converting the data frequency, the full sample covers periods from Q1 1991 to Q4 2011 (N = 84 observations). The left figure reports the ADF t-statistics obtained from forward recursive regression with $r_0 = 0.09$ (the first ADF t-statistic is obtained based on the first 8 samples). The ADF testing results are from Q4 1994 to Q4, 2011. The figure in the right side plots the ADF t-statistics obtained by rolling regressions with a time window of 16 observations (four-year data). The results are from Q4, 1994 to Q4, 2011. As can be seen, the estimated dates of start and ending points of housing bubbles are very similar to the estimation based on monthly data. If looking at the ADF t-statistic for the price-rent ratio get by forward recursive regressions, the first occurrence date for exuberance is Q3, 2002, and the evidence of the explosiveness is maintained until Q2, 2006. If using rolling regressions to estimate, explosive behavior is detected from Q1, 2002 to Q2, 2005. *Figure 3.7* plots the ADF t-statistics for city-level data.

*** *Figure 3.7* ***

3.4.3 Testing Bubble Migration

Based on the model for testing bubble migration and the estimated timeline of housing bubbles in the eight U.S. metropolitan areas, now I report the testing and results for bubble migration in the U.S. housing market. I assume that that the housing bubble started from Detroit; then, it spread to Chicago, Boston, New York, and Miami in between late 1998 and early 2000; finally it reached the west coast in around 2002. To check the assumptions, I test whether the housing bubble in Detroit spreads to other cities first. As shown in *Table 3.4*, I find significant evidence of bubble migration from Detroit to the east coast cities, Boston, Miami, and New York. As expected, negative signs indicate that the boom of the housing markets in the east coast cities can be explained by the collapse of the Detroit housing market. Then I test the how housing bubbles migrated to the west coast cities. The results are reported in *Table 3.5*. The significant findings are as follows. First, no significant evidence can be found to support the bubble migration from Miami and Chicago to Los Angeles and San Francisco. Positive signs show the positive relationship only. Second, there are strong evidence to support the migration from Boston and New York to the west coast. It can be seen that estimated coefficient β_{1n} for the two cities are found to be significantly negative in the 1% level. Therefore, the tests show significant evidence to support the trend of U.S. housing bubbles moving from Detroit to the east coast cities and finally to the west coast cities.

*** *Table 3.4* and *Table 3.5* ***

3.5 Summary

This essay studies the U.S. housing market with focus on locating the timeline of housing bubbles and testing the bubble migration. A forward recursive regression and a rolling regression based on augmented Dickey-Fuller (ADF) tests are applied to investigate the explosive behavior for the series of housing prices, rents, and price-to-rent ratios. I discuss the supply and demand equilibrium in the rental market, and analyze the market status when the rent is followed by an explosive process. The situations of considering a time invariant discount rate and a time varying discount rate in the housing market is also discussed. I apply the test procedure on price-to-rent series as an alternative when the housing price and rent are both found to be explosive. Moreover, I test the migration of housing bubbles in U.S. cities to verify the assumption which is made based on the timeline estimation of the bubbles.

The empirical findings are summarized as follows. First, the test confirms the existence of explosive rents in most of the tested U.S. cities, including Los Angeles, San Francisco, Chicago, Boson, New York, Miami, and Detroit, but no significant evidence to support the exuberance of rent for the U.S. city-average data. Second, the timeline of housing bubbles are estimated by using the price-to-rent data. The result shows that the exuberance or housing bubbles can be detected from February 2002 and January 2006 for the U.S. city-average data. Based on the results for the city-level data, I assume a trend of the U.S. housing bubble moving from the east coast to the west coast. Finally, I apply tests to verify this assumption on bubble migration. As expected, I find significant evidence to support the migration of housing bubbles proceeding from Detroit to the east coast cities, and then to the west coast.

Figure 3.1 The Series of Real Housing Price Index, Real Rent Index, and Real Price-Rent Ratios

The time series trajectories of the U.S average real housing price indices, the real rent indices, and the price-rent ratios are plotted. By normalizing the real housing price index and rent index to 100 at the beginning of the sample, it can be seen that the price series was flat until the year 2000. Then it explosively grew to the peak in around June 2006 when the subprime crisis happened. After that, it suddenly started to drop until the end of the sample (January 2012). Regarding the rent, the series grew smoothly from the beginning to the end of the sample. Also, the movements of the price-rent ratio series are similar to the price series.

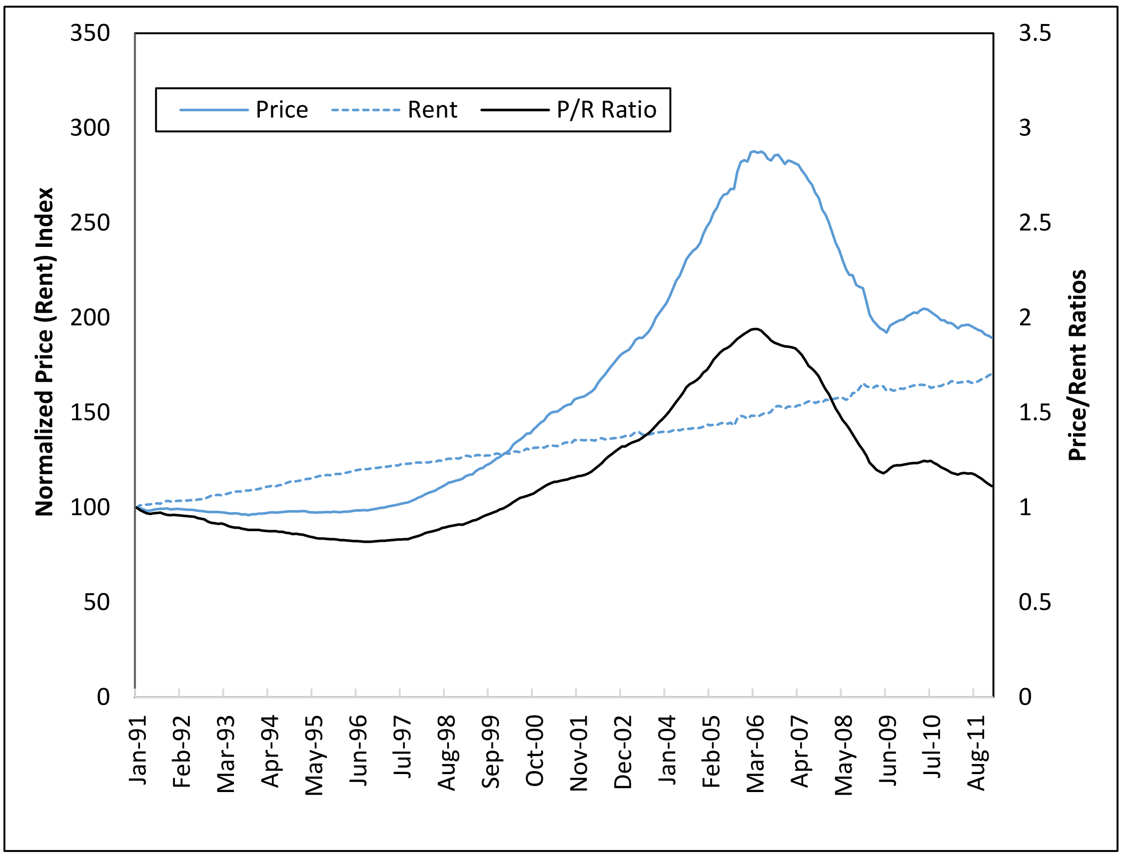


Figure 3.2 The Series of ADF t-statistic for the Logarithmic Real Housing Price Index, the Logarithmic Real Housing Rent Index, and Real Price-Rent Ratio (Forward Recursive Regressions)

This figure plots the ADF t-statistics of three tested series: log price, log rent and the ratio of price-to-rent. The data is from January 1991 to January 2012 (N = 252 observations). Every ADF t-statistic is obtained from the forward recursive regression with $r_0 = 0.06$ (the first ADF t-statistic is obtained based on the first 17 samples). The obtained ADF t-statistics are from May 1992 to January 2012. Based on ADF t-statistic for the price-rent ratio, the first occurrence date for exuberance is February 2002, and the evidence of the explosiveness is detected until June 2006.

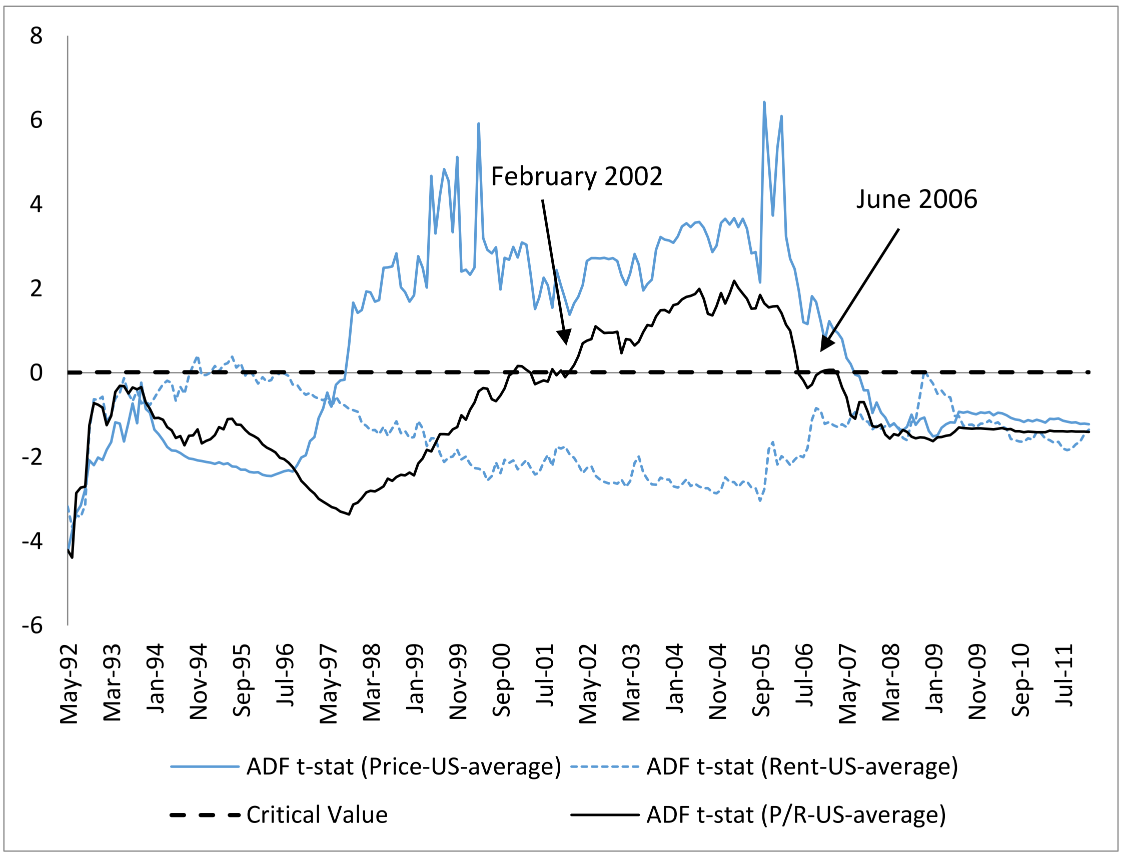


Figure 3.3 The Series of ADF t-statistic for the Logarithmic Real Housing Price Index, the Logarithmic Real Housing Rent Index, and Real Price-Rent Ratio (Moving Window)

The data is from January 1991 to January 2012 (N = 252 observations). Every ADF t-statistic is obtained from the rolling regression with the first sample start from January 1991 to December 1994 (48 observations, 4-year data). The obtained ADF t-statistics are from December 1994 to January 2012. Based on ADF t-statistic for the price-rent ratio, the first occurrence date for exuberance is August 2003, and the evidence of the explosiveness is detected until June 2006.

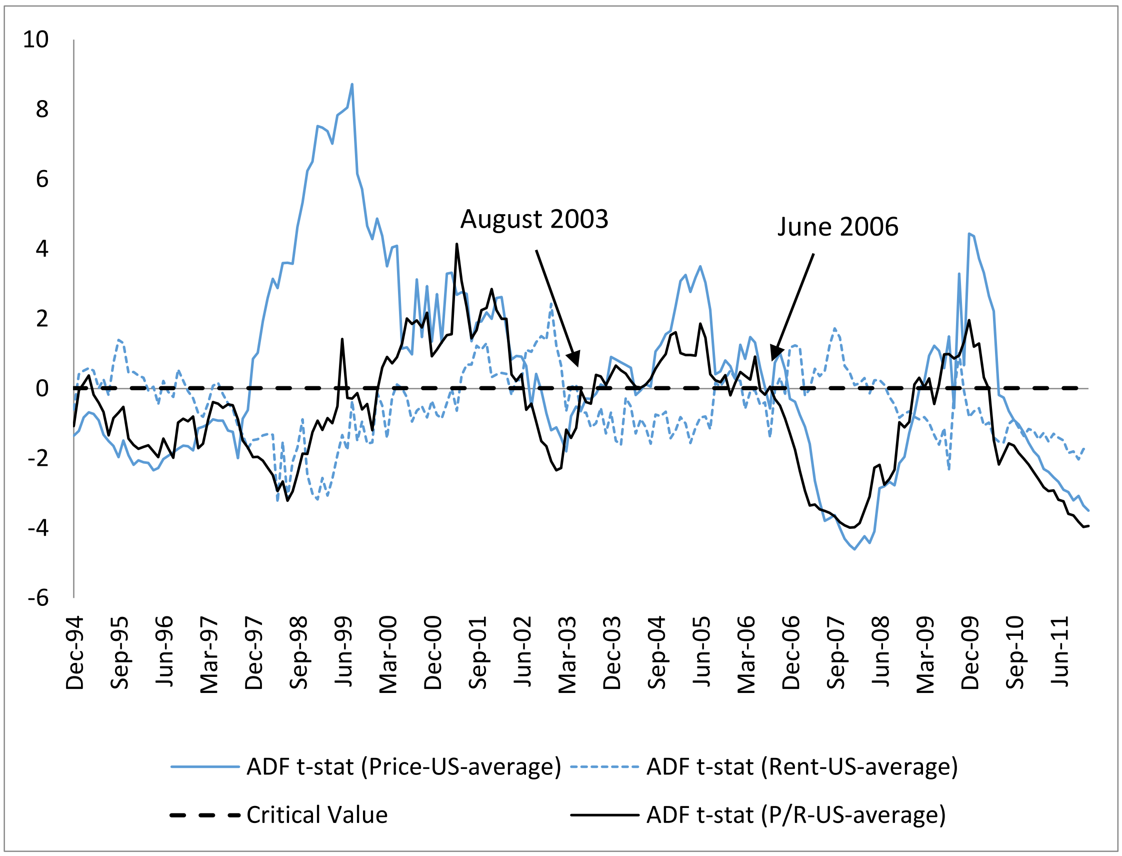


Figure 3.4 The Series of ADF t-statistic for the City-Level Data (Monthly Data)

The figures plot the ADF t-statistics of three series for each tested city: log price, log rent and the ratio of price-to-rent. The data is from January 1991 to January 2012 ($N = 252$ observations). Every ADF t-statistic is obtained from the forward recursive regression with $r_0 = 0.06$ (the first ADF t-statistic is obtained based on the first 17 samples).

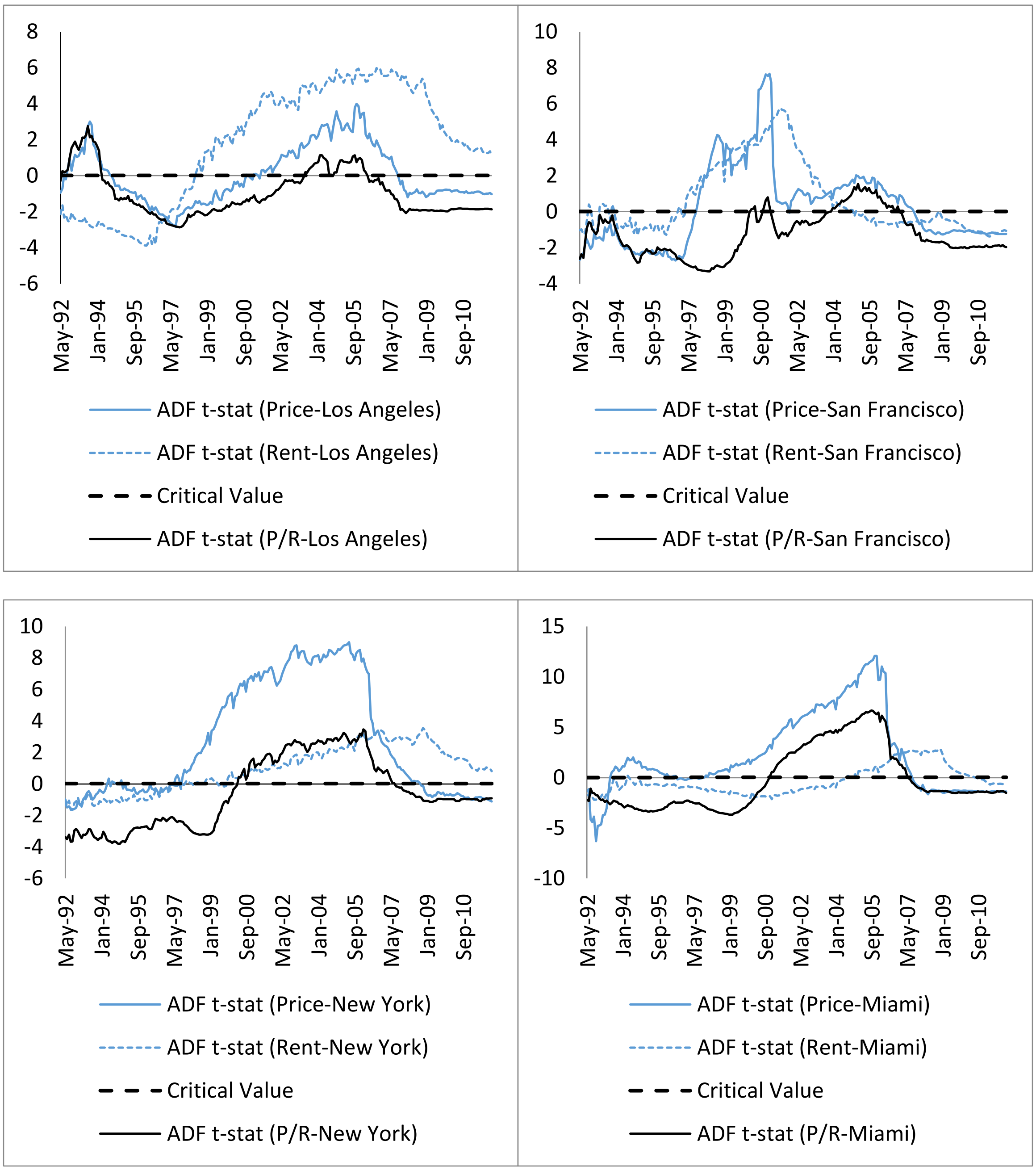


Figure 3.4 (Cont.) The Series of ADF t-statistic for the City-Level Data (Monthly Data)

The figures plot the ADF t-statistics of three series for each tested city: log price, log rent and the ratio of price-to-rent. The data is from January 1991 to January 2012 ($N = 252$ observations). Every ADF t-statistic is obtained from the forward recursive regression with $r_0 = 0.06$ (the first ADF t-statistic is obtained based on the first 17 samples).

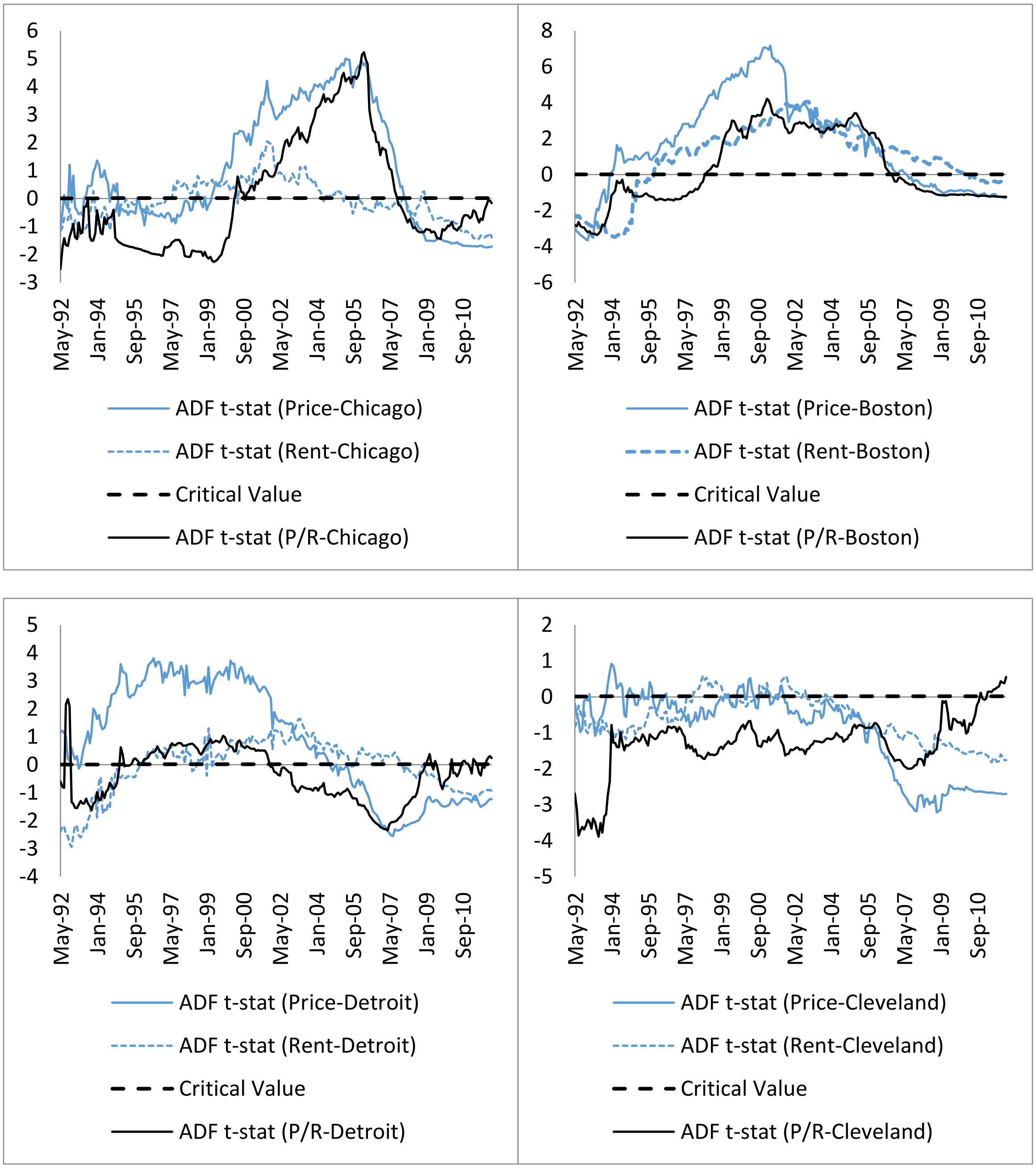
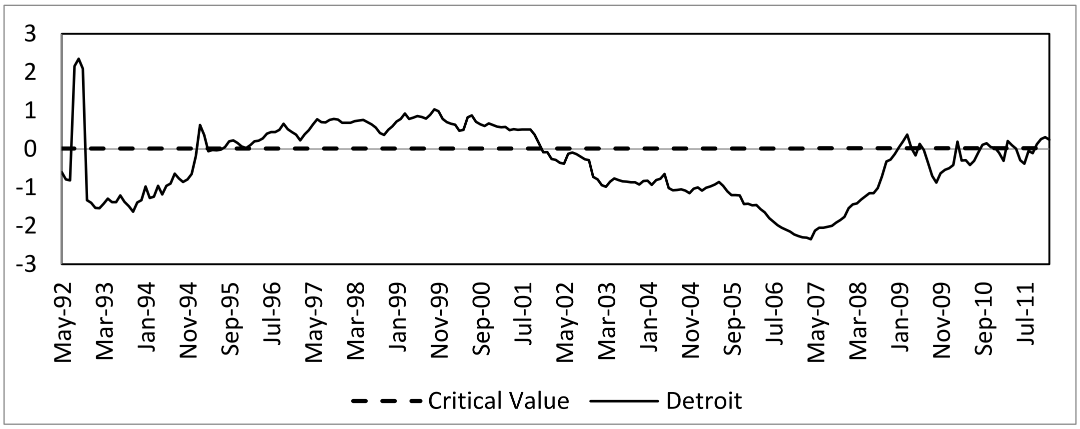
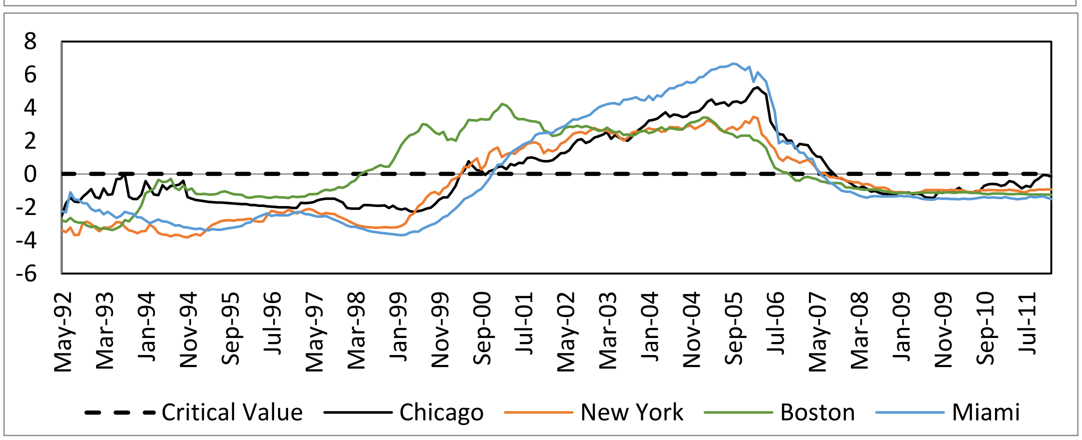


Figure 3.5 Timeline of Housing Bubbles in Difference Urban Areas

The ADF t-statistics obtained by forward recursive regressions based on the price-to-rent series in different urban areas are plotted. It can be seen that a housing bubble is found in Detroit in February 1995 as the first occurrence of housing bubbles in the U.S. After that, the second group of urban areas where housing bubbles are found to start in between July 1998 and November 2000 include Chicago, New York, Boston, and Miami. Finally, housing bubbles are detected to start in the end of 2003 in Los Angeles and San Francisco.





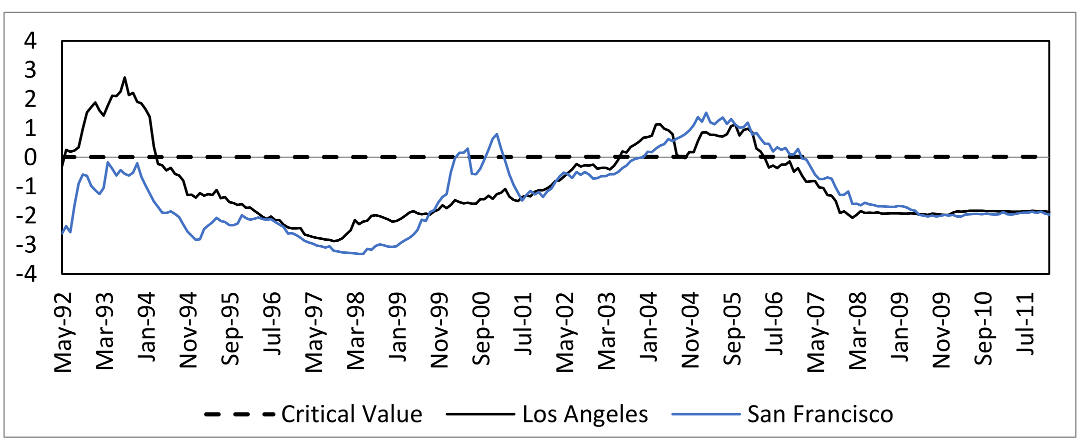


Figure 3.6 The Series of ADF t-statistic for the Logarithmic Real Housing Price Index, the Logarithmic Real Housing Rent Index, and Real Price-Rent Ratio (U.S. City-Average Quarterly Data)

This figure plots the ADF t-statistics of three tested series: log price, log rent and the ratio of price-to-rent. The full sample is from Q1 1991 to Q4 2011 ($N = 84$ observations). The left figure reports the ADF t-statistics obtained from forward recursive regression with $r_0 = 0.09$ (the first ADF t-statistic is obtained based on the first 8 samples). The obtained ADF t-statistics are from Q4 1994 to Q4, 2011. Based on ADF t-statistic for the price-rent ratio, the first occurrence date for exuberance is Q3, 2002, and the evidence of the explosiveness is maintained until Q2, 2006. The figure in the right side plot the ADF t-statistics obtained by rolling regressions with a time window of 16 observations (four year data). The reported results are from Q4, 1994 to Q4, 2011. Based on the ADF t-statistics of the price-to-rent series, explosive behavior is detected from Q1, 2002 to Q2, 2005.

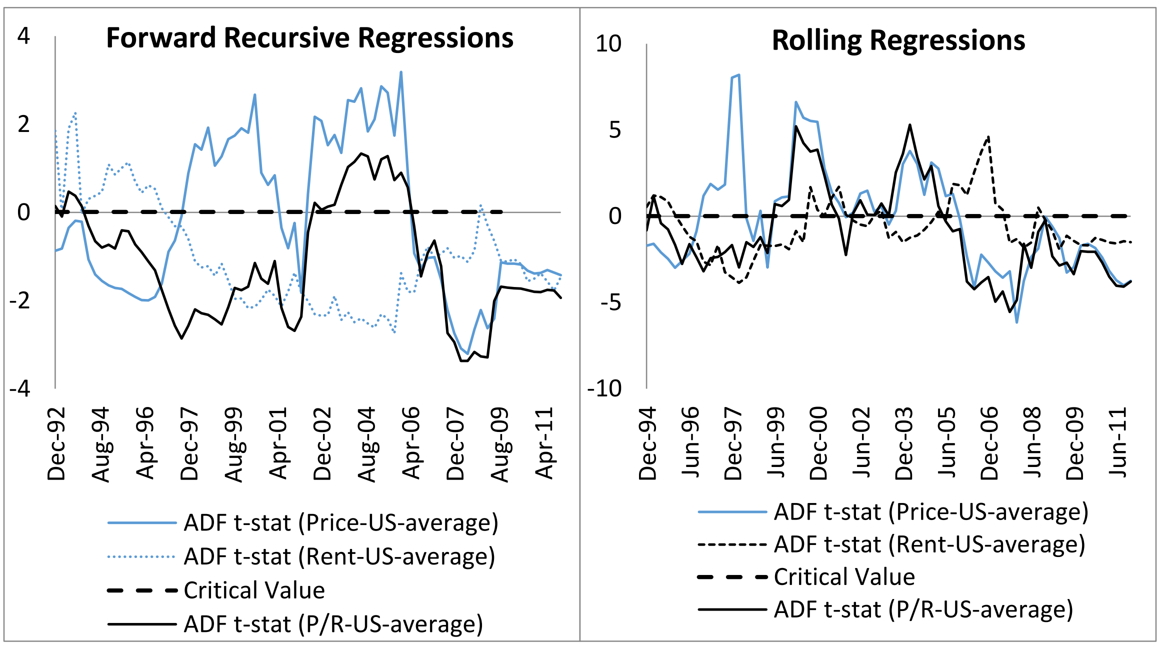


Figure 3.7 The Series of ADF t-statistic for the City-Level Data (Quarterly Data)

The figures plot the ADF t-statistics of three series for each tested city: log price, log rent and the ratio of price-to-rent. The data is from Q1 1991 to Q4 2011 ($N = 84$ observations). Every ADF t-statistic is obtained from the forward recursive regression with $r_0 = 0.09$ (the first ADF t-statistic is obtained based on the first 8 samples). .

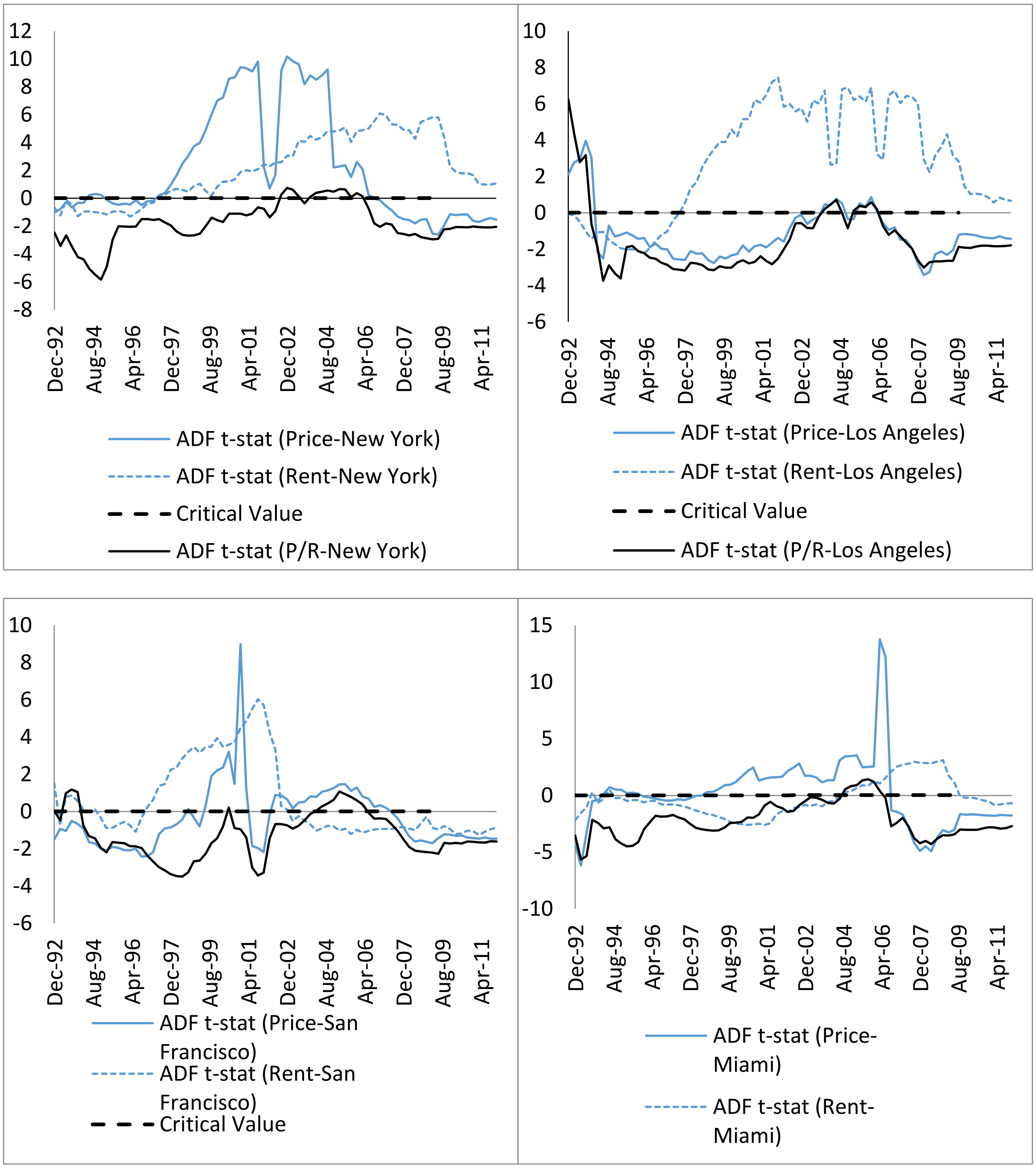


Figure 3.7 (Cont.) The Series of ADF t-statistic for the City-Level Data (Quarterly Data)

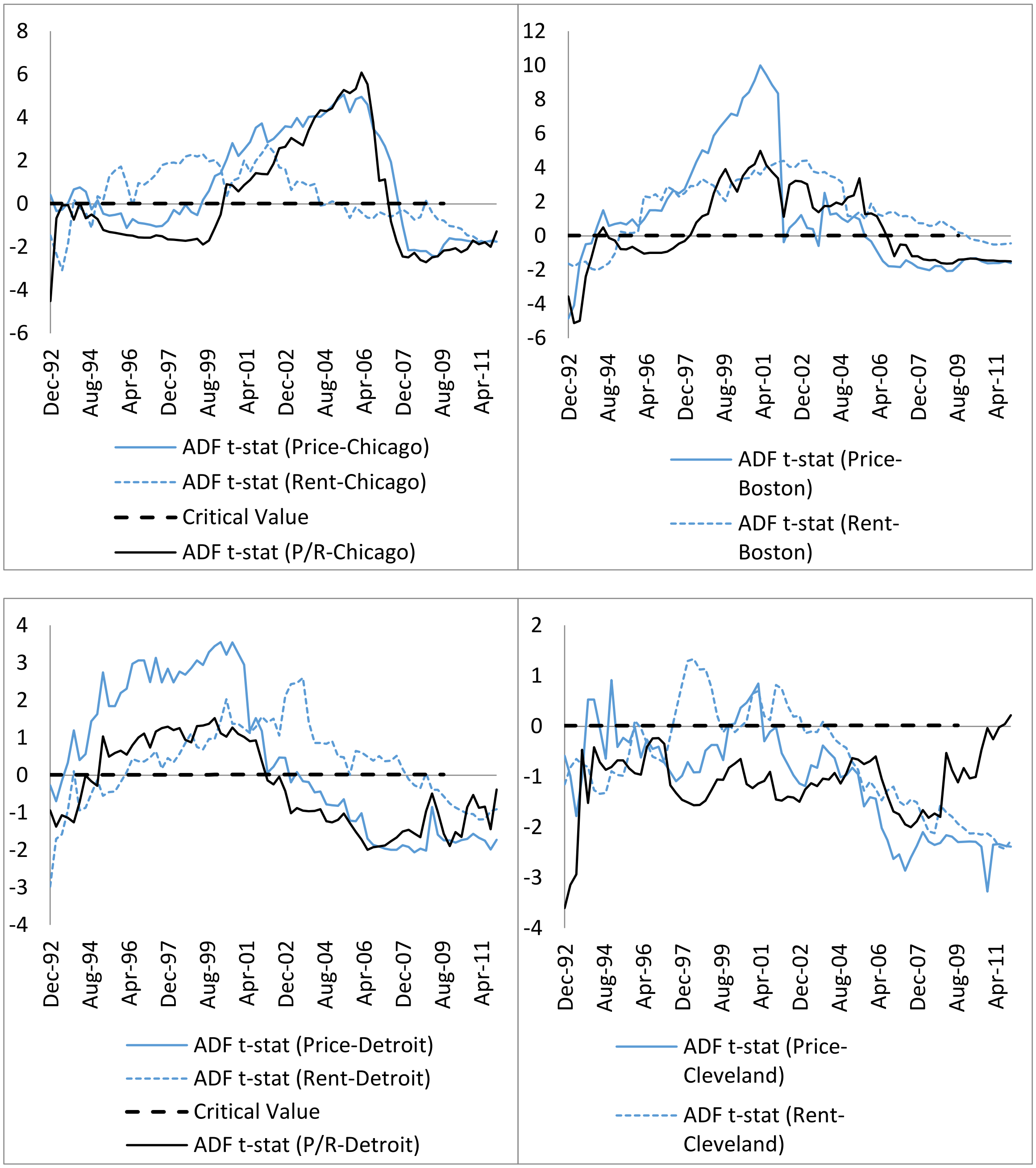


Table 3.1 The ADF t-statistic Summary for the Explosive Behavior Testing (U.S. City-Average)

This table reports the key statistics for the tests of the null hypothesis of non-stationary against the alternative of stationary for the U.S. city-average data. Two types of tests are implemented, one is the forward recursive regressions, and the other one is the rolling regressions (moving window). The tests are based on the data from January 1991 to January 2012 (254 observations). The optimal lag length for the ADF test is determined by a top-down sequential significance testing with 12 as the maximum lag (the maximum lag is set to be 6 in the rolling regressions) and 5% as the significant value. The model for the ADF tests contains a constant but no trend. For the forward recursive regression, the first sample is from January 1991 to May 1992 (17 observations, $r_0 = 0.06$). For the rolling regression, the moving window size is set to 48 observations, 4-year data. The first sample is from January 1991 to December 1994.

Forward Recursive Regression			
	ADF_{r_0}	ADF_1	$\sup_{r\in[r_0,1]} ADF_r$
Log price p_t	-4.257	-1.217	6.430
Log rent r_t	-3.173	-1.352	0.411
Rolling Regression (Moving Window)			
	ADF^{n-m+1}	ADF^n	$\sup_{i\in[n-m+1,n]} ADF^i$
Log price p_t	-2.120	-3.637	9.965
Log rent r_t	-0.054	-1.091	1.821
Critical Values for the Explosive Alternative			
1%	0.60		
4%	0.01		
5%	-0.08		
10%	-0.44		

Table 3.2 The ADF t-statistic Summary for the Explosive Behavior Testing (City Level)

This table reports the key statistics for the tests of the null hypothesis of non-stationary and the alternative of stationary for the city-level data. The same as the tests for the U.S. city-average data, both the forward recursive regressions and rolling regressions are implemented. The testing data is from January 1991 to January 2012. The optimal lag length for the ADF test is determined by a top-down sequential significance testing with 12 as the maximum lag (the maximum lag is set to be 6 in the rolling regressions) and 5% as the significant value. The model for the ADF tests contains a constant but no trend. For the forward recursive regressions, the first sample is from January 1991 to May 1992 (17 observations, $r_0 = 0.06$). For the rolling regressions, the moving window size is set to 48 observations, 4-year data. The first sample is from January 1991 to December 1994.

		Forward Recursive Regression			Rolling Regression (Moving Window)		
		ADF_0	ADF_1	$\sup_{r\in[r_0,1]} ADF_r$	ADF^0	ADF^T	$\sup_{t\in[0,T]} ADF^t$
Los Angeles	p_t	-0.953	-1.031	3.986	-0.555	-4.795	5.234
	r_t	-2.142	1.095	5.972	-3.080	2.535	2.768
San Francisco	p_t	-2.682	-1.231	7.669	-2.333	-4.039	4.062
	r_t	-1.021	-1.059	5.756	-1.309	-0.891	3.192
Miami	p_t	-1.387	-1.388	12.078	1.053	-7.954	6.241
	r_t	-1.714	-0.747	2.846	-0.812	-2.381	3.433
Chicago	p_t	-0.994	-1.715	5.012	-0.614	-1.434	3.890
	r_t	-1.152	-1.388	2.048	-0.599	-1.847	2.068
Boston	p_t	-2.957	-1.284	7.163	0.811	-2.233	4.633
	r_t	-2.257	-0.434	4.065	-1.039	-1.542	1.994
Detroit	p_t	1.171	-1.233	3.817	2.902	-3.109	4.504
	r_t	-2.404	-0.931	1.649	-0.447	-2.561	1.397
New York	p_t	-1.244	-1.114	9.002	-0.302	-1.733	5.055
	r_t	-0.986	0.817	3.552	-1.122	1.737	2.094
Clevelan d	p_t	-0.190	-2.700	0.918	0.076	-1.175	2.115
	r_t	-0.842	-1.761	0.587	-0.796	-1.022	1.456

Table 3.3 The Estimated Dates of Origination and Collapse of Housing Bubbles

This table reports the estimations of the origination and collapse of housing bubbles in eight U.S. cities. The data used in the estimations are the housing price-rent ratios from January 1991 to January 2012. This estimations are obtained by ADF tests using forward recursive regressions with the first sample from January 1991 to May 1992 which includes 17 samples ($r_0 = 0.06$). As shown in the table, the evidences of the earliest occurrence and collapse of housing bubbles are both detected in Detroit respectively in February 1995 and November 2001. The second group of cities, where the evidences of housing bubbles are detected to start in between 1998 and 2000, contains all the tested east coast cities, Boston, New York, and Miami, and one mid-west city, Chicago. Also, the third group of cities are found to have housing bubbles start in around 2003 includes the two tested west coast cities, Los Angeles and San Francisco. Furthermore, Cleveland is the only city where the evidence of housing bubbles are identified in 2009, after the subprime crisis. In addition, another important finding in this table is that even if the start points of housing bubbles are distinct in different cities, the collapse points of these cities are all in 2006 and 2007, expect for Detroit and Cleveland. It also shows that the lengths of the existence of housing bubbles are longer in the east coast cities than in the west coast cities.

Rank	**Cities**	**Origination**	**Rank**	**Cities**	**Collapse**	**Rank**	**Cities**	**Length (Months)**
1	Detroit	Feb-95	1	Detroit	Nov-01	1	Boston	100
2	Boston	Jun-98	2	Boston	Apr-06	2	Miami	92
3	Chicago	May-00	3	Chicago	Sep-06	3	Chicago	89
4	New York	Jun-00	4	New York	Dec-06	4	New York	86
5	Miami	Dec-00	5	Miami	May-07	5	Detroit	82
6	Los Angeles	Jul-03	6	Los Angeles	Jul-07	6	San Francisco	38
7	San Francisco	Nov-03	7	San Francisco	Sep-07	7	Los Angeles	34
8	Cleveland	Jan-09	8	Cleveland	N/A	8	Cleveland	N/A

Table 3.4 Testing Bubble Migration from Detroit to Other Cities in the U.S.

This table reports the estimated coefficients, β_{0n} and β_{1n}, and related t-statistics obtained by the model, $\delta_Y(\tau) = \beta_{0n} + \beta_{1n}\delta_X(\tau)\frac{\tau - \tau_p X}{m} + error$, for testing the bubble migration from X_t to Y_t ($H_0 : \beta_{1n} = 0$, no bubble migration; $H_a : \beta_{1n} < 0$, bubble migration). The sample is from the point when the ADF coefficient of X_t reaches the peak to the point when the ADF coefficient of Y_t reaches the peak. *m* is the sample size for each regression. i.e., 35 observations are used to test the bubble migration from Detroit to Boston. As reported, significant evidence is found to support bubble migration from Detroit to the east coast cities, such as Boston, New York and Miami at the 1% level, while there is no significant evidence of housing bubble migration from Detroit to Chicago.

	X = Detroit		
	β_{0n}	β_{1n}	R^2
Boston ($m = 35$)	-0.075*	-17.567*	0.833
	(13.025)	(12.625)	
New York ($m = 53$)	-0.069*	-11.136*	0.368
	(9.772)	(5.396)	
Chicago ($m = 72$)	-0.041*	-0.781	0.000
	(3.557)	(0.161)	
Miami ($m = 78$)	-0.011	-13.251*	0.165
	(1.529)	(3.846)	

Notes: The numbers in brackets are the absolute values of t-statistics; * indicates the number is statistically significant at the 1% level.

Table 3.5 Testing Bubble Migration from the East to the West in the U.S.

This table reports the estimated coefficients, β_{0n} and β_{1n}, and related t-statistics obtained by the model, $\delta_Y(\tau) = \beta_{0n} + \beta_{1n}\delta_X(\tau)\frac{\tau - \tau_p X}{m} + error$, for testing the bubble migration from X_t to Y_t ($H_0 : \beta_{1n} = 0$, no bubble migration; $H_a : \beta_{1n} < 0$, bubble migration). The sample is from the point when the ADF coefficient of X_t reaches the peak to the point when the ADF coefficient of Y_t reaches the peak. *m* is the sample size for each regression. As reported, significant evidence is found to support the bubble migration moving from New York and Boston to Los Angeles and San Francisco, while no evidence can be found to support the bubble migration from other cities to the west coast.

	X = New York			X = Boston			X = Miami			X = Chicago		
	(m=40)			(m=58)			(m=15)			(m=21)		
Los Angeles	β_{0n}	β_{1n}	R^2	β_{0n}	β_{1n}	R^2	β_{0n}	β_{1n}	R^2	β_{0n}	β_{1n}	R^2
	-0.006*	**-0.762***	0.821	-0.010*	**-0.872***	0.521	-0.002*	0.177*	0.966	-0.002*	0.188*	0.842
	(15.845)	(13.211)		(10.698)	(7.798)		(11.440)	19.331		(6.838)	(10.074)	
	(m=49)			(m=67)			(m=24)			(m=30)		
San Francisco	β_{0n}	β_{1n}	R^2	β_{0n}	β_{1n}	R^2	β_{0n}	β_{1n}	R^2	β_{0n}	β_{1n}	R^2
	0.007*	**-0.733***	0.283	-0.015*	**-1.687***	0.529	-0.003*	0.265*	0.977	-0.004*	0.364*	0.935
	(3.365)	(4.304)		(10.037)	(8.544)		(24.014)	(30.541)		(18.596)	(20.009)	

Notes: The numbers in brackets are the absolute values of t-statistics; * indicates the number is statistically significant at the 1% level.

CHAPTER 4

CONCLUSIONS

This dissertation investigates the characteristics of the U.S. housing market based on the studies of bubbles. While the first essay focuses on the estimation of housing bubbles and the fundamental analysis, the second essay considers more about the dating approach of bubbles and the migration test. Some interesting findings can be summarized as follows. First, I find significant evidence of housing bubbles for major U.S. cities in between 2004 and 2007, except Detroit and Cleveland. Second, new evidence is found to support that the population and income are the major factors to the housing fundamental. Third, housing bubbles are found to be highly associated with the expected present value capital gain, rental housing vacancy rates, and homeowner housing vacancy rates. Fourth, the estimating model is verified to have a good fit to the U.S. housing data based on the study. Fifth, I locate the timeline of the housing bubbles in the U.S. during the subprime crisis. While the U.S. city-average data show that the bubble started in February 2002 and ended in June 2006, the bubbles in cities are also detected. Sixth, the study confirms the hypothesis that the exuberance in rent data can be found in most of the cities in the sample, but not in the U.S. aggregate data. Finally, I find evidence to support a bubble migration moving from the east to the west during the subprime crisis. All the findings serve as additional evidence for the study of market activities during financial crisis, and a lot of related future work can be applied based on them.

I summarize the future work as follows. First of all, in addition to completing the city-level fundamental and bubble analysis, the trading behavior during the bubble bursting time of a housing bubble is also important to study. The research will lead to find the major factors which trigger the burst of bubbles. Second, after the verifying the housing bubble migration, it still needs to be linked to economic factors. So that the research will help in understanding the reasons of the bubble migration, hence the related financial market activity will be better studied. Finally, assuming a time varying discount rate will help in building improved models to fit the real-world data, estimate bubbles and explain market activities.

REFERENCES

Abreu, D., & Brunnermeier, M. K. (2002). Synchronization Risk and Delayed Arbitrage. *Journal of Financial Economics*, 66:341-360.

Abreu, D., & Brunnermeier, M. K. (2003). Bubbles and Crashes. *Econometrica*, 71(1):173-204.

Ackert, L. F., Charupat, N., Church, B. K., & Deaves, R. (2005). Bubbles in Experimental Asset markets: Irrational Exuberance No more. *Federal Reserve Bank of Atlanta Working Paper*.

Ackert, L. F., Charupat, N., Deaves, R., & Kluger, B. (2006). The Origins of Bubbles in Laboratory Asset Markets. *Federal Reserve Bank of Atlanta Working Paper*.

Allen, F., & Gorton, G. (1993). Churning Bubbles. *Review of Economic Studies*, 60:813-36.

Blanchard, O., & Kahn, C. (1980). The Solution of Linear Difference Models Under Rational Expectations. *Econometrica*, 48 (1980), 1305-1311.

Blanchard, O., & Watson, M. (1982). *Bubbles, Rational Expectations, and Financial Markets.* D.C.: Crisis in the Economic and Financial Structure, ed. by P. Wachtel, pp.295-316.

Brunnermeier, M. K. (2007). Bubbles. *The New Palgrave Dictionary of Economics. Oxford University Press*.

Brunnermeier, M. K., & Nagel, S. (2004). Hedge Funds and the Technology Bubble. *Journal of Finance*, 59(5):2013-2040.

Caballero, J. R., Farhi, E., & Gourinchas, P.-O. (2008a). Financial Crash, Commodity Price, and Global Imbalances. *Brookings Papers on Economic Activity*, Vol. 2008, 1-55.

Caballero, J. R., Farhi, E., & Gourinchas, P.-O. (2008b). An Equilibrium Model of "Global Imbalances" and Low Interest Rates. *American Economic Association*, Vol.98, No. 1, 358-393.

Cambell, S. D., Davis, M., Gallin, J., & Martin, R. F. (2009). What moves housnig markets: A variance decomposition of the rent-price ratio. *Journal of Urban Economics*, 66:90-102.

Campbell, J., & Shiller, R. J. (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies*, 5(2):243-280.

Case, K. E., & Shiller, R. J. (1988). The Behavior of Home Buyers in Boom and Post-Boom Markets. *New England Economic Review*, 80:29-46.

Case, K. E., & Shiller, R. J. (2003). Is There A Bubble in the Housing Market? *Brookings Papers on Economic Activity*, 2:299-362.

Diba, B. T., & Grossman, H. I. (1988). The Theory of Rational Bubbles in Stock Prices. *Economic Journal*, 98(392):746-754.

Downs, D. H., & Guner, Z. (2013). Commercial Real Estate, Information production and Market Activity. *Journal of Real Estate Finance and Economics*, 46:282-298.

Dufwenberg, M., Lindqvist, T., & Moore, E. (2005). Bubbles and Experience: An Experiment. *American Economic Review*, Vol. 95, 1731-1737.

Evens, G. (1991). Pitfalls in Testing for Explosive Bubbles in Asset Prices. *American Economic Review*, 81(4),992-930.

Flood, R. P., & Garber, P. M. (1980). Market Fundamentals versus Price-Level Bubbles: The First Tests. *Journal of Political Economy*, vol. 88, no. 4. pp. 745-770.

Gallin, J. (2004). The Long-Run Relationship between House Prices and Rents. *Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System*.

Garber, P. M. (2000). Famous First Bubbles: The Fundamentals of Early Manias. *MIT Press*.

Himmelberg, C., Mayer, C., & Todd, S. (2005). Assessing High House Prices: Bubbles, Fundamentals and Misperceptions. *Journal of Economic Perspectives*, Vol.19, Num 4.

Jorgenson, D. W. (1963). Capital Theory and Investment Behavior . *American Economic Review*, 53(2):247-259.

Kim, K.-H., & Suh, S.-H. (1993). Speculation and Price Bubbles in the Korean and Japanese real Estate Markets. *Journal of Real estate Finance and economics*, 6:73-87.

Kindleberger, C. P. (1992). Bubbles. *The New Palgrave Dictionary of Money and Finance, ed. by P. Newman et al.*

Kindleberger, C. P. (1996). *Manias, Panics, and Crashes: A History of Financial Crises, 3rd ed.* New York: John Wiley & Sons.

Lei, V., Noussair, C. N., & Plott, C. R. (2001). Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality. *Econometrica*, Vol. 69, No. 4, 831-859.

Levin, E. J., & Wright, R. E. (1997). Speculation in the Housing Market. *Urban Studies*, 34:1419-1437.

Muellbauer, J., & Murphy, A. (1997). Booms and Busts in the UK Housing Market. *The Economic Journal*, 107:1701-1727.

O'Hara, M. (2008). Bubbles: Some Perspectives (and Loose Talk) from History. *the Review of Financial Studies*, 21(1):11-17.

Phillips, P., & Yu, J. (2011). Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics 2*, pp.455-491.

Phillips, P., Wu, Y., & Yu, J. (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? . *International Economic Review*, 52:201-226.

Scheinkman, J., & Xiong, W. (2003). Overconfidence and Speculative Bubbles. Journal of Political Economy. *Journal of Political Economy*, 111:1183–219.

Shiller, R. J. (2000). Irrational Exuberance. *Princeton University Press, Princeton, NJ*.

Shleifer, A., & Vishny, R. W. (1997). The Limits of Arbitrage. *Journal of Finance*, 52(1):35-55.

Smith, V. L., Suchankek, G. L., & Williams, A. W. (1998). Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. *Econometrica*, 56, 1119-1151.

Stevenson, S. (2008). Modeling Housing Market Fundamentals: Empirical Evidence of Extreme Market Conditions. *Real Estate Economics*, 1:1-29.

Tirole, J. (1982). On the Possibility of Speculation under Rational Expectations. *Econometrica*, 1163-1182.

West, K. D. (1987). A Specification Test for Speculative Bubbles. *Quarterly Journal of Economics*, 102(3), 533-580.

Wu, Y. (1995). Are there rational bubbles in foreign exchange markets? Evidence from an alternative test. *Journal of International Money and Finance*, 14(1):27-46.

Wu, Y. (1997). Rational bubbles in the stock market: accounting for the U.S. stock-price volatility. *Economic Inquiry*, 35(2):309-319.

APPENDIX: THE EXPLOSIVENESS OF (3.21)

According to Phillips and Yu (2011), the discount rate is assumed to follow the process:

$$i_{t+s} = \begin{cases} g_R + \dfrac{t_b - t - s}{s} C_a + \dfrac{\chi_1}{s}, & \text{for } 0 \le s < t_b - t, \\ g_R + C_a + \dfrac{\chi_2}{s}, & \text{for } s \ge t_b - t. \end{cases} \tag{A.1}$$

The price-to-rent ratio can be written as follows,

$$\begin{aligned} \frac{F_t}{R_t} &= \int_0^{t_b - t} \exp(-C_a(t_b - t - s) - \chi_1) ds + \int_{t_b}^{\infty} \exp(-c_a s - \chi_2) ds \\ &= e^{-\chi_1} \left[\frac{e^{-C_a(t_b - t - s)}}{C_a} \right]_0^{t_b - t} + e^{-\chi_2} \left[\frac{e^{-C_a s}}{-C_a} \right]_{t_b - t}^{\infty} \\ &= \frac{e^{-\chi_1}}{C_a} \left[1 - e^{-C_a(t_b - t)} \right] + \frac{e^{-\chi_2}}{C_a} e^{-C_a(t_b - t)} \\ &= \frac{e^{-\chi_1}}{C_a} + \frac{(e^{-\chi_2} - e^{-\chi_1})}{C_a} e^{-C_a(t_b - t)} := \sigma_t \end{aligned} . \tag{A.2}$$

Multiply the both sides of (A.2) by C_a , yields

$$C_a \frac{F_t}{R_t} = e^{-\chi_1} + (e^{-\chi_2} - e^{-\chi_1}) e^{-C_a(t_b - t)}. \tag{A.3}$$

On one hand, if $t \in (0, t_b]$, the differential equation of F_t,

$$\begin{aligned} dF_t &= (e^{-\chi_2} - e^{-\chi_1}) e^{-C_a(t_b - t)} D_t dt + \sigma_t dD_t \\ &= \frac{(e^{-\chi_2} - e^{-\chi_1}) e^{-C_a(t_b - t)}}{e^{-\chi_1} + (e^{-\chi_2} - e^{-\chi_1}) e^{-C_a(t_b - t)}} C_a F_t d_t + \sigma_t dD_t \end{aligned} \tag{A.4}$$

Assume t is very closed to t_b, (A.4) can be rewritten as follows,

$$\begin{aligned} dF_t &\approx \frac{e^{-\chi_2}-e^{-\chi_1}}{e^{-\chi_1}+(e^{-\chi_2}-e^{-\chi_1})} C_a F_t d_t + \sigma_t dD_t \\ &= \left[1-e^{-(\chi_1-\chi_2)}\right] C_a F_t dt + \sigma_t dD_t. \end{aligned} \tag{A.5}$$

Since $\left[1-e^{-(\chi_1-\chi_2)}\right] C_a > 0$, $dF_t > 0$, hence F_t is proved to be explosive.

On the other hand, if $t \in (t_b, \infty]$, $r_{t+s} = g_R + C_a + \dfrac{\chi_2}{s}$,then

$$\begin{aligned} \frac{F_t}{R_t} &= \int_0^\infty \exp(-c_a s - \chi_2) ds \\ &= e^{-\chi_2} \left[\frac{e^{-C_a s}}{-C_a} \right]_0^\infty = \frac{e^{-\chi_2}}{C_a}. \end{aligned} \tag{A.6}$$

Therefore, $\dfrac{F_t}{R_t}$ is proved to be constant for $t \in (0, t_b]$, and it is non-explosive.

VITA

Keli Xiao

1985	Born April 1 in Pingxiang, Jiangxi Province, China
2002	Graduated from Nanchang No.2 High School, Nanchang, Jiangxi Province, China
2002-2006	Studied in Wuhan University of Technology, Wuhan, Hubei Province, China
2006	Bachelor’s Degree in Computer Science and Technology, Wuhan University of Technology
2006-2007	Studied in the Department of Computer Science, City University of New York – Queens College, Flushing, New York, U.S.A.
2007-2013	Studied in Rutgers Business School-Newark and New Brunswick, Rutgers, The State University of New Jersey, U.S.A.
2008	Master’s Degree in Computer Science, City University of New York – Queens College
2009	Master’s Degree in Quantitative Finance, Rutgers Business School – Newark and New Brunswick
2011-2013	Instructor at Rutgers Business School – Newark and New Brunswick
2013	Ph.D. in Finance, Rutgers Business School – Newark and New Brunswick