UNDERSTANDING THE 2007-08 FINANCIAL CRISIS, THE STORY OF THE HOUSING BUBBLE AND BUST: AN EMPIRICAL ANALYSIS

by

Shweta Khosla

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Economics

Charlotte

2013

Approved by:

____
Dr. Rob Roy McGregor

____
Professor Azhar Iqbal

____
Dr. Benjamin Russo

ABSTRACT

SHWETA KHOSLA. Understanding the 2007-08 Financial Crisis, the story of the housing bubble and bust: an empirical analysis (Under direction of DR. ROB ROY MCGREGOR)

The 2007-08 financial crisis has been claimed to be the most severe one in modern history and possibly the worst one in the post-World War II period. Several factors have been criticized that led to this financial debacle including easy monetary policy pursued by the Greenspan Federal Reserve post 2000, global financial imbalances, Fannie, Freddie and subprime lenders inflating the housing bubble, securitized debt obligations receiving AAA ratings and the utter recklessness of some of the top banking houses. What followed were forceful responses by central banks around the world in a frantic attempt to avoid the utter collapse of the global financial system and save the world from entering into a deep global recession.

In light of the severity of this crisis and being regarded as the most virulent global financial crisis ever, this paper seeks to understand what were the underlying causes of the financial crisis and the many lessons that need to be learned to prevent such a crisis from happening again.

ACKNOWLEGEMENTS

I would like to express my appreciation to Professor Azhar Iqbal for the debt of gratitude I owe him for the help and support provided in the development of this thesis. His active role and guidance was present from the development of the literature, to the econometric methods and computer code, to the interpretation of results, among many other things and without him this thesis would not be possible in its present form. He would always make time from his busy schedule to discuss my analysis / findings and in turn would provide with key insights given his high understanding of all aspects of the financial crisis.

A special thanks goes to Dr. Rob Roy McGregor and Dr. Benjamin Russo for their role in this thesis and their teaching which proved very valuable in the understanding of financial and economic phenomena studied in this thesis. In addition, I would like to express my gratitude to the Department of Economics and Dr. Carol Stivender for the opportunity and experience of being a graduate assistant in both research and teaching. Finally, I would like to express my gratitude to my family and friends who supported and encouraged me during my work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

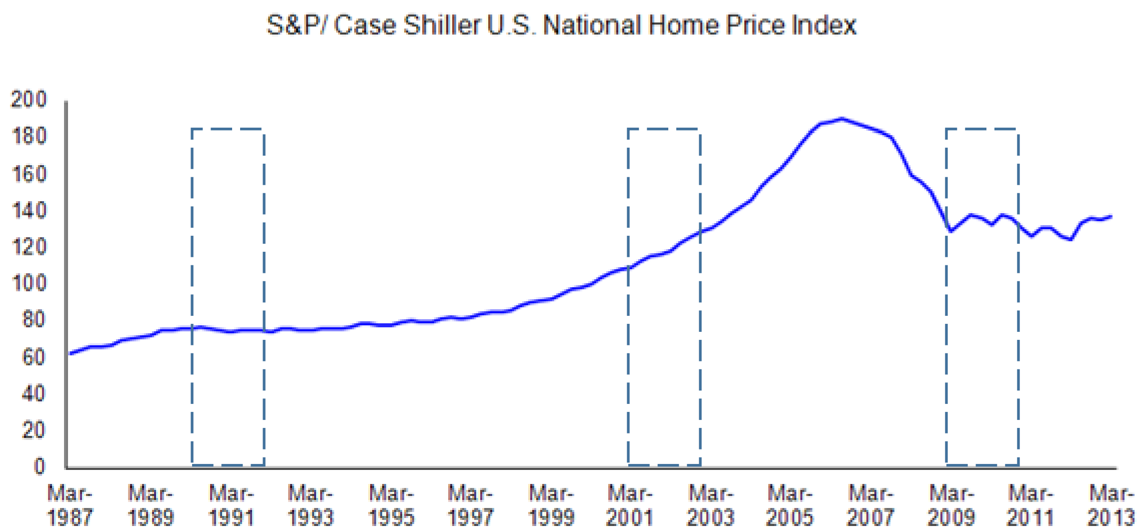
LTV	Loan-to-value ratio
GSE	Government sponsored enterprise
FY	Fiscal year
CDO	Collateralized debt obligation
HUD	U.S. Department of Housing and Urban Development
FHFA	Federal Housing Finance Agency
GDP	Gross domestic product
OCC	Office of the Comptroller of the Currency
MBS	Mortgage backed securities
E-G	Engle and Granger
VAR	Vector autoregressive
ECM	Error Correction Model

CHAPTER 1: INTRODUCTION

September 2008 marked a complete turning point in global history. On September 7, Freddie Mac and Fannie Mae, two large government sponsored agencies, were seized by the U.S. government and nationalized so as to make them financially stable. On September 15, the 158 year old investment bank Lehman Brothers was turned down support from the Federal Reserve and thus filed for bankruptcy. The Dow fell 504 points that day, the largest drop in over seven years. Later that day, Merrill Lynch was bought by Bank of America and thus saved from bankruptcy. On September 16, one of the world's largest insurance companies AIG met with federal regulators to raise desperately needed cash and received $85 billion from the Federal Reserve to bail them out of financial trouble. By Wednesday that week, stock markets around the globe saw huge swings in value. Over the following two-week period banks continued to fail. By September most banks stopped lending to each other at anything but extremely high rates, credit and financial markets froze. On October 3, the Bush administration stepped in offering $700 billion in financial assistance.[1] The financial crisis continued to spread worldwide. In Germany the government guaranteed all consumer bank deposits and bailed out Hypo Real Estate (Germany's second largest commercial property lender). In the United Kingdom the government nationalized Bradford & Bingley (UK's eighth largest mortgage lender). Iceland bailed out its third largest bank and Ireland guaranteed

[1] Exact dates from Federal Reserve Bank of St. Louis website.

deposits and debt of its six major financial institutions. The governments of the Netherlands, Belgium and Luxembourg injected $16.37 billion into Fortis NV. India's central bank promised to pump in cash in ICICI bank (India's second largest bank) to prevent a bank run. From here on started the global recession which ended up with near global economic collapse.



Dashed lines indicate periods of US recession

FIGURE 1: S&P / Case Shiller U.S. National Home Price Index
Source: S&P Dow Jones Indices, author's illustration

The biggest factor leading up to the financial crisis was cited as the collapse of the housing sector. Several economists concede that the 1998-2006 housing boom led to the global financial crisis and the 2007-09 global recession. However, what is a matter of considerable debate are the causes of the housing boom. One group of economists suggests that easy monetary policy during the 2002-04 period fueled housing speculation, resulting in an unprecedented bubble. Another group of economists rejects this claim. Both groups have justifications and arguments in order to support their views. The purpose of this paper is to provide a high level understanding of factors that resulted in the global financial crisis and empirically quantify which were the primary drivers and which merely played a supporting role. Specifically, this study seeks to answer three fundamental questions:

1. How did the housing bubble originate and under what circumstances did the bubble begin to grow? What role did monetary policy play in creating the housing bubble?

2. What factors other than the housing bubble caused the global financial crisis? Was monetary policy combined with additional factors of regulations, financial innovations, GSEs, structural change, and international capital flows popularly known as the global saving glut hypothesis responsible for the bubble?

3. What are the key takeaways for reform of the global financial system?

The first part of the study theoretically explains and seeks to address the above questions. In the second part of the study we will econometrically quantify the role of each of these factors in the housing bubble and hence their overall impact on the financial crisis. Specifically, we test the relationship between the housing bubble and monetary policy, the housing bubble and the global savings glut, the housing bubble and financial innovations, and the possibility of a structural break (as the housing bubble impacted some of the US states more than others). Next we examine whether a combination of all these factors contributed to the housing bubble or not. In conclusion, even as we continue working to stabilize our global financial system and revitalize the global economy, it is essential to learn the lessons from the crisis so that moving forward we can appropriately hedge against the risk of it happening again.

CHAPTER 2: LITERATURE REVIEW

2.1 The Bubble Trouble

Before we start analyzing the factors behind the housing bubble let us first clarify what the term “bubble” really means. Bubble implies a situation in which excessive public expectations of future price increases cause prices to be temporarily elevated. Say in a housing price bubble, homebuyers start thinking that a home that ordinarily was too expensive for them is now an acceptable purchase as they will be compensated by significant future price increases. So homeowners now think that they will not need to save as much as they otherwise would have, because they expect the increased value of their home to do the saving for them. First-time homebuyers also start worrying midst a housing bubble that if they do not buy now, they may not be able to afford a home later. Moreover, these expectations of home price appreciations can have a strong impact on demand if people think that home prices are highly unlikely to fall, and definitely not likely to fall for long, so this amounts to barely any perceived risk associated with a home investment. Prices cannot keep going up forever, and when individuals perceive that prices have stopped going up, this support for their acceptance of high home prices could break down. The bubble now bursts as prices start falling as a result of reduced demand. Several top economists including Cynthia Angell and Norman Williams; Ben Bernanke; had reassuringly, downplayed concerns about possible home price falls. Angell and Williams (2005), for example, concluded that “In over 80 percent of the metro-area price booms we examined between 1978 and 1998 the boom ended in a period of stagnation… the expectation would be that metro-area home price busts will continue to be relatively rare. (p. 5)."

On the other hand economists like Robert Shiller warned over many years that house prices might fall. Shiller, a Yale economist, is co-creator of the widely followed Standard & Poor's/Case-Shiller home price index. He has been widely ranked among the most influential economists in the world. Shiller (2008, p. 30) considered ratios of home prices to rent, personal income and building costs, and found that home prices in 2004 were looking very anomalous at that time, like a rocket taking off. As per Shiller real home prices for the United States as a whole increased 85% between 1997 and the peak in 2006. He wrote that it looked like the rocket might come crashing back down to earth. Shiller believed that the boom was driven by an epidemic of irrational public enthusiasm for housing investments. He precisely describes this as (p. 35), "the social contagion of boom thinking, mediated by the common observation of rapidly rising prices' which feeds into a narrative of a new era." As evidence of these raised expectations of home prices, he cites his work on median expected price increases in several regions in 2005 which he found to be truly extravagant.

2.2 Monetary Policy and the Housing Bubble

A combination of factors, including a moderate recession in 2001 due to the dot-com bubble bust and the resulting sharp decline in stock prices, geopolitical uncertainty due to the terrorist attacks of September 2001, several corporate scandals in 2002, and the Iraq invasion in March 2003 clouded the US economy in the early part of the decade. The Federal Reserve System under Chairman Alan Greenspan began aggressively expanding the U.S. money supply (M2). The target federal funds (interbank overnight) rate was lowered quickly in response to the 2001 recession, from 6.5% in late 2000 to 1.75% in December 2001 and later to 1% in June 2003. After reaching the then record low of 1%, the target rate remained at that level for a year. Finally in June 2004, the Federal Open Market Committee (FOMC) began to raise the target rate, reaching 5.25% in June 2006. After some years of slow growth, U.S. house prices began to rise more rapidly in the late 1990s. Prices grew at a 7 to 8% annual rate in 1998 and 1999, and in the 9 to 11% range from 2000 to 2003. Thus, the beginning of the run-up in housing prices predates the period of highly accommodative monetary policy. Shiller (2007) dates the beginning of the boom in 1998. On the other hand, the most rapid price gains were in 2004 and 2005, when the annual rate of house price appreciation was between 15 and 17%. Thus, the timing of the housing bubble does not rule out some contribution from monetary policy. Several economists claim that the loose monetary policy was appropriate for the macroeconomic conditions that prevailed then, and that it was neither a principal cause of the housing bubble nor the right tool for controlling the increase in house prices. The Fed's Chairman Bernanke (2010) defended the Federal Reserve's policy and suggested the lower fed funds rate during the 2002-04 period was a necessary step due to several factors: weak recovery/ jobless recovery from the 2001 recession, the uncertainty that followed the September 11 terrorist attack and the fear of deflation as had occurred in Japan in the 1990's. He also pointed out that housing did not boom simply in the United States. Most other industrialized countries also experienced housing booms during this period.

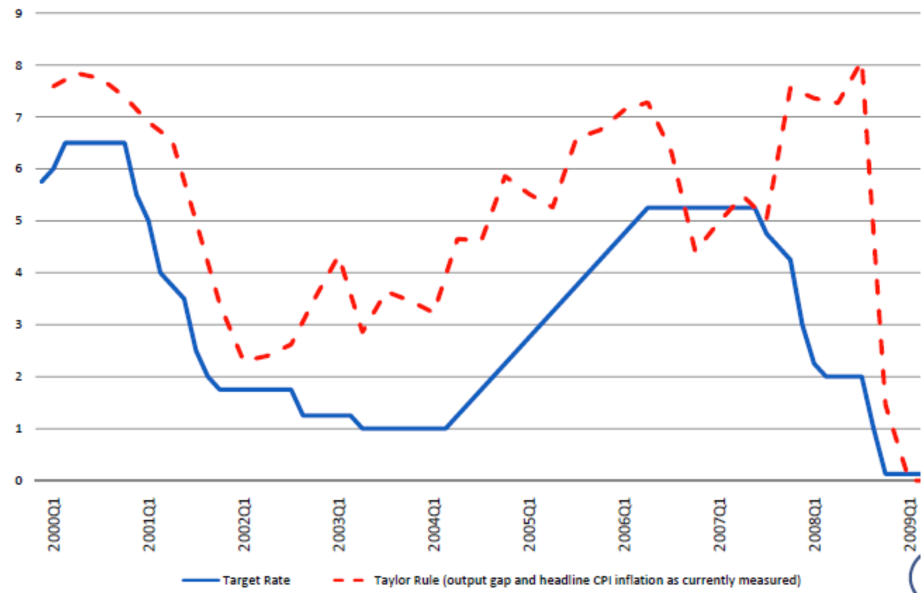


FIGURE 2: The target federal funds rate and the Taylor (1993) rule
Source: Monetary policy and the housing bubble, Bernanke (2010)

In contrast other economists have assigned monetary policy a central role in the crisis that caused a bubble in house prices in the United States, a bubble whose inevitable bust resulted in the crisis. Taylor (2008; 2010) suggested that the excessively easy monetary policy was the major cause of the housing bubble. As per Taylor, the 2007-08 financial crisis was a result of the housing boom and bust which in turn was caused by monetary excesses. Taylor presented evidence that the actual interest rate decisions of the Federal Reserve from 2004-07 fell well below what they would have been if the Fed had continued to follow the policy that it had in the previous 20 years of good economic performance.

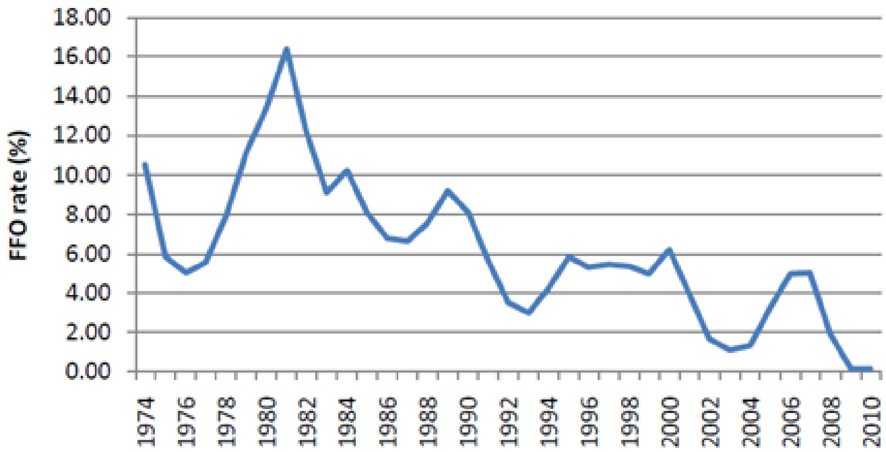


FIGURE 3: Federal funds rate since 1974
Source: Federal Reserve Bank of St. Louis' website, author's illustration

Taylor in his paper on Housing and Monetary Policy (Sep 2007) focusses on the role of monetary policy in creating the 2007 turmoil in the housing market and how it can be used to resolve similar crises in future:

> During the period from 2003 to 2006 the federal funds rate was well below what experience during the previous two decades of good economic macroeconomic performance — the Great Moderation — would have predicted. There have been other periods during the Great Moderation where the federal funds rate veered off the typical policy rule responses—in particular during the fall of 1998—but this was the biggest deviation, comparable to the turbulent 1970s. With low borrowing rates housing demand increased tremendously which led to a surge in housing price inflation. With housing prices rising rapidly, delinquency and foreclosure rates on sub-prime mortgages also fell, which led to more favorable credit ratings than could ultimately be sustained. As the short term interest rate returned to normal levels, housing demand rapidly fell bringing down both construction and housing price inflation. Delinquency and foreclosure rates then rose sharply, ultimately leading to the meltdown in the subprime market and on all securities that were derivative from the subprimes. (p. 2-3)

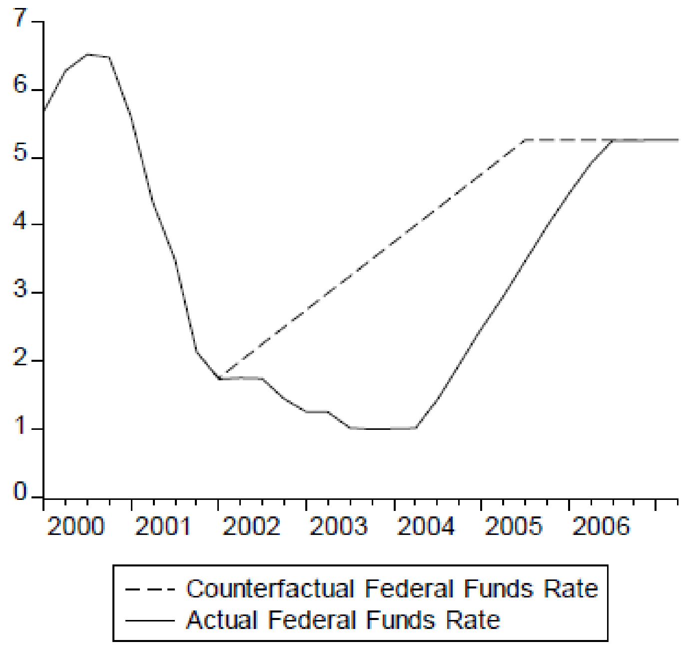


FIGURE 4: Loose-fitting monetary policy
Source: Taylor (2007)

To test the impact of monetary policy Taylor used a simple housing starts equation with only one explanatory variable the federal funds rate. He used quarterly data over a 50 year period from 1959 to 2007 and found that there was a strong, statistically significant effect that federal funds rate has on housing starts and it occurs with a lag. Next, he simulated his model under two scenarios: (1) the federal funds rate follows its actual path and (2) the federal funds rate follows a Taylor rule, smoothed to have the 25 basis point increment rather than what the rate was actually from 2002-2006. He compared the dynamic simulation where housing boom occurred as a function of actual interest rates with the counterfactual simulation where he found that there was a much smaller increase in housing starts with a higher federal funds rate being used in the model. Figure 4 above provides an empirical measure that monetary policy was too loose fitting or easy during this period. It also shows how it was an unusually big deviation from the Taylor rule and thus the actual interest rate decisions fell well below what policy suggested based on historical experience.

2.3 Securitization of Sub-Prime Mortgages

Securitization of sub-prime mortgages has been held responsible as another major factor responsible for the financial crisis. In order to understand the complexity of the securitization process we must re-visit the old school model of banking.

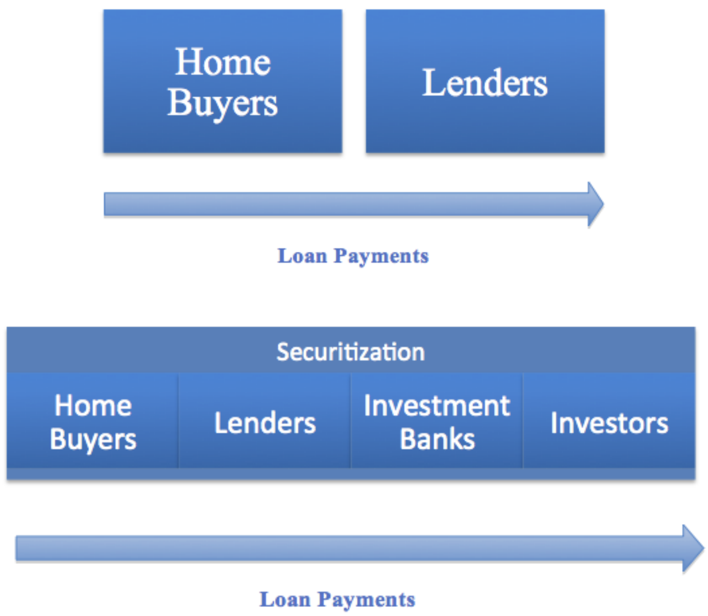


FIGURE 5: Home financing – old and new model
Source: Author's illustration

Historically, banks acted as financial intermediaries transferring funds from depositors (savings) to borrowers (loans) at two different rates of interest which accounted for the banks' profit (Diamond, 1984). Securitization altered this original idea of banking as banks then became intermediaries between investors (rather than just depositors) and borrowers. This evolution of the banks' business model resulted in transforming the nature of their operations from a traditional business model, wherein banks make loans to customers and hold these loans on their balance sheets (buy-and-hold strategy), to a business model wherein banks originate loans and then securitize them (originate-to-distribute, OTD). This OTD model has its positive and negative connotations. On the positive front, it means that assets that ordinarily would have been held up to maturity are now sold to many operators, providing a high volume of funding, giving a boost to the economy as a whole. On the negative front, it enables a high degree of leverage and the possibility of reduced intermediaries' incentives to monitor their portfolios' quality.

The process of mortgage securitization began with a home purchase. To get the much needed cash to purchase the house the home buyer signed a mortgage loan agreement and gave it to an 'originator' (specialized mortgage brokers). In exchange for the cash the originator then sold these mortgages to a securitizing firm who bundled a large number of such mortgages into pools to create new securities, private label ABS (asset backed securities), that had claims on the pools' cash flows. These were 'structured' securities meaning that mortgage pools were divided into 'tranches' according to the predicted riskiness of the loan in the entire pool. Lastly, in exchange for cash, the securitizing firms sold these newly created securities to individual investors, hedge funds, college endowments and insurance companies etc. If all goes according to the plan then the picture looks quite good, but this chain was broken by a wave of mortgage defaults. Two important points to notice now are the following: First, this process did not change the total amount of risk embedded in the mortgages but it did result in making some securities less risky than average and some more. Second, each time a new security was created or rated fees were being earned by investment banks and rating agencies. There is nothing inherently wrong with securitization if we have affirmative answers to the following two questions: Were home buyers making sounding decisions regarding their ability to repay home loans? The second question is did the ultimate investors recognize the additional risk? The answer to the first question is a resounding no, as the majority of these home buyers were subprime borrowers defined as one who displays a range of credit risk characteristics, including two or more 30-day delinquencies in the last 12 months, or one or more 60-day delinquencies in the last 24 months; judgment, foreclosure, repossession, or charge-off in the prior 24 months; bankruptcy in the last 5 years; relatively high default probability as evidenced by a credit bureau risk score (FICO) of 660 or below; and debt service-to-income ratio of 50% or greater; or, otherwise limited ability to cover family living expenses after deducting total debt-service requirements from monthly income. The answer to the second question is also a resounding no. From 2002-07 Wall Street saw a staggering growth in structured securities. The largest component of these securities were mortgage-related securities while other securities were backed by assets such as commercial mortgages, leveraged loans, student loans and corporate bonds (IMF 2008).

2.4 Role of GSE's in the Housing Bubble

As losses at the two companies mounted, the government placed Fannie Mae and Freddie Mac into conservatorship on September 6, 2008. At that point the two GSEs together held in portfolio or guaranteed through mortgage-backed securities (MBS) some $5.2 trillion of mortgages, or over 40 percent of the $12 trillion residential mortgage market. (Financial Crisis Inquiry Commission Staff Report 2010)

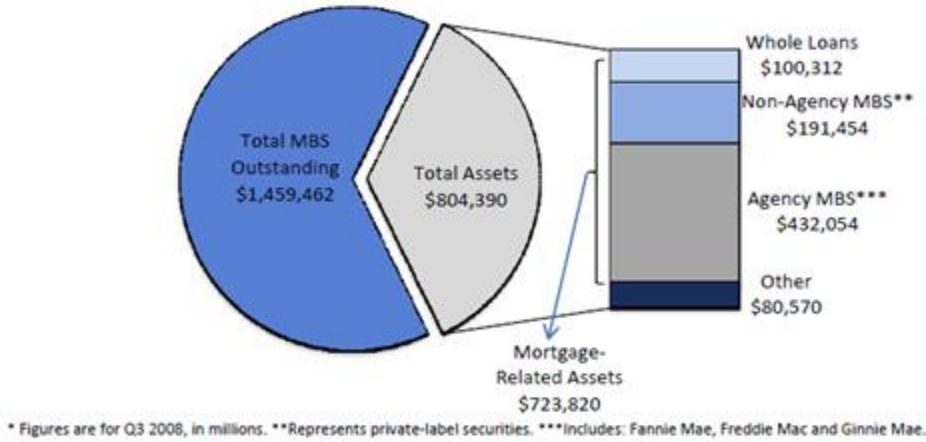
In 1938 the Federal National Mortgage Association (FNMA or Fannie Mae) was started as a wholly owned government corporation, a federal agency, to purchase loans insured by the Federal Housing Administration (FHA) and thus in the process provide liquidity to lenders in the mortgage market. In 1968, Congress chartered Fannie Mae as a GSE, a publicly traded private corporation. In 1970, Congress chartered a second GSE, Freddie Mac. Playing the key role of implementing new HUD policies, the two big GSE's essentially repurchased mortgages from originators of all types. For years both the GSE's operated under congressionally mandated goals of extending mortgage credit to specific income groups and fulfilling certain housing goals together with competing for higher profits in the U.S. mortgage market. What led to the gradual deterioration of lending standards in the entire U.S. mortgage market were the federally initiated affordable housing goals beginning in the 1990s.

In 1995, the Department of Housing and Urban Development (HUD) came out with a target goal of 70 percent homeownership rate among low-income groups. In 1999 HUD directed Fannie Mae and Freddie Mac to relax their requirement standards on mortgage loans, including a move toward sub- and non-prime loan approval. Thus the GSE share of mortgage loans with high loan-to-value (LTV) ratios increased from about 6% of purchases in 1992 to about 19% in 1995. By 1997, Fannie was purchasing 97% LTV mortgages and by 2001, it was purchasing mortgages with no down payment whatsoever.[2]

The mortgages bought by these two GSEs were either held in portfolios or securitized into mortgage-backed securities (MBS). From 2002, the private mortgage market saw an aggressive move toward non-conforming and jumbo mortgage loans. Likewise, Fannie Mae and Freddie Mac shifted their portfolio allocations toward private label mortgage-backed securities to fulfill Congress' affordable housing goals. From

[2] Figures from Wallison (2009).

2002 and 2006, total MBS holdings for them decreased by nearly half while their combined holdings of private label MBS increased substantially in this time. The figures below indicate how, by third quarter of 2008, both Fannie Mae and Freddie Mac were securitizing many more mortgages than they were funding on their balance sheets.



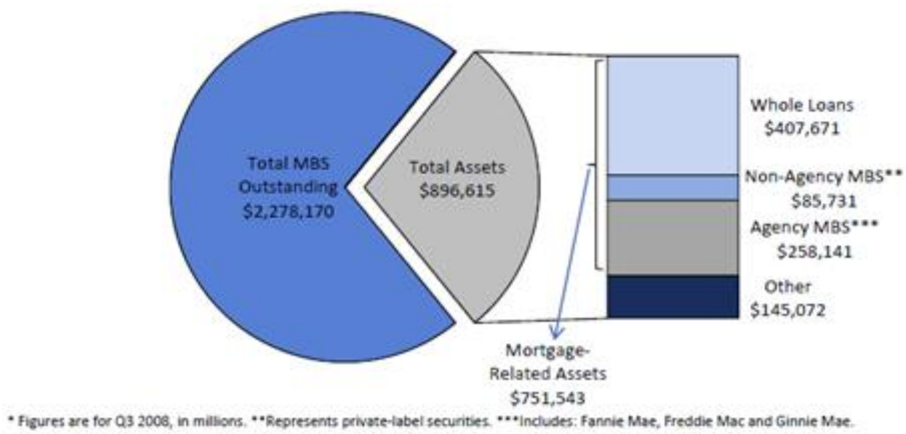


FIGURE 6: Assets and MBS – Fannie Mae and Freddie Mac
Source: Financial Crisis Inquiry Commission (2010)

The decline in underwriting standards that swamped the U.S. mortgage market from 2000 to 2006 weakened the quality of holdings even in the GSEs' portfolios as a huge portion of their MBS holdings were securitized from sub-prime and non-prime mortgages. By 2007 Fannie Mae held 18% of its loans with credit scores less than 660, 25% of its total loans with LTV above 80%, 23% in sub-prime and high-risk mortgages and about 15% in interest-only loans. The figure below shows growth in assets in the Fannie Mae and Freddie Mac portfolio from 1970 through the boom in mortgage refinancing in 2003. Most of these assets were funded by issuing debt.

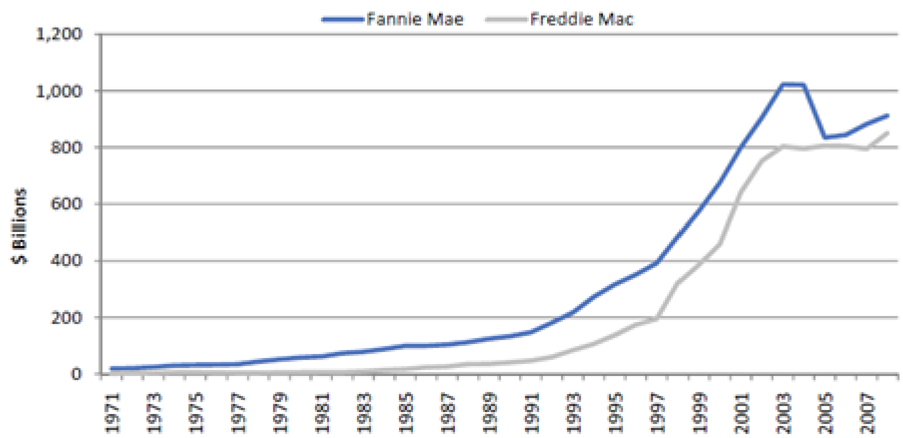


FIGURE 7: Total assets: Fannie Mae and Freddie Mac
Source: Financial Crisis Inquiry Commission (2010)

Fannie Mae stated in its 2007 10K filing:

> We are experiencing high serious delinquency rates and credit losses across our conventional single-family mortgage credit book of business, especially for loans to borrowers with low credit scores and loans with high loan-to-value ratios. In addition, in 2007 we experienced particularly rapid increases in serious delinquency rates and credit losses in some higher risk loan categories, such as alt-A loans, adjustable-rate loans, interest-only loans, negative amortization loans, loans made for the purchase of condominiums and loans with second liens. Many of these higher risk loans were originated in 2006 and the first half of 2007.

Similarly Freddie Mac stated in its 2007 10K filing:

> The proportion of higher risk mortgage loans that were originated in the market during the last four years increased significantly. We have increased our securitization volume of non-traditional mortgage products, such as interest-only loans and loans originated with less documentation in the last two years in response to the prevalence of these products within the origination market. Total nontraditional mortgage products, including those designated as alt-A and interest-only loans, made up approximately 30%

and 24% of our single-family mortgage purchase volume in the years ended December 31, 2007 and 2006, respectively.

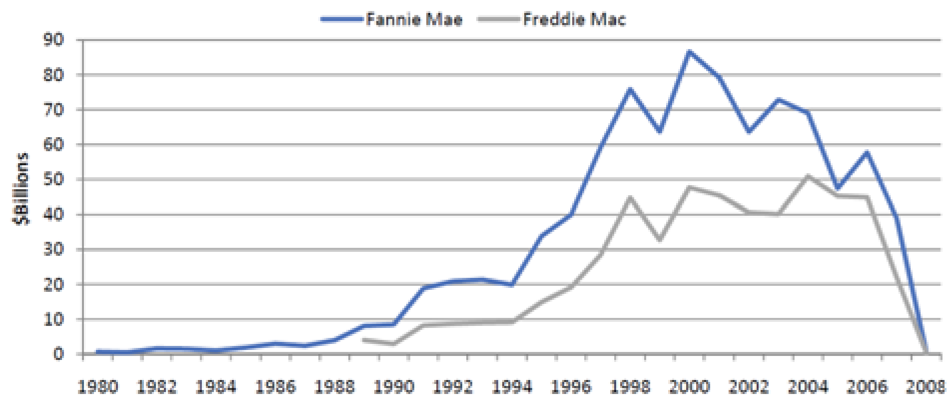


FIGURE 8: Market Capitalization: Fannie Mae and Freddie Mac
Source: Financial Crisis Inquiry Commission (2010)

Between 2005-2007 Freddie and Fannie purchased around $1 trillion in sub-prime and Alt-A loans adding up to 40% of their mortgage purchases. As subprime delinquencies increased, both GSEs took significant losses on their mortgage holdings. Compared to GSEs' losses their total capital was much less, which had a huge impact on the GSEs' net worth. By end of 2007, the sum of liabilities and MBS that the two GSEs' had guaranteed equaled national U.S. debt. It is said that for every $100 they had guaranteed or lent via securities all they had was $1.20 of equity. However, Paul Krugman (2008) attacking critics of the two GSEs', was of the view that Fannie and Freddie had nothing to do with risky lending and had not made any subprime loans.[3] True, Fannie and Freddie did not lend to subprime borrowers (in fact they did not lend at all), but they did buy loans, a huge proportion of which were subprime loans.

The insolvency of these two GSEs resulted in the failure of market leading mortgage lenders, such as Countrywide Financial and IndyMac Bank, which the market

[3] Norberg (2009a)

had earlier considered to be successful. Thus notably, the housing finance GSEs played a central role in the systemic nature of the collapse of the financial market. The key lessons to learn here are to understand the failure of this institutional model and restore properly aligned incentives to the U.S. housing and housing finance market. Congressional leaders need to wind down the GSEs and thereby create a U.S housing finance market free of the distortions generated out of this institutional arrangement.[4]

2.5 Global Savings Glut

> I will argue that over the past decade a combination of diverse forces has created a significant increase in the global supply of saving - a global saving glut - which helps to explain both the increase in the U.S. current account deficit and the relatively low level of long-term real interest rates in the world today. (Bernanke 2005)

As per Bernanke (2005) the global savings glut phenomenon was singlehandedly responsible for the movement from moderate deficits to large surpluses in emerging-market countries, a huge increase in the current account deficit in the U.S., and the significant decline in long-term real interest rates. These low long-term real interest rates in turn gave rise to the housing market asset price inflation whose inevitable collapse precipitated the financial crisis.

To illustrate and understand Bernanke's argument I will use a two country version of the Metzler diagram. The current account identity for country i can be written as a function of savings (S) minus investment (I) as per the equation: $CA^{i} = S^{i} - I^{i}$. Now in the global economy the sum of current accounts must equal zero i.e. $CA^{U.S.} + CA^{Rest\ of\ the\ World} = 0$. The two-country Metzler diagram is based on these

[4] Jeffrey (2010a)

two equations. Say in 1996 both the U.S. and the Rest of the World had balanced current accounts. Figure 9 shows that in 1996 the world interest rate (r^w) equaled the autarky interest rate (r^a) in both economies. Thus neither had a current account deficit and S = I in both the economies.

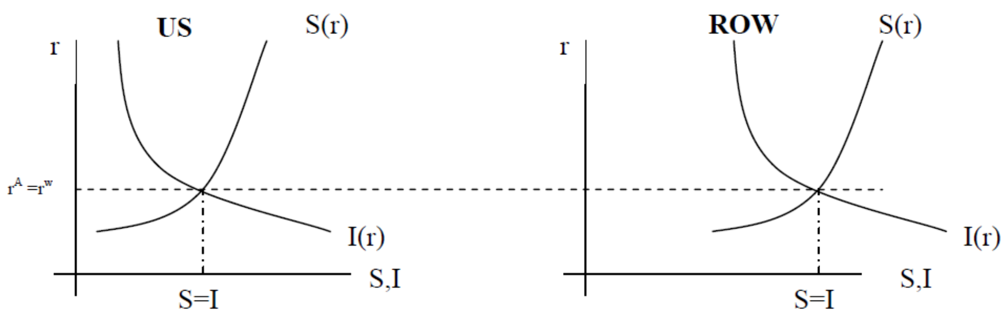


FIGURE 9: Savings and investment world in1996
Source: Two-country Metzler diagram

However, as shown in Figure 10, due to the East Asian financial crisis in 1997, investment in East Asian and oil producing countries dropped significantly in comparison to the savings in these countries. We can represent this decline in investment relative to savings using the two country Metzler diagram by a left-ward shift of the ROW investment schedule (Figure 11, investment in ROW is lower after 1997, at every given interest rate, corresponding to a higher investment in U.S. after 1997).

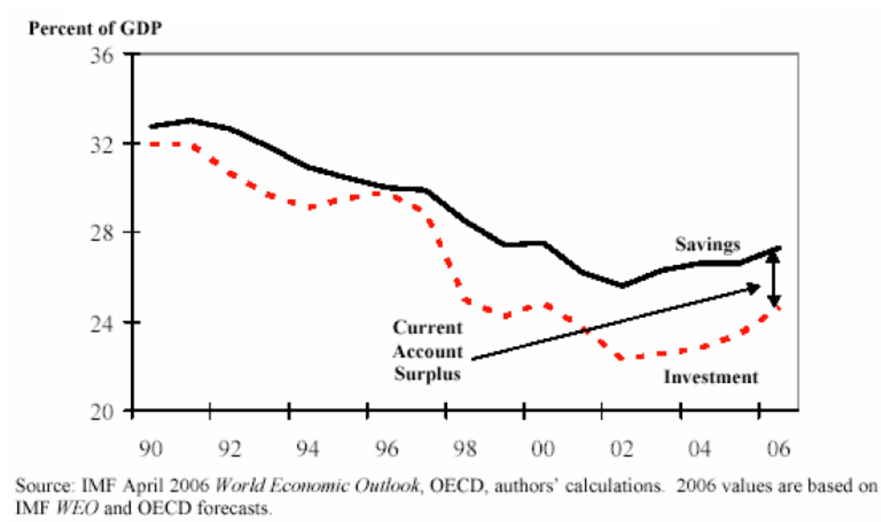


FIGURE 10: Savings and investment in Asia (excluding China)

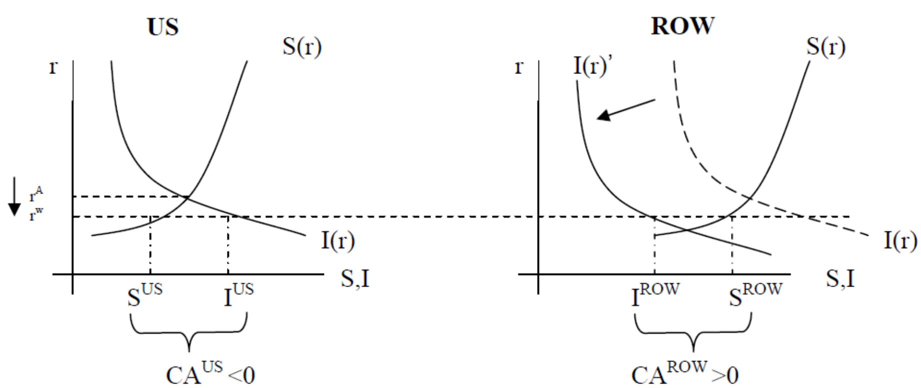


FIGURE 11: The world after 1997
Source: Two-country Metzler diagram

The leftward shift of the ROW investment schedule in Figure 11 results in the downward shift of the world interest rate r^{w} until the World current account identity holds, $CA^{U.S.} + CA^{Rest\ of\ the\ World} = 0$. Thus, Bernanke argued that this saving glut (excess of domestic saving over investment) in Asia post East Asian financial crisis of 1997 caused the world interest rate to fall and the U.S. current account deficit to increase.

Greenspan (2010) said that the saving rates of the developing world soared from 24% of nominal GDP in 1999 to 34% by 2007, far outstripping its investment rate.

Bernanke (2010) also pointed out that capital inflows from the emerging markets to industrial countries can help to explain asset price appreciation and low long-term global real interest rates. Bernanke provided empirical evidence of the relationship between capital inflows and house prices appreciation for 20 industrial countries. The results suggested that the relationship is highly significant, both statistically and economically, and about 31% of the variability in house price appreciation across countries is explained.

Taylor (2008), however, disagreed with the global saving glut hypothesis and suggested that there is actually no evidence for a global saving glut. He discusses whether low interest rates in 2002–04 were a result of the global savings glut and so possibly beyond the control of the monetary authorities. As per Taylor, US was running a current account deficit so savings were a lot less than investments. So the positive savings gap outside the US was offset by an equal-sized negative gap in the US, so clearly no global gap exists between savings and investment and thus there should be no extra impact on world interest rates. He used the figure below, produced by the International Monetary Fund in 2005, to show that there is in fact a savings shortage. It shows that from 2002-2004 the global saving rate, defined as world saving as a fraction of world GDP, was very low when compared with the 1970s and 1980s.

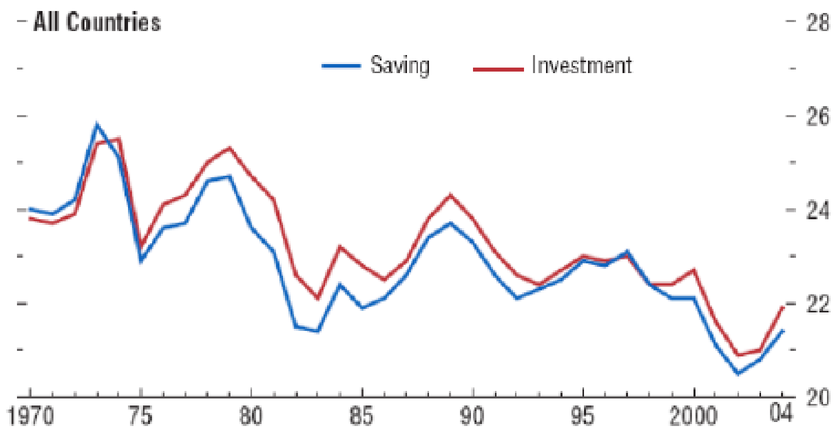


FIGURE 12: Global savings and investment as a share of world GDP
Source: Taylor (2008)

As a result, there is mixed evidence that global saving glut fueled the housing bubble.

2.6 Basel II – What Happened?

The Basel II framework on bank capital adequacy has been accused by several economists, policy-makers and market operators to be a major cause of the financial crisis, so this warrants an analysis of the degree to which we can ascribe the financial crisis to Basel II. To understand Basel II's contribution let us discuss the rationale behind the Basel II rules and what led to the adoption of the Basel II framework. The basic idea of public regulation stems from banks being involved in payment systems, their role in the allocation of financial resources in the economy and the fact that insolvency of even a single intermediary may entail social costs affecting the stability of the entire financial system. Francesco Cannata and Mario Quagliariello (2009) discuss how the 1988 Basel Accord (Basel I) was based on the central idea that a well-designed incentives structure, with banks themselves free to choose their preferred risk/return profile, is more effective than structural controls. Basel I endorsed this approach at the international level. The authors further believe that Basel I principles suffered from several shortcomings, such as focus on mere credit risk, non-existence of risk-sensitive measures of the creditworthiness of counterparties, and barely any incentives for banks to develop and strengthen risk management systems. These shortcomings combined with a dynamic landscape of financial innovation, created significant opportunities for regulatory arbitrage. Players started exploiting flaws of Basel I in terms of widespread use of the OTD model, increasing off balance-sheet exposures, a significant role being played by non-regulated and highly-leveraged intermediaries. This paved the way for Basel II, which aimed at overcoming Basel I shortcomings, strengthening prudential regulation and keeping it updated and consistent with evolving financial markets and the multitude of banking risks. Finally the authors discuss the major innovations brought about by Basel II which included a ‘three-pillar’ structure, monitoring a wide range of risks and the possibility for banks to choose risk measurement methods more suitable for their operations and complexity. The ‘three-pillars’ structure is based on the interaction between minimum capital requirements (Pillar 1), a supervisory review process (Pillar 2) and market discipline (Pillar 3). Regarding the wide range of risk, Basel II requires banks to meet minimum capital requirements in line with new types of risk, financial innovations and banking practices, thereby keeping the incentives for regulatory arbitrage to a minimum. Intermediaries can choose from different methodologies, ranging from simple methods to complex tools for measuring and quantifying each risk type. Basel II ensures the proportionality principle is fully implemented, there is no longer a level playing field with one-size-fits-all type of regulation, but there are several rules depending on the activity, size and risk-profile of each bank.

So how is Basel II blamed for the crisis? One accusation is that the average level of capital required by Basel II is inadequate, so this undercapitalization of banks compared to their risk exposures caused the collapse of many banks. Among others Onado (2008), while recognizing that Basel II aims at overcoming the insufficient correlation between risk and capital of Basel I, claimed that regulators did not properly exploit the Basel II reform for raising the capital base of the international banking system; rather, they committed to set up Pillar 1 rules aimed at keeping the level of capital on an aggregate basis unchanged with respect to Basel I. Another accusation is that Basel II interacting with the fair-value accounting for trading book assets caused remarkable losses in the portfolios of intermediaries. Thus the implementation of Basel II made banks' balance sheets more vulnerable to asset value fluctuations. Another blame on Basel II regulations is that capital requirements are cyclical and thus reinforce business cycle fluctuations. Goodhart and Persaud (2008, p. 1) are of the opinion that "the main problem is not the structure of regulatory oversight, either national or international, but the lack of counter-cyclical control mechanisms or instruments. Having foreseen the danger, the regulatory authorities did not have the instruments to do much about it. The Basel regime for capital adequacy does nothing to constrain credit booms. Its effect, if any, on the crunch will be to deepen it further". Basel II also gets the blame for assuming that banks' internal models for measuring risk exposure are superior to any other because during the crisis banks' internal models for measuring risk revealed serious shortcomings. Another charge against Basel II is that it provides incentives to intermediaries to deconsolidate some very risky exposures from their balance sheets.

It became quite lucrative for banks to remove items from their balance sheets and have businesses that do not show up as a part of their assets and liabilities. Banks would indirectly invest in structured products without building adequate capital reserves, which otherwise would have been necessary in the case of on-balance-sheet investments. The banks would do this by creating a special company called a "structured investment vehicle" (SIV) which would then make investments and finance its operations by taking cheap short-term loans directly in the financial market. This SIV is a separate legal entity not owned by the bank but the bank made a commitment to fund this company if capital could not be raised in the market. What this translates into saying is if the bank had invested $100 it would be required to keep $8 in capital but with the SIV in place the bank needs only 80 cents of capital. As a $100 investment made by the bank this is not a big deal but if the bank is indirectly investing $100 billion it is saving about $7.2 billion by this simple operation. This provided banks with an incentive to bundle mortgage loans into packages and keep them off the balance sheets via the SIV's.[5]

Though Basel II was ready in 2004 and banks started reviewing their credit standards to make them consistent with the incoming Basel II discipline, what is often forgotten is that Basel II was ultimately postponed to 2010 in the U.S. and involved a limited number of banks. At the same time it is also true that many banks in the attempt to transform well established credit processes and risk management methodologies, may have misjudged the actual exposures to new risk types (or new manifestations of traditional risks). The financial crisis at best highlighted some of the aspects of Basel II that need rework and adjustments to be made.[6] In conclusion Basel II was not one of the

[5] Norberg (2009b)
[6] Jeffrey (2010b)

major drivers of the crisis and thus should not be abandoned, yet it is imperative to strengthen those aspects of Basel II that have not worked properly during the financial crisis.

2.7 Structured Finance Meltdown

> There is broad consensus that structured finance played an important role in the development and propagation of the financial turmoil. For example, the IMF has concluded that … the proliferation of new complex structured finance products, markets, and business models exposed the financial system to a funding disruption and a breakdown in confidence and that certain structured finance products… likely exacerbated the depth and duration of the crisis by adding uncertainty relating to their valuation as the underlying fundamentals deteriorated. (IMF 2008b, p. 54)

Securitization had enabled banks to make huge amounts of loans, repackage them, and sell them for profit to investors. This was in turn resulting in an exponential increase in the amount of cash generated from a simple mortgage transaction. Based on sophisticated quantitative financial engineering, the concept of the then successful securitization model started getting applied to other financial instruments. Figure 13 below shows the several main types of structured finance instruments that came up before 2007 in a big way.

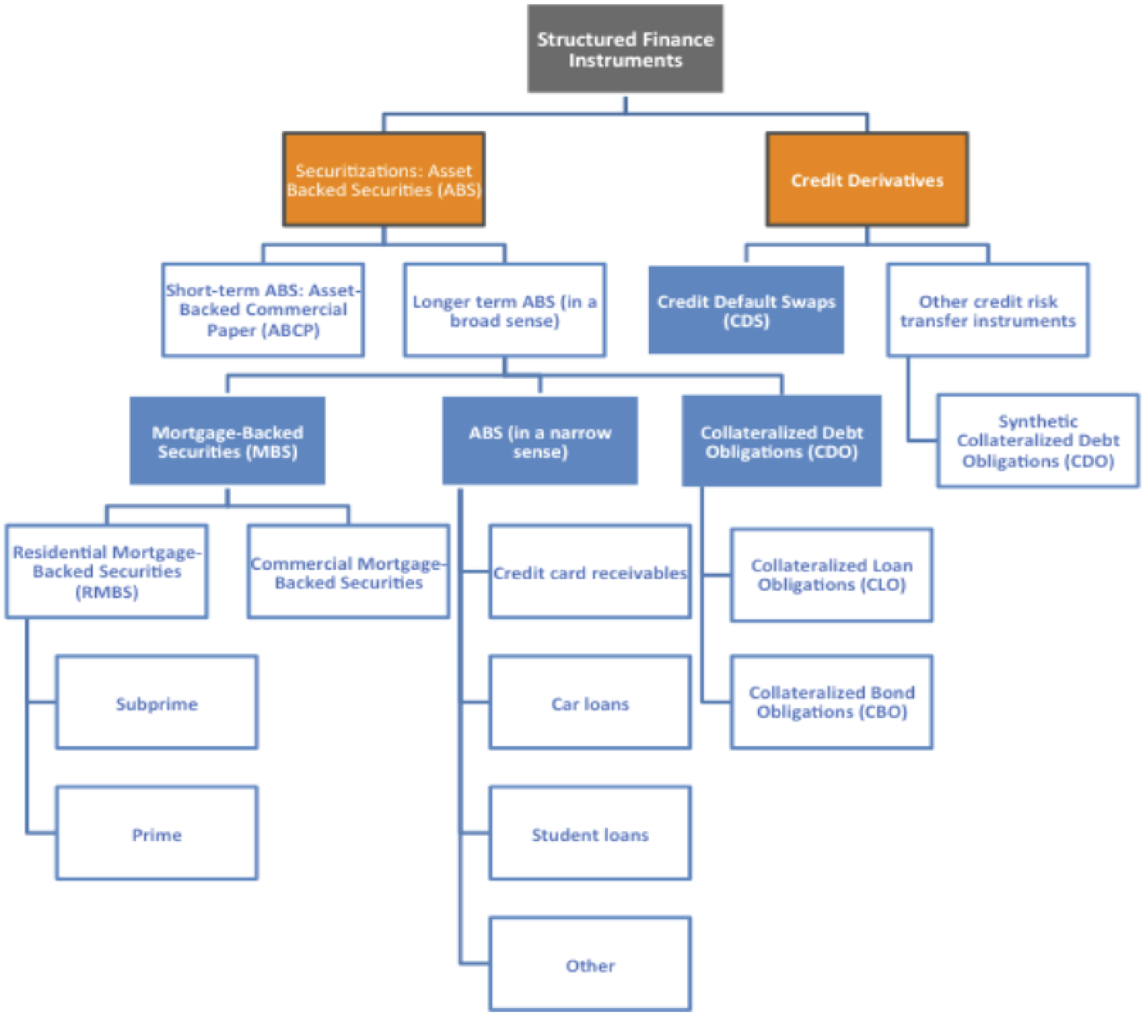


FIGURE 13: Structured finance instruments
Source: Criado, S. and Adrian Van Rixtel, 2008 (p. 14)

Developed in the mid-1980s, asset - backed securities are financial securities backed by a loan, lease, receivables, or credit card debt. Essentially they are the same as a mortgage-backed security, except that the securities backing it are not mortgage-based securities. Structured into different classes or tranches most asset-backed securities have different risks and yields. They are usually divided into three tranches: class A which has the highest credit rating (usually triple A), class B which has a lower rating than class A and class C which has the lowest rating. However, due to higher inherent risk in class B and class C tranches they pay a higher yield. Senior tranches hold a first lien on the assets of the asset pool, are generally safer investments and mostly insurance companies, and pension funds invest in these types of securities. Junior notes are tranches with either a second lien or no lien and are more risky investments as they are not backed by assets.[7] Hedge funds and investors seeking higher risk/return profiles are the typical investors in these types of securities. Under the sequential pay method, the highest tranche first receives all the payments of principal and after all investors of this tranche have been paid off, the next tranche receives the principal payments, then the next, and so on. The size of mortgage-related securities grew substantially from $100 billion in 2000 to $790 billion by end 2006.[8]

MBSs are basically ABSs that have securitized a pool of mortgages. It is assumed that real estate prices are less correlated as the geographical distance between the real estates grow, so mortgages securitized in an MBS are mostly dispersed geographically in an attempt to create diversification. The collateral for MBSs are sensitive to changing home prices and as the U.S. home prices started to fall in 2007 so did the value of the MBSs. As home prices continued to decline, the probability of default on mortgages increased as borrowers were having a hard time selling their real estate at or above the mortgage value. The two risks that are unique to MBS are prepayment risk and default risk. If interest rates move down more mortgage holders will probably pay off their loan and as the mortgage has been paid off, the investors now need to reinvest the prepaid capital. This is known as reinvestment risk as there is no assurance that the investor can reinvest at the same higher rate. Default risk is the risk that the borrower will not repay

---

[7] Criado and Rixtel (2008).
[8] SIFMA U.S. Research Quarterly (2007, p. 5).

on time and in full. Mostly, the default risk on MBSs that are backed by subprime mortgages as collateral is higher than MBSs backed by prime mortgages as collateral.

The common theme connecting all these complex financial securities and derivatives contracts was their impact on the financial crisis and most of them being based on housing as the underlying asset. However, the role of collateralized debt obligations (CDOs) needs to be particularly highlighted. Collateralized Debt Obligations, or CDOs, were structured ABS that purchased and pooled various financial assets. The most common type of CDO is the cash flow CDO whereby the special purpose vehicle owns the underlying debt posted as collateral in the CDO. A synthetic CDO is one where the trust does not own the underlying debt, and instead invests in credit default swaps (CDSs) to synthetically track their performance. A hybrid CDO is one which combines cash flow CDOs with an artificial CDO. A structured CDO is one which securitizes another securitization, using the tranches of another ABS as collateral. A CDO squared is one which securitizes the tranches of another CDO. The CDO market really took off from early 2000s when MBS were spreading like wildfire.[9] How CDOs facilitated a market for riskier tranches was that a CDO manager pooled together several MBS rated less than AAA with other assets or MBS from different geographical regions of the country to create a new CDO which by virtue of its diversification carried lower risk, and thereby could garner AAA rating despite the fact that the underlying assets carried higher risk. In the crisis CDOs were held responsible for approximately $420 billion in write-downs at financial institutions, 65% of the original issuance balance.[10] These problems in the CDO market were primarily due to three reasons: poorly constructed CDOs that provided

[9] Mengle (2007).
[10] Cordell, Huang and Williams (2012, p. 2).

a false sense of security about their safety, irresponsible underwriting practices, and flawed credit rating procedures wherein Moody's, S&P, and Fitch initially gave a rating of AAA to these securities most of which were later marked down. Several observers have also argued that credit default swaps (CDS) contributed immensely to the credit crisis. A credit default swap is a type of derivative contract in which one party (the protection seller) agrees to reimburse another party (the protection buyer) against default on a financial obligation by a third party (the reference entity).

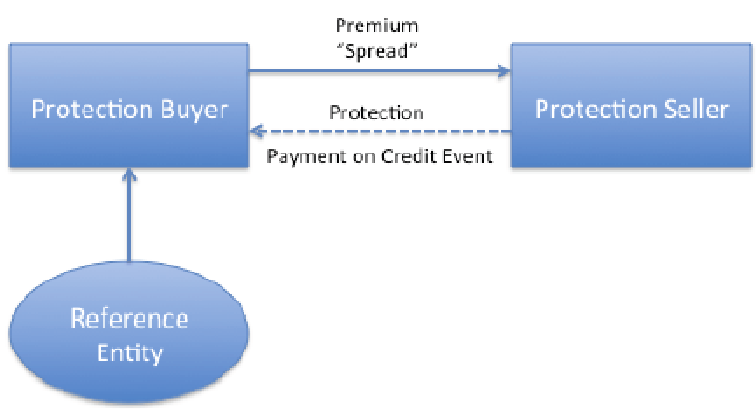


FIGURE 14: Credit default swap
Source: Criado, S. and Adrian Van Rixtel, 2008 (p. 32)

In the above figure is a simple CDS transaction where say the Protection Buyer (a Bank B) buys $10 million worth of bonds from the Reference Entity (a Company A). Now say if B wants to transfer the risk of a possible loss in value of the bonds purchased from A (say A's future prospects are declining) then B can transfer this credit risk of keeping the bond by entering into a CDS contract with the Protection Seller (a counterparty C, another Bank). The Protection Seller agrees to pay $10 million (or the negotiated notional amount) to the Protection Buyer if the Reference Entity defaults and the Protection Buyer in turn agrees to make an ongoing Premium payment called a "spread" to party C. CDS instruments were largely traded in the over-the-counter (OTC) market rather than on an exchange. What happens if any one participant in the CDS chain defaults, basically it can trigger off a cascade of losses through the highly interconnected structure, with the only question being who ultimately pays the $10 million loss.[11] As per the statistics provided by The Bank for International Settlements (BIS) on the CDS market size, the CDS market grew from $180 billion in 1998 to $6 trillion in 2004 to $57 trillion by June 2008. How the CDS market interconnections were capable of jeopardizing the entire financial system is clearly demonstrated by how Bear Sterns was acquired by JPMorgan Chase and AIG was rescued by the Fed (both Bear Sterns and AIG were claimed to have extensive exposures to CDS). AIG had provided insurance on a very large scale on AAA tranches in securitization, most of which lost their value in the face of declining U.S. home prices, and as the crisis progressed AIG's liabilities became very large and eventually it was taken over by the government.

[11] Gibson (2007).

CHAPTER 3: RELATED DATA, EVENTS AND EMPIRICAL WORK

Before describing the data used in the analysis, this paper will show data related to the financial crisis, adjoining events and some popular empirical research work done in this field.

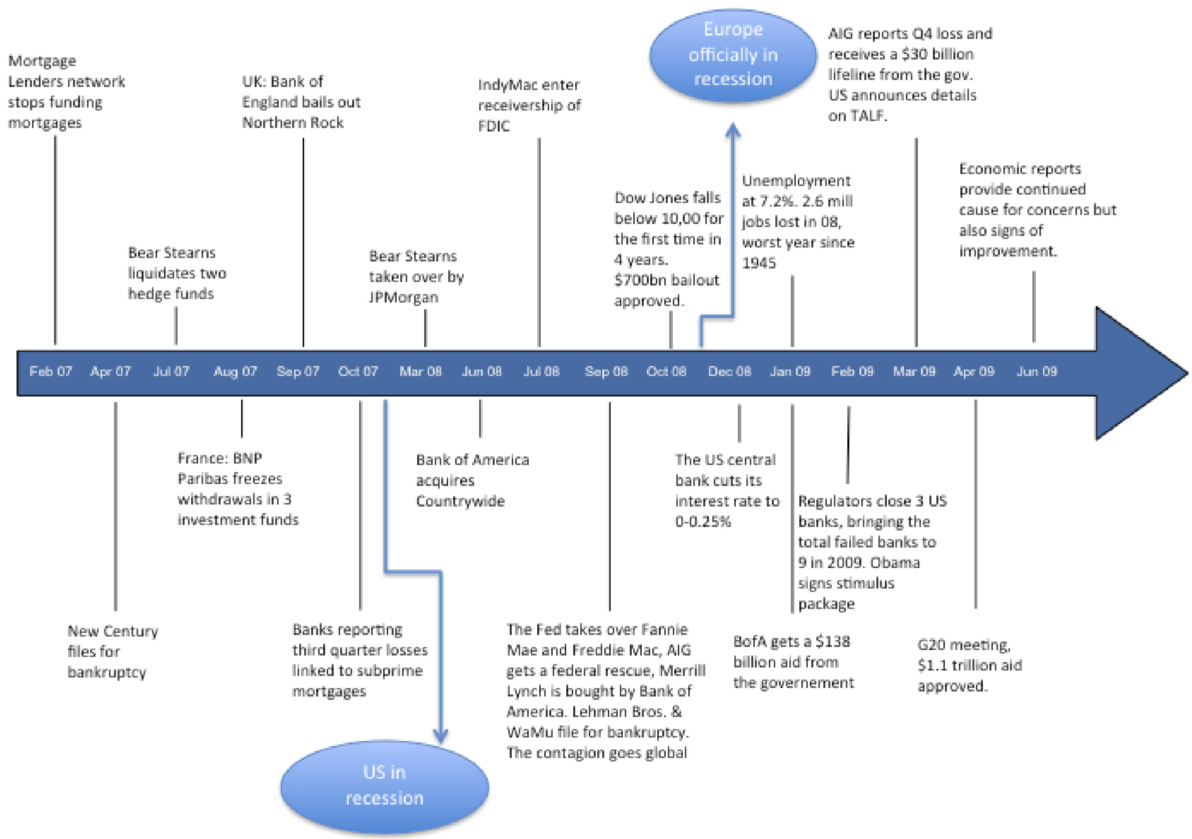


FIGURE 15: Timeline of major events during 2007- 2009
Source: Federal Reserve Bank of St. Louis' website, author's illustration

The above timeline of events surrounding the sub-prime credit crisis provides a high level summary of the events and helps us understanding the repercussions of the financial crisis. It is important to delve on the chain of events as that helps put in perspective how macro prudential regulations, federal funds rate, home prices bubble, global savings glut, securitization and financial innovation acted as the precursors to the crisis and what was the impact of the crisis.

It is important to emphasize and analyze whether the U.S. was alone facing the housing bubble or whether the housing bubble was happening in many countries around the globe at that time. Greenspan (2010) analyzes the global housing crisis and was of the opinion that by 2006 long-term rates in all developed economies and major developing economies had converged to single digits probably for the first time. This had resulted in asset prices, particularly house prices, moving dramatically higher globally. He believed that a fall in long-term mortgage interest rates occurred due to geo-political events which ultimately led, with a lag, to an unsustainable boom in house prices globally. Gonzalo Bernardos (2009) cites the existence of great liquidity between 2001 and 2005 as the prime factor behind the global real estate bubble. This liquidity in turn was a result of unbeatable financing conditions and historically very low interest rates. So from 2001-02 housing adopted a new role as a haven asset similar to the role shared by other assets such as bonds or gold. Eventually he says that between 2006 and 2008, the different real estate bubbles burst. The bubble first burst in the U.S., then in Spain, Ireland and the U.K., and eventually burst in remaining parts of Eastern Europe. He also believes that even without the financial crisis in the U.S. the bubble would have still burst in some countries such as Spain, the U.K. and Ireland on account of considerable excess supply of housing offers in these countries.

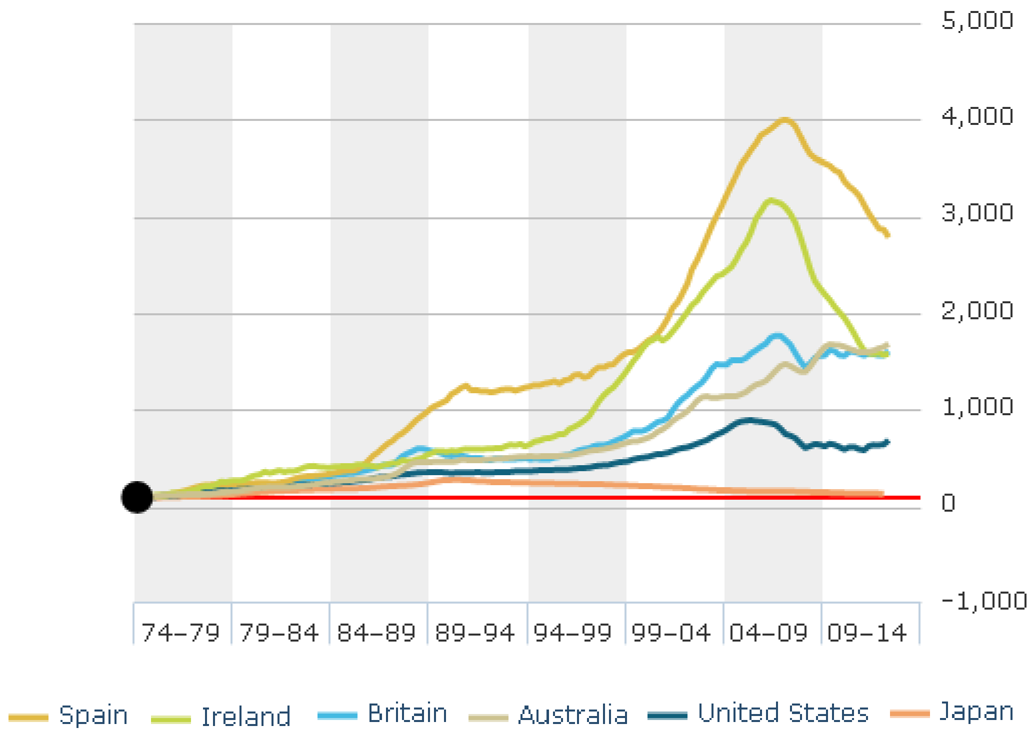


FIGURE 16: Nominal house-price index in several countries rebased to 100
Source: The Economist global house prices interactive graph detail

The figure above clearly illustrates that increases in house prices in the U.S. lagged behind some of the other countries like Spain, Ireland, Britain and Australia. This is also in line with Bernardos' and Greenspan's analysis that many countries other than the U.S. were having their own homegrown real estate bubbles. Thus, clearly the housing bubble was global in nature, not just a U.S. phenomenon.

It is also imperative to understand that while some cities within the U.S. suffered an enormous housing crisis, other cities were almost entirely spared. As per Federal Deposit Insurance Corporation (FDIC 2009), four states—Arizona, California, Florida, and Nevada had both highest rates of home price appreciation when the bubble was being formed and suffered from the most acute housing downturn when the bubble burst. In the 1980s and 1990s Arizona, Florida, and Nevada experienced very high population growth rates about two to four times the national rate due to their favorable weather and relatively affordable housing. Rapid population growth continued from 2004 to 2007 and all this resulted in higher than average rates of home construction. The increasing number of new residents in these states led to higher job creation and strong employment growth. With demand for housing increasing at an unprecedented level there occurred a housing market imbalance with home price appreciation far outpacing income growth. A huge factor contributing to this imbalance was the availability of new mortgage "affordability" products such as hybrid ARMs, interest-only loans, negative-amortization loans, balloon payment loans wherein for most of these products first-time home buyers and investors were not required to make a down payment.

By 2006, Arizona, California, Florida, and Nevada accounted for nearly half of total U.S. originations of privately securitized affordability mortgages. The majority of these mortgages originated were nontraditional mortgages. Las Vegas, West Palm Beach, Miami, and Phoenix also saw increased presence of speculators, investors and second-home purchases. Housing starts increased an average of 11% annually in these four states. From 2003-06 annual home price appreciation rates in these four states consistently exceeded the national average; Nevada saw a house price appreciation in 2004 peak at 37%; Arizona and Florida saw a house price appreciation in 2005 peak at rates more than twice the national average.

The Sand States Account for a Disproportionately High Share of Foreclosure Activity		
	National Share of Foreclosures Started	National Share of Mortgages Serviced
California	19.2%	12.9%
Florida	16.2%	7.8%
Arizona	4.4%	2.7%
Nevada	2.7%	1.2%
Sand States Total:	**42.5%**	**24.6%**

FIGURE 17: Sand states foreclosure activity
Based on data from Q1:2008-Q4:2008
Source: FDIC, Mortgage Bankers Association

Ultimately the imbalances led to the housing sector collapse and by 2008 average home prices in Arizona, California, Florida, and Nevada declined between 27 - 38% from their peak. With decelerating rates of home price appreciation foreclosure activity rose and in 2008 as exhibited by the above figure, these four States accounted for more than 40% of all mortgage foreclosures.

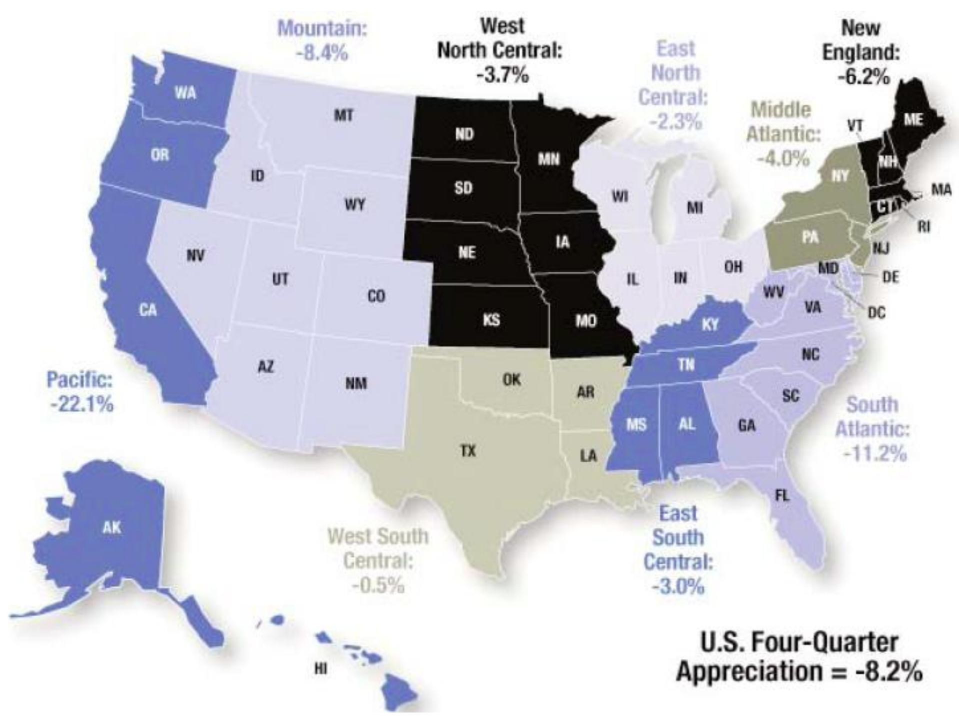


FIGURE 18: Four-quarter price change purchase-only index
For period ending Q4:2008
Source: Federal Housing Finance Agency report to Congress 2008

In Federal House Finance Agency report to Congress in 2008 they acknowledge that home prices were falling throughout the U.S. in 2008 but the magnitude of regional price declines exhibited vast differences. The Pacific Census Division comprising of the West Coast states suffered the worst house price decreases with 22.1% fall in prices. Homes in California experienced an average price decline of 25.5%. Of the Mountain Census Division conditions in Arizona and Nevada markets were particularly weak and of the South Atlantic Division Florida market house prices sustained the greatest declines. Home foreclosures continued to rise in 2008, with California and Florida experiencing the largest increases. All this reinforces our main point that bubbles did not appear everywhere in the U.S. and some states were more affected than others.

We now briefly analyze some of the quantitative approaches that have been applied by researchers in the area of study related to the aim of this thesis. It would also be helpful in selecting the final empirical model for this thesis. Kohn and Bryant (2010) present a synopsis of the debate concerning the definition of a bubble, methods of detection of the bubble, and root causes of the bubble. They use standard multiple regression analysis with median asking prices as the dependent variable and both supply and demand factors as independent variables for housing consumption. They use six independent variables with housing inventory and vacancy rates representing supply factors and consumer price index, 30-year fixed conventional mortgage rates, personal income and population representing demand factors. They used three versions of the model for multiple regression analysis: pre-bubble model (1988: Q1 – 1996: Q4), bubble model (1997: Q1- 2007: Q4) and full model (1988: Q1 – 2007: Q4).

$$\text{MAP} = F(\text{CPI, Housing Inventory, Mortgage rates, PI, Population, Vacancy Rates})$$

The authors determine that a bubble did occur as they found significant differences in results between two examined periods 1988 to 1996 (pre-bubble) and 1997 to 2007 (bubble). They found that in the pre-bubble period, median housing prices were impacted by only two variables, personal income and vacancy rates, whereas in the bubble period, all six variables included were significant.

Mikhed and Zemcik (2009) use panel data unit root tests and Pedroni test statistic for testing cointegration between house prices and rents, with and without a bubble term. They utilize two datasets for house price and rent indices in their study covering 1978-2006 time period. The first dataset comprises of house price index (HPI) and rent of primary residence index for 23 Metropolitan Statistical Areas from 1978 to 2006 semi-annually. The second dataset contains the HPI and the fair market rent (FRM) data on 273 Metropolitan Statistical Areas from 1986 to 2006 at a yearly frequency. The former dataset used by the authors has a longer time span and a smaller cross-sectional dimension while the latter dataset has a large cross-section and a shorter time series. Using these two different datasets the authors test for bubbles in housing markets and find that house prices and rents are not cointegrated or have a different order of integration. The authors suggest that there is a bubble if either the price-level is non-stationary while the rent-level is stationary, or both series are of first order of integration but they are not cointegrated. As they find no cointegration in the two different datasets the authors conclude that this indicates the presence of a bubble term in their asset pricing model and it implies that an error correction model for house prices with rents as the fundamental factor cannot be used. Finally, they use Granger causality tests on their panel data which suggests that house price changes are helpful in predicting changes in rents and vice versa.

Case and Shiller (2003) promote a completely alternative approach to the analysis of housing bubbles. The authors view housing bubbles as a result of unrealistic expectations of future prices sustained by speculative feedback and social contagion. They claim that when current prices depend on expectations of future price increases an asset bubble appears and when these expectations change current prices may decline rapidly. So they believe that measuring expectations is extremely relevant to housing bubble analysis. They also analyze fundamentals including personal per capita income, population, and employment from 1985-2003 using linear and log-linear regressions with three dependent variables: the level of home prices, the quarter-to-quarter change in home prices, and the price-to-income ratio. They conclude that for 42 states income is sufficient to explain price movements and for the other 8 states the fundamental economic variables of population and employment add explanatory power. The authors also report the results of a survey mailed out to 500 people who bought houses in 2002 in four different metropolitan areas of Los Angeles, San Francisco, Boston, and Milwaukee. They found that expectations of future prices and market perceptions play an influential role in people's decision to buy a home. They also document that recent homebuyers in 2003 expected a substantial price increase during their first year of occupancy. Furthermore, if people think that home prices are highly unlikely to fall, and certainly not likely to fall for long, then a future expectation of large price increases may have a strong impact on housing demand, as people now foresee barely any risk associated with home purchase. Case and Shiller also report that the term "housing bubble" had essentially no popularity prior to 2002 while the term “housing boom” had been in much more frequent use since the 1980s.

CHAPTER 4: DATA DESCRIPTION

In my empirical analysis the first dependent variable is U.S. housing starts, a measure and a proxy of the housing bubble. Housing starts is the number of new privately owned houses/ residential construction projects that have begun during any particular month. From Jan 2000-Jan 2006 when the housing bubble reached to a peak, housing starts increased by more than 38% (author's calculation) compared to a mere 10.1% (author's calculation) increase during the 1990s. In the post housing bubble period, from Jan 2006-Apr 2009 housing starts dropped by about 78% (author's calculation).

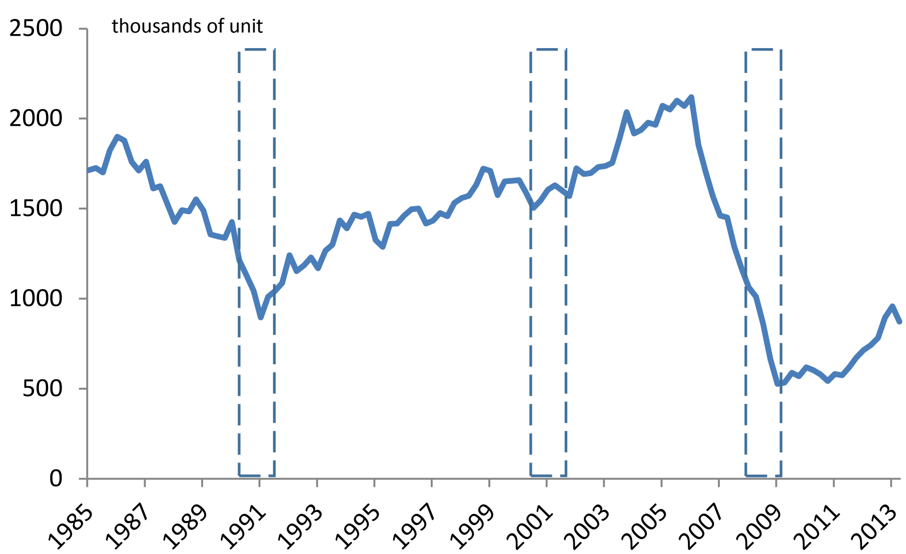


FIGURE 19: Housing starts total: new privately owned housing units started
Quarterly average, levels data
Source: Federal Reserve Bank of St. Louis' website, author's illustration
Dashed lines indicate periods of US recession

New housing starts data are collected and reported monthly by the U.S. Census Bureau and the U.S. Department of Housing and Urban Development. The Census Bureau conducts the Survey of Construction which provides national and regional statistics on starts, completions, and characteristics of new, privately-owned single-family and multifamily housing units. Data are available monthly and annually for housing starts since 1959 to the public on the Census Bureau's website.

The second dependent variable in my analysis and a measure of the housing bubble is home prices. Home prices dramatically increased during the first half of this decade (particularly during 2002-06 period). There are a few key points that we must stress about home prices. There is no single widely used standard of measuring U.S. home prices, and there are several different theoretically sound measures of home prices. Fortunately all of the home price measures are positively correlated over time, so in terms of direction, they all tell the same story. I will use two measures of home price indices: Federal Housing Finance Agency (FHFA) Purchase-only index, and the S&P/Case-Shiller national house price index.

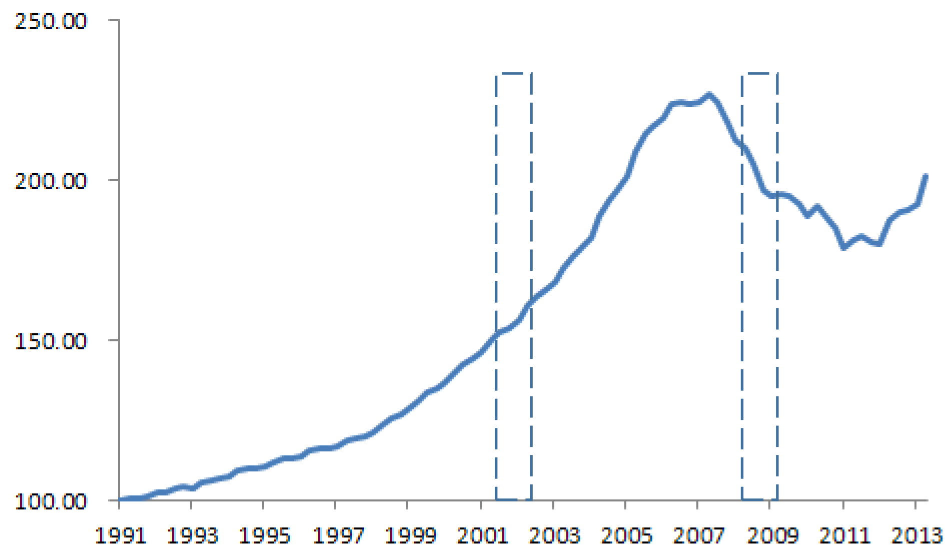


FIGURE 20: FHFA purchase only index, quarterly average
Levels data 1991Q1=100
Source: FHFA website, author's illustration
Dashed lines indicate periods of US recession

The FHFA HPI measures the price of single-family housing, using data provided by Fannie Mae and Freddie Mac. It is a weighted, repeat-sales index, which measures average price changes from repeat sales and refinancing of the same properties in 363 cities. FHFA produces a purchase only home price index, FHFA Purchase-only, which excludes refinancing and markets focus primarily on this purchase-only index. The benefit of using the FHFA HPI over the all-transactions HPI is that it excludes home appraisals, which are used in the construction of the all-transactions index and are considered to be systematically biased measures of true home values. The implication of using home appraisal data is that short-term price patterns may be obscured (FHFA website).

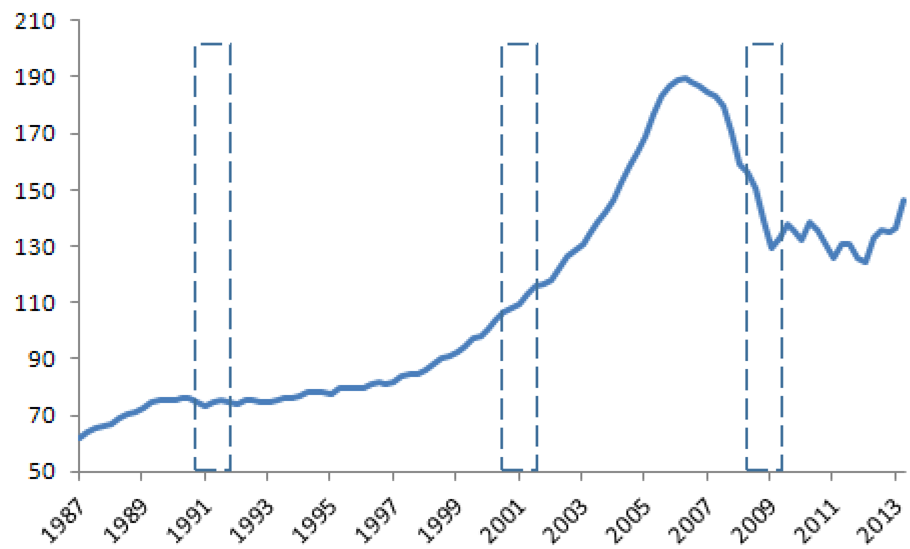


FIGURE 21: S&P/Case-Shiller national house price index, quarterly average
Source: S&P Dow Jones Indices website, author's illustration
Dashed lines indicate periods of US recession

The S&P/Case-Shiller national house price index is a leading measure of U.S. residential real estate prices and is calculated quarterly. The index is a composite of single-family home price indices for the nine U.S. Census divisions and tracks on a national level changes in the value of residential real estate. It is designed to be a reliable and consistent benchmark of home prices in the United States. The FHFA and S&P/Case-Shiller national house price index cover about the same geographical area and follow the same fundamental repeat-valuation approach but the key difference is that the latter includes foreclosed homes and shows a much larger decline in national home prices (S&P Dow Jones Indices website).

My third dependent variable is the ratio of S&P/Case-Shiller HPI to FHFA HPI which will serve as a proxy for the financial innovation/ non-conventional mortgages and the role played by structured finance in the crisis. To calculate this ratio we will divide S&P/Case-Shiller HPI year-over-year percent change by FHFA HPI year-over-year percent change. The reason why this ratio might capture financial innovation is that S&P/Case-Shiller includes non-conventional mortgages as well as foreclosed homes while FHFA excludes foreclosed homes.

Moving onto the explanatory variables of the model, my first independent variable and one of the many possible contributors to the housing bubble is the federal funds rate which will serve as a proxy for monetary policy. The federal funds rate is the interest rate at which depository institutions lend balances held at the Federal Reserve, called federal funds to each other overnight on an uncollateralized basis. Institutions with surplus balances in their accounts lend those balances to institutions in need of larger balances. The Federal Open Market Committee (FOMC) meets eight times a year to determine the federal funds target rate.

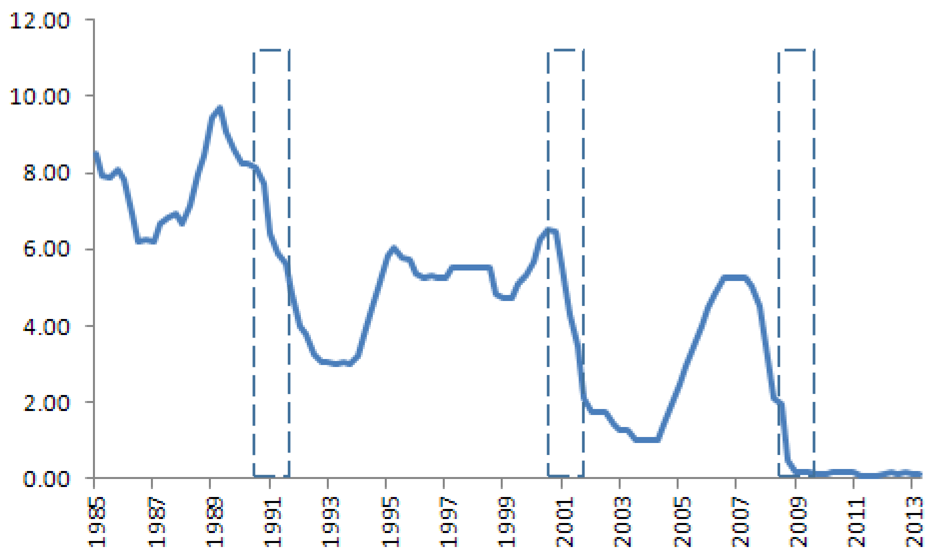


FIGURE 22: Effective federal funds rate (in %), quarterly average
Source: Federal Reserve Bank of St. Louis' website, author's illustration
Dashed lines indicate periods of US recession

The federal funds rate is the benchmark interest rate in the U.S. financial market which influences the prime rate, the rate banks charge their customers while making loans. The federal funds rate also influences longer- term interest rates such as mortgages, loans, and savings, key to consumer wealth and confidence. It also affects inflation, economic activity and other short term interest rates. Raising the fed funds rate makes it more expensive to borrow, increases other short term interest rates and thus helps lower inflation. We will use the quarterly year-over-year percent change for the federal funds rate.

My second explanatory variable is the U.S. current account balance as percent of GDP (CA_GDP) which will serve as a proxy for the global saving glut. Bernanke (2010) also used CA_GDP as an explanatory variable and I will follow his approach. The World Bank's definition of the current account balance is the sum of net exports of goods and services, net primary income, and net secondary income. Data on this variable were provided by Professor Azhar Iqbal via IHS Global Insight database.

Lastly, I will add several exogenous variables to improve the fit of my model and avoid mis-specification error including variables that measure the ability of consumers to handle mortgage debt (household debt ratio, housing affordability index, mortgage rate) and a few economic variables (unemployment rate and GDP).

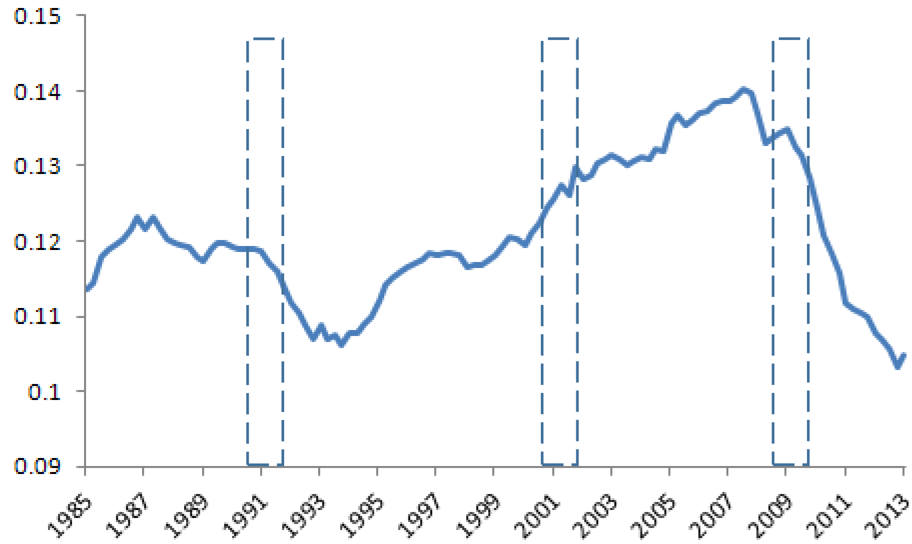


FIGURE 23: Household debt service ratio, quarterly average
Source: Federal Reserve Board website, author's illustration
Dashed lines indicate periods of US recession

The Federal Reserve Board defines household debt service ratio (DSR) as an estimate of the ratio of debt payments to disposable personal income. Debt payments consist of the estimated required payments on outstanding mortgage and consumer debt. Quarterly DSR data are available from the Fed website.

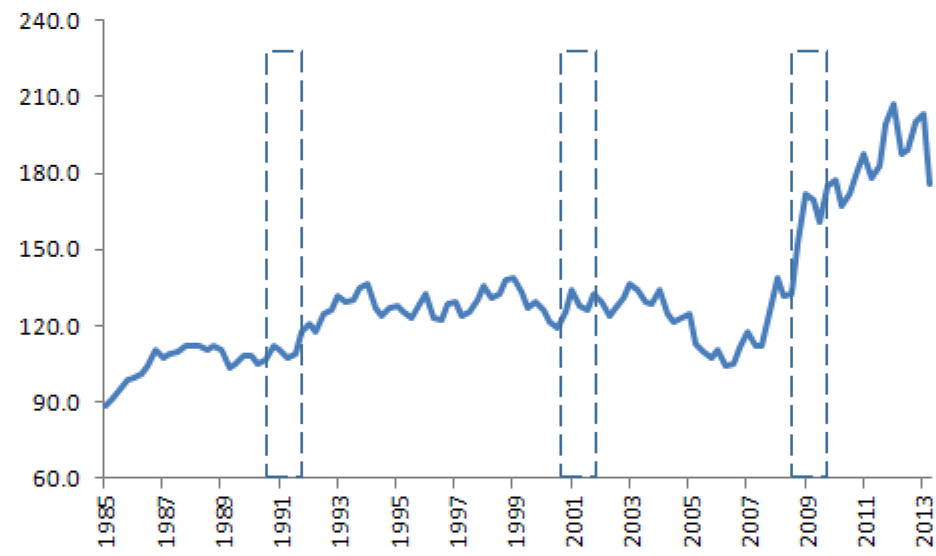


FIGURE 24: Housing affordability index, quarterly average
Source: Federal Reserve Bank of St. Louis' website, author's illustration
Dashed lines indicate periods of US recession

The National Association of Realtors (NAR) defines the Housing Affordability Index as a measure of whether or not a typical family can qualify for a mortgage loan on a typical home. They define a typical home as the national median-priced, existing single-family home as calculated by NAR; a typical family as one earning the median family income as reported by the U.S. Census Bureau; and the prevailing mortgage interest rate as the effective rate on loans closed on existing homes from the Federal Housing Finance Board. An index value of 100 means that a family with the median income has exactly enough income to qualify for a mortgage on a median-priced home while an index value above 100 implies that a family earning the median income has more than enough income to qualify for a mortgage loan on a median-priced home. Data on the monthly and quarterly housing affordability index are available from the NAR website.

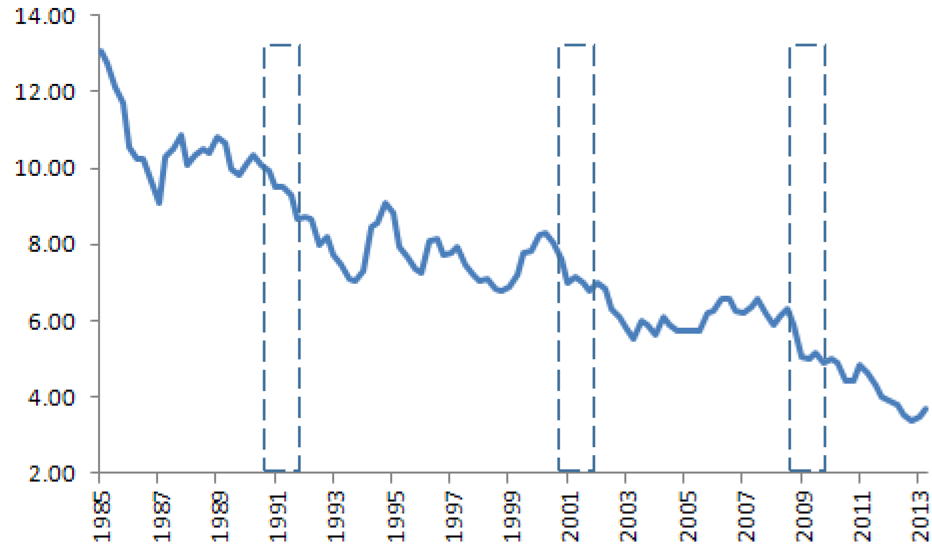


FIGURE 25: 30-year conventional mortgage rate (in %), quarterly average
Source: Federal Reserve Bank of St. Louis' website, author's illustration
Dashed lines indicate periods of US recession

Another exogenous variable that I will use is the Mortgage Rate as it helps determine consumer demand for housing. To model this variable I will use the 30-year conventional mortgage rate available in a monthly, quarterly, semi-annual and annual format from the Federal Reserve Bank of St. Louis' website. They define the 30-year conventional mortgage rate as the contract interest rate on commitments for fixed-rate first mortgages.

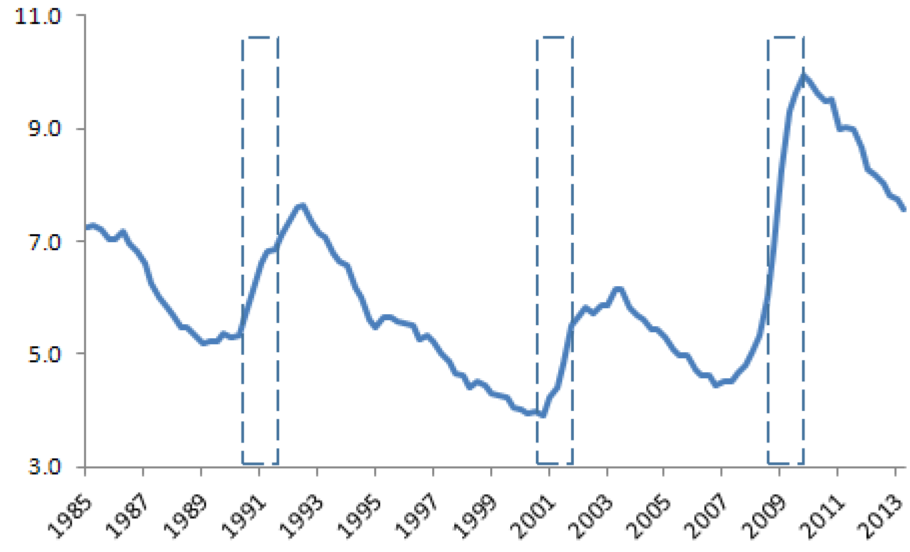


FIGURE 26: Civilian unemployment rate (in %), quarterly average
Source: U.S. Department of Labor: Bureau of Labor Statistics website, author's illustration
Dashed lines indicate periods of US recession

The final explanatory variables that I will use in my model are the unemployment rate and GDP. Monthly data on the unemployment rate were obtained from the U.S. Bureau of Labor Statistics. The monthly data were converted to quarterly data by a simple mean. The U.S. Bureau of Labor Statistics defines the unemployment rate as the number of unemployed as a percentage of the labor force. Labor force data are restricted to people 16 years of age and older, who currently reside in one of the 50 states or the District of Columbia, who do not reside in institutions (e.g., penal and mental facilities, homes for the aged), and who are not on active duty in the Armed Forces.

FIGURE 27: GDP growth rate, quarterly percent change from year ago
Source: Federal Reserve Bank of St. Louis' website, author's illustration
Dashed lines indicate periods of US recession

Quarterly Nominal GDP data in percent change from year ago form were procured from the Federal Reserve Bank of St. Louis' website. They provide GDP data from 1947 to current period. They define GDP, the featured measure of U.S. output, as the market value of the goods and services produced by labor and property located in the United States.

CHAPTER 5: MODEL SPECIFICATION AND EMPIRICAL METHODOLOGY

The purpose of this paper is multi-fold: understanding the contributing factors to the financial crisis, quantifying the contributions of these factors to the housing bubble, and the key takeaways for policy reform to avoid another global financial meltdown. Generally speaking, traditional econometric methodology proceeds in a very methodical way to address an economic problem: statement of hypothesis; specification of the mathematical, statistical, or econometric model; collecting data; estimation of the parameters of the model; hypothesis testing; forecasting or prediction; using the model for policy purposes.[12] Proceeding on the same lines, I will address my research question by analyzing the relation between the dependent and independent variables in my model in overall three time periods: pre-bubble period (1987:Q1 – 2001:Q2), bubble period (2001:Q3 – 2006:Q2), and the entire period (1987:Q1 – 2013:Q1). We chose 2001:Q3 – 2006:Q2 as the bubble period as one of the aims of this thesis is to analyze the extent to which factors such as monetary policy, global savings glut, financial innovations and other economic fundamentals can be held accountable for the overly appreciating home prices (a proxy measure of the bubble). My econometric analysis will comprise of:

- Looking at trends in the base series of dependent and independent variables
- Correlation analysis between dependent and independent variables

[12] For details see Gujarati (2004, chapter 1).

- Building a regression model to estimate to estimate each explanatory variables contribution to the housing bubble (first at an individual level and then using a combination of factors)
- Use cointegration methodology to check for a bubble in the housing market
    - The first step in establishing cointegration is to test the variables for stationarity. A time series is stationary if the series fluctuates with a constant variation about a mean, which remains constant over time. A stationary variable has a propensity to return to and frequently cross the mean. Since a stationary variable shows a tendency to revert to a constant mean it exhibits no trend. Thus, only variables which are non-stationary may be cointegrated. To establish nonstationarity we employ the ADF (Augmented Dickey Fuller) test. The second step employs Johansen-Juselius procedure (1990) to test for cointegration.
- If we find cointegration over the entire period, then we can apply an error correction model to estimate the acceleration speed of the short-run deviation to the long-run equilibrium or in other words estimate the magnitude of the error correction term. If no cointegration is found over the entire period then error correction model would not be appropriate for these non-stationary, non-cointegrated variables.
- For time periods where we find cointegration (pre-bubble / bubble / entire period) we will apply Granger causality test to understand the direction of the relationship

Simple Statistical Tools

We will begin our econometric analysis using simple statistical tools of trend analysis, correlation, regression and then move onto time series econometrics and test for cointegration, Granger causality and see if we can apply an error correction model.

Trend Analysis

The first step of econometric analysis in my research is to analyze the simple trend exhibited by the base time series data on all variables of interest using the graphical method and plotting observations over time.

Correlation Analysis

The second step of my econometric research is to run a simple correlation analysis among all the variables of interest. Correlation helps determine the strength or degree of linear association between two variables in terms of a number from -1.0 to +1.0. It is a measure of the statistical relationship, either positive or negative between two variables. It indicates the extent to which two or more variables fluctuate together. In correlation analysis there is no distinction between the dependent and explanatory variables so both variables are treated symmetrically. In general, if x and y have a strong positive linear correlation, as values for x increase, values for y also increase, value ranging from .00 to 1.00. If x and y have a strong negative linear correlation, as values for x increase, values for y decrease, value ranging from -1.00 to .00. If x and y have no correlation then there is a random or nonlinear relationship between the two variables. When we use correlation analysis, we aim only to understand the correlation between the two variables and not explain what the specific relationship is.[13] I test correlation between

[13] For details see Gujarati (2004, chapter 1).

the dependent and independent variables in overall three time periods: pre-bubble period, bubble period, and the entire period.

Regression Analysis

Regression analysis is a statistical tool to estimate the quantitative effect that causal variables/explanatory variables have upon the variable that they influence. It is a far superior estimation tool than correlation as correlation merely explains the association between variables, nevertheless, correlation analysis serves as a good starting point to understand how underlying variables interact with each other.

Regression helps estimate or predict the average value of one variable on the basis of the fixed values of other variables. It also helps ascertain the statistical significance of the estimated relationship, meaning what is the degree of confidence that the estimated relationship is close to the true relationship. I will first run a simple bivariate/ two variable regression in which the dependent variable/ regressand is related to a single explanatory variable/ regressor. After that I will use multiple regression analysis where the regressand is related to more than one regressor.[14] I will estimate the individual factor's contribution to the dependent variables only in the bubble phase. Consider a simple two variable time series regression model:

$$Y_t = \alpha + \beta X_t + \varepsilon_t \quad \text{(where t= 1,2,3,….,T)} \qquad (1)$$

Based on the above regression analysis we can decide to reject or to not reject the null hypothesis H_0: β = 0, based on a specific level of statistical significance. Mostly, we use a 5% level of significance; that is, we reject the null hypothesis that the estimated coefficient of the independent variable is zero and conclude that the dependent and the

[14] For details on bi-variate and multi-variate regression model see Gujarati (2004, chapters 2 & 3).

independent variables have a relationship with each other during the sample period tested with a 5% probability of error in our decision. A 5% significance level also means that our confidence level is 95% implying that, we are 95% confident that the estimated coefficient is close to the true (but unknown) parameter linking our x and y variables. SAS statistical procedures produce p-values (probability values) and if a p-value is less than or equal to 0.05, then we will reject the null hypothesis at a 5% level of significance, which means that we are 95% confident in our results that our x variable does influence our y variable. Similarly, a 2% significance level will correspond to a 98% confidence level, while a 10% significance level will correspond to a 90% confidence level.

A regression model also outputs a t-value or t-test statistic which also determines whether the x variable is statistically associated with the y variable. We have the same null hypothesis H_0: β = 0 meaning that X_t does not statistically explain the variation in Y_t. The rule of thumb is that if a t-value is greater than or equal to 2 in absolute value, we reject the null hypothesis meaning that X_t is statistically useful in explaining the variation in Y_t. SAS statistical procedures produce a p-value for a t-test, and if the p-value is less than or equal to 0.05 then we can say that the null is rejected at 5% level of significance.

In a multiple regression model, a useful measure to test the model's overall goodness of fit is the F-test. Consider a three-variable time series regression model:

$$Y_t = \alpha + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t \quad \text{(where t= 1,2,3,….,T)} \qquad (2)$$

The F-test jointly tests the null $H_0: \beta_1 = \beta_2 = 0$ examining whether the explanatory variables X and Z are statistically useful in explaining the variation in the dependent variable Y. SAS statistical procedures output a F-test value as well as a p-value for that F-test value. If the p-value is less than or equal to 0.05, then we reject the null hypothesis and can say that at 5% level of significance X and Z are statistically useful variables to explain movements in Y.

To analyze the fit of the regression model which includes an intercept/ constant term, we look at the R-squared (R^2) value of the model which indicates how much variation in the dependent variable is explained by the explanatory variable(s). For a two variable regression model a higher R-squared is considered better than a lower R-squared. However for a multiple-variable regression model, an adjusted R-squared would provide a better insight about a model's goodness of fit as it imposes a penalty each time an additional variable is added to the regression model compared to the simple R-squared.[15] SAS software produces both R-squared and adjusted R-squared values.

Time Series Econometrics

Stationarity

In time series analysis we study data collected over a period of time. A time series is a set of observations on values that a variable takes over time such as daily, monthly, quarterly or annually. Majority of the empirical work carried out on time series data assumes that the underlying time series is stationary. Thus the first step of time series analysis is to test whether a series is stationary or not, also known as testing for a unit root. Broadly speaking a stationary time series is one whose statistical properties remain constant over time. Specifically, this implies the mean and variance of the series are constant over time. A stationary series is mean reverting meaning that it fluctuates around a constant, long-run mean with a finite / constant variance which is not a function of time. When the series is far from the mean, it is more likely to move towards its mean rather than away from it. Thus a stationary variable exhibits no trend as it shows a tendency to

[15] For details on R-squared vs. adjusted R-squared see Gujarati (2004, chapter 7, p. 217-18).

revert to a constant mean. On the other hand, if the mean and/or the variance of a series are not constant over time, in which case the series is non-stationary, it can be characterized as containing a unit root.[16] A non-stationary series has no tendency to return to its long-run mean and the variance of the series is time dependent. We must note that if one or more time series are non-stationary, the OLS method cannot be employed as it makes the fundamental assumption of stationarity. All tests and conclusions based on t-values, F statistics, R-squared and forescasts will be not valid if the underlying data series is non-stationary. If the OLS method is applied to non-stationary data then the results will be spurious. Non-stationarity can lead to two time series appearing to be related when they are not, a problem called spurious regression.[17] So when we run a regression between two non-stationary time series variables we might get highly significant t-statistics, high adjusted R-squared values and coefficient signs consistent with economic theory, but the problem will be with the low Durbin-Watson d statistic value.[18] Thus the presence of these conditions together might signal spurious regression. For a non-stationary series, its behavior or change from one period to the next is random. Thus its future values are unpredictable. If a forecast is made assuming that the series is stationary when it is actually non-stationary, the result will be not only a misleading forecast, but also an inaccurate forecast interval.

Non-stationary behavior comes from two major sources: difference-stationary (DS) and trend stationarity (TS). As both these reasons for non-stationarity have different implications and solutions we must be able to identify whether the series follows DS or TS patterns. If a series follows the DS pattern, then the effect of any shock will be

---

[16] For a detailed discussion about the unit root concept see Maddala and Kim (1998).
[17] For details see Gujarati (2004, chapter 21).
[18] For details about the Durbin-Watson “d” test see Maddala and Lahiri (2009, chapter 6).

permanent. To convert this kind of a non-stationary process into a stationary one we need to generate the difference of the series. TS behavior means that the series has a deterministic trend (upward or downward) over time. In this case, we need to detrend the series by regressing the series on a time trend in order to remove the trend and make it stationary. There are several different ways to test for a unit root as described below.

The Dickey-Fuller Test

In 1979, Dickey and Fuller introduced the first standard process for unit root testing the DF test. Consider a time series y_t below following a simple autoregressive process of order one (AR (1)):

$$y_t = \rho y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim wn(0, \sigma^2) \qquad (3)$$

where ρ is the parameter to be estimated, and ε_t is the error term, assumed to be white noise, which means it has a zero mean, constant variance, and serially uncorrelated error terms. Our null hypothesis is H_o: $\rho = 1 \Rightarrow y_t$ has a unit root / is non-stationary. The alternative hypothesis is H_1: $|\rho| < 1 \Rightarrow y_t$ is stationary. On subtracting y_{t-1} from both sides of the above equation we get $\Delta y_t = \alpha y_{t-1} + \varepsilon_t$ where $\alpha = \rho - 1$. Our null hypothesis now becomes H_o: $\alpha = 0 \Rightarrow \rho = 1 \Rightarrow y_t$ has a unit root / is non-stationary. The alternative hypothesis now becomes H_1: $\alpha < 0 \Rightarrow |\rho| < 1 \Rightarrow y_t$ is stationary.

The test statistic is $t_{\rho=1} = \dfrac{\hat{\rho} - 1}{SE(\hat{\rho})}$, where $\hat{\rho}$ is the least square estimate and SE ($\hat{\rho}$) is the estimated standard error. Since y_t does not have a constant mean and variance, the t-statistic here does not follow the conventional Student's t-distribution as the t-distribution assumes a constant mean and variance. Dickey and Fuller derived the asymptotic distribution of $t_{\rho=1}$, called a Dickey-Fuller distribution. So we now decide whether to reject or fail to reject the null hypothesis of H_o: $\alpha = 0$ on the basis of the Dickey-Fuller statistic (tau statistic), and we compare this to the critical values from the DF table values provided by Fuller (1976). SAS statistical procedures automatically produce these table values along with the estimated tau values.

The Augmented Dickey-Fuller Test

The DF test for a unit root suffers from one critical issue. It is valid only if the series follows an AR(1) process. So if the series is correlated at a higher order lag and follows an AR(p) > AR(1) process, then the error terms are no longer white noise. So in 1981 Dickey and Fuller extended the DF test into the ADF (Augmented Dickey-Fuller) test by adding up to p lagged differenced terms of the dependent variable Δy_t on the right hand side of the Dickey Fuller test regression as below:

$$\Delta y_t = \gamma + \alpha y_{t-1} + \beta time + \sum_{j=1}^{p} \phi_j \Delta y_{t-j} + \varepsilon_t \qquad (4)$$

The standard ADF unit root test consists of three different cases with the same null hypothesis of non-stationarity and alternative hypothesis of stationarity: the random walk case, the random walk with drift case and the linear deterministic trend with a constant case. It is important to highlight that the ADF unit root test also has some limitations. It is accurate only about 31% of the time.[19] If the underlying time series has an ARMA process then the ADF results are not so accurate. Schwert (1989) found that if Δy_t has an ARMA representation with a large and negative MA component then the ADF test results reject non-stationarity too often even if it is true, so the test is severely size

[19] For details refer Stock and Watson (2007, p. 652).

distorted. For relatively short time series data ADF may not have enough power to detect the underlying long term change. Moreover, Maddala and Kim (1998) are of the opinion that due to the lack of power against the meaningful alternatives we should not use the DF and ADF tests.

DF-GLS / ERS Test

As there are limitations to the ADF test, Elliot, Rothenberg and Stock (1996) (hereafter ERS) proposed a simple modification of the ADF test called DF-GLS test in which the data are de-trended such that the explanatory variables are "taken out" of the data prior to running the test regression. ERS developed a feasible point optimal test, called P-test, which takes serial correlation of the error term into account and a DF-GLS test, which is an ADF-type test applied to the de-trended data without intercept. The ERS point optimal test calculates the quasi-difference of y_t as per the following equation:

$$d(y_t / a) = \begin{cases} y_t & if \quad t = 1 \\ y_t - a y_{t-1} & if \quad t > 1 \end{cases} \qquad (5)$$

where "a" represents the specific point alternative against which we wish to test the null. Then we run an OLS regression as per the equation: $d(y_t / a) = d(x_t / a)' \delta(a) + \varepsilon_t$. Now to get a value for "a" ERS recommend that a= $\bar{a}$ where $\bar{a} = \begin{cases} 1 - 7/T & if \quad x_t = (1) \\ 1 - 13.5/T & if \quad x_t = (1, t) \end{cases}$.

We now simply use Generalized Least Squares (GLS) with de-trended data, y_t^d, using the $\bar{a}$ estimates $y_t^d \equiv y_t - x_t' \hat{\delta}(\bar{a})$. On this GLS de-trended data DF-GLS test is performed with the null hypothesis that y_t is a random walk, possibly with drift. There are two possible alternative hypotheses: y_t is stationary about a linear time trend or y_t is stationary with a possibly nonzero mean but without a time trend. Elliott, Rothenberg, and Stock and later studies have shown that this test has significantly greater power than the previous versions of the augmented Dickey–Fuller test.[20]

Cointegration Analysis

In recent times, one of the most influential approaches in time series econometrics has been that of cointegration analysis. We have learned that a spurious or nonsensical relationship may result when one non-stationary time series is regressed against one or more other non-stationary time series. The best way to guard against spurious regressions is to check whether the variables used in our time series modeling are cointegrated. The purpose of the cointegration test is to determine whether groups of non-stationary series are cointegrated. So if the series are non-stationary in level form, we must apply cointegration tests on the variables of interest to obtain reliable econometric results. Cointegration analysis indicates whether there is existence of any long-run equilibrium and statistically significant relationship among the time series variables of interest. Cointegration is a property of two or more variables moving together through time; though these variables follow their own individual trends they will never drift too far apart as they are linked together in some sense.

In 1987 Engle and Granger highlighted that a linear combination of two or more non-stationary series may be stationary. If such a stationary linear combination exists, the non-stationary time series are said to be cointegrated. The stationary linear combination is called the cointegrating equation and it represents the long-run statistically significant relationship between the variables of interest. The first cointegration test was provided by Engle and Granger (E-G). This test calls for performing a unit root test on the residuals

[20] For details refer Stock and Watson (2007, p. 650–55).

obtained from regressing one nonstationary variable on another. If the residuals are stationary then there exists a co-integrating relationship. However, a major limitation to the E-G test is it can test for only one cointegrating relationship between variables of interest. So in the case of models with more than two variables, the E-G test cannot determine if there is more than one cointegrating relationship.

In 1990 Johansen and Juselius introduced a much more popular and widely used approach to test for cointegration. Consider a multi-variate model as below:

$$Y_t = \alpha + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t.$$

The E-G test will indicate whether these three variables are cointegrated or not, but will not tell how many cointegrating relationships exist. On the other hand, the Johansen approach will help test for more than one cointegrating relationship. The Johansen approach is a multiple equation approach, which is a vector autoregressive (VAR) based method. It provides two test statistics which are the trace test and the maximum eigenvalue test for cointegration. Of these two, Johansen and Juselius (1990) suggest that the maximum eigenvalue test may be better. SAS statistical procedures provide us with both the E-G test and the Johansen test, including the trace and maximum eigenvalue tests.

The Granger Causality Test

Though we have already carried out correlation and regression analyses, we now need to establish cause-and-effect relationships between our variables of interest. Correlation does not necessarily imply causation and regression analysis, which helps establish if there is existence of any relationship between variables, also does not prove causality or the direction of influence. So it is very important to distinguish between statistical correlation and statistical causality. While statistical correlation will tell us if two variables move together over time, statistical causality will indicate which variable leads and which variable lags. In 1969 Granger introduced the Granger causality test, a test for determining a causal relationship between two or more variables. The principle of Granger causality implies that if a variable X_t "Granger-causes” a variable Y_t , then past values of X_t should contain information that helps predict Y_t above and beyond the information contained in past values of Y_t alone (Granger 1969).

The Granger causality test is a statistical hypothesis test which helps determine if there is a statistical causal relation between the variables of interest, so can we use one time series to forecast another. Consider a simple bilateral causality regression equation:

$$\begin{aligned} X_t &= \alpha_1 + \sum_{i=1}^{p} \beta_i Y_{t-i} + \sum_{j=1}^{p} \gamma_j X_{t-j} + u_{1t} \\ Y_t &= \alpha_1 + \sum_{i=1}^{p} \lambda_i Y_{t-i} + \sum_{j=1}^{p} \delta_j X_{t-j} + u_{2t} \end{aligned} \qquad (6)$$

where u_{1t} and u_{2t} are assumed to be uncorrelated.

Test 1 Null hypothesis H_o : β_i = 0 (i=1,…..,p) i.e. Y does not Granger cause X. Test 2 Null hypothesis H_o : δ_j = 0 (j=1,…..,p) i.e. X does not Granger cause Y. As per the OLS coefficients generated for above equations, there are three possibilities:

1. A one-way Granger causality, unidirectional Granger causality ($\beta_i \neq 0$, $\delta_j = 0$ or $\beta_i = 0$, $\delta_j \neq 0$)
2. A two-way Granger causality, bidirectional causality ($\beta_i \neq 0$, $\delta_j \neq 0$)
3. No Granger causality, independence ($\beta_i = \delta_j = 0$)

The SAS statistical procedure provides us with Granger causality test results. The software also produces an F-test together with a p-value for the F-test. If the p-value is less than or equal to 0.05, then at 5% significance level, we can reject the null hypothesis of no Granger causality.

The Granger causality test also has certain limitations. It assumes that the underlying data series are stationary, it assumes that the error terms are normally distributed i.e. are white noise and lastly it does not tell us how many lags of X and Y variables to be used in the model. The lag order of variables is determined using the Schwartz Bayesian Criterion and that particular model's lag length (p order) is picked which has the lowest SBC value. We can also extend the Granger causality concept to multivariate causality by applying the technique of vector autoregression (VAR).

Error Correction Model

The error correction mechanism helps reconcile the static long-run equilibrium relationship of cointegrated time series with their dynamic short run disequilibrium. Once we have determined there is a long-run relationship between two or more variables via cointegration tests, we next need to see how much deviation is possible from a long-run equilibrium relationship in the short run.[21] The error correction model (ECM) identifies the degree of deviation and helps determine the short-run dynamics of the relationship. Thus, cointegration and ECM analysis will determine a statistical relationship between variables of interest, with cointegration revealing the long-run properties and the ECM showing the short-run properties of the relationship. Engle and Granger (1987) provided the ECM approach where the first step is to run an OLS regression by using the level form of the variables and saving the estimated residuals. In the next step, use the first

[21] For details on non-stationarity, cointegration and ECM refer to Maddala and Kim (1998).

difference form of all the variables and include the first lag of the estimated residuals from the first step as an independent variable in the model. Thus is called the Engle-Granger methodology. The reason for using the lag of the residual instead of the current period residual is that the estimated ε_t from the first step may correlate with current error term in the second step and if this happens then OLS results will be unreliable as one of the OLS assumptions is that the value of the independent variable should not be correlated with the error term.[22] By using the lag of the estimated ε_t we know that it may not correlate with the error term in the second step equation below as lag of the ε_t occurred before the error term v_t:

$$\Delta Y_t = c + \beta_3 \Delta X_t + \beta_4 \Delta Z_t + \beta_5 \hat{\varepsilon}_{t-1} + V_t \qquad (7)$$

where Δ represents the first difference form of the variable and the first lag of estimated ε_t is used in level form. Since all variables are stationary in above equation we can use OLS to obtain estimated coefficients. Clearly ECM is a single equation method and is related to the Engle Granger approach of co-integration. There is another method for analyzing the short-run properties of a relationship through a system of equations known as the vector error correction model (VECM). The VECM is related to the Johansen co-integration approach as both are multiple-equation methods. So the VECM model helps understand whether the deviation from long run path is statistically significant and what is the magnitude of this deviation? While checking for disequilibrium in the short run we can use the error term to tie the short-run behavior of the series to their long-run value.[23] The VECM(p) model is constructed based on information provided in the VAR(p) model from the cointegration step. The SAS system provides five different functional form

[22] For details see Gujarati (2004, chapter 3).
[23] For details see Gujarati (2004, chapter 21).

options for the VECM(p) model: Case 1 - No separate drift in the VECM(p) form; Case 2 - No separate drift in the VECM(p) form, but a constant enters only via the error correction term; Case 3 - Separate drift in the VECM but none in the cointegrating relation; Case 4 - Separate drift and no separate linear trend in the VECM(p) form; Case 5 - Separate linear trend in the VECM(p) form.[24]

[24] For more details about these cases see SAS/ ETS 9.2, VARMAX procedure.

CHAPTER 6: DISCUSSION OF RESULTS

Trend Analysis

At the very beginning it is important to understand the format in which the data on the variables is being used in our empirical analysis. Data on the following variables are being used in a year-over-year (yoy) percentage change format: housing starts, home prices, housing affordability index. Data on the following variables are being used in a rate/ percent format: effective fed funds rate, mortgage rate, unemployment rate, GDP growth rate. Data on the following variables are being used in a ratio format: current account as percentage of GDP, debt service ratio, ratio of financial innovation (S&P Case-Shiller home price index yoy / FHFA home price index yoy). We take the yoy percentage change form of some variables and some in percent form as we do not expect an instantaneous effect of the independent variables on the dependent variables but probably an effect that shows in a year or more. This kind of representation is very commonly used in empirical research and a classic example is that of the Phillips curve where the dependent variable inflation rate is denoted as the percentage change in the price level format while the independent variable unemployment rate is in a percent format.[25] Figure 28 shows graphs of all variables used in this analysis in their levels format from 1987-2013 Q1, except FHFA purchase only home price index, data on which are available only from 1991.

---

[25] For details see Gujarati (2004, chapter 6, p. 186-8).

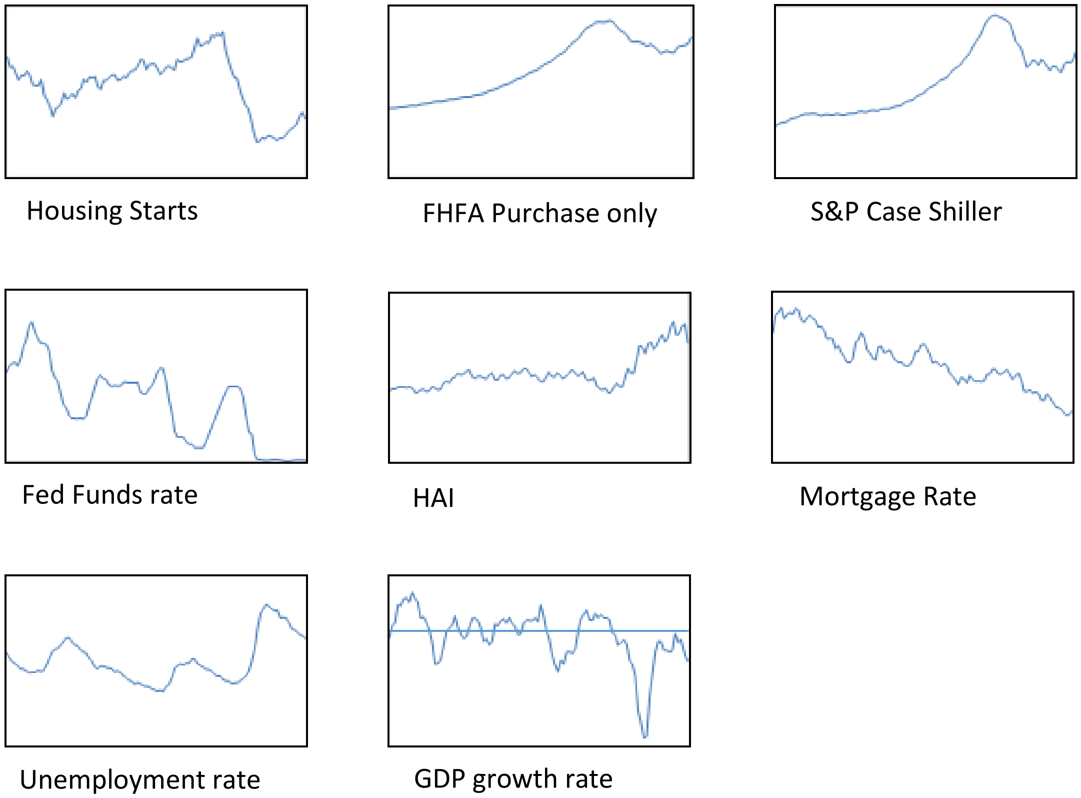
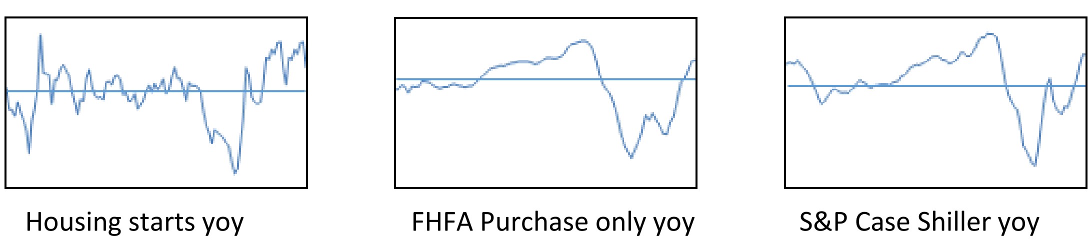
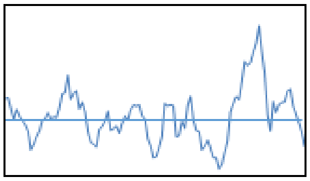


FIGURE 28: Graphs for levels data

—— Represents the mean of the series in the overall time period

Housing starts and home prices were steadily increasing from 1990 until they peaked in 2006. Most of the graphs above exhibit non-constant mean as well as variability over 1987-2013 Q1, suggesting that the time series are not stationary. The GDP growth rate is an exception and seems to be mean reverting and thus probably stationary. As housing starts and home prices data in levels seem non-stationary we convert them to a yoy growth format graphs for which look like below:



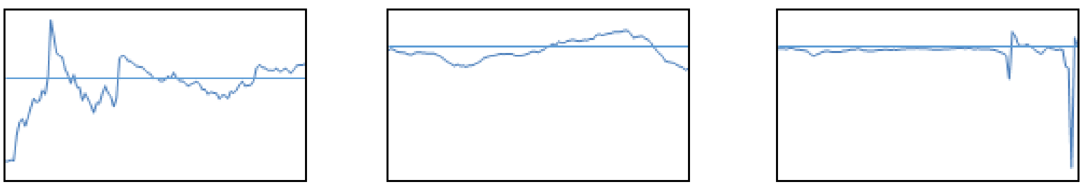


HAI yoy

FIGURE 29: Graphs for variables used in yoy growth rate format

Represents the mean of the series in the overall time period

Housing starts yoy seems to be stationary around single mean of -1.12. Home prices even in yoy growth format do not appear to be stationary. HAI yoy also seems to be stationary around single mean of 2.61.



CA/GDP ratio   Debt service ratio   Ratio CS/ FHFA

FIGURE 30: Graphs for variables used in ratio format

Represents the mean of the series in the overall time period

CA/GDP ratio graph seems to be mean reverting around -1.2. The debt service ratio over 26 years does not seem to be mean reverting and exhibits non-constant variability, so it is probably non-stationary. The ratio CS/FHFA seems to be mean reverting to zero and overall stationary.

Correlation Analysis

I run a simple correlation analysis among all the variables of interest in overall three time-periods. Correlation analysis based on the pre-bubble period produced the below results in Table 1:

TABLE 1: The correlation analysis (pre-bubble period)

Variables	1987:Q1 – 2001:Q2			
	Housing Starts	Case-Shiller HPI	FHFA-P HPI	Ratio (CS/FHFA)
Fed Funds	0.06	0.25***	-0.16	0.52*
CA_GDP	0.14	-0.31**	0.29***	-0.50*
Debt / Income	-0.19	0.30**	-0.36**	0.36*
HAI	0.27**	-0.38*	0.25	-0.44*
Mortgage Rate	0.23***	0.15	-0.01	0.28**
Unemployment Rate	-0.24***	-0.15	0.22	-0.46*
GDP Growth Rate	0.05	0.49*	-0.03	0.67*

* Significant at 1%   ** Significant at 5%   *** Significant at 10%

CA_GDP = Current Account balance as a percent of GDP.
Ratio (CS/FHFA) = Ratio of S&P/Case-Shiller(YoY) and FHFA(YoY).
HAI = Housing Affordability Index.
FHFA-P HPI = FHFA-Purchases only home price index

In the pre-bubble phase, as shown in results presented in Table 1, the fundamental variables exhibited a lowest correlation coefficient of -0.01 between the Mortgage rate and the FHFA-P HPI and a highest correlation coefficient of 0.67 between the GDP growth rate and the ratio (CS/FHFA) (both in absolute terms). In the pre-bubble period of all the independent variables three were found to have a statistically significant correlation with Housing Starts. Housing Starts (a proxy for the housing bubble) and HAI share 27% of their movements and a positive correlation coefficient sign indicates if HAI moves up then housing starts increase. Housing Starts and Unemployment Rate share 24% of their movements and a negative sign indicates if Unemployment Rate goes down then housing starts increase. Surprisingly, we get a positive correlation coefficient between mortgage rate and housing starts but it is a weak positive correlation. With Case Shiller HPI (another proxy for the housing bubble) almost all independent variables have a statistically significant correlation except the mortgage rate and the unemployment rate.

With FHFA-P HPI, CA_GDP and DSR have a statistically significant relation. All independent variables were found to have a statistically significant correlation with the financial innovation ratio.

TABLE 2: The correlation analysis (bubble period)

Variables	2001:Q3 – 2006:Q2			
	Housing Starts	Case-Shiller HPI	FHFA-P HPI	Ratio (CS/FHFA)
Fed Funds	-0.22	0.46**	-0.19	-0.17
CA_GDP	0.48**	-0.32	0.46**	0.42***
Debt / Income	-0.08	0.01	-0.03	-0.29
HAI	-0.09	-0.45**	0.18	0.35
Mortgage Rate	0.03	0.11	-0.20	-0.06
Unemployment Rate	-0.14	-0.41***	-0.04	-0.06
GDP Growth Rate	0.08	0.60*	-0.01	-0.01

* Significant at 1%   ** Significant at 5%   *** Significant at 10%

CA_GDP = Current Account balance as a percent of GDP.
Ratio (CS/FHFA) = Ratio of S&P/Case-Shiller(YoY) and FHFA(YoY).
HAI = Housing Affordability Index.
FHFA-P HPI = FHFA-Purchases only home price index

In the bubble phase, as shown in results presented in Table 2, the fundamental variables exhibited a lowest correlation coefficient of -0.01 between DSR and Case Shiller HPI, the GDP growth rate and FHFA-P HPI and the GDP growth rate and the financial innovation ratio, while we see a highest correlation coefficient of 0.60 between the GDP growth rate and the Case Shiller HPI (both in absolute terms). We clearly see that the number of statistically significant correlations between our variables of interest go down in the bubble phase as compared to the pre-bubble phase which can mean several things. It could mean that correlation between any two variables has probably gone down as other omitted variables have started impacting the dependent variable in the bubble phase and so perhaps a multiple regression analysis can help throw light on these reduced correlations between just two stand-alone variables. It could also mean that the relation between the sets of variables is no longer linear.[26] Nevertheless, we do see several statistically significant correlation coefficients. Also CA_GDP is found to have a statistically significant correlation coefficient with all proxy variables for the housing bubble except Case Shiller HPI.

TABLE 3: The correlation analysis (entire period)

Variables	1987:Q1 – 2013:Q1			
	Housing Starts	Case-Shiller HPI	FHFA-P HPI	Ratio (CS/FHFA)
Fed Funds	0.22**	0.41*	0.17	-0.04
CA_GDP	0.11	-0.24**	0.36*	-0.13
Debt / Income	-0.04	0.36*	-0.05	0.15
HAI	-0.39*	-0.80*	0.06	0.05
Mortgage Rate	0.16***	0.17***	-0.09	0.03
Unemployment Rate	-0.57*	-0.51*	0	0.11
GDP Growth Rate	0.38*	0.63*	-0.16	0.01

* Significant at 1% ** Significant at 5% *** Significant at 10%

CA_GDP = Current Account balance as a percent of GDP.
Ratio (CS/FHFA) = Ratio of S&P/Case-Shiller(YoY) and FHFA(YoY).
HAI = Housing Affordability Index.
FHFA-P HPI = FHFA-Purchases only home price index

In the overall period, as shown in results presented in Table 3, the fundamental variables exhibited a lowest correlation coefficient of 0.0 between the unemployment rate and FHFA-P HPI, while we see a highest correlation coefficient of 0.80 between the HAI and Case Shiller HPI (both in absolute terms). The number of statistically significant correlation coefficients again goes up in the overall period but what stands out is that in there is no statistically significant correlation coefficient with the ratio of financial innovation, probably implying that in the run-up to the bubble and in the bubble phase, variables were correlated with this ratio, but as the bubble burst and we saw a huge decline in MBS/CDS, the variables in the overall period are not correlated with this variable.

[26] For details see Gujarati (2004, chapter 3, p. 86-7). He explains how even a zero correlation coefficient between two variables does not imply that there is no relationship between the two variables, it could simply imply a non-linear relation.

The conclusion of this section is that the global saving glut (as measured by current account balance as a percent of GDP) was correlated with housing starts, FHFA-P home prices and non-conventional mortgages. Correlations existed between most of the variables under analysis in the run-up to the bubble and the overall time period but not so much in the bubble phase, probably implying that a combination of factors were more closely correlated with the proxy for bubble variables than any one individual factor.

Regression Analysis

For the regression analysis we specify our choice of independent and dependent variables and then run first bivariate and then multivariate regressions to see if there is any statistically significant relationship, the sign of the estimates and how much variation in the dependent variable is explained by the independent variable/s (R-squared of the regression). For the purpose of this research we discuss the regression results only for the bubble period (2001:Q4 - 2006:Q3 period). We exclude DSR (debt service ratio) from both individual and multiple regression analysis to avoid multicollinearity problem, as we believe DSR and HAI are very closely correlated which was leading to increased standard error of the DSR estimates.[27]

[27] We consider this to be a case of imperfect multicollinearity, where two or more regressors are highly correlated but not a perfect linear combination of the other. For details on imperfect multicollinearity refer Stock and Watson (2007, chapter 6).

TABLE 4: Bivariate regression analysis: individual factors (bubble period)

Variables	Time Period: 2001:Q3-2006:Q2	Dependant			
		Housing Starts	Case-Shiller HPI	FHFA-P HPI	Ratio (CS/FHFA)
Independent					
Fed Funds	Estimate	-3.72	3.41	-0.34	-0.11
	t-value	-0.95	2.2	-0.81	-0.73
	R-squared	0.05	0.21	0.04	0.03
CA_GDP	Estimate	17.76	-5.11	1.8	0.59
	t-value	2.32	-1.42	2.17	1.95
	R-squared	0.23	0.1	0.21	0.17
HAI	Estimate	0.12	-0.27	0.03	0.01
	t-value	0.36	-2.13	0.75	1.56
	R-squared	0	0.2	0.03	0.12
Mortgage Rate	Estimate	0.76	1.28	-0.57	-0.06
	t-value	0.12	0.47	-0.86	-0.26
	R-squared	0	0.01	0.04	0
GDP Growth Rate	Estimate	0.39	1.28	0	0
	t-value	0.34	3.2	-0.03	-0.03
	R-squared	0	0.36	0	0
Unemployment Rate	Estimate	-4.93	-6.29	-0.13	-0.08
	t-value	-0.59	-1.89	-0.15	-0.25
	R-squared	0.02	0.17	0	0

CA_GDP = Current Account balance as a percent of GDP.
Ratio (CS/FHFA) = Ratio of S&P/Case-Shiller(YoY) and FHFA(YoY).
HAI = Housing Affordability Index.
FHFA-P HPI = FHFA-Purchases only home price index

We first analyze each individual factor's contribution to the housing bubble. As shown in results presented in Table 4, the statistically significant relationships are highlighted in yellow. The fed funds is explaining 21% (R-squared=0.21) of variation in the Case Shiller HPI. CA_GDP (a proxy for the global saving glut) is responsible for 23% (R-squared=0.23) of the variation in the housing starts. Also, the positive sign of the estimate implies that the global saving glut did boost housing starts in the U.S. The relationship between the CA_GDP and FHFA purchase-only home price index is statistically significant with R-squared =0.21. The positive coefficient indicates that the global saving glut did contribute to the boost in home prices in the United States.

TABLE 5: Multivariate regression analysis: dependent variable – housing starts

Time Period: 2001:Q3-2006:Q2	Housing Starts						
	FF	C/G	MR	UR	HAI	GDPR	R^2
Estimate	13.37	42.91					0.38
t-value	2.06	3.04					
Estimate	17.07	55.15	10.6				0.51
t-value	2.72	3.83	2				
Estimate	7.59	45.84	9.27	-15.67			0.57
t-value	0.85	2.99	1.77	-1.44			
Estimate	6.03	47.26	7.24	-15.01	-0.29		0.58
t-value	0.65	3	1.2	-1.36	-0.72		
Estimate	7.71	56.67	4.62	-1.78	-0.18	3.18	0.69
t-value	0.91	3.82	0.83	-0.15	-0.5	2.09	

TABLE 6: Multivariate regression analysis: dependent variable – Case Shiller HPI

Time Period: 2001:Q3-2006:Q2	Case Shiller HPI						
	FF	C/G	MR	UR	HAI	GDPR	R^2
Estimate	5.47	5.18					0.24
t-value	1.75	0.76					
Estimate	5.97	6.83	1.43				0.25
t-value	1.78	0.89	0.5				
Estimate	6.14	6.99	1.45	0.28			0.25
t-value	1.2	0.8	0.48	0.04			
Estimate	5	8.03	-0.02	0.76	-0.21		0.29
t-value	0.95	0.9	-0.01	0.12	-0.93		
Estimate	6.1	14.17	-1.74	9.4	-0.14	2.08	0.53
t-value	1.36	1.8	-0.59	1.5	-0.73	2.56	

TABLE 7: Multivariate regression analysis: dependent variable – FHFA-P HPI

Time Period: 2001:Q3-2006:Q2	FHFA-P HPI						
	FF	C/G	MR	UR	HAI	GDPR	R^2
Estimate	1.5	4.61					0.38
t-value	2.16	3.07					
Estimate	1.6	4.95	0.29				0.39
t-value	2.15	2.9	0.46				
Estimate	1.35	4.7	0.26	-0.42			0.39
t-value	1.2	2.43	0.39	-0.3			
Estimate	1.22	4.81	0.1	-0.36	-0.02		0.4
t-value	1.03	2.4	0.12	-0.26	-0.45		
Estimate	1.41	5.89	-0.21	1.15	-0.01	0.37	0.52
t-value	1.28	3.03	-0.28	0.75	-0.22	1.83	

TABLE 8: Multivariate regression analysis: dependent variable – ratio (CS/FHFA)

Time Period: 2001:Q3-2006:Q2	Ratio (CS/FHFA)						
	FF	C/G	MR	UR	HAI	GDPR	R^2
Estimate	0.5	1.5					0.31
t-value	1.9	2.7					
Estimate	0.58	1.82	0.25				0.37
t-value	2.15	2.91	1.1				
Estimate	0.47	1.7	0.24	-0.19			0.37
t-value	1.14	2.41	0.98	-0.38			
Estimate	0.62	1.57	0.43	-0.26	0.03		0.47
t-value	1.53	2.31	1.66	-0.53	1.59		
Estimate	0.68	1.9	0.34	0.21	0.03	0.11	0.55
t-value	1.76	2.8	1.34	0.38	1.89	1.6	

where

FF = Fed Funds
C/G = Current Account balance as a percent of GDP
MR = Mortgage Rate
UR = Unemployment Rate
HAI = Housing Affordability Index
GDPR = GDP Growth Rate

We now run several multivariate regression models testing the role of a combination of these potential factors to the bubble. We want to analyze whether a combined effect of the potential factors on the bubble is statistically significant. In Table 5, we test whether the above shown possible combinations of the explanatory variables are statistically significant in explaining the variation in housing starts which serves as the dependent variable. In Tables 6, 7, and 8 we carry out the same process except that the dependent variable keeps changing to Case Shiller HPI, FHFA-P HPI and the financial innovation ratio (CS/FHFA) respectively. In Table 5, we find evidence that monetary policy and the global saving glut have a statistically significant relationship with housing starts. Also, the first two explanatory variables combined with mortgage rates have a statistically significant relationship with housing starts. In Table 6, we find that there is no combine effect of any of these explanatory factors on Case Shiller HPI. In Table 7, our statistical analysis suggests that monetary policy and the global saving glut have a combined effect on FHFA-P HPI. Table 8 provides evidence that none of the combinations of the potential contributors to the housing bubble has a statistically significant relationship with the financial innovation ratio.

Summing up, regression analysis concludes that during the bubble period monetary policy is associated with Case Shiller HPI—a proxy for the housing bubble. The global savings glut (measured by current account as a percent of GDP) is associated with housing starts and FHFA-P HPI—both proxies for the housing bubble. HAI and the GDP growth rate are also associated with Case Shiller HPI. So while a number of these factors individually contributed to the housing bubble, at a combined level only monetary policy, the global savings glut, and mortgage rates contributed to housing starts, and monetary policy and the global savings glut contributed to FHFA-P home prices.

It is worth mentioning that in developing the multivariate regression empirical models there might have been a model mis-specification error as we are only hypothesizing a possible housing bubble model and our understanding of what this model looks like, it is quite possible that there is a more appropriate true model and so we have model specification errors such as omission of a relevant variable(s), inclusion of an unnecessary variable(s), or adopting the wrong functional form.[28] The result of this under- fitting of a model is biased estimates and thus a biased estimation of the strength of the relationship. Thus, we will use conclusions from correlation and regression analysis in conjunction with time series econometrics conclusions to derive final results.

Time Series Econometrics
Stationarity/ Augmented Dickey-Fuller Test[29]

For each variable we applied the ADF test on the overall time period from 1987Q1-2013Q1.[30] Within the ADF test, the first case we consider is the "Zero-Mean" case which implies that the dependent variable is regressed on its own lag and an error term (no intercept and trend in the test regression). This is also known as a random walk model. The second case we consider is that of "Single Mean" which includes a constant /drift parameter in the test regression. It is also known as a random walk with drift model. The third case we consider is the "Trend" case which includes a constant and a linear trend in the test regression.[31] In each case the null hypothesis is that the time series has a unit root (not mean-reversion) and the alternative hypothesis is that the time series is stationary (mean-reversion). For the "Zero Mean" case we can decide to reject/fail to reject the null hypothesis based on the "Pr < Tau" value; if it is smaller than 0.05 then at 5 percent level of significance we can reject the null hypothesis and conclude that the data series is stationary. The null hypothesis of the F-test is that the coefficients of the underlying model are simultaneously zero and the alternative hypothesis is that the

[28] For details see Gujarati (2004, chapter 13).
[29] It is a SAS 9.2 limitation that we cannot perform the DF-GLS / ERS test; however, for the benefit of the reader the SAS code is provided in appendix B.
[30] Performed in SAS using PROC ARIMA, see appendix B for details.
[31] For SAS output refer to Appendix A: Results.
For details refer Stock and Watson (2007, chapter 14).

coefficients are not zero. For the "Single Mean" and "Trend" case we look at both "Pr < Tau" and "Pr > F" probabilities and conclude accordingly.[32] The ADF test results based on the form of variables used in the analysis are reported in Table 9.

TABLE 9: Results: unit root testing, data in levels (entire period)

Variables	1987:Q1 – 2013:Q1		
	Mean	S.D.	Unit Root
Housing Starts	-1.3	16.74	No
Case-Shiller HPI	3.31	7.24	No
FHFA-P HPI	3.20	4.84	Yes
Ratio (CS/FHFA)	0.42	5.83	No
Fed Funds	3.96	2.63	Yes
CA_GDP	-1.2	0.67	No
Debt / Income	0.12	0.01	Yes
HAI	2.62	7.60	No
Mortgage Rate	7.15	1.89	Yes
Unemployment Rate	6.05	1.53	Yes
GDP Growth Rate	4.97	2.13	No

S.D = Standard Deviation
Yes = Unit Root, Non-Stationary
No = No Unit Root, Stationary

Thus variables found to be non-stationary/ containing a unit root at lag order 1 at 5% significance level were: FHFA-purchases only home price index (year/year percent change), the federal funds rate (percent), the debt service ratio, the mortgage rate and the unemployment rate. To make these time series stationary we calculate their first difference and apply the same ADF Unit root test on the first difference form of the series and get results reported in Table 10:

[32] In this analysis we use the 10 percent significance level; therefore, any probability value less than 0.10 identifies a significant relationship, implying that the series is stationary. To identify stationarity we also examine the ACF and PACF (partial auto-correlation plot) for the variables. If ACF decays slowly it means that the process has a long memory.

TABLE 10: Unit root testing, first difference of data (entire period)

Variables	1987:Q1 – 2013:Q1 Mean	S.D.	Unit Root
FHFA-P HPI	0.06	1.07	No
Fed Funds	-0.06	0.47	No
Debt / Income	0.00	0.00	No
Mortgage Rate	-0.05	0.36	No
Unemployment Rate	0.01	0.29	No

S.D = Standard Deviation
Yes = Unit Root, Non-Stationary
No = No Unit Root, Stationary

Thus results now indicate that the data are stationary. For sub-samples of the entire period the unit root results are not expected to be robust and thus we only conduct the unit root tests for the overall sample period. Su Zhou (2001, p.914-5) discusses how models with zero lag order ADF tests have relatively large size distortion unless the time span is 60 years or longer and so the power of these tests depends on time span more than on mere number of observations. So we test for stationarity only in the overall sample period.

Cointegration Analysis

We carry out the bivariate and multivariate cointegration tests using the Johansen trace and maximum eigenvalue tests. It is worth mentioning that although Johansen's methodology is used in a setting with all I(1) variables in the system, having I(0) or stationary variables is theoretically not an issue, and Johansen (1995) states that there is little need to pre-test the variables in the system to establish their orders of integration. If time series are integrated of different orders over a sample period then we will not find cointegration between them. Keeping this in mind and the fact that unit root test results over smaller sub-sample periods are not robust we apply the Johansen methodology to check for cointegration for pre-bubble, bubble and overall time period. We will first check for bivariate cointegrating relationships between housing bubble proxy variables and monetary policy and the global savings glut. Knowing that there exists a possibility that these individual variables might not be cointegrated with housing bubble proxy variables, we will combine the individual variables into a multivariate model and then test for possible cointegration between their linear combinations.

We shall be considering only those cointegrating relationships for the VECM model that are supported by Granger causality. We also know that for m variables there can be no more than m-1 cointegrating relationships. The rank order of cointegration and the lag length of the VAR model determined from this step will be used in the VECM step.

TABLE 11: Lag length selection: pre-bubble, bubble and entire period

PRE BUBBLE: 1987:Q1 – 2001:Q2

CI	No. of lags	SBC	AIC	AR
Housing Starts vs. Fed Funds				
Y	1	7.4	7.1	N
N	2	7.1	6.7	N
N	3	7.3	6.8	N
Housing Starts vs. CA_GDP				
N	1	6.6	6.4	N
N	2	6.8	6.5	N
N	3	7.0	6.5	N
Housing Starts vs. Fed Funds, CA_GDP & Others				
Y	1	0.4	-1.6	N
Y	2	2.4	-1.4	N
N	3	4.3	-1.2	Y

CI	No. of lags	SBC	AIC	AR
Case Shiller HPI vs. Fed Funds				
N	1	-1.4	-1.6	Y
N	2	-2.0	-2.4	Y
Y	3	-1.9	-2.4	N
Case Shiller HPI vs. CA_GDP				
N	1	-2.1	-2.3	Y
N	2	-2.2	-2.6	Y
N	3	-2.1	-2.6	N
Case Shiller HPI vs. Fed Funds, CA_GDP & Others				
N	1	-8.0	-10.0	Y
N	2	-6.2	-10.0	Y
N	3	-5.2	-10.8	Y

CI	No. of lags	SBC	AIC	AR
FHFA-P HPI vs. Fed Funds				
N	1	-2.1	-2.3	Y
N	2	-2.6	-3.0	Y
N	3	-3.1	-3.7	N
FHFA-P HPI vs. CA_GDP				
N	1	-3.4	-3.7	Y
N	2	-3.2	-3.6	Y
N	3	-3.8	-4.4	N
FHFA-P HPI vs. Fed Funds, CA_GDP & Others				
N	1	-10.2	-12.5	Y
N	2	-8.1	-12.5	Y
Y	3	-7.6	-14.1	N

CI	No. of lags	SBC	AIC	AR
Ratio (CS/FHFA) vs. Fed Funds				
N	1	-3.6	-3.9	Y
N	2	-4.0	-4.4	N
N	3	-3.9	-4.4	N
Ratio (CS/FHFA) vs. CA_GDP				
N	1	-4.3	-4.5	Y
N	2	-4.2	-4.5	N
N	3	-3.9	-4.4	N
Ratio (CS/FHFA) vs. Fed Funds, CA_GDP & Others				
N	1	-10.3	-12.4	Y
N	2	-8.3	-12.2	N
N	3	-6.2	-12.0	N

CI = Cointegration; Y = Yes, N = No
SBC = Schwarz-Bayes information criterion
AIC = Akaike's information criterion
AR = Autocorrelation in residuals
Others = Mortgage Rate, Unemployment Rate, Housing Affordability Index, GDP Growth rate

BUBBLE: 2001:Q3 – 2006:Q2

CI	No. of lags	SBC	AIC	AR
Housing Starts vs. Fed Funds				
N	1	6.1	5.8	N
N	2	6.3	5.8	N
N	3	6.3	5.6	N
Housing Starts vs. CA_GDP				
N	1	4.6	4.3	N
N	2	4.8	4.4	N
N	3	5.1	4.4	N
Housing Starts vs. Fed Funds, CA_GDP & Others				
Y	1	-7.1	-9.9	N

CI	No. of lags	SBC	AIC	AR
Case Shiller HPI vs. Fed Funds				
N	1	-2.4	-2.7	Y
N	2	-2.2	-2.7	N
N	3	-2.5	-3.2	N
Case Shiller HPI vs. CA_GDP				
N	1	-3.3	-3.6	Y
N	2	-3.5	-4.0	N
N	3	-2.9	-3.6	N
Case Shiller HPI vs. Fed Funds, CA_GDP & Others				
N	1	-15.7	-18.5	Y

CI	No. of lags	SBC	AIC	AR
FHFA-P HPI vs. Fed Funds				
N	1	-2.3	-2.6	Y
N	2	-1.7	-2.2	Y
N	3	-2.5	-3.2	N
FHFA-P HPI vs. CA_GDP				
N	1	-3.9	-4.2	N
N	2	-3.3	-3.8	N
N	3	-2.8	-3.5	N
FHFA-P HPI vs. Fed Funds, CA_GDP & Others				
N	1	-14.9	-17.7	Y

CI	No. of lags	SBC	AIC	AR
Ratio (CS/FHFA) vs. Fed Funds				
N	1	-5.9	-6.2	Y
N	2	-6.4	-6.9	Y
N	3	-6.2	-6.9	Y
Ratio (CS/FHFA) vs. CA_GDP				
N	1	-8.4	-8.7	Y
N	2	-7.7	-8.2	Y
N	3	-7.2	-7.9	N
Ratio (CS/FHFA) vs. Fed Funds, CA_GDP & Others				
N	1	-20.9	-23.6	Y

CI = Cointegration; Y = Yes, N = No
SBC = Schwarz-Bayes information criterion
AIC = Akaike's information criterion
AR = Autocorrelation in residuals
Others = Mortgage Rate, Unemployment Rate, Housing Affordability Index, GDP Growth rate

TABLE 11: (continued)

TABLE 11: (continued)

Entire Time-period: 1987:Q1 – 2013:Q1

CI	No. of lags	SBC	AIC	AR
Housing Starts vs. Fed Funds				
N	1	7.5	7.3	Y
N	2	7.0	6.8	N
N	3	7.2	6.8	N
Housing Starts vs. CA_GDP				
N	1	6.2	6.1	Y
N	2	6.3	6.1	N
N	3	6.4	6.1	N
Housing Starts vs. Fed Funds, CA_GDP & Others				
N	1	1.1	-0.3	N
N	2	1.7	-0.9	N
N	3	3.0	-1.0	N

CI	No. of lags	SBC	AIC	AR
Case Shiller HPI vs. Fed Funds				
N	1	1.0	0.9	Y
N	2	0.2	-0.1	Y
N	3	0.4	0.0	Y
Case Shiller HPI vs. CA_GDP				
N	1	-0.1	-0.2	Y
N	2	-0.3	-0.6	Y
N	3	-0.1	-0.5	Y
Case Shiller HPI vs. Fed Funds, CA_GDP & Others				
N	1	-5.2	-6.6	Y
N	2	-4.6	-7.3	Y
N	3	-3.5	-7.4	Y

CI	No. of lags	SBC	AIC	AR
FHFA-P HPI vs. Fed Funds				
N	1	0.6	0.4	Y
N	2	-0.4	-0.7	Y
N	3	-0.2	-0.6	Y
FHFA-P HPI vs. CA_GDP				
N	1	-1.1	-1.3	Y
N	2	-1.4	-1.7	Y
N	3	-1.2	-1.6	Y
FHFA-P HPI vs. Fed Funds, CA_GDP & Others				
N	1	-6.5	-8.1	Y
N	2	-5.5	-8.5	Y
N	3	-4.1	-8.5	Y

CI	No. of lags	SBC	AIC	AR
Ratio (CS/FHFA) vs. Fed Funds				
N	1	2.3	2.2	Y
N	2	1.7	1.5	N
N	3	1.9	1.5	N
Ratio (CS/FHFA) vs. CA_GDP				
N	1	1.0	0.9	Y
N	2	1.1	0.9	N
N	3	1.3	0.9	N
Ratio (CS/FHFA) vs. Fed Funds, CA_GDP & Others				
N	1	-3.8	-5.2	N
N	2	-2.7	-5.5	N
N	3	-1.2	-5.3	N

CI = Cointegration; Y = Yes, N = No
SBC = Schwarz-Bayes information criterion
AIC = Akaike's information criterion
AR = Autocorrelation in residuals
Others = Mortgage Rate, Unemployment Rate, Housing Affordability Index, GDP Growth rate

The tables above show how we select the lag order to test for cointegration between each variable that stands as a proxy for housing bubble and all other variables that possibly contributed to the bubble. The lag length for each VAR model is chosen based on minimum SBC values and the one which makes the estimated residuals stationary.[33] To correct for autocorrelation in residuals we keep adding lags in the model until the AR disturbances in the residuals cannot reject the null hypothesis that the residuals are uncorrelated. SAS presents this output as an F test for AR disturbances.[34] For consistency we test for upto three lags for each of the models in our system. The actual results from the Johansen trace and maximum eigenvalue tests where cointegrating

[33] In cases where there was a conflict between minimum lag length selection and lag order at which variables are cointegrated, for this study priority was given to lag length at which cointegration occurs so that we can study the relation between variables.
[34] Refer to appendix A for actual results.

relations were found are presented in Table 12. Toda (1994) compares the small-sample properties of the trace and maximum eigenvalue using Monte Carlo methods and concludes that though neither test is uniformly superior, the trace test does perform better in certain situations where the power is low.[35]

TABLE 12: Johansen cointegration approach[36] (pre-bubble period)

1987:Q1 – 2001:Q2 Cointegration between	Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Eigenvalue	Maximum Statistic	0.05 Critical Value
Housing Starts vs. Fed Funds	$H_0: r = 0$ $H_1: r > 0$	0.38 0.01	28.32 0.37	15.34 3.84	0.20 0.11	12.59 6.51	14.07 3.76
Case Shiller HPI vs. Fed Funds	$H_0: r = 0$ $H_1: r > 0$	0.33 0.05	25.76 2.96	15.34 3.84	0.38 0.01	27.94 0.37	14.07 3.76
Housing Starts vs. Fed Funds, CA_GDP & Others	$H_0: r = 0$ $H_1: r > 0$	0.76	190.92	123.04	0.76	82.30	45.28
	$H_0: r = 1$ $H_1: r > 1$	0.59	108.62	93.92	0.59	51.96	39.37
	$H_0: r = 2$ $H_1: r > 2$	0.32	56.66	68.68	0.32	22.75	33.46
FHFA-P HPI vs. Fed Funds, CA_GDP & Others	$H_0: r = 0$ $H_1: r > 0$	0.92	267.50	123.04	0.90	94.79	45.28
	$H_0: r = 1$ $H_1: r > 1$	0.74	167.47	93.92	0.65	44.63	39.37
	$H_0: r = 2$ $H_1: r > 2$	0.66	113.72	68.68	0.46	25.79	33.46
	$H_0: r = 3$ $H_1: r > 3$	0.54	70.35	47.21	0.38	19.89	27.07
	$H_0: r = 4$ $H_1: r > 4$	0.45	38.95	29.38	0.31	15.47	20.97
	$H_0: r = 5$ $H_1: r > 5$	0.22	14.80	15.34	0.14	6.13	14.07

As shown in Table 12 for the trace/maximum eigenvalue test, the first row tests r = 0 against r > 0; the second row tests r = 1 against r > 1 and so on. At each row we compare the trace/maximum eigenvalue test statistic with the critical value at 5% significance level, and if found greater we reject the null in favor of the alternative. The results presented in Table 12 indicate that in the pre-bubble period we can reject the null of no cointegration at 5% level of significance for four particular models and thus find cointegration between Housing Starts vs. Fed Funds, Case Shiller HPI vs. Fed Funds and multivariate models of Housing Starts/ FHFA-P HPI vs. Fed Funds, CA_GDP, Mortgage Rate, Unemployment Rate, Housing Affordability Index and GDP growth rate. Thus for the period preceding the run up in home prices there appears to have been a strong link between these variables. Also, monetary policy is found to have a close link with two

[35] In accordance with Toda for the purpose of this study , if conflict arises between the trace and maximum eigenvalue tests we give priority to the trace test results.
[36] For trace/ maximum eigenvalue test SAS output where cointegration holds, refer to appendix A.

housing bubble proxy variables, while the global savings glut only has an impact via the multivariate model and not at an individual level. It must also be noted that as no cointegration exists with the financial innovation ratio it could imply two things: eiher the role of non-conventional mortgages and related structured finance securities was dominant only in the bubble phase or the variables that were influencing this ratio have not been accounted for comprehensively in our models.

TABLE 13: Johansen cointegration approach (bubble period)

2001:Q3 – 2006:Q2 Cointegration between	Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Eigenvalue	Maximum Statistic	0.05 Critical Value
Housing Starts vs. Fed Funds, CA_GDP & Others	H_0 : r = 0 H_1 : r > 0	0.98	188.98	123.04	0.98	70.38	45.28
	H_0 : r = 1 H_1 : r > 1	0.90	118.60	93.92	0.90	43.46	39.37
	H_0 : r = 2 H_1 : r > 2	0.80	75.14	68.68	0.80	30.52	33.46

The results presented in Table 13 indicate that from 2001:Q3 – 2006:Q2 the null of no cointegration can be rejected at the 5% significance level for only one of the model cases of Housing Starts vs. Fed Funds, CA_GDP, Mortgage Rate, Unemployment Rate, Housing Affordability Index and GDP growth rate.

What is extremely insightful and what will drive the major interpretations from the cointegration test is the fact that in the overall time period from 1987-2013 as illustrated in Table 11, we do not find any cointegrating relationships between our various model cases. Thus we see that in the bubble period the cointegrating relationships that existed in the pre-bubble period have suddenly become unstable; in fact there is a drastic severing of these relations in the bubble and post bubble period. This finding is consistent with the presence of a bubble which is why the variables are no longer cointegrated.

The Granger Causality Test

We now use Granger causality test to identify statistical causal relationships between our variables wherever cointegration was found in the prior analysis, identifying if our a priori estimation of leading and lagging variables is correct.[37] We apply this test on the stationary form of all our variables, so variables found to be non-stationary by the ADF test are first differenced to make them stationary to avoid spurious causality, and then the Granger causality test is applied. Lag-length for each model is chosen based on lowest SBC/AIC value.[38]

TABLE 14: The Granger Causality Wald test

Pre-Bubble Period: 1987:Q1 – 2001:Q2		
Group 1 Variables	**Group 2 Variables**	**Pr > ChiSq**
Housing Starts	Fed Funds	0.06***
Housing Starts	Fed Funds, CA_GDP, HAI, Mortgage Rate, GDP Rate, Unemployment Rate	0.002*
Case-Shiller HPI	Fed Funds	0.08***
FHFA-P HPI	Fed Funds, CA_GDP, HAI, Mortgage Rate, GDP Rate, Unemployment Rate	0.001*
Bubble Period: 2001:Q3 – 2006:Q2		
Group 1 Variables	**Group 2 Variables**	**Pr > ChiSq**
Housing Starts	Fed Funds, CA_GDP, HAI, Mortgage Rate, GDP Rate, Unemployment Rate	0.004*

* Significant at 1%
** Significant at 5%
*** Significant at 10%

In the pre-bubble and bubble periods we test for Granger causality for models in which cointegration was found. In Table 14, we find that the federal funds rate Granger causes housing starts and the Case Shiller HPI. It is also interesting to note how the multivariate model shown in the table above Granger causes both housing starts and

37 Performed in SAS using PROC VARMAX, see appendix B for details.
38 We select the same lag order for Granger causality test as used in Johansen cointegration tests.

FHFA-P HPI. In the bubble period, as shown in Table 14, only one multivariate VAR model of group2 variables as shown Granger causes housing starts. In the overall sample period in which no cointegrating relationships were found, we do not test for Granger causality. Also because no cointegration was found between the financial innovation ratio and monetary policy, the global savings glut and other variables used in our models, we can conclude that it is very hard to analyze and quantify measures for the financial innovation ratio that is essentially capturing the role of securitization of mortgages.

In conclusion, a combination of monetary policy and the global savings glut were Granger causing the housing bubble in the period of rising home prices in the pre-bubble period and the housing bubble period. Other variables including HAI, the mortgage rate, the GDP growth rate and the unemployment rate were added to improve the fit of the model. This confirms our a priori hypothesis about some of the potential contributors to the housing bubble.

Error Correction Model

Once we have used cointegration to test for long-run relationships among variables we need to check whether these series deviate from the long run-path. Based on the findings of this study in our full sample period, an error-correction model is not appropriate. The detailed quantitative analysis carried out in this thesis found no cointegrating relationships (and thus no convergence to a stable equilibrium relationship) between the variables used in the bivariate and multivariate models in the overall sample period. The results of this study are also confirmed by several other academic studies done in this field.

Gallin (2006) for instance uses panel-data tests for cointegration in a panel of 95 metro areas over 23 years and could not reject the hypothesis of no cointegration, confirming that error correction models relating house prices to fundamental economic variables are likely to be inappropriate. Another study by Zhou and Sornette (2006) presents evidence supporting a bubble. They analyze the northeast, midwest, south, and west regions quarterly average sale prices of new houses sold in the U.S. and conclude that for 22 states, mostly in the northeast and the west, there was a clear indication of a fast growing bubble. Theoretically as well our conclusion about the housing bubble and thus not applying a VECM model is in line with the Granger Representation Theorem (Engle and Granger 1987 and Johansen 1991) which can be summarized as stating that, in a model with two or more non-stationary variables, cointegration and error correction are statistically equivalent so if one exists then the other must also exist. In the absence of cointegration if an error correction model is applied then the results will be spurious. Thus the fact that some of the variables in our model were cointegrated with housing prices before 2001, but do not remain cointegrated in the overall time period clearly points out that the U.S. housing market experienced a bubble after 2000. Thus, an error correction model is not appropriate to derive conclusive results.[39]

[39] Though eventually this thesis could not employ a VECM, for the benefit of readers the SAS code for a VECM model is provided in appendix B.

CHAPTER 7: SUGGESTIONS FOR FUTURE RESEARCH / LIMITATIONS

This study has employed nationwide aggregate time series data for understanding the housing bubble and which factors played a major contributing role to the bubble. However, a limitation of time series analysis is the low power for smaller sample periods. For future studies, researchers should attempt to use panel data and cover longer study duration of above 25 years and use more indicator variables. Due to data and time limitations this study does not include every potential, or even plausible, variable that had influenced the 2007-08 financial crisis/ housing bubble. To fully understand the social, economic and political conditions surrounding the financial crisis/ housing bubble time period we certainly need more comprehensive models to avoid another such episode of financial fiasco. Some additional economic variables which might make the model more comprehensive include Adjustable Mortgage Rate/ 12-month LIBOR – Adjustable rate mortgages are often pegged to the LIBOR rate, so this ratio can serve as an indicator of the effect of the LIBOR rate on the adjustable rate and to an extent help gauge the riskiness of mortgage loans and thus serve as a proxy for the role played by securitization of sub-prime mortgages. In addition including and analyzing data on credit scores and mortgage applications (subject to availability of such data) could further shed light on this. Some other variables that can be included are real median family income (an indicator of wealth available to purchase a house), homeowner rate (an indicator of the strength of economy and local housing market), real house price/ rent (an indicator of the attractiveness of home-ownership versus renting), and the real Lehman mortgage-backed securities index (an indicator of the level and strength of the mortgage bond market).

This study uses national housing indexes to analyze the housing bubble at a national level. It will probably be very interesting to develop the same model using local housing indices and see if there were a number of local bubbles for the variables under consideration using state home prices data. Since there are housing indexes for local markets, this extension should prove feasible. It would also be interesting to analyze if the results found on a national level in the pre-bubble and bubble period in this thesis hold true for local housing market bubbles.

For the multiple regression analysis used in this thesis there is a possibility that problems are created by multicolinearity effects. To overcome this, we recommend using structural equation modeling (SEM) which is particularly useful in resolving problems created by multicolinearity effects. Furthermore, SEM would help deal with issues of confirmatory versus exploratory analysis for verification of the bubble effect adding weight to cointegration/ Granger causality results.[40]

Lastly, it is hard to empirically quantify the role played by some of the factors held responsible for the financial crisis including Basel II, role of Govt. sponsored entities and structured finance, and so that was not undertaken in this research paper. However, it might be a very interesting extension of this thesis to come up with suitable proxy variables for these factors and analyze their contribution to the financial crisis.

[40] Hoyle (1995).

CHAPTER 8: CONCLUSION

The purpose of this thesis is multifold. This thesis is an attempt to get a high level understanding of the factors that caused the 2007-08 financial fiasco and specifically seeks to analyze the role of monetary policy, the global savings glut and financial innovations in the crisis. This paper provides information on some of the existing academic literature on causes of the financial crisis and housing bubble and also presents research on data and events related to the financial crisis. The literature included how a combination of factors, including a bubble in the housing market, the Fed's monetary policy, how securitization concentrated risk in the financial sector, housing initiatives and other policies followed by Fannie Mae and Freddie Mac, Basel rules and banks' leverage and credit default swaps, were held responsible for the financial fiasco.

This thesis presented a detailed econometric analysis of the relation between housing bubble proxy variables and a number of causal factors related to the housing bubble/ financial crisis via trend analysis, correlation, bivariate regression and multivariate regression, followed by time series statistical analysis including unit root testing, cointegration, Granger causality and the possibility of an error correction model. We divided our data into three sub-periods: the pre-bubble period (1987:Q1 – 2001:Q2), the bubble period (2001:Q3 – 2006:Q2), and the entire period (1987:Q1 – 2013:Q1). The statistical evidence presented in this thesis supports the presence of a real estate bubble in the housing market and how monetary policy and the global savings glut caused the housing bubble at an individual level in the pre-bubble period and at a cumulative level with other economic variables in the bubble period. The variables making up the model in this thesis behaved in more complex ways in the bubble period when compared to the pre-bubble period. Eventually no evidence of cointegration between proxy variables of the housing bubble and possible factors causing it was found in the overall time period.

There are several studies which attempt to understand the 2007-08 financial crisis, so the academic literature on this subject is not new. However, this thesis adds to the existing literature in several ways. First, the time period that this thesis focusses on includes the start of the rise in U.S. house prices up until its peak and a considerable period after the peak and bust until the most recent 2013 time period. Second, this study provides a detailed theoretical understanding of potential explanatory factors for the housing bubble which was held responsible for the 2007-08 financial crisis. Third, this thesis further supports the evidence provided by other studies on the existence of a housing bubble. Fourth, the econometric methodology employed in this thesis can be used to test for housing bubbles in countries across the globe where similar fundamental situations might exist. Lastly, the technique employed in this research can help to verify bubbles in other asset markets with their corresponding set of variables.

REFERENCES

Acharya, Viral V., and Matthew Richardson, eds. 2009. "Causes Of The Financial Crisis." New York University Stern School of Business, John Wiley & Sons. Retrieved from:http://pages.stern.nyu.edu/~sternfin/vacharya/public_html /acharya_richardson_critical.pdf

Angell, Cynthia, and Norman Williams. 2005. "US Home Prices: Does Bust Follow Boom?" Washington DC: Federal Deposit Insurance Corporation. May 2, 2005. Retrieved from: http://www.fdic.gov/bank/analytical/fyi/2005/050205fyi .html

Ashcraft, A. and Schuermann T. 2008. "Understanding the Securitization of Subprime Mortgage Credit." Federal Reserve Bank of New York Staff Reports, no. 318, March 2008, p. 10-11

BIS Committee on the Global Financial System. 2008. "Ratings in structured finance: what went wrong and what can be done to address shortcomings?" Bank for International Settlements. July 2008.

Bernanke, Ben. 2010. "Monetary Policy and the Housing Bubble." American Economic Association meeting, Atlanta, GA. January 3, 2010. Retrieved from: http://www.federalreserve.gov/newsevents/speech/bernanke20100103a.htm

Bernanke, Ben. 2007. "Global Imbalances: Recent Developments and Prospects." at the Bundesbank Lecture, Berlin, Germany. September 11, 2007. Retrieved from: http://www.federalreserve.gov/newsevents/speech/bernanke20070911a.htm

Bernanke, Ben. 2005. "The Global Saving Glut and the U.S. Current Account Deficit." Sandridge Lecture, Virginia Association of Economists, April 14, 2005, Richmond, Virginia. Retrieved from: http://www.federalreserve.gov /boarddocs / speeches/2005/200503102/

Bernardos, Gonzalo. 2009. "The rise and fall of the first global real estate bubble." June, 2009. Retrieved from: http://www.raco.cat/index.php/Paradigmes/article/ download/223810/304675

Bluhm, Christian. 2003. "CDO Modeling: Techniques, Examples and Applications." HypoVerinsbank. December 10, 2003

Campbell, John and Robert Shiller. 1987. "Cointegration and Tests of Present Value Models." Journal of Political Economy. October 1987. p. 1062-88.

Cannata F., Quagliariello M. 2009. “The role of Basel II in the subprime financial crisis: guilty or not guilty?” January 2009. Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1330417

Case, K.E., & Shiller, R.J. 2003. “Is there a bubble in the housing market?” (with comments and discussion). Brookings Paperson Economic Activity, 2, p. 299–342. Retrieved from: http://www.econ.yale.edu/~shiller/pubs/p1089.pdf

Clark,S.P., & Coggin,T.D. 2011. “Was there a U.S. house price bubble? An econometric analysis using national and regional panel data.” The Quarterly Review of Economics and Finance. Retrieved from: www.horizoninvestments.com/mydocuments/clark_coggin_qref.pdf

Cordell, Larry Huang, Yilin Williams, Meredith. 2012. “Sizing And Assessing The Subprime CDO Crisis.” May 2012. p. 2. Retrieved from: http://www.philadelphiafed.org/research-and-data/publications/working-papers/2011/wp11-30.pdf

Criado, S. and Adrian Van Rixtel. 2008. “Structured Finance and the Financial Turmoil of 2007-2008: An Introductory Overview.” Banco De Espana, Documentos Ocasionales, No 0808. Retrieved from: http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosOcasionales/10/Fic/do0808e.pdf

Diamond, Douglas W. 1984. “Financial Intermediation and Delegated Monitoring.” Review of Economic Studies. July, 1984. 51:p. 393-414.

Dickey, David and Wayne Fuller. 1981. “Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root.” Econometrica, July1981. p. 1057-72.

Dickey, D., and Fuller, W. 1979. “Distribution of the Estimators for Autoregressive Time Series with a Unit Root.” Journal of American Statistical Association, Vol. 74, p. 427-4311.

Economist House Price Index data, Global House Prices. 2013. Retrieved from: http://www.economist.com/blogs/dailychart/2011/11/global-house-prices

Peter J. Wallison. 2009. “Critical Review Economics, Fiscal Policy.” Retrieved from: http://www.aei.org/article/economics/fiscal-policy/cause-and-effect/

Engle, Robert and B. Yoo. 1987. “Forecasting and Testing in Co-integrated Systems.” Journal of Econometrics, p. 143-59.

Engle, Robert and Granger, Clive. 1987. “Co-integration and Error Correction: Representation, estimation and Testing.” Econometrica, Vol 55, issue 2.

Elliott, G., Rothenberg, T. & Stock, J. 1996. “Efficient Tests for an Autoregressive Unit Root.” Econometrica 64, 813-836.

Federal Deposit Insurance Corporation. 2009. “The 2009 Economic Landscape; The Sand States: Anatomy of a Perfect Housing-Market Storm.” Retrieved from: http://www.fdic.gov/bank/analytical/quarterly/2009_vol3_1/AnatomyPerfectHousing.html

Federal Housing Finance Agency. 2013. About FHFA. Retrieved from: http://www.fhfa.gov/Default.aspx?Page=81

Federal Housing Finance Agency Annual Report to Congress. 2008. Retrieved from: http://www.fhfa.gov/Default.aspx?Page=240

Federal Reserve Bank of St. Louis. 2009. The Financial Crisis: A Timeline of Events and Policy Actions. Retrieved from: <http://timeline.stlouisfed.org/index.cfm?p=timeline>

Federal Reserve Bank of St. Louis. 2013a. Housing Starts: Total: New Privately Owned Housing Units Started . Retrieved from: http://research.stlouisfed.org/fred2/series/HOUST/downloaddata?cid=32302

Federal Reserve Bank of St. Louis. 2013b. Housing Affordability Index (Composite). Retrieved from: http://research.stlouisfed.org/fred2/series/COMPHAI/downloaddata

Federal Reserve Bank of St. Louis. 2013c. 30-Year Conventional Mortgage Rate. Retrieved from: http://research.stlouisfed.org/fred2/series/MORTG/downloaddata?cid=114

Federal Reserve Bank of St. Louis. 2013d. Gross Domestic Product Statistics. Retrieved from: http://research.stlouisfed.org/fred2/series/GDP/downloaddata

Federal Reserve Board. 2013. Household Debt Service and Financial Obligations Ratios. Retrieved from: http://www.federalreserve.gov/releases/housedebt/default.htm

Financial Crisis Inquiry Commission. 2010. “Government Sponsored Enterprises And The Financial Crisis.” Retrieved from: http://fcic-static.law.stanford.edu/cdn_media/fcic-reports/2010-0409-GSEs.pdf

Friedman, Jeffrey (2010a). What Caused the Financial Crisis: Housing Initiatives and Other Policy Factors. Chapter 6. Philadelphia: University of Pennsylvania Press. p. 172-82

Friedman, Jeffrey. 2010b. What Caused the Financial Crisis: A regulated Meltdown, The Basel rules and banks' leverage. Chapter 8. Philadelphia: University of Pennsylvania Press. p. 200-27

Gallin, J. 2006. "The long-run relationship between house prices and income." Real Estate Economics, 34, 417–438. Retrieved from: http://www.federalreserve.gov/pubs/

Goodhart C., Persaud A. 2008. "How to avoid the next crash?" Financial Times, January 30, 2008. Retrieved from: http://www.ft.com/intl/cms/s/0/fed495d4-cf4b-11dc-854a-0000779fd2ac.html#axzz2jDreqACQ

Gibson, Michael. 2007. "Credit Derivatives and Risk Management." Finance and Economics Discussion Series, Federal Reserve Board, May 22, 2007. Retrieved from: http://www.federalreserve.gov/pubs/feds/2007/200747/200747pap.pdf

Granger, Clive and Engle, Robert. 1987. "Cointergration and the Error-Correction Representation: Estimation and Testing." Econometrica, p. 251-76

Granger, C.W., and Newbold, P. 1974. ''Spurious Regressions in Econometrics.'' Journal of Econometrics, p. 111–20

Greenspan, Alan. 2010. "The Crisis." March 2010. Retrieved from: http://www.brookings.edu/~/media/Files/Programs/ES/BPEA/2010_spring_bpea_papers/spring2010_greenspan.pdf

Greene, William H. 2007. Econometric Analysis. Prentice Hall, New Jersey, USA.

Gujarati, D.N. 2004. Basic Econometrics. Fourth Edition, The McGraw−Hill Irwin, NY, USA

Household Debt Service and Financial Obligations Ratios 2013. Federal Reserve Board. Retrieved from: http://www.federalreserve.gov/releases/housedebt/default.htm

Hoyle, R. H. 1995. "The structural equation modeling approach: Basic concepts and fundamental issues." Thousand Oaks, CA: Sage Publications, Inc., p. 1-15

International Monetary Fund. 2008a. "Structured finance: Issues of valuation and disclosure." Retrieved from: http://www.imf.org/External/Pubs/FT/GFSR/2008/01/pdf/chap2.pdf

International Monetary Fund. 2008b. "Global Financial Stability Report." Retrieved from: http://www.imf.org/External/Pubs/FT/GFSR/2008/01/pdf/text.pdf

Johansen, S.1991. “Estimation and hypothesis testing on cointegrating vectors in Gaussian vector autoregressive models.” Econometrica, 59, 1551–1580. Retrieved from: http://www.jstor.org/discover/10.2307/2938278?uid=3739808&uid=2&uid=4&uid=3739256&sid=21102786284397

Johansen, S. 1995. “Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.” Oxford, UK: Oxford University Press

Kohn, Jonathan and Sarah K. Bryant. 2010. “An Econometric Interpretation of the Recent US Housing Boom.” Research in Business and Economics Journal. Retrieved from: http://www.aabri.com/manuscripts/09381.pdf

Maddala, G.S. and Lahiri, Kajal. 2009. Introduction to Econometrics. Wiley, 4th Edition, New Jersey,USA.

Maddala, G.S., and Kim, In-Moo. 1998. “Unit Roots, Cointegration, and Structural Change.” Cambridge, UK: Cambridge University Press

Mengle, David. 2007. “Credit Derivatives: An Overview.” Paper presented at Financial Markets Conference, Federal Reserve Bank of Atlanta. Retrieved from: http://www.frbatlanta.org/news/conferen/07fmc/07FMC_mengle.pdf

Mikhed,V., & Zemcik, P. 2009. “Testing for bubbles in housing markets: A panel data approach.” Journal of Real Estate Finance and Economics, 38, 366–386. Retrieved from: http://www.cerge-ei.cz/pdf/wp/Wp338.pdf

Norberg, Johan. 2009a. Financial Fiasco, How America’s infatuation with home ownership and easy money created the economic crisis. Chapter 2. Washington D.C.: Cato Institute.

Norberg, Johan. 2009b. Financial Fiasco, How America’s infatuation with home ownership and easy money created the economic crisis. Chapter 3. Washington D.C.: Cato Institute. p. 45-55

Onado, M. 2008. “Capital is a problem.” Retrieved from: http://www.lavoce.info

Phillips, Peter and Pierre Perron. 1988. “Testing for a Unit Root in Time Series Regression.” Biometrika, June, 1988. p. 335-46

SAS/ETS(R) 9.2 User's Guide (2013). Vector Error Correction Modeling. Retrieved from:http://support.sas.com/documentation/cdl/en/etsug/63348/HTML/default/viewer.htm#etsug_varmax_sect035.htm

Schwert, W. 1989. “Test for Unit Roots: A Monte Carlo Investigation.” Journal of Business and economic Statistics, 7, p. 147-59. Retrieved from: http://schwert.ssb.rochester.edu/jbes89.pdf

Shiller, R. 2008. The Subprime Solution, How Today's Global Financial Crisis Happened, and What to Do about It. Princeton: Princeton University Press. Retrieved from: http://www.afi.es/eo/The%20Subprime%20Solution%20-%20Shiller,%202008.pdf

Shiller, Robert J. 2007. "Understanding Recent Trends in House Prices and Homeownership." Kansas City: Federal Reserve Bank of Kansas City, p.89- 123

Securities Industry and Financial Markets Association (SIFMA). 2007. U.S. Research Quarterly. Retrieved from: http://www.sifma.org/research/item.aspx?id=21578

Stock, James H. and Watson Mark W. 2007. Introduction to Econometrics. Second Edition, Pearson Addison-Wesley Series, Boston Massachusetts

Stulz R. 2009. "Credit Default Swaps And The Credit Crisis?" September 2009. Retrieved from: http://fisher.osu.edu/fin/faculty/stulz/publishedpapers/jep%2024%201.pdf

S&P/Case-Shiller. 2013. U.S. National Home Price Index Statistics data. Retrieved from: http://us.spindices.com/indices/real-estate/sp-case-shiller-us-national-home-price-index

Taylor, John. 2010. "The Fed and the Crisis: A Reply to Ben Bernanke." The Wall Street Journal, January 10, 2010

Taylor, John. 2008. "The Financial Crisis and the Policy Responses: An Empirical Analysis of What Went Wrong." A Festschrift in Honor of David Dodge, November 2008. Bank of Canada. Retrieved from: http://www.stanford.edu/~johntayl/FCPR.pdf

Taylor, John. 2007. "Housing and Monetary Policy." NBER Working Paper Series 13682. Cambridge, Mass.: National Bureau of Economic Research, December, 2007. Retrieved from: www.nber.org/papers/w13682.pdf

Taylor, John. 1999a. "An Historical Analysis of Monetary Policy Rules." in John B. Taylor, ed., Monetary Policy Rules. Chicago: University of Chicago Press

Taylor, John B., ed. 1999b. "Monetary Policy Rules." Chicago: University of Chicago Press

The Financial Crisis 2013, A Timeline of Events and Policy Actions, Federal Reserve Bank of St. Louis website. Retrieved from:http://timeline.stlouisfed.org/pdf/CrisisTimeline.pdf

Toda, H.Y. 1995. “Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions.” Retrieved from: http://www.jstor.org/discover/10.2307/2109827?uid=3739808&uid=2&uid=4&uid=3739256&sid=21102777937417

United States Department of Labor. 2013. Bureau of Labor Statistics. Labor Force Statistics. Retrieved from: http://data.bls.gov/pdq/SurveyOutputServlet

Using the Metzler Diagram to explain Ben Bernanke’s Savings Glut Speech. 2005. Retrieved from: http://econ-server.umd.edu/~mendoza/econ442/files/Metzler%20Diagram%20and%20the%20savings%20glut.pdf

Wallison J. Peter. 2009. “Cause and Effect: Government Policies and the Financial Crisis.” Retrieved from: http://www.aei.org/article/economics/fiscal-policy/cause-and-effect/

Zhou, Su. 2001. “The Power of Cointegration Tests Versus Data Frequency and Time Spans.” Southern Economic Journal, 67(4), p. 906-921

Zhou, W.X., & Sornette D. 2006. “Is there a real-estate bubble in the USA?” Physica A, 361, 297–308. Retrieved from: http://arxiv.org/pdf/physics/0506027v1.pdf

APPENDIX A: RESULTS

CORRELATION ANALYSIS – PRE BUBBLE PERIOD

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	CAGDP_Q	UHS_Q_yoy	HPICSCUS_ Q_yoy	HAI_Q	GDP_Q	Ratio_ CSFHFA
CAGDP_Q	1.00000	0.14053	-0.30611	0.05595	-0.46270	-0.50349
CAGDP#Q		0.2884	0.0230	0.6738	0.0002	<.0001
	59	59	55	59	59	55
UHS_Q_yoy	0.14053	1.00000	0.00834	0.27320	0.05460	0.24997
	0.2884		0.9518	0.0363	0.6813	0.0657
	59	59	55	59	59	55
HPICSCUS_Q_yoy	-0.30611	0.00834	1.00000	-0.38442	0.49077	0.84008
	0.0230	0.9518		0.0038	0.0001	<.0001
	55	55	55	55	55	55
HAI_Q	0.05595	0.27320	-0.38442	1.00000	-0.37903	-0.43567
HAI_Q	0.6738	0.0363	0.0038		0.0031	0.0009
	59	59	55	59	59	55
GDP_Q	-0.46270	0.05460	0.49077	-0.37903	1.00000	0.67014
GDP#Q	0.0002	0.6813	0.0001	0.0031		<.0001
	59	59	55	59	59	55
Ratio_CSFHFA	-0.50349	0.24997	0.84008	-0.43567	0.67014	1.00000
	<.0001	0.0657	<.0001	0.0009	<.0001	
	55	55	55	55	55	55
FFTQ_Qd1	-0.45546	0.05857	0.24513	-0.45151	0.62603	0.52120
	0.0003	0.6595	0.0713	0.0003	<.0001	<.0001
	59	59	55	59	59	55

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	FFTQ_Qd1	XPOIUS_ Q_yoyd1	DSR_Qd1	MTG_Qd1	URT_Qd1
CAGDP_Q	-0.45546	0.28632	-0.03801	-0.23155	0.44688
CAGDP#Q	0.0003	0.0814	0.7750	0.0776	0.0004
	59	38	59	59	59
UHS_Q_yoy	0.05857	0.41987	-0.19148	0.22674	-0.24239
	0.6595	0.0087	0.1463	0.0842	0.0644
	59	38	59	59	59
HPICSCUS_Q_yoy	0.24513	0.15435	0.30262	0.15454	-0.15185
	0.0713	0.3548	0.0247	0.2599	0.2684
	55	38	55	55	55
HAI_Q	-0.45151	0.24543	-0.44779	-0.19983	0.13556
HAI_Q	0.0003	0.1375	0.0004	0.1292	0.3060
	59	38	59	59	59
GDP_Q	0.62603	-0.03260	0.08692	0.26413	-0.49037
GDP#Q	<.0001	0.8459	0.5127	0.0432	<.0001
	59	38	59	59	59
Ratio_CSFHFA	0.52120	-0.02615	0.36093	0.27556	-0.46429
	<.0001	0.8762	0.0068	0.0417	0.0004
	55	38	55	55	55
FFTQ_Qd1	1.00000	-0.15949	0.21786	0.43249	-0.63839
		0.3388	0.0974	0.0006	<.0001
	59	38	59	59	59

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	CAGDP_Q	UHS_Q_yoy	HPICSCUS_ Q_yoy	HAI_Q	GDP_Q	Ratio_ CSFHFA
XPOIUS_Q_yoyd1	0.28632	0.41987	0.15435	0.24543	-0.03260	-0.02615
	0.0814	0.0087	0.3548	0.1375	0.8459	0.8762
	38	38	38	38	38	38
DSR_Qd1	-0.03801	-0.19148	0.30262	-0.44779	0.08692	0.36093
	0.7750	0.1463	0.0247	0.0004	0.5127	0.0068
	59	59	55	59	59	55
MTG_Qd1	-0.23155	0.22674	0.15454	-0.19983	0.26413	0.27556
	0.0776	0.0842	0.2599	0.1292	0.0432	0.0417
	59	59	55	59	59	55
URT_Qd1	0.44688	-0.24239	-0.15185	0.13556	-0.49037	-0.46429
	0.0004	0.0644	0.2684	0.3060	<.0001	0.0004
	59	59	55	59	59	55

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	FFTQ_Qd1	XPOIUS_ Q_yoyd1	DSR_Qd1	MTG_Qd1	URT_Qd1
XPOIUS_Q_yoyd1	-0.15949 0.3388 38	1.00000 38	-0.35510 0.0287 38	-0.01065 0.9494 38	0.21718 0.1903 38
DSR_Qd1	0.21786 0.0974 59	-0.35510 0.0287 38	1.00000 59	0.03137 0.8135 59	-0.12358 0.3511 59
MTG_Qd1	0.43249 0.0006 59	-0.01065 0.9494 38	0.03137 0.8135 59	1.00000 59	-0.38537 0.0026 59
URT_Qd1	-0.63839 <.0001 59	0.21718 0.1903 38	-0.12358 0.3511 59	-0.38537 0.0026 59	1.00000 59

CORRELATION ANALYSIS – BUBBLE PERIOD

The CORR Procedure

Pearson Correlation Coefficients, N = 20
Prob > |r| under H0: Rho=0

	CAGDP_Q	UHS_Q_yoy	HPICSCUS_ Q_yoy	HAI_Q	GDP_Q	Ratio_ CSFHFA
CAGDP_Q CAGDP#Q	1.00000	0.47907 0.0326	-0.31727 0.1729	0.72625 0.0003	-0.72489 0.0003	0.41686 0.0675
UHS_Q_yoy	0.47907 0.0326	1.00000	0.50062 0.0246	0.08501 0.7216	0.08084 0.7347	0.81269 <.0001
HPICSCUS_Q_yoy	-0.31727 0.1729	0.50062 0.0246	1.00000	-0.44888 0.0471	0.60162 0.0050	0.49253 0.0274
HAI_Q HAI_Q	0.72625 0.0003	0.08501 0.7216	-0.44888 0.0471	1.00000	-0.69943 0.0006	0.34514 0.1361
GDP_Q GDP#Q	-0.72489 0.0003	0.08084 0.7347	0.60162 0.0050	-0.69943 0.0006	1.00000	-0.00708 0.9764
Ratio_CSFHFA	0.41686 0.0675	0.81269 <.0001	0.49253 0.0274	0.34514 0.1361	-0.00708 0.9764	1.00000
FFTQ_Qd1	-0.86555 <.0001	-0.21807 0.3557	0.45977 0.0414	-0.71326 0.0004	0.75137 0.0001	-0.17052 0.4723
XPOIUS_Q_yoyd1	0.45610 0.0433	0.81318 <.0001	0.62171 0.0034	0.17504 0.4604	-0.00600 0.9800	0.83910 <.0001
DSR_Qd1	0.08268 0.7289	-0.08198 0.7311	0.00913 0.9695	0.05850 0.8065	-0.08220 0.7305	-0.29496 0.2068

The CORR Procedure

Pearson Correlation Coefficients, N = 20
Prob > |r| under H0: Rho=0

	FFTQ_Qd1	XPOIUS_ Q_yoyd1	DSR_Qd1	MTG_Qd1	URT_Qd1
CAGDP_Q CAGDP#Q	-0.86555 <.0001	0.45610 0.0433	0.08268 0.7289	-0.36299 0.1157	0.56272 0.0098
UHS_Q_yoy	-0.21807 0.3557	0.81318 <.0001	-0.08198 0.7311	0.02828 0.9058	-0.13829 0.5609
HPICSCUS_Q_yoy	0.45977 0.0414	0.62171 0.0034	0.00913 0.9695	0.10933 0.6464	-0.40619 0.0755
HAI_Q HAI_Q	-0.71326 0.0004	0.17504 0.4604	0.05850 0.8065	-0.50278 0.0239	0.56929 0.0088
GDP_Q GDP#Q	0.75137 0.0001	-0.00600 0.9800	-0.08220 0.7305	0.39902 0.0814	-0.76114 <.0001
Ratio_CSFHFA	-0.17052 0.4723	0.83910 <.0001	-0.29496 0.2068	-0.06205 0.7950	-0.05958 0.8030
FFTQ_Qd1	1.00000	-0.18781 0.4278	-0.19352 0.4136	0.17560 0.4590	-0.78843 <.0001
XPOIUS_Q_yoyd1	-0.18781 0.4278	1.00000	-0.02877 0.9042	-0.19931 0.3995	-0.03527 0.8826
DSR_Qd1	-0.19352 0.4136	-0.02877 0.9042	1.00000	-0.30223 0.1953	0.20865 0.3773

The CORR Procedure

Pearson Correlation Coefficients, N = 20
Prob > |r| under H0: Rho=0

	CAGDP_Q	UHS_Q_yoy	HPICSCUS_ Q_yoy	HAI_Q	GDP_Q	Ratio_ CSFHFA
MTG_Qd1	-0.36299 0.1157	0.02828 0.9058	0.10933 0.6464	-0.50278 0.0239	0.39902 0.0814	-0.06205 0.7950
URT_Qd1	0.56272 0.0098	-0.13829 0.5609	-0.40619 0.0755	0.56929 0.0088	-0.76114 <.0001	-0.05958 0.8030

Pearson Correlation Coefficients, N = 20
Prob > |r| under H0: Rho=0

	FFTQ_Qd1	XPOIUS_ Q_yoyd1	DSR_Qd1	MTG_Qd1	URT_Qd1
MTG_Qd1	0.17560 0.4590	-0.19931 0.3995	-0.30223 0.1953	1.00000	-0.12613 0.5962
URT_Qd1	-0.78843 <.0001	-0.03527 0.8826	0.20865 0.3773	-0.12613 0.5962	1.00000

CORRELATION ANALYSIS – ENTIRE PERIOD

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	CAGDP_Q	UHS_Q_yoy	HPICSCUS_ Q_yoy	HAI_Q	GDP_Q	Ratio_ CSFHFA
CAGDP_Q CAGDP#Q	1.00000 106	0.10878 0.2670 106	-0.24374 0.0136 102	0.17273 0.0766 106	-0.42041 <.0001 106	-0.12920 0.1956 102
UHS_Q_yoy	0.10878 0.2670 106	1.00000 106	0.56860 <.0001 102	-0.39424 <.0001 106	0.38299 <.0001 106	-0.22911 0.0205 102
HPICSCUS_Q_yoy	-0.24374 0.0136 102	0.56860 <.0001 102	1.00000 102	-0.79960 <.0001 102	0.62580 <.0001 102	-0.06502 0.5162 102
HAI_Q HAI_Q	0.17273 0.0766 106	-0.39424 <.0001 106	-0.79960 <.0001 102	1.00000 106	-0.67789 <.0001 106	0.05201 0.6037 102
GDP_Q GDP#Q	-0.42041 <.0001 106	0.38299 <.0001 106	0.62580 <.0001 102	-0.67789 <.0001 106	1.00000 106	0.01396 0.8893 102
Ratio_CSFHFA	-0.12920 0.1956 102	-0.22911 0.0205 102	-0.06502 0.5162 102	0.05201 0.6037 102	0.01396 0.8893 102	1.00000 102
FFTQ_Qd1	-0.38049 <.0001 106	0.21877 0.0243 106	0.40430 <.0001 102	-0.50562 <.0001 106	0.46780 <.0001 106	-0.04211 0.6743 102

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	FFTQ_Qd1	XPOIUS_ Q_yoyd1	DSR_Qd1	MTG_Qd1	URT_Qd1
CAGDP_Q CAGDP#Q	-0.38049 <.0001 106	0.36353 0.0006 85	-0.15536 0.1135 105	-0.23032 0.0175 106	0.28857 0.0027 106
UHS_Q_yoy	0.21877 0.0243 106	0.42686 <.0001 85	-0.03666 0.7104 105	0.16424 0.0925 106	-0.57061 <.0001 106
HPICSCUS_Q_yoy	0.40430 <.0001 102	0.17846 0.1022 85	0.36492 0.0002 101	0.16825 0.0910 102	-0.50165 <.0001 102
HAI_Q HAI_Q	-0.50562 <.0001 106	0.06133 0.5771 85	-0.37757 <.0001 105	-0.24369 0.0118 106	0.52844 <.0001 106
GDP_Q GDP#Q	0.46780 <.0001 106	-0.15696 0.1514 85	0.24720 0.0110 105	0.22099 0.0228 106	-0.66655 <.0001 106
Ratio_CSFHFA	-0.04211 0.6743 102	-0.18654 0.0874 85	0.14907 0.1368 101	0.03012 0.7638 102	0.10826 0.2788 102
FFTQ_Qd1	1.00000 106	0.16855 0.1231 85	0.15608 0.1118 105	0.36484 0.0001 106	-0.53751 <.0001 106

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	CAGDP_Q	UHS_Q_yoy	HPICSCUS_ Q_yoy	HAI_Q	GDP_Q	Ratio_ CSFHFA
XPOIUS_Q_yoyd1	0.36353 0.0006 85	0.42686 <.0001 85	0.17846 0.1022 85	0.06133 0.5771 85	-0.15696 0.1514 85	-0.18654 0.0874 85
DSR_Qd1	-0.15536 0.1135 105	-0.03666 0.7104 105	0.36492 0.0002 101	-0.37757 <.0001 105	0.24720 0.0110 105	0.14907 0.1368 101
MTG_Qd1	-0.23032 0.0175 106	0.16424 0.0925 106	0.16825 0.0910 102	-0.24369 0.0118 106	0.22099 0.0228 106	0.03012 0.7638 102
URT_Qd1	0.28857 0.0027 106	-0.57061 <.0001 106	-0.50165 <.0001 102	0.52844 <.0001 106	-0.66655 <.0001 106	0.10826 0.2788 102

The CORR Procedure

Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	FFTQ_Qd1	XPOIUS_ Q_yoyd1	DSR_Qd1	MTG_Qd1	URT_Qd1
XPOIUS_Q_yoyd1	0.16855 0.1231 85	1.00000 85	-0.04886 0.6589 84	-0.09340 0.3952 85	0.00254 0.9816 85
DSR_Qd1	0.15608 0.1118 105	-0.04886 0.6589 84	1.00000 105	-0.01601 0.8712 105	0.01787 0.8564 105
MTG_Qd1	0.36484 0.0001 106	-0.09340 0.3952 85	-0.01601 0.8712 105	1.00000 106	-0.29727 0.0020 106
URT_Qd1	-0.53751 <.0001 106	0.00254 0.9816 85	0.01787 0.8564 105	-0.29727 0.0020 106	1.00000 106

REGRESSION ANALYSIS – BUBBLE PERIOD

The AUTOREG Procedure

Dependent Variable UHS_Q_yoy

Ordinary Least Squares Estimates

SSE	1133.8882	DFE	18
MSE	62.99379	Root MSE	7.93686
SBC	143.502518	AIC	141.511054
Regress R-Square	0.0476	Total R-Square	0.0476
Durbin-Watson	0.7175		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	4.5176	1.8076	2.50	0.0223
FFTQ_Qd1	1	-3.7200	3.9240	-0.95	0.3557

SSE	917.268411	DFE	18
MSE	50.95936	Root MSE	7.13858
SBC	139.262363	AIC	137.270899
Regress R-Square	0.2295	Total R-Square	0.2295
Durbin-Watson	0.7533		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	26.9293	9.9482	2.71	0.0144	
CAGDP_Q	1	17.7596	7.6697	2.32	0.0326	CAGDP#Q

SSE	1182.50262	DFE	18
MSE	65.69459	Root MSE	8.10522
SBC	144.342127	AIC	142.350663
Regress R-Square	0.0067	Total R-Square	0.0067
Durbin-Watson	0.7018		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	4.4864	1.9989	2.24	0.0376
DSR_Qd1	1	-481.5664	1380	-0.35	0.7311

SSE	1189.55204	DFE	18
MSE	66.08622	Root MSE	8.12934
SBC	144.461002	AIC	142.469537
Regress R-Square	0.0008	Total R-Square	0.0008
Durbin-Watson	0.6481		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	4.2078	1.8224	2.31	0.0330
MTG_Qd1	1	0.7635	6.3617	0.12	0.9058

SSE	1167.73592	DFE	18
MSE	64.87422	Root MSE	8.05445
SBC	144.090801	AIC	142.099337
Regress R-Square	0.0191	Total R-Square	0.0191
Durbin-Watson	0.6239		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	4.1429	1.8029	2.30	0.0338
URT_Qd1	1	-4.9268	8.3164	-0.59	0.5609

Ordinary Least Squares Estimates

SSE	1181.90054	DFE	18
MSE	65.66114	Root MSE	8.10316
SBC	144.331941	AIC	142.340477
Regress R-Square	0.0072	Total R-Square	0.0072
Durbin-Watson	0.7303		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	4.5917	2.1217	2.16	0.0441	
HAI_Q	1	0.1159	0.3201	0.36	0.7216	HAI_Q

Ordinary Least Squares Estimates

SSE	1182.72306	DFE	18
MSE	65.70684	Root MSE	8.10598
SBC	144.345855	AIC	142.354391
Regress R-Square	0.0065	Total R-Square	0.0065
Durbin-Watson	0.6368		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	2.0862	6.3826	0.33	0.7476	
GDP_Q	1	0.3961	1.1510	0.34	0.7347	GDP#Q

SSE	733.837779	DFE	17
MSE	43.16693	Root MSE	6.57015
SBC	137.795853	AIC	134.808656
Regress R-Square	0.3836	Total R-Square	0.3836
Durbin-Watson	1.0762		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	57.9559	17.6174	3.29	0.0043	
FFTQ_Qd1	1	13.3698	6.4858	2.06	0.0549	
CAGDP_Q	1	42.9077	14.0946	3.04	0.0073	CAGDP#Q

SSE	588.587955	DFE	16
MSE	36.78675	Root MSE	6.06521
SBC	136.380352	AIC	132.397423
Regress R-Square	0.5056	Total R-Square	0.5056
Durbin-Watson	0.9699		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	73.5224	18.0519	4.07	0.0009	
FFTQ_Qd1	1	17.0697	6.2702	2.72	0.0151	
CAGDP_Q	1	55.1495	14.3962	3.83	0.0015	CAGDP#Q
MTG_Qd1	1	10.6000	5.3345	1.99	0.0643	

SSE	516.843397	DFE	15
MSE	34.45623	Root MSE	5.86994
SBC	136.776356	AIC	131.797694
Regress R-Square	0.5659	Total R-Square	0.5659
Durbin-Watson	1.0469		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	62.2435	19.1395	3.25	0.0054	
FFTQ_Qd1	1	7.5904	8.9431	0.85	0.4094	
CAGDP_Q	1	45.8356	15.3552	2.99	0.0093	CAGDP#Q
MTG_Qd1	1	9.2740	5.2439	1.77	0.0973	
URT_Qd1	1	-15.6725	10.8612	-1.44	0.1696	

SSE	498.378942	DFE	14
MSE	35.59850	Root MSE	5.96645
SBC	139.044504	AIC	133.07011
Regress R-Square	0.5814	Total R-Square	0.5814
Durbin-Watson	0.9573		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	63.1653	19.4962	3.24	0.0059	
FFTQ_Qd1	1	6.0286	9.3452	0.65	0.5293	
CAGDP_Q	1	47.2623	15.7329	3.00	0.0095	CAGDP#Q
MTG_Qd1	1	7.2421	6.0308	1.20	0.2497	
URT_Qd1	1	-15.0118	11.0778	-1.36	0.1968	
HAI_Q	1	-0.2918	0.4052	-0.72	0.4833	HAI_Q

SSE	373.229695	DFE	13
MSE	28.70998	Root MSE	5.35817
SBC	136.256902	AIC	129.286777
Regress R-Square	0.6865	Total R-Square	0.6865
Durbin-Watson	1.0089		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	58.6034	17.6444	3.32	0.0055	
FFTQ_Qd1	1	7.7121	8.4311	0.91	0.3770	
CAGDP_Q	1	56.6638	14.8292	3.82	0.0021	CAGDP#Q
MTG_Qd1	1	4.6161	5.5601	0.83	0.4214	
URT_Qd1	1	-1.7827	11.7949	-0.15	0.8822	
HAI_Q	1	-0.1831	0.3676	-0.50	0.6268	HAI_Q
GDP_Q	1	3.1793	1.5228	2.09	0.0571	GDP#Q

The AUTOREG Procedure

Dependent Variable HPICSCUS_Q_yoy

Ordinary Least Squares Estimates

SSE	177.154338	DFE	18
MSE	9.84191	Root MSE	3.13718
SBC	106.374787	AIC	104.383322
Regress R-Square	0.2114	Total R-Square	0.2114
Durbin-Watson	0.3995		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	10.8553	0.7145	15.19	<.0001
FFTQ_Qd1	1	3.4069	1.5510	2.20	0.0414

Ordinary Least Squares Estimates

SSE	202.026885	DFE	18
MSE	11.22372	Root MSE	3.35018
SBC	109.002376	AIC	107.010912
Regress R-Square	0.1007	Total R-Square	0.1007
Durbin-Watson	0.3067		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	4.6124	4.6687	0.99	0.3363	
CAGDP_Q	1	-5.1091	3.5994	-1.42	0.1729	CAGDP*Q

SSE	221.954976	DFE	18
MSE	12.33083	Root MSE	3.51153
SBC	110.883851	AIC	108.892387
Regress R-Square	0.0120	Total R-Square	0.0120
Durbin-Watson	0.2746		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	11.1797	0.7872	14.20	<.0001
MTG_Qd1	1	1.2823	2.7480	0.47	0.6464

Ordinary Least Squares Estimates

SSE	187.575682	DFE	18
MSE	10.42087	Root MSE	3.22814
SBC	107.518008	AIC	105.526544
Regress R-Square	0.1650	Total R-Square	0.1650
Durbin-Watson	0.4667		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	11.0906	0.7226	15.35	<.0001
URT_Qd1	1	-6.2860	3.3331	-1.89	0.0755

Ordinary Least Squares Estimates

SSE	179.3774	DFE	18
MSE	9.96541	Root MSE	3.15680
SBC	106.6242	AIC	104.632735
Regress R-Square	0.2015	Total R-Square	0.2015
Durbin-Watson	0.3030		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	10.2369	0.8266	12.38	<.0001	
HAI_Q	1	-0.2658	0.1247	-2.13	0.0471	HAI_Q

SSE	143.333333	DFE	18
MSE	7.96296	Root MSE	2.82187
SBC	102.137819	AIC	100.146354
Regress R-Square	0.3619	Total R-Square	0.3619
Durbin-Watson	0.3550		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	4.3458	2.2219	1.96	0.0662	
GDP_Q	1	1.2803	0.4007	3.20	0.0050	GDP*Q

SSE	171.325395	DFE	17
MSE	10.07796	Root MSE	3.17458
SBC	108.701385	AIC	105.714189
Regress R-Square	0.2373	Total R-Square	0.2373
Durbin-Watson	0.5335		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	17.3058	8.5124	2.03	0.0580	
FFTQ_Qd1	1	5.4698	3.1338	1.75	0.0990	
CAGDP_Q	1	5.1793	6.8103	0.76	0.4574	CAGDP#Q

SSE	168.685608	DFE	16
MSE	10.54285	Root MSE	3.24698
SBC	111.386558	AIC	107.403629
Regress R-Square	0.2491	Total R-Square	0.2491
Durbin-Watson	0.5204		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	19.4043	9.6640	2.01	0.0618	
FFTQ_Qd1	1	5.9685	3.3567	1.78	0.0944	
CAGDP_Q	1	6.8297	7.7069	0.89	0.3887	CAGDP#Q
MTG_Qd1	1	1.4290	2.8558	0.50	0.6236	

SSE	168.663265	DFE	15
MSE	11.24422	Root MSE	3.35324
SBC	114.379642	AIC	109.40098
Regress R-Square	0.2492	Total R-Square	0.2492
Durbin-Watson	0.5229		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	19.6033	10.9336	1.79	0.0932	
FFTQ_Qd1	1	6.1358	5.1088	1.20	0.2484	
CAGDP_Q	1	6.9940	8.7718	0.80	0.4377	CAGDP#Q
MTG_Qd1	1	1.4524	2.9956	0.48	0.6348	
URT_Qd1	1	0.2766	6.2045	0.04	0.9650	

SSE	158.845488	DFE	14
MSE	11.34611	Root MSE	3.36840
SBC	116.175929	AIC	110.201535
Regress R-Square	0.2929	Total R-Square	0.2929
Durbin-Watson	0.5146		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	20.2755	11.0067	1.84	0.0867	
FFTQ_Qd1	1	4.9970	5.2759	0.95	0.3596	
CAGDP_Q	1	8.0343	8.8821	0.90	0.3810	CAGDP#Q
MTG_Qd1	1	-0.0293	3.4047	-0.01	0.9933	
URT_Qd1	1	0.7584	6.2541	0.12	0.9052	
HAI_Q	1	-0.2128	0.2288	-0.93	0.3680	HAI_Q

SSE	105.476071	DFE	13
MSE	8.11354	Root MSE	2.84843
SBC	110.982704	AIC	104.012578
Regress R-Square	0.5305	Total R-Square	0.5305
Durbin-Watson	0.9217		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	17.2965	9.3798	1.84	0.0881	
FFTQ_Qd1	1	6.0964	4.4820	1.36	0.1969	
CAGDP_Q	1	14.1738	7.8833	1.80	0.0954	CAGDP#Q
MTG_Qd1	1	-1.7441	2.9558	-0.59	0.5653	
URT_Qd1	1	9.3973	6.2702	1.50	0.1578	
HAI_Q	1	-0.1418	0.1954	-0.73	0.4811	HAI_Q
GDP_Q	1	2.0762	0.8095	2.56	0.0235	GDP#Q

The AUTOREG Procedure

Dependent Variable XPOIUS_Q_yoyd1

Ordinary Least Squares Estimates

SSE	12.963851	DFE	18
MSE	0.72021	Root MSE	0.84865
SBC	54.0776562	AIC	52.0861917
Regress R-Square	0.0353	Total R-Square	0.0353
Durbin-Watson	0.2537		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.0810	0.1933	-0.42	0.6800
FFTQ_Qd1	1	-0.3404	0.4196	-0.81	0.4278

SSE	10.6424046	DFE	18
MSE	0.59124	Root MSE	0.76892
SBC	50.1312896	AIC	48.139825
Regress R-Square	0.2080	Total R-Square	0.2080
Durbin-Watson	0.3208		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	2.1890	1.0716	2.04	0.0560	
CAGDP_Q	1	1.7963	0.8261	2.17	0.0433	CAGDP#Q

SSE	12.9040306	DFE	18
MSE	0.71689	Root MSE	0.84669
SBC	53.9851547	AIC	51.9936902
Regress R-Square	0.0397	Total R-Square	0.0397
Durbin-Watson	0.3823		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.1225	0.1898	-0.65	0.5267
MTG_Qd1	1	-0.5718	0.6626	-0.86	0.3995

SSE	13.4211271	DFE	18
MSE	0.74562	Root MSE	0.86349
SBC	54.7709627	AIC	52.7794982
Regress R-Square	0.0012	Total R-Square	0.0012
Durbin-Watson	0.2245		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	-0.1122	0.1933	-0.58	0.5689
URT_Qd1	1	-0.1335	0.8916	-0.15	0.8826

SSE	13.0261423	DFE	18
MSE	0.72367	Root MSE	0.85069
SBC	54.1735261	AIC	52.1820615
Regress R-Square	0.0306	Total R-Square	0.0306
Durbin-Watson	0.3066		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	-0.0234	0.2227	-0.11	0.9175	
HAI_Q	1	0.0253	0.0336	0.75	0.4604	HAI_Q

SSE	13.4373643	DFE	18
MSE	0.74652	Root MSE	0.86401
SBC	54.7951445	AIC	52.8036799
Regress R-Square	0.0000	Total R-Square	0.0000
Durbin-Watson	0.2307		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	-0.0942	0.6803	-0.14	0.8914	
GDP_Q	1	-0.003123	0.1227	-0.03	0.9800	GDP#Q

SSE	8.34758347	DFE	17
MSE	0.49103	Root MSE	0.70074
SBC	48.2695345	AIC	45.2823377
Regress R-Square	0.3788	Total R-Square	0.3788
Durbin-Watson	0.8249		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	5.6593	1.8790	3.01	0.0079	
FFTQ_Qd1	1	1.4954	0.6917	2.16	0.0452	
CAGDP_Q	1	4.6092	1.5033	3.07	0.0070	CAGDP#Q

SSE	8.23733032	DFE	16
MSE	0.51483	Root MSE	0.71752
SBC	50.999351	AIC	47.0164219
Regress R-Square	0.3870	Total R-Square	0.3870
Durbin-Watson	0.7706		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	6.0882	2.1355	2.85	0.0116	
FFTQ_Qd1	1	1.5974	0.7418	2.15	0.0469	
CAGDP_Q	1	4.9465	1.7031	2.90	0.0103	CAGDP#Q
MTG_Qd1	1	0.2920	0.6311	0.46	0.6498	

SSE	8.18664571	DFE	15
MSE	0.54578	Root MSE	0.73877
SBC	53.8716423	AIC	48.8929809
Regress R-Square	0.3908	Total R-Square	0.3908
Durbin-Watson	0.7670		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	5.7884	2.4088	2.40	0.0296	
FFTQ_Qd1	1	1.3454	1.1255	1.20	0.2505	
CAGDP_Q	1	4.6989	1.9325	2.43	0.0280	CAGDP#Q
MTG_Qd1	1	0.2568	0.6600	0.39	0.7027	
URT_Qd1	1	-0.4166	1.3669	-0.30	0.7648	

SSE	8.06967924	DFE	14
MSE	0.57641	Root MSE	0.75921
SBC	56.5795642	AIC	50.6051705
Regress R-Square	0.3995	Total R-Square	0.3995
Durbin-Watson	0.7694		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	5.8618	2.4808	2.36	0.0331	
FFTQ_Qd1	1	1.2211	1.1892	1.03	0.3219	
CAGDP_Q	1	4.8124	2.0020	2.40	0.0306	CAGDP#Q
MTG_Qd1	1	0.0951	0.7674	0.12	0.9032	
URT_Qd1	1	-0.3640	1.4096	-0.26	0.8000	
HAI_Q	1	-0.0232	0.0516	-0.45	0.6593	HAI_Q

Ordinary Least Squares Estimates

SSE	6.42299705	DFE	13
MSE	0.49408	Root MSE	0.70291
SBC	55.0107186	AIC	48.0405927
Regress R-Square	0.5220	Total R-Square	0.5220
Durbin-Watson	1.1767		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	5.3385	2.3147	2.31	0.0382	
FFTQ_Qd1	1	1.4142	1.1060	1.28	0.2234	
CAGDP_Q	1	5.8909	1.9453	3.03	0.0097	CAGDP#Q
MTG_Qd1	1	-0.2061	0.7294	-0.28	0.7819	
URT_Qd1	1	1.1535	1.5473	0.75	0.4692	
HAI_Q	1	-0.0107	0.0482	-0.22	0.8271	HAI_Q
GDP_Q	1	0.3647	0.1998	1.83	0.0910	GDP#Q

The AUTOREG Procedure

Dependent Variable Ratio_CSFHFA

Ordinary Least Squares Estimates

SSE	1.69537478	DFE	18
MSE	0.09419	Root MSE	0.30690
SBC	13.3924369	AIC	11.4009724
Regress R-Square	0.0291	Total R-Square	0.0291
Durbin-Watson	0.3784		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	1.3765	0.0699	19.69	<.0001
FFTQ_Qd1	1	-0.1114	0.1517	-0.73	0.4723

SSE	1.44271562	DFE	18
MSE	0.08015	Root MSE	0.28311
SBC	10.1649041	AIC	8.1734396
Regress R-Square	0.1738	Total R-Square	0.1738
Durbin-Watson	0.4931		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	2.1244	0.3945	5.38	<.0001	
CAGDP_Q	1	0.5918	0.3042	1.95	0.0675	CAGDP#Q

Ordinary Least Squares Estimates

SSE	1.73942249	DFE	18
MSE	0.09663	Root MSE	0.31086
SBC	13.9054235	AIC	11.913959
Regress R-Square	0.0039	Total R-Square	0.0039
Durbin-Watson	0.4077		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	1.3654	0.0697	19.59	<.0001
MTG_Qd1	1	-0.0642	0.2433	-0.26	0.7950

Ordinary Least Squares Estimates

SSE	1.73994775	DFE	18
MSE	0.09666	Root MSE	0.31091
SBC	13.911462	AIC	11.9199975
Regress R-Square	0.0035	Total R-Square	0.0035
Durbin-Watson	0.3506		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|
Intercept	1	1.3659	0.0696	19.63	<.0001
URT_Qd1	1	-0.0813	0.3210	-0.25	0.8030

SSE	1.53814268	DFE	18
MSE	0.08545	Root MSE	0.29232
SBC	11.4458732	AIC	9.45440863
Regress R-Square	0.1191	Total R-Square	0.1191
Durbin-Watson	0.5496		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	1.4289	0.0765	18.67	<.0001	
HAI_Q	1	0.0180	0.0115	1.56	0.1361	HAI_Q

SSE	1.74605826	DFE	18
MSE	0.09700	Root MSE	0.31145
SBC	13.9815769	AIC	11.9901124
Regress R-Square	0.0001	Total R-Square	0.0001
Durbin-Watson	0.3453		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	1.3738	0.2452	5.60	<.0001	
GDP_Q	1	-0.001328	0.0442	-0.03	0.9764	GDP#Q

SSE	1.19062396	DFE	17
MSE	0.07004	Root MSE	0.26464
SBC	9.31964281	AIC	6.33244598
Regress R-Square	0.3181	Total R-Square	0.3181
Durbin-Watson	0.9448		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	3.2747	0.7096	4.61	0.0002	
FFTQ_Qd1	1	0.4956	0.2612	1.90	0.0749	
CAGDP_Q	1	1.5241	0.5677	2.68	0.0157	CAGDP#Q

SSE	1.10727009	DFE	16
MSE	0.06920	Root MSE	0.26307
SBC	10.8637772	AIC	6.88084809
Regress R-Square	0.3659	Total R-Square	0.3659
Durbin-Watson	0.7573		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	3.6476	0.7830	4.66	0.0003	
FFTQ_Qd1	1	0.5843	0.2720	2.15	0.0473	
CAGDP_Q	1	1.8174	0.6244	2.91	0.0102	CAGDP#Q
MTG_Qd1	1	0.2539	0.2314	1.10	0.2887	

SSE	1.09650379	DFE	15
MSE	0.07310	Root MSE	0.27037
SBC	13.664092	AIC	8.68543068
Regress R-Square	0.3720	Total R-Square	0.3720
Durbin-Watson	0.8212		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	3.5094	0.8816	3.98	0.0012	
FFTQ_Qd1	1	0.4682	0.4119	1.14	0.2736	
CAGDP_Q	1	1.7033	0.7073	2.41	0.0293	CAGDP#Q
MTG_Qd1	1	0.2377	0.2415	0.98	0.3407	
URT_Qd1	1	-0.1920	0.5003	-0.38	0.7065	

SSE	0.92812886	DFE	14
MSE	0.06629	Root MSE	0.25748
SBC	13.3255956	AIC	7.35120192
Regress R-Square	0.4685	Total R-Square	0.4685
Durbin-Watson	0.9281		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	3.4214	0.8413	4.07	0.0012	
FFTQ_Qd1	1	0.6173	0.4033	1.53	0.1481	
CAGDP_Q	1	1.5670	0.6789	2.31	0.0368	CAGDP#Q
MTG_Qd1	1	0.4317	0.2603	1.66	0.1194	
URT_Qd1	1	-0.2551	0.4781	-0.53	0.6020	
HAI_Q	1	0.0279	0.0175	1.59	0.1333	HAI_Q

SSE	0.77582342	DFE	13
MSE	0.05968	Root MSE	0.24429
SBC	12.7364151	AIC	5.76628921
Regress R-Square	0.5557	Total R-Square	0.5557
Durbin-Watson	1.0722		

Variable	DF	Estimate	Standard Error	t Value	Approx Pr > |t|	Variable Label
Intercept	1	3.2622	0.8045	4.06	0.0014	
FFTQ_Qd1	1	0.6760	0.3844	1.76	0.1021	
CAGDP_Q	1	1.8950	0.6761	2.80	0.0149	CAGDP*Q
MTG_Qd1	1	0.3401	0.2535	1.34	0.2027	
URT_Qd1	1	0.2064	0.5378	0.38	0.7073	
HAI_Q	1	0.0317	0.0168	1.89	0.0814	HAI_Q
GDP_Q	1	0.1109	0.0694	1.60	0.1342	GDP*Q

UNIT ROOT ANALYSIS – ENTIRE PERIOD

The ARIMA Procedure

Name of Variable = UHS_Q_yoy

Mean of Working Series	-1.12763
Standard Deviation	16.74597
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-15.6371	0.0049	-2.85	0.0048		
	1	-21.0240	0.0009	-3.18	0.0018		
Single Mean	0	-15.6160	0.0289	-2.82	0.0589	4.02	0.0874
	1	-20.9939	0.0066	-3.14	0.0266	5.00	0.0388
Trend	0	-15.6393	0.1443	-2.81	0.1961	3.98	0.3812
	1	-20.9827	0.0450	-3.13	0.1054	4.90	0.1990

Name of Variable = HPICSCUS_Q_yoy

Mean of Working Series	3.311956
Standard Deviation	7.237841
Number of Observations	102

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-2.5015	0.2758	-1.05	0.2649		
	1	-18.3534	0.0021	-2.97	0.0033		
Single Mean	0	-3.1646	0.6314	-1.20	0.6712	0.73	0.8838
	1	-22.3388	0.0045	-3.27	0.0190	5.35	0.0314
Trend	0	-2.7306	0.9456	-1.01	0.9371	1.02	0.9650
	1	-22.8361	0.0289	-3.22	0.0864	5.32	0.1297

Name of Variable = XPOIUS_Q_yoy

Mean of Working Series	3.204823
Standard Deviation	4.841047
Number of Observations	86

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-0.7395	0.5183	-0.43	0.5235		
	1	-5.8619	0.0926	-1.56	0.1106		
Single Mean	0	-1.7229	0.8076	-0.84	0.8016	0.48	0.9542
	1	-9.1900	0.1519	-2.05	0.2652	2.14	0.5314
Trend	0	-1.5151	0.9792	-0.70	0.9697	0.40	0.9900
	1	-10.0467	0.4085	-2.07	0.5529	2.18	0.7427

Name of Variable = XPOIUS_Q_yoy

Period(s) of Differencing	1
Mean of Working Series	0.058708
Standard Deviation	1.067626
Number of Observations	85

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-31.2378	<.0001	-4.35	<.0001		
	1	-20.4335	0.0010	-3.15	0.0020		
Single Mean	0	-31.3412	0.0008	-4.34	0.0008	9.40	0.0010
	1	-20.4898	0.0070	-3.13	0.0280	4.91	0.0443
Trend	0	-31.4501	0.0030	-4.32	0.0048	9.33	0.0010
	1	-20.6249	0.0457	-3.13	0.1053	4.96	0.2005

Name of Variable = Ratio_CSFHFA

Mean of Working Series	0.419901
Standard Deviation	5.825134
Number of Observations	102

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-93.6708	<.0001	-9.29	<.0001		
	1	-82.1367	<.0001	-6.30	<.0001		
Single Mean	0	-94.1146	0.0009	-9.29	<.0001	43.12	0.0010
	1	-83.0359	0.0009	-6.31	<.0001	19.91	0.0010
Trend	0	-95.7741	0.0003	-9.38	<.0001	43.99	0.0010
	1	-87.3978	0.0003	-6.41	<.0001	20.54	0.0010

Name of Variable = FFTQ_Q

Mean of Working Series	3.95934
Standard Deviation	2.625211
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1.3588	0.4139	-1.35	0.1632		
	1	-3.7808	0.1794	-1.69	0.0855		
Single Mean	0	-0.9658	0.8875	-0.52	0.8814	0.94	0.8322
	1	-8.7660	0.1716	-2.04	0.2693	2.33	0.4779
Trend	0	-4.7801	0.8311	-1.60	0.7866	1.45	0.8887
	1	-24.4571	0.0200	-3.44	0.0512	5.93	0.0716

Name of Variable = FFTQ_Q

Period(s) of Differencing	1
Mean of Working Series	-0.0581
Standard Deviation	0.466874

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-34.2089	<.0001	-4.54	<.0001		
	1	-32.8464	<.0001	-4.03	<.0001		
Single Mean	0	-34.8598	0.0010	-4.57	0.0003	10.47	0.0010
	1	-33.7042	0.0010	-4.07	0.0017	8.28	0.0010
Trend	0	-34.8643	0.0014	-4.55	0.0021	10.36	0.0010
	1	-33.7288	0.0019	-4.04	0.0102	8.19	0.0058

Name of Variable = CAGDP_Q

Mean of Working Series	-1.2036
Standard Deviation	0.670972
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-4.8962	0.1259	-2.54	0.0114		
	1	-5.5687	0.1021	-2.56	0.0107		
Single Mean	0	-13.2351	0.0545	-3.41	0.0128	6.40	0.0037
	1	-17.3245	0.0181	-3.76	0.0044	7.57	0.0010
Trend	0	-14.0267	0.1997	-3.25	0.0798	5.86	0.0752
	1	-18.7165	0.0748	-3.66	0.0297	7.21	0.0262

Name of Variable = DSR_Q

Mean of Working Series	0.121547
Standard Deviation	0.00975
Number of Observations	105

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-0.1360	0.6502	-1.07	0.2554		
	1	-0.1600	0.6448	-0.83	0.3550		
Single Mean	0	0.2275	0.9663	0.14	0.9674	0.59	0.9239
	1	-1.7049	0.8104	-0.68	0.8468	0.53	0.9433
Trend	0	1.6360	0.9995	0.95	0.9999	2.23	0.7328
	1	-0.7400	0.9904	-0.27	0.9905	0.50	0.9900

Name of Variable = DSR_Q

Period(s) of Differencing	1
Mean of Working Series	-0.00016
Standard Deviation	0.001513

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-61.9061	<.0001	-6.61	<.0001		
	1	-33.3331	<.0001	-4.04	<.0001		
Single Mean	0	-62.7767	0.0010	-6.65	<.0001	22.09	0.0010
	1	-33.9588	0.0010	-4.05	0.0018	8.22	0.0010
Trend	0	-64.7176	0.0003	-6.70	<.0001	22.54	0.0010
	1	-35.9996	0.0010	-4.15	0.0075	8.60	0.0010

Name of Variable = HAI_Q

Mean of Working Series	2.619863
Standard Deviation	7.596719
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-11.3280	0.0176	-2.45	0.0144		
	1	-18.0059	0.0024	-2.99	0.0031		
Single Mean	0	-12.1188	0.0730	-2.46	0.1276	3.10	0.2810
	1	-19.8022	0.0092	-3.05	0.0339	4.71	0.0476
Trend	0	-12.4363	0.2710	-2.47	0.3443	3.05	0.5670
	1	-20.6412	0.0487	-3.08	0.1168	4.75	0.2276

Name of Variable = MTG_Q

Mean of Working Series	7.154245
Standard Deviation	1.893086
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-0.7558	0.5157	-1.51	0.1226		
	1	-1.0264	0.4665	-1.90	0.0555		
Single Mean	0	-0.7800	0.9038	-0.39	0.9058	1.13	0.7833
	1	-2.2754	0.7421	-1.04	0.7379	1.96	0.5723
Trend	0	-18.4829	0.0789	-3.27	0.0762	5.64	0.0870
	1	-23.6700	0.0241	-3.38	0.0596	5.72	0.0829

Name of Variable = MTG_Q

Period(s) of Differencing	1
Mean of Working Series	-0.05162
Standard Deviation	0.36145

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-88.6818	<.0001	-9.25	<.0001		
	1	-101.462	0.0001	-7.26	<.0001		
Single Mean	0	-91.0655	0.0010	-9.47	<.0001	44.93	0.0010
	1	-109.807	0.0001	-7.56	<.0001	28.59	0.0010
Trend	0	-91.0821	0.0003	-9.40	<.0001	44.44	0.0010
	1	-109.957	0.0001	-7.51	<.0001	28.30	0.0010

Name of Variable = URT_Q

Mean of Working Series	6.049057
Standard Deviation	1.534585
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	0.0709	0.6973	0.15	0.7276		
	1	-0.3332	0.6054	-0.29	0.5781		
Single Mean	0	-1.3950	0.8448	-0.72	0.8361	0.32	0.9872
	1	-10.6572	0.1063	-2.22	0.1998	2.50	0.4336
Trend	0	-2.4725	0.9549	-1.21	0.9022	1.51	0.8754
	1	-12.1052	0.2879	-2.43	0.3596	2.99	0.5808

Name of Variable = URT_Q

Period(s) of Differencing	1
Mean of Working Series	0.009524
Standard Deviation	0.287692

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-31.1162	<.0001	-4.30	<.0001		
	1	-25.1063	0.0002	-3.56	0.0005		
Single Mean	0	-31.1833	0.0010	-4.28	0.0008	9.17	0.0010
	1	-25.2086	0.0021	-3.56	0.0084	6.33	0.0053
Trend	0	-31.3833	0.0035	-4.25	0.0054	9.11	0.0010
	1	-25.2637	0.0164	-3.51	0.0435	6.26	0.0543

Name of Variable = GDP_Q

Mean of Working Series	4.967195
Standard Deviation	2.127438
Number of Observations	106

Augmented Dickey-Fuller Unit Root Tests

Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Zero Mean	0	-1.4961	0.3940	-0.94	0.3057		
	1	-3.3325	0.2078	-1.40	0.1488		
Single Mean	0	-7.7465	0.2214	-1.94	0.3144	1.90	0.5875
	1	-20.0443	0.0086	-3.10	0.0298	4.84	0.0436
Trend	0	-11.3302	0.3315	-2.48	0.3381	3.15	0.5475
	1	-27.8246	0.0088	-3.67	0.0286	6.75	0.0389

COINTEGRATION TEST – PRE BUBBLE PERIOD

Cointegration Rank Test Using Trace

Variable	H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process
UHS_Q	0	0	0.3823	28.3156	15.34	Constant	Linear
FFTQ_Q	1	1	0.0064	0.3717	3.84		

Information Criteria

AICC	7.149963
HQC	7.221704
AIC	7.138678
SBC	7.351827
FPEC	1259.995

Univariate Model AR Diagnostics

Variable	AR1 F Value	AR1 Pr > F	AR2 F Value	AR2 Pr > F	AR3 F Value	AR3 Pr > F	AR4 F Value	AR4 Pr > F
UHS_Q	0.28	0.5971	0.11	0.8999	0.08	0.9680	0.71	0.5907

Cointegration Rank Test Using Maximum Eigenvalue

H0: Rank=r	H1: Rank=r+1	Eigenvalue	Maximum	5% Critical Value
0	1	0.1951	12.5888	14.07
1	2	0.1061	6.5061	3.76

The VARMAX Procedure

Granger-Causality Wald Test

Test	DF	Chi-Square	Pr > ChiSq
1	1	3.57	0.0590

Test 1: Group 1 Variables: UHS_Q_yoy
Group 2 Variables: FFTQ_Qd1

Variable	H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process
HPICSCUS_Q	0	0	0.3345	25.7637	15.34	Constant	Linear
FFTQ_Q	1	1	0.0514	2.9555	3.84		

Information Criteria

AICC	-2.32891
HQC	-2.20403
AIC	-2.40034
SBC	-1.894
FPEC	0.090926

Univariate Model AR Diagnostics

Variable	AR1 F Value	AR1 Pr > F	AR2 F Value	AR2 Pr > F	AR3 F Value	AR3 Pr > F	AR4 F Value	AR4 Pr > F
HPICSCUS_Q	0.95	0.3346	0.40	0.6739	0.25	0.8591	13.18	<.0001

Cointegration Rank Test Using Maximum Eigenvalue

H0: Rank=r	H1: Rank=r+1	Eigenvalue	Maximum	5% Critical Value
0	1	0.3823	27.9439	14.07
1	2	0.0064	0.3717	3.76

The VARMAX Procedure

Granger-Causality Wald Test

Test	DF	Chi-Square	Pr > ChiSq
1	3	6.83	0.0777

Test 1: Group 1 Variables: HPICSCUS_Q_yoy
Group 2 Variables: FFTQ_Qd1

Variable
UHS_Q
FFTQ_Q
CAGDP_Q
MTG_Q
URT_Q
HAI_Q
GDP_Q

H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process	Information Criteria	
0	0	0.7580	190.9238	123.04	Constant	Linear		
1	1	0.5918	108.6218	93.92				
2	2	0.3245	56.6589	68.68			AICC	-1.2682
3	3	0.2609	33.9046	47.21			HQC	-0.80226
4	4	0.1769	16.3695	29.38			AIC	-1.57717
5	5	0.0705	5.0763	15.34			SBC	0.412226
6	6	0.0143	0.8338	3.84			FPEC	0.209134

Univariate Model AR Diagnostics

Variable	AR1 F Value	AR1 Pr > F	AR2 F Value	AR2 Pr > F	AR3 F Value	AR3 Pr > F	AR4 F Value	AR4 Pr > F
UHS_Q	0.12	0.7271	0.02	0.9830	0.01	0.9978	0.41	0.7981

Cointegration Rank Test Using Maximum Eigenvalue

H0: Rank=r	H1: Rank=r+1	Eigenvalue	Maximum	5% Critical Value
0	1	0.7580	82.3019	45.28
1	2	0.5918	51.9630	39.37
2	3	0.3245	22.7542	33.46
3	4	0.2609	17.5351	27.07
4	5	0.1769	11.2932	20.97
5	6	0.0705	4.2425	14.07
6	7	0.0143	0.8338	3.76

Test	DF	Chi-Square	Pr > ChiSq
1	6	20.71	0.0021

Test 1: Group 1 Variables: UHS_Q_yoy
Group 2 Variables: FFTQ_Qd1 CAGDP_Q MTG_Qd1 URT_Qd1 HAI_Q GDP_Q

Variable	H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process	Information Criteria	
XPOIUS_Q	0	0	0.9180	267.5034	123.04	Constant	Linear		
FFTQ_Q	1	1	0.7392	167.4709	93.92				
CAGDP_Q	2	2	0.6618	113.7168	68.68			AICC	-4.70814
MTG_Q	3	3	0.5439	70.3545	47.21			HQC	-11.7683
URT_Q	4	4	0.4533	38.9534	29.38			AIC	-14.1193
HAI_Q	5	5	0.2219	14.7991	15.34			SBC	-7.61707
GDP_Q	6	6	0.1122	4.7613	3.84			FPEC	1.922E-6

Test	DF	Chi-Square	Pr > ChiSq
1	6	22.89	0.0008

Test 1: Group 1 Variables: XPOIUS_Q_yoyd1
Group 2 Variables: FFTQ_Qd1 CAGDP_Q MTG_Qd1 URT_Qd1 HAI_Q GDP_Q

COINTEGRATION TEST –BUBBLE PERIOD

Granger-Causality Wald Test

Test	DF	Chi-Square	Pr > ChiSq
1	18	37.89	0.0040

Test 1: Group 1 Variables: UHS_Q_yoy
Group 2 Variables: FFTQ_Qd1 CAGDP_Q MTG_Qd1 URT_Qd1 HAI_Q GDP_Q

Variable	H0: Rank=r	H1: Rank>r	Eigenvalue	Trace	5% Critical Value	Drift in ECM	Drift in Process
UHS_Q	0	0	0.9754	188.9767	123.04	Constant	Linear
FFTQ_Q	1	1	0.8985	118.6012	93.92		
CAGDP_Q	2	2	0.7993	75.1436	68.68		
MTG_Q	3	3	0.6577	44.6278	47.21		
URT_Q	4	4	0.5200	24.2587	29.38		
HAI_Q	5	5	0.2938	10.3129	15.34		
GDP_Q	6	6	0.1771	3.7041	3.84		

Information Criteria	
AICC	-5.5592
HQC	-9.37519
AIC	-9.84628
SBC	-7.06267
FPEC	0.000078

Univariate Model AR Diagnostics

	AR1		AR2		AR3		AR4	
Variable	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F
UHS_Q	0.00	0.9735	0.10	0.9077	0.07	0.9740	0.07	0.9899

Cointegration Rank Test Using Maximum Eigenvalue

H0: Rank=r	H1: Rank=r+1	Eigenvalue	Maximum	5% Critical Value
0	1	0.9754	70.3755	45.28
1	2	0.8985	43.4576	39.37
2	3	0.7993	30.5158	33.46
3	4	0.6577	20.3691	27.07
4	5	0.5200	13.9458	20.97
5	6	0.2938	6.6088	14.07
6	7	0.1771	3.7041	3.76

APPENDIX B: COMPUTER CODE

```
*=====================================================;
*November 2013  Shweta Khosla:SAS code-THESIS;
* ECON 6999, UNC-Charlotte                     ;
*=====================================================;

*=================================================================;
*Variables included:                                              ;
*=================================================================;
*======================;
*Dependent variables 2: Housing Starts (UHS_M), Case-Shiller HPI
(HPICSCUS_Q), FHFA HPI (HPIFHFA_Q),ratio of S&P/Case-Shiller HPI to
FHFA HPI ;
*Independent variables 7: Federal Funds rate (FFTQ_Q), U.S. Current
account balance as percent of GDP (ca_gdp), Household Debt Ratio
(DSR_Q), Housing Affordability Index (HAI_Q),Mortgage Rate (MTG_Q),
Unemployment Rate (URT_Q), GDP growth rate (GDP_Q);
*======================;

* FFTQ_Q  CAGDP_Q  UHS_Q  HPICSCUS_Q XPOIUS_Q HPIFHFA_Q  DSR_Q  MTG_Q
URT_Q HAI_Q GDP_Q
Ratio_CSFHFA;

*Macro;

%let regression_start_date ='1jan1980'd;
%let regression_end_date   ='1apr2013'd;

%let var_lst =FFTQ_Q  CAGDP_Q  UHS_Q  HPICSCUS_Q XPOIUS_Q HPIFHFA_Q
DSR_Q MTG_Q URT_Q HAI_Q GDP_Q
Ratio_CSFHFA FFTQ_Qd1 XPOIUS_Q_yoyd1 DSR_Qd1 MTG_Qd1 URT_Qd1;

%macro define_variable_lst(variable_lst_real_names=);
%global y1 y2 y3 y4 y5 y6 y7 y8 y9 y10 y11 y12 y13 y14 y15 y16 y17;
   %do i=1 %to 17;
       %let y&i=%scan(&variable_lst_real_names, &i);
   %end;
%mend;

%define_variable_lst(variable_lst_real_names=&var_lst);
%let variable_lst=&y1 &y2 &y3 &y4 &y5 &y6 &y7 &y8 &y9 &y10 &y11 &y12
&y13 &y14 &y15 &y16 &y17;

%let data_dir=C:\SK2\Model;

*Import base data on variables from XL into PC SAS;

Proc import datafile="&data_dir\fed_housing_Shweta_2.xls"
            out     =data1 replace;
            sheet   =hs_ff;
            range   ="A5:F806";
```

```
            getnames=yes;
run;

Data data1 (rename=(GII_Mnemonic=date));
   Set data1;
run;

Proc Import datafile="&data_dir\fed_housing_Shweta_2.xls"
            out     =data2 replace;
            sheet   =HPI;
            range   ="A5:H159";
            getnames=yes;
Run;

Data data2 (rename=(GII_Mnemonic=date));
   Set data2;
Run;

Proc import datafile="&data_dir\fed_housing_Shweta_2.xls"
            out     =data3 replace;
            sheet   =CA_GDP;
            range   ="A5:E271";
            getnames=yes;
Run;

Data data3 (rename=(GII_Mnemonic=date));
   Set data3;
Run;

Proc import datafile="&data_dir\fed_housing_Shweta_2.xls"
            out     =data4 replace;
            sheet   =Others;
            range   ="A5:F119";
            getnames=yes;
Run;

Data data4;
   Set data4;
Run;

Proc Sort Data=data1          ; by date; run;
Proc Sort Data=data2          ; by date; run;
Proc Sort Data=data3          ; by date; run;
Proc Sort Data=data4          ; by date; run;

data estimate;
   merge data2 data3 data4 data1;
   by date;
run;

proc sort data=estimate     ; by date; run;

*Creating ratio of S&P/Case-Shiller HPI to FHFA HPI;
%macro yoy (var=,);
data estimate;
set estimate;
 &var._yoy=((&var-lag4(&var))/lag4(&var))*100;
```

```
run;
%mend yoy;

%yoy (var=&y3);
%yoy (var=&y4);
%yoy (var=&y5);
%yoy (var=&y6);

data estimate;
set estimate ;
Ratio_CSFHFA   =&y4._yoy/&y6._yoy;
FFTQ_Qd1=dif(FFTQ_Q);
XPOIUS_Q_yoyd1=dif(&y5._yoy);
DSR_Qd1=dif(DSR_Q);
MTG_Qd1=dif(MTG_Q);
URT_Qd1=dif(URT_Q);
run;

data final;
   set estimate ;
   if date >=&regression_start_date and date <=&regression_end_date
then output;
run;

%let current= &y1 &y2 &y3 &y4 &y5 &y6 &y7 &y8 &y9 &y10 &y11 &y12 &y13
&y14 &y15 &y16 &y17;

%let yoy    = &y1 &y2 &y3._yoy &y4._yoy &y5._yoy &y6._yoy &y7 &y8 &y9
&y10 &y11 &y12 &y13 &y14 &y15 &y16 &y17;

%let stationary = &y2 &y3._yoy &y4._yoy &y10 &y11 &y12 &y13 &y14 &y15
&y16 &y17;

*====================================;
* (1) Correlation Analysis;
*====================================;

%macro dates (startdate=, enddate=,);
 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;
%macro correlation (lag=,) ;

proc corr data=final2;
var &lag ;
title "Sample is &startdate and &enddate:Pre-Bubble period";
*Keep changing in the title: Pre-Bubble - Bubble- Overall;
run;
quit ;

%mend correlation ;

%correlation (lag=&stationary)     ;

%mend dates;
```

```
%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1apr2013'd); *Overall time
period ;
%dates (startdate='1jan2000'd, enddate='1jul2006'd); *Bubble period ;

*=====================================================;
* (2) Regression  Analysis: Bivariate & Multivariate
*=====================================================;

%macro regression (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro estimation (dependent=,);

proc autoreg  data=final2;

model  &dependent    =  &y13            ;
model  &dependent    =  &y2            ;
model  &dependent    =  &y16            ;
model  &dependent    =  &y17            ;
model  &dependent    =  &y10            ;
model  &dependent    =  &y11            ;
model  &dependent    =  &y13 &y2 ;
model  &dependent    =  &y13 &y2 &y16 ;
model  &dependent    =  &y13 &y2 &y16 &y17 ;
model  &dependent    =  &y13 &y2 &y16 &y17 &y10 ;
model  &dependent    =  &y13 &y2 &y16 &y17 &y10 &y11        ;
title "Sample is &startdate and &enddate";
run;
quit;

%mend estimation ;

%estimation (dependent=&y3._yoy);
%estimation (dependent=&y4._yoy);
%estimation (dependent=&y14);
%estimation (dependent=Ratio_CSFHFA);

%mend regression;

%regression (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble
period ;
%regression (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke
used this time period ;
%regression (startdate='1jan1987'd, enddate='1apr2013'd); *Overall time
period ;
%regression (startdate='1jan2000'd, enddate='1jul2006'd); *Bubble
period ;
```

```
*===============================================================;
* (3) Unit Root Testing                                          ;
*     Testing Unit Root for level & first difference of series ;
*===============================================================;
* The Augmented Dickey-Fuller (ADF) Test;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro unitroot (current=,) ;
Proc ARIMA Data= final2;
Identify var = &current stationarity=(ADF=1);
Identify var = &current(1) stationarity=(ADF=1);
title "Sample is &startdate and &enddate";
Run;
Quit;

%mend unitroot ;

* Level of Series ;
%unitroot (current=&y1)     ;
%unitroot (current=&y2)     ;
%unitroot (current=&y3)     ;
%unitroot (current=&y4)     ;
%unitroot (current=&y5)     ;
%unitroot (current=&y6)     ;
%unitroot (current=&y7)     ;
%unitroot (current=&y8)     ;
%unitroot (current=&y9)     ;
%unitroot (current=&y10)     ;
%unitroot (current=&y11)     ;
%unitroot (current=&y12)     ;

* Growth Rate of Series ;
%unitroot (current=&y3._yoy)     ;
%unitroot (current=&y4._yoy)     ;
%unitroot (current=&y5._yoy)     ;
%unitroot (current=&y6._yoy)     ;

%mend dates;

%dates (startdate='1jan1987'd, enddate='1apr2013'd); *Overall time
period ;

*=============================================;
* (4) Cointegration  Analysis;
*=============================================;

* Test variables for cointegration using Johansen-Juselius procedure:
Trace test
and Maximum Eigen value test;
```

```
*============================================== ;
* (5) Johansen : Trace Test - Just Monetary Policy;
*============================================== ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Trace (dependent=,) ;

Proc varmax data=final2;
model &dependent &y1 / p=3 cointtest =
(johansen=(normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
quit;

%mend Trace ;

%Trace (dependent=&y3);
%Trace (dependent=&y4);
%Trace (dependent=&y5);
%Trace (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2006'd); *Extended period ;
%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*============================================== ;
* (6) Johansen : Trace Test - Just GSG;
*============================================== ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Trace (dependent=,) ;

Proc varmax data=final2;
model &dependent &y2 / p=3 cointtest =
(johansen=(normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
```

```
quit;

%mend Trace ;

%Trace (dependent=&y3);
%Trace (dependent=&y4);
%Trace (dependent=&y5);
%Trace (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2006'd); *Extended period ;
%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*================================================= ;
* (7) Johansen : Trace Test - Monetary Policy & Others;
*================================================= ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Trace (dependent=,) ;

Proc varmax data=final2;
model &dependent &y13 &y16 &y17 &y10 &y11 / p=1 cointtest =
(johansen=(normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
quit;

%mend Trace ;

%Trace (dependent=&y3._yoy);
%Trace (dependent=&y4._yoy);
%Trace (dependent=&y14);
%Trace (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;
```

```
*================================================ ;
* (8) Johansen : Trace Test - GSG & Others;
*================================================ ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Trace (dependent=,) ;

Proc varmax data=final2;
model &dependent &y13 &y2 &y17 &y10 &y11 / p=1 cointtest =
(johansen=(normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
quit;

%mend Trace ;

%Trace (dependent=&y3._yoy);
%Trace (dependent=&y4._yoy);
%Trace (dependent=&y14);
%Trace (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*============================================;
* (9) Johansen Cointegration Test: MP,GSG & Others    ;
*============================================;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Trace (dependent=,) ;

Proc varmax data=final2;
model &dependent &y1 &y2 &y8 &y9 &y10 &y11 / p=2 cointtest =
(johansen=(normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
quit;
```

```
%mend Trace ;
%Trace (dependent=&y3);
%Trace (dependent=&y4);
%Trace (dependent=&y5);
%Trace (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2006'd); *Extended period ;
%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*============================================;
* (10) Johansen Cointegration Test: Maximum Test  ;
*============================================;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Max (dependent=,) ;
*EXTRA;
Proc varmax data=final2;
model &y1 &y4 / p=1 cointtest = (johansen=(type=max normalize=&y4));
title "Sample is &startdate and &enddate";
run;

Proc varmax data=final2;
model &dependent &y13 &y15 &y16 &y17 &y10 &y11 / p=1 cointtest =
(johansen=(type=max normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
quit;

%mend Max ;

%Max (dependent=&y3._yoy);
%Max (dependent=&y4._yoy);
%Max (dependent=&y14);
%Max (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1apr2013'd); *Overall time
period ;
```

```
*============================================= ;
* (11) Johansen : Max Test - Just Monetary Policy;
*============================================= ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Trace (dependent=,) ;

Proc varmax data=final2;
model &dependent &y1 &y2 &y8 &y9 &y10 &y11/ p=1 cointtest =
(johansen=(type=max normalize=&dependent));
title "Sample is &startdate and &enddate";
run;
quit;

%mend Trace ;

%Trace (dependent=&y3);
%Trace (dependent=&y4);
%Trace (dependent=&y5);
%Trace (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*================================================;
* (12) The Granger causality Test: Just Monetary Policy ;
*================================================;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Grangercausality (dependent=, independent=,) ;

Proc varmax data=final2;
model &dependent &y13 / p=1 noprint;
      causal group1=(&dependent) group2=(&y13);
      causal group1=(&y13) group2=(&dependent);
title "Sample is &startdate and &enddate";
run;
```

```
quit;

%mend Grangercausality ;
%Grangercausality (dependent=&y3._yoy);
%Grangercausality (dependent=&y4._yoy);
%Grangercausality (dependent=&y14);
%Grangercausality (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*==================================================;
* (13) The Granger causality Test: Just GSG             ;
*==================================================;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Grangercausality (dependent=, independent=,) ;

Proc varmax data=final2;
model &dependent &y2 / p=3 noprint;
      causal group1=(&dependent) group2=(&y2);
title "Sample is &startdate and &enddate";
run;
quit;

%mend Grangercausality ;

%Grangercausality (dependent=&y3._yoy);
%Grangercausality (dependent=&y4._yoy);
%Grangercausality (dependent=&y14);
%Grangercausality (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*========================================================;
* (14) The Granger causality Test: Monetary Policy and Others ;
*========================================================;
```

```
%macro dates (startdate=, enddate=,);
 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Grangercausality (dependent=, independent=,) ;

Proc varmax data=final2;
model &dependent &y13 &y16 &y17 &y10 &y11 / p=1 noprint;
      causal group1=(&dependent) group2=(&y13 &y16 &y17 &y10 &y11);
title "Sample is &startdate and &enddate";
run;
quit;

%mend Grangercausality ;

%Grangercausality (dependent=&y3._yoy);
%Grangercausality (dependent=&y4._yoy);
%Grangercausality (dependent=&y14);
%Grangercausality (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;

*=========================================================  ;
* (15) The Granger causality Test: Global Savings Glut and Others;
*=========================================================  ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Grangercausality (dependent=, independent=,) ;

Proc varmax data=final2;
model &dependent &y2 &y16 &y17 &y10 &y11 / p=1 noprint;
      causal group1=(&dependent) group2=(&y2 &y16 &y17 &y10 &y11);
title "Sample is &startdate and &enddate";
run;
quit;

%mend Grangercausality ;

%Grangercausality (dependent=&y3._yoy);
%Grangercausality (dependent=&y4._yoy);
%Grangercausality (dependent=&y14);
%Grangercausality (dependent=Ratio_CSFHFA);

%mend dates;
```

```
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;

*==================================================================  ;
* (16) The Granger causality Test: Both Fed Funds & Global Savings Glut
& Others;
*==================================================================  ;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro Grangercausality (dependent=, independent=,) ;

Proc varmax data=final2;
model &dependent &y13 &y2 &y16 &y17 &y10 &y11 / p=3 noprint;
      causal group1=(&dependent) group2=(&y13 &y2 &y16 &y17 &y10 &y11);
      causal group1=(&y13 &y2 &y16 &y17 &y10 &y11) group2=(&dependent);
title "Sample is &startdate and &enddate";
run;
quit;

%mend Grangercausality ;

%Grangercausality (dependent=&y3._yoy);
%Grangercausality (dependent=&y4._yoy);
%Grangercausality (dependent=&y14);
%Grangercausality (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2006'd); *Extended period ;
%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;

*=============================================;
* (17) Vector Error Correction Model ;
*=============================================;

%macro dates (startdate=, enddate=,);

 data final2;
 set final;
  if date >=&startdate and date <=&enddate then output;
  run;

%macro VECM (dependent=,) ;

Proc varmax data=final2;
model &dependent &y1 / p=2 ecm = (rank=2 normalize=&dependent ectrend)
```

```
print=(parcoef);
run;
quit;
%mend VECM ;

%VECM (dependent=&y3);
%VECM (dependent=&y4);
%VECM (dependent=&y5);
%VECM (dependent=&y6);
%VECM (dependent=Ratio_CSFHFA);

%mend dates;

%dates (startdate='1jan1987'd, enddate='1jul2001'd); *Pre-Bubble period
;
%dates (startdate='1oct2001'd, enddate='1jul2006'd); *Bernanke used
this time period ;
%dates (startdate='1jan1987'd, enddate='1jan2013'd); *Overall time
period ;
```