Dynamic Effects of the Housing Price in Different City Tiers in China on the Shanghai Stock Exchange

Christian Pettersen

Thesis submitted for the degree of Master in Economics
30 credits

Department of Economics
Faculty of Social Sciences

UNIVERSITY OF OSLO

10.05.2017
Dynamic Effects of the Housing Price in Different City Tiers in China on the Shanghai Stock Exchange

Christian Pettersen
© Christian Pettersen

Dynamic effects of the housing price in different city tiers in China on the Shanghai Stock Exchange

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo
Preface

I would like to thank my advisor, Yikai Wang for helpful advices along the way while I was writing my thesis. I would also like to thank Ragnar Nymoen for answering questions about methodology and usage of the Oxmetrics software. Finally, I would like to thank my beloved girlfriend, Chen Qian, for proofreading and giving me support during the time I have worked on the thesis.

Any inaccuracies and errors in this thesis are solely my responsibility.

May, 2017

Christian Pettersen
Abstract

In China, due to strong governmental capital controls, investors have few alternatives to invest their savings outside of the domestic equity markets, the real estate market and to deposit in the banks. Since there are so few investment options available for the Chinese investor, vigilance must be directed towards how these few alternatives interact with one another, both for investors and policy makers. In this thesis, I analyze whether or not housing prices in different city categories in China has any effect on the Shanghai Stock Exchange (SSE). The novelty of my approach is that I categorize representative Chinese cities into different city categories based on the city tier system that is commonly used in China. Previous studies on this subject have not taken into account the differences of price growth in different city tiers in China, which is important to address both for investors and policy makers alike.

I examine the relationship between these city tiers and the SSE through an Autoregressive Distributed-Lag Model (ADL) and the reparameterization of the ADL into an Error Correction Model (ECM), where the month by month growth rate of the SSE Composite Index is the dependable variable, and month by month housing price growth rate in four different city categories are independent variables. The housing price data was collected from the housing price indices on 70 large and medium sized cities in China, also called the NBS 70-city index, which is developed by the National Bureau of Statistics in China.

I found significant lagged effects of tier 1 and 2 cities’ price growth on the composite stock exchange. Tier 3 and tier 4 & 5 cities’ price growth not only had lagged effects but also long-run effects. The negative relationship between the stock market and the housing prices were found to be stronger with tier 3 and tier 4 & 5 cities. I theorize possible reasons for these results to be related to the high housing prices in tier 1 and tier 2 cities compared to tier 3 and tier 4 & 5, which may act as a barrier for retail investors with limited capital. Increases in housing prices in these lower tier cities may then attract investors away from the stock exchanges, which results in lower demand for stocks and subsequently lower stock prices.
# Table of contents

1 Introduction

1.1 Literature Review .......................................................... 3

2 Stylized Facts of Chinese Investment Options

2.1 High Savings Rate but Few Investment Alternatives ...................... 5
2.2 Bank Deposit ........................................................................... 6
2.3 The Capital Market .................................................................. 7
   2.3.1 Brief History ........................................................................... 7
   2.3.2 Investors on the Stock Exchanges in China ......................... 9
   2.3.3 Recent Developments ......................................................... 10
2.4 The Residential Real Estate Market ........................................ 12
   2.4.1 Brief History ........................................................................... 12
   2.4.2 Buyers of Residential Real Estate ....................................... 13
   2.4.3 Recent Developments ......................................................... 14

3 Methodology

3.1 The Model .............................................................................. 18
3.2 Unit Root ................................................................................. 18
3.3 Model Specification and Goodness of Fit .................................... 19
   3.3.1 Information Criteria ......................................................... 20
   3.3.2 Error Autocorrelation test .................................................. 20
   3.3.3 Normality test ........................................................................ 21
   3.3.4 Heteroscedasticity test ....................................................... 21
   3.3.5 Functional form test .......................................................... 21
3.4 Weak Exogeneity ..................................................................... 22
3.5 Error Correction Model .......................................................... 22
3.6 Invariance of the variables and constancy of parameters .......... 25

4 Empirical Analysis

4.1 Data Description ..................................................................... 26
4.2 Empirical Results ................................................................. 29
4.3 Discussion ................................................................................. 37
5 Conclusion 38
Reference list 40
Appendix 43
List of Tables

1  City Category ........................................................................................................ 28
2  Descriptive Statistics............................................................................................ 29
3  ADF Test.................................................................................................................. 30
4  ADL results with DWH-test.................................................................................... 31
5  Misspecification tests of ADL regression............................................................... 33
6  ECM results.............................................................................................................. 35

List of Figures

1  Gross savings (% of GDP)........................................................................................ 5
2  Deposit Interest rate (%)........................................................................................... 7
3  Stock Market Capitalization as percentage of GDP............................................... 10
4  SSE Annual Stock Turnover Value (100 million RMB) ........................................ 10
5  SSE Composite Index Monthly values ................................................................. 11
6  SSE Composite Index month by month growth rate.............................................. 11
7  Fixed Assets Investment in percent of GDP ........................................................... 14
8  China’s GDP (100 million RMB).............................................................................. 14
9  Monthly Housing Price Index divided by city tiers (%)......................................... 15
10 Monthly city price level divided by city tiers, and SSE Composite Index with January as base ........................................................................................................ 15
A1  Residual QQ-plot against N(0,1)............................................................................ 43
A2  Graphs of recursive estimated coefficients and Chow tests for tier 1 model......... 43
A3  Graphs of recursive estimated coefficients and Chow tests for tier 2 model......... 44
A4  Graphs of recursive estimated coefficients and Chow tests for tier 3 model........ 44
A5  Graphs of recursive estimated coefficients and Chow tests for tier 4 & 5 model.... 45
Dynamic Effects of the Housing Price in Different City Tiers in China on the Shanghai Stock Exchange

Christian Pettersen

1. Introduction

Most Chinese investors do not enjoy the same range of investment alternatives that their western counterparts do. Not only are there restrictions on cross-border portfolio investments, but the domestic equity markets are highly speculative and volatile. Retail investors thus have few alternatives to invest their savings outside of the domestic equity markets, the real estate market and the banks. Because of having so few investment alternatives, the need for determining the possible effects these markets might have on each other is of great importance. Could investors have a hand in both the equity markets and the real estate market to hedge against risk? Either the markets don’t affect each other or they are negatively correlated with each other, or the markets move in together the same direction such that a negative shock in one market could lead to negative movements in the other.

There are three stock markets in China: the Shanghai Stock Exchange (SSE), the Shenzhen Stock Exchange (SZSE) and the Hong Kong Stock Exchange (HKEX). The stock markets in mainland China (SSE and SZSE) are infamous for their volatility and speculative nature through their short existence and has gone through three major booms and busts until today. The first being during the Asian financial crisis in the late 90’s. The second and largest bust was set off due to spill-over effects from the American subprime crisis in 2007, where two-thirds of the market value on Chinese stock markets were wiped out (Schmidt, 2009, p. 1). The latest boom and bust took place in the summer of 2015 when the Shanghai Composite Index fell from its peak of 5,166 down to 3700s during June and July with aftershocks lasting out the year. The Chinese stock exchanges have significant differences from open and liberal stock markets in developed economies in that it is heavily regulated and that the Chinese government can and will intervene directly and indirectly in the stock market should they deem it necessary, as it did during the summer of 2015 when the bust was setting in. Firm directors and shareholders owning over 5% of the shares in a firm were prohibited from selling, while state-owned enterprises (SOE) were told by the government not to sell any shares. Stimulation packages and more relaxation of rules pertaining to margin trading were
introduced through the China Securities Regulatory Commission in order to let SOE’s and pension funds to stimulate the stock markets (Schell, 2015).

Real estate in China has become an important engine to the Chinese economy, making out an increasing proportion of fixed investments through the last two decades until the recent years, reaching a peak in 2014. In later years, there has been increasing worries of housing price bubbles emerging in the larger metropolitan cities, to which certain cities has implemented restrictions to try to dampen the price growth. Buying real estate is not only seen as an investment, but also carries with it important connotations to culture. There is for example often an informal requirement for men in China to own housing to be considered eligible for marriage. Household registration system (hukou) is also an aspect buyers of housing may consider, as having household registration in a city may give you more rights to local welfare services and social security. Several cities in Guangdong province implemented a system that would grant non-hukou individuals points based on house ownership and other factors towards granting hukou (Zhang, 2012, p. 511-518). These types of point based systems exist in other cities in China, but the requirements vary and not all of them include housing ownership as a point-giving factor. Motivation for purchasing housing thus varies greatly among heterogenous agents in the market as compared to the motivation of investors on the stock markets.

This paper’s objective is to examine possible effects that the residential real estate price growth in different city tiers of China has on the SSE. As far as I know, previous studies on the subject have not taken into account the great differences that exists within the Chinese real estate market by differentiating between cities of different size and importance. The distinction is important, since some of the larger cities arguably shows signs of bubbling housing markets, while other cities might have excess supply of housing and subsequently low housing prices. The model used in the paper is an Autoregressive Distributed-lag model (ADL) to analyse dynamics, which later is reparametrized into an Error-Correction Model (ECM) to look at long-run dynamic effects. Using these methods, I found significant lagged effects of tier 1 and tier 2 cities’ price growth on the composite stock exchange, whereas tier 3 and tier 4 & 5 cities’ price growth not only had lagged effects but also long-run effects. The negative relationship between the stock market and the housing prices were found to be stronger with tier 3 and tier 4 & 5 cities. I theorize possible reasons for these results to be related to the high housing prices in tier 1 and tier 2 cities compared to tier 3 and tier 4 & 5, which may act as a barrier for retail investors with limited capital. Increases in housing prices in these lower tier cities may then attract investors away from the stock exchanges, which
results in lower demand for stocks and subsequently lower stock prices. The effects may be lagged because of the less liquid nature of the real estate market compared to the capital market.

The paper is structured in the following way: section one consists of the introduction and a literature review where the viewer may find other studies on the subject or with relevance to the subject that is examined in this paper. Following that, in section two, more thorough background information and recent developments on the SSE and the real estate markets will be given so that the reader get a more comprehensive picture of the situation of these two markets in China. In the third section, the methodology used to analyse the subject is presented and briefly explained, followed by a brief description of the data used in the analysis. In part four, the empirical results are presented and discussed. In part five the I conclude on the basis of the results of the empirical analysis. There is also included an appendix at the very end, which contains some graphic plots used in the empirical analysis.

1.2 Literature Review

There have been conducted numerous studies on the question of whether the stock market and the real estate market are integrated as opposed to segmented has been an ongoing discussion. Studies that suggest segmentation of the markets: Liu et al. (1990) found that the commercial real estate market and the stock market was segmented. However, they point to cost, amount, and quality of information as being the major source of segmentation. Lu et al. (2007) looked at the relationship between the real estate and stock markets in Taiwan using quarterly data from 1986 until 2006 and found no evidence of cointegration of these markets by using Johansen test and the Engle and Granger test for cointegration. Studies that suggest integration of the markets: Okunev et al. (2000) examined the causal relationship between the US real estate and S&P 500 stock markets with observations from 1972 to 1998. They found a linear unidirectional relationship from the real estate market to the stock market when performing linear causality tests. However, they also conducted non-linear causality tests and found a strong unidirectional relationship from the stock market to the real estate market. Apergis and Lambrinidis (2007) examined the case of the US and the UK markets through cointegration and Error Correction (EC) causality methods. They found causality from the stock market to the real estate market which supports a wealth effect in which gains from the stock market leads to increased real estate investment. Su (2011) investigated long-run non-linear relationship in Western European countries using the non-parametric rank test. He
found strong support of the existence of long-run equilibrium relationship with non-linear adjustment. By the use of threshold error-correction model (TECM) and Granger causality tests, asymmetric price transmissions were identified among all the countries in the study. As for the case of China, several studies have been done on the subject: Quan and Titman (1999) used cross-sectional and time series tests to analyse the relationship between stock returns and commercial real estate indexes in 17 countries, including China. They observed significant positive contemporaneous and lagged effects between these two markets. However, after controlling for fundamental factors such as rental rates and other macro variables this relation weakened considerably. Liu and Su (2010) employed the Momentum Threshold Autoregressive model (M-TAR) and from the corresponding TECM, found there to be a short-run unidirectional causality from the Shenzhen stock exchange (Shenzhen Composite Index) to the real estate market (Real Estate Price Index). This gives weight to the wealth effect hypothesis and also the portfolio adjustment mechanism because of investors who makes gains in the stock markets would want to reinvest some of their gain in real estate to adjust risk. They also found a long-run bidirectional feedback causality relationship between the stock and real estate markets. Lin and Fuerst (2014) found there to be partial cointegration between Shanghai Composite Index and the quarterly real estate index (released by Oxford Economics from the period of 1980 to 2012). Their empirical results suggest that the stock market and direct real estate markets are segmented in China. This implies that both direct real estate and stocks should be included in a portfolio to reduce risk. They list possible reasons for cointegration in Taiwan, Singapore and Hong Kong to be because of the densely-populated nature of these economies. This might lead to reduced transaction costs and less information asymmetry which increases the liquidity of real estate. These factors, they suggest, may lead to less segmentation from stock-markets. First tier cities of China are also densely populated, and share similarities with these economies. Ding et al. (2014) found a strongly significant causal relationship between real estate and stock markets in China when running a quantile causality test, especially in the tail quantiles. Their results imply that when the return in the stock market is very high or very low, it becomes more vulnerable to shocks in the housing market and vice versa.

To my knowledge there has not yet been done a study that looks at how the residential housing price growth rates of different types of cities may have an effect on the stock market. Previous studies that examined the relationship between the stock and real estate markets did not take this into account may thus have skipped over important differences that exist in a huge and diverse country such a China.
2. Stylized Facts of Chinese Investment Options

2.1 High Savings Rate but Few Investment Alternatives

China has a very high savings rate compared to many economies and the world on average (figure 1). The World Bank puts the gross domestic savings as a percentage of the GDP in China at around 50%. Compare that with the EU and the US which have not breached 30% in the last two decades. Households in China saved 37.99% of their income in 2014, having increased steadily from the lowest point of 27.21% in 2002 according to OECD. Not only is the households savings rate very high, but so is the saving rates for the government and corporate sectors. The high savings rate in China have been examined through several studies during the last decades and several possible reasons for the increase in savings rate has been presented. According to Yang et al. (2011), government saving in China is high to due to increased tax revenue where the largest contributor stemming from increased production. Increased tax income was also combined with less expenditure on social welfare payments and other transfers. Corporate saving is also high due to increased profitability of enterprises due to low labour costs of rural laborers and focus on export industry. The transition from largely having public education, health care and pension systems to increasingly privatized institutions may have contributed to increased precautionary household savings. Increase in income uncertainty due to economic shocks, and pension reforms reducing the pension of workers who retire after 1997 could have contributed to the rise in savings (Chamon et al. 2013). A study by Zhang et al. (2012) suggested that increased profitability of State-owned enterprises (SOE) has lead the government to invest more in SOEs which has crowded out certain types of public expenditures leading to increased private savings rate as well as the increase in public investment.

Figure 1, Gross savings (% of GDP)

Source: The World Bank
Even though China is relatively open to Foreign Direct Investment (FDI), Chinese private investors cannot trade in foreign denominated assets, and the same for foreign private investors that would want to invest in Chinese denominated assets. According to Song et al. (2014, p. 8) capital controls in China are highly asymmetric in that there are limited restrictions on direct investment, but tight restrictions on portfolio investment. As a result of the strong capital control imposed by the Chinese government, the domestic stock exchange markets and the real estate markets in China has become attractive investment options due to the comparatively high rate of return compared to other options.

In the following, I will expand a bit on the main investment alternatives available to the majority of Chinese citizens. Chinese citizens can invest their savings in real estate, in the capital market or simply deposit in the bank. Since the theme for this paper is about the residential housing market and the stock exchange, this section will mainly be about those markets, while the possibility of saving in a bank will get less attention.

### 2.2 Bank Deposit

Though the focus of this paper is not clearly related to savings deposited in banks, some background information is provided here to give the reader a clearer view of the investment/saving options available to Chinese investors.

The interest rates in China have historically been regulated by the government, however in recent years this has been increasingly liberalized. The liberalization of the interest rate began in 1983 when the Peoples Bank of China (PBOC) were allowed to adjust the benchmark interest rate by a maximum of 20%. Later, several reforms eased the governments control over the various interest rates. Big steps towards liberalization were taken in 2004, when the floor on deposit interest rates was removed, giving banks more flexibility in setting their deposit interest rates. The upper limit to the deposit rates were gradually raised in the period 2012-15 and finally, in 2015 the ceiling on deposit interest rates were removed (Tan et al. 2016, p. 14). Though the interest rate has been liberalized in China, the government still has ways to influence the banks into following policy (through window guidance).
The deposit interest rate in China is historically low compared to average returns from the stock market or the real estate market. For investors looking to get high returns on their savings, depositing in a bank is not the best option. However, the majority of Chinese savings are saved in deposit accounts in banks. According to China Household Finance Survey (CHFS) in 2012, 60.9% of the households in the survey had at least one bank deposit account (Gan et al. 2014).

2.3 The Capital Market

2.3.1 Brief history

In China, the capital market is mainly represented by the Shanghai Stock Exchange (SSE) and the Shenzhen Stock Exchange (SZSE). There are also other exchanges, such as the China Financial Futures Exchange (CFFEX) for trading in financial futures. For trading commodities: Shanghai Futures Exchange (SHFE), Dalian Commodity Exchange (DCE) and Zhengzhou Commodity Exchange (ZCE). The SSE has become the most important stock market in mainland China since its inception in 1990, and is a non-profit organization directly governed by the China Securities Regulatory Commission (CSRC). Directly quoted from the SSE webpage: “SSE bases its development on the principles of legitimacy, regulation, self-discipline, and compliance in order to create a transparent, open, reliable and efficient marketplace. SSE endeavours to perform a variety of functions such as providing marketplace and facilities for the securities trading, formulating business rules, accepting and arranging listings, organizing and monitoring securities trading, regulating members and listed companies, and managing and disseminating market information”. The main products traded on the SSE are shares, bonds, funds and other financial securities. The SSE Composite Index together with the CSI 300 Index represents the main benchmark of the Chinese market. From when the stock markets in China were established in 1990, the stock-issue system has gone through several reforms. The most notable change was when the stock-issue system went
from being an “Approval system” to becoming a “Verification system” in 2001 (Wan, 2016). Under the Approval System, the stock-issuing regulation showed strong characteristics of a heavily planned economy such as during the “Quote Management”-phase from 93-95, when the total volume of stocks-issue was approved by the State Council. Provincial governments, autonomous zones, municipalities directly controlled by the central government and municipalities with separate economic planning were given stock-issuing quotas to recommend firms for approval by the CSRC based on the development of the economic situation in the regions and whether or not the firm was part of a “key industry” promoted by the government (Su and Green 2003, p. 56). The issuing of the stocks themselves were thoroughly planned to determine suitable volume, price, method and timing. Later, from 1996 up until 2001, more framework and clarification on procedures regarding stock-issuing were introduced. This phase was called the “target management”-phase, and shifted more of the responsibility of stock quota approval to the CSRC. The firms recommended by local governmental institutions were also subject to pre-selection examination to further try to ensure the quality of the recommended firms. From 2001 until the present, the stock-issuing system has been a “Verification System”. The “Verification System” brought in reforms that changed the way firms were recommended to be listed, increased transparency of stock-issue regulation, introduced the requirement of lead underwriters to take some of the risk of stock-issuing by verifying the issuer’s materials and abolishing of targets and quotas of stocks. In 2005, “Book Building” was adopted, which is the procedure of determining the offering price of the stocks for IPO by collecting demand information from institutional investors. Responsibilities and accountability of sponsors during IPO were later emphasized and clarified with the intent to further protect the rights of investors in 2009. Even though sponsors like investment banks could help stock-issue new banks, CSRC had and still has the final say on which firms actually gets listed. The CSRC may also influence IPO pricing through “Window Guidance”, which is a sort of off-the-paper persuasion-tool used to make the IPO-pricing follow official policy through meetings with government officials or other through means (Song et al. 2014. p. 5). The stock-issuing system in China is moving in the direction of the “Registration System”, that stock exchange markets in economies like USA and Japan enjoy.

The stock markets in China have some peculiar characteristics that is necessary to note here. First: stocks on the SSE are divided into A-shares which are traded in the Chinese yuan (CNY), and B-shares which are traded in USD or HKD. A-shares used to be reserved for trade by Chinese nationals, but as of 2003 could also be traded by approved foreign funds and
institutions through the so-called Qualified Foreign Institutional Investors (QFII) system. B-shares were originally only to be tradable by foreign investors, but in 2006 with the introduction of the Qualified Domestic Institutional Investor (QDII) system, could also be purchased by Chinese nationals through fund management institutions, insurance companies, securities companies and other assets management institutions which have been preapproved by the CSRC. Second: many of the firms that get listed on the stock exchanges that used to be fully state-owned are still partially owned by the government which may cause conflict of interests because the state is both the regulator and administrator of the stock markets and/or partly owners of the firms that get listed. This in turn may lead to underperforming firms getting listed on the stock exchange, which can result in market inefficiency and volatility. Firms that aspire to get listed also need CSRC’s final approval even if the firms meet all the other conditions required. The current listing requirements for IPO’s on SSE, taken from the English webpage of SSE: “Its shares have been issued to the public pursuant to CSRC approval. Its total capital stock is no less than 50 million RMB. Its publicly issued shares account for no less than 25% of its total shares; where its total capital stock is more than RMB 400 million, its publicly issued shares shall account for no less than 10% of its total shares. It has no record of major legal violations within the last three years and its financial and accounting reports do not contain any false records. Other conditions required by the SSE.”

2.3.2 Investors on the Stock Exchanges in China
On the stock exchanges in China, there is a very large presence of retail investors (individual investors). According to China Security Depository and Clearing Co (CSDC), which is the government organ that supervises accounts, in March this year there were 122.6 million individual investor accounts and 369,000 institutional investor accounts in China. The number of investor accounts is quite high and might not truly reflect the true number of stock market investors as individuals may own multiple accounts. According to the CHFS, in 2012 less than 9% of households in the survey answered that they had a stock account and that urban households were far more likely than rural households to have a stock account (16.54% and 1.89% respectively). In 2015, 23.15% of retail investors investment through the stock exchanges were less than 10,000 yuan, 48% in the range of 1000-100,000 yuan, and 21.65% from 100,000-500,000 yuan. Of institutional investors, the largest groups were: 12.26% in the range of 10,000-100,000 yuan, 16.50% in the range of 100,000-500,000 yuan, 18.11% in the range of 1-5 million yuan, and 17.88% in the range of 10-100 million yuan. It is made clear from these numbers that individual investors on average invest less than institutional
Investors, but a minority of retail investors owns disproportionately large part of the market capitalization. Over 60,000 individual investors have invested between 10-100 million yuan, which only make up 0.12% of individual investors. A survey by State Street (2014) showed that in China, 81% of retail investors trade at least monthly. The global average in that same report was 57%. Chinese retail investors appear to be willing to take a bit more risk than the world average in order to have a chance of bigger gains. When asked about the reasons for investing their savings, 42% of Chinese investors answered that they did it because they enjoyed investing, which is a huge step away from the global average at 9%. 43% answered that the motivation for investing further savings was if the markets rose significantly, compared to the global average of 27%.

2.3.3 Recent developments

According to SSE and NBS, the stock market capitalization in China reached its highest in 2007, right before the financial crisis which decimated the stock exchanges in China, when it reached 121% of GDP. The market stock capitalization has still not recovered, and made out 77.1% of GDP in 2015, of which the SSE constituted 42.8% and SZSE 34.3%. The stock market capitalization in China is still relatively small in proportion to the GDP compared to that of developed economies like the US, where the stock market capitalization to GDP is 139% in 2015 (The World Bank). High (greater than 100%) stock market capitalization to GDP ratios may hint of an overvalued market, whereas low (around 50%) ratios may suggest undervalued markets. The stock market capitalization increased significantly in 2014 and 2015 after staying stable at around 40% from 2011 until 2014. Annual stock turnover at the
SSE was 133.1 trillion RMB in 2015, which was an increase of 353% from 2014. The number of listed firms on the SSE went up from 293 in 1996 to 1081 in 2015. Whereas the number of listed firms on the SZSE went up from 505 in 2003 to 1756 in 2015 (SSE and SZSE).

Figure 5, SSE Composite Index monthly values

![Figure 5, SSE Composite Index monthly values](source: Shanghai Stock Exchange)

Figure 6, SSE Composite Index month by month growth rate

![Figure 6, SSE Composite Index month by month growth rate](source: Shanghai Stock Exchange)

In the time period where this paper is concerned, there have been two stock bubble bursts. The first one during the financial crisis of 2007/08, the second in the summer of 2015. In figure 4, one can clearly see how the stock index was hugely inflated during the boom periods only to fall greatly when the bubbles busted. As a reaction to the financial crisis in 2007/08, the Chinese government made adjustments to its monetary policy; stimulating economic recovery in the aftermath of the crisis. Zhang et al. (2011) however, found a significant causal link between monetary policy and stock market volatility, and argue that the accommodating monetary policy also made the stock markets more volatile through large amounts of stimulus money that found its way into the stock markets.
2.4 The Residential Real Estate Market

2.4.1 Brief history

Real estate in China has become an increasingly important part of the economic growth since the Chinese government again began to experiment with allowing individuals to own private property again in 1979. Before 1953, private real estate in China was recognized by the Constitution of China. But from that date onwards, private ownership of real estate began to be transferred over to joint cooperatives which were part private part public, and by 1955 most industrial and commercial firms had been turned into cooperatives. However, before the cultural revolution in 1966 the private parts of these cooperatives were completely transferred into the public’s hands, resulting in China becoming a complete socialist country (Wang, 2016, p. 55). Housing in cities were allocated by the government, state owned enterprises (SOE) or Danwei, which was a type of working unit which supervised much of a workers’ life at the time, and could not be traded. Throughout the 80’s and 90’s the scope of allowing privately owned real estate increased through reforms with the goal of housing commercialisation. After the Asian Financial crisis in 1997, which slowed down China’s exports, the real estate market was further reformed to help pick up the pace for China’s economic growth by developing a domestic real estate market. Not only would the real estate market stimulate economic growth, but it would also lessen the burden on the Danwei’s and local governments who formerly administered housing allocation (Wang, 2016, p. 57). In 1998, Danwei’s had to grant housing allowances to workers for them to buy public-owned housing in order to further lessen burdens on the welfare system, which marked the revival of the Chinese housing market (Zhu, 2012, p. 247). The Law on Urban Real Estate Administration (LUREA) was enacted in 1994. This law specified two ways in which individuals could acquire rights to use land; either by land lease or through government allocation (mainly intended for projects with the public in interest such as housing for low-income families etc.). According to the Property Rights Law (wuquanfa) enacted in 2007, for residential purposes, individuals or firms can acquire land lease rights lasting for a maximum of 70 years. For industrial and tourism purposes, the land lease terms are respectively 50 and 40 years. Thus, in the strictest sense, individuals cannot own land in China, but only obtain rights to use the land by the government. However, when the land lease term for land with residential purposes expires, it will be automatically renewed. The law does however not specify whether such renewal would require a fee. Even though individuals and firms cannot own land, they can according to the Property Rights Law own “… immovable and movable
properties as their lawful incomes, houses, articles for daily use, tools of production and raw materials.”. This system does however have strong caveats since it in essence involves the government as the ultimate owner and distributor of land. Corruption in the real estate sector in China is a problem since local governments and various lower bureaucratic governmental units can collude with resourceful firms to make land allocation opaque even though there are regulations and directives in place to combat just that phenomenon. The real estate corruption increased through decentralization of the allocation process from the government to smaller bureaucratic units with rent seeking behaviour, and with most corruption in the real estate sector occurring during the process of transferring land use rights (Zhu, 2012, p. 251).

2.4.2 Buyers of residential housing
It can be argued that there are two types of buyers when it comes to residential housing, namely consumers and investors. Consumers would buy housing not out of main concern for expectations of future price increase, but to buy a place to live. Pure investors, on the other hand, would buy housing with the intent of renting it out and/or with expectations of future price increase. Rational investors would invest their capital in the venture they could gain the most return based on their willingness to risk, which would for many Chinese investors be a choice between the real estate market and the stock markets. Zhou et al. (2016) found that homeowners are less likely to invest in the stock market but owners of multiple housing invest more in the stock market than those who only own one residential housing. Expectations of housing price increase were found to have a negative effect on investments on the stock market, and especially so for the homeowners with multiple housing. The opposite was true when there was an expectation of price decrease in housing price. Their result could mean that when housing prices are increasing, which brings about expectations of further price increase, then investors might invest more in the real estate market rather than the stock market. Buyers of residential housing in China has been found to have herding behaviour when the return of housing is increasing, while a decrease in market return only leads to herding during market turbulence (Lan, 2014).

2.4.3 Recent developments
According to official numbers from the National Bureau of Statistics of China, real estate investments accounted for 23.7% of total fixed assets investments in 2003 and accounted for 23.9% in 2015. Real estate made out the biggest part of fixed investments in 2013, when it reached 26.6% of total fixed assets investments. Real estate investment has since the
Beginning of the millennium become an increasingly important part of the GDP of China, where it made out 4% of GDP in 1997 to 19.5% in 2015. Out of this, investment in residential real estate made out 6.7% in 1997 and grew to 11.64% in 2015. Urban residential real estate investments increased from 4.2% in 1997 to 10.6% in 2015.

By looking at the data of real estate as a part of the economy, it appears that the portion of real estate that accounts for GDP reached a top in 2014 and has decreased in 2015. The growth rate in total fixed assets investments has slowed down from 23.8% as the highest annual growth rate in 2011, after the financial crisis to 9.8% in 2015. Excess production capacity and inventory accumulation can be attributing to the growth decline in the manufacturing and real estate investments according to China’s Macroeconomic Outlook: Quarterly Forecast and Analysis Report released by Center for Macroeconomic research of Xiamen in February 2016. The report also noted that investment has slowed down all across the board, and pointed to decreased demand as the key factor. The decreased domestic demand also being the result of weaker foreign demand and structural deflation. The authors of the report do however state that the slowdown in investment and the unbalanced supply-demand relationship should be short term because the excess supply is due to China transitioning from an average- to above average-income economy. Excess supply of real estate can be explained by local governments relying too heavily on land rights sales to finance public spending. Whereas demand for real estate is driven by positive historical gains, low real interest rates, lack of alternative financial assets and capital account restrictions.
In figure 9, the growth rates shown in figure 8 is converted into price levels with January 2006 as starting date. Examining figure 9, one can see that the housing prices in the different city tiers vary significantly. The Tier 1 cities show the highest price increases, followed by Tier 2, Tier 3, and Tier 4 & 5.

Figure 9, Monthly Housing price index divided by city tiers (%)

Source: National Bureau of Statistics of China

Figure 10, Monthly city price level divided by city tiers and SSE Composite Index with january 2006 as base

Note: the secondary y-axis to the right are the values of the SSE Comp Index with january 2006 as base
don source: National Bureau of Statistics of China

In figure 9, the growth rates shown in figure 8 is converted into price levels with January 2006 as starting date. Examining figure 9, one can see that the housing prices in the different
tiers grew at similar rates up until 2015, when first-tier housing prices suddenly shot started to rise very quickly. Second-tier city housing prices also increased, but to a lesser extent than that of the first-tier cities. Third-tier city housing prices increased somewhat and reached a similar level as the fourth- and fifth-tier housing prices. The financial crisis in 2007 may indeed to have had a negative impact on the housing prices in all of the categories shown by low or negative monthly growth rates until the beginning of 2009, when the housing prices began to rise again, however it is not the only reason. Prior to the financial crisis in 2007, in the period 2004-2007, housing prices were rising fairly quickly which lead the Chinese government to tighten loan-to-value ratios and raised mortgage interest in 2007 (Ahuja et al. 2010, p. 6). These restrictive measures lasted until the end of 2008, after which the rent-to-price ratio again began to increase. Jin (2014, p. 1878-1879) stated that the decrease in real estate prices in 2008 upset the price expectations for real estate owners which did not start rising again until 2009. He further concluded that stimulus money introduced by the Chinese government after 2008 went not only into the capital markets but also found its way to the real estate market which drove up the prices of real estate. The Chinese government tried to contain the rapid rise of real estate prices by influencing the interest rates which again upset future real estate expectations resulting in volatile housing price growth in bigger cities in 2010 and 2011. His findings appear to match the graphs in figure 8, where one can see very unstable growth rates from tier 1 and tier 2 cities in the period of the financial crisis until 2011.

The stock bubble burst which began in the summer of 2015 appears to have had a negative effect on the housing prices of the first-tier cities as the housing prices growth decreased in the initial period of the stock market bubble burst (or new restrictions?). It is reasonable to think that the apparent co-movements of the two markets during the financial crisis might be due to the spillover-effects from the situation of the macroeconomy as a whole. According to Schmidt (2009, p. 2), exports fell by 2.2% in November 2008 from 20% in the previous month, while imports fell by over 21% in the following month. The annual GDP growth in 2008 was at 9% which was a substantial decrease from the 2007 GDP growth of 13%. Shu et al. (2011, p. 38-39) found feedback effects between the leading index of macroeconomic prosperity, which among other indicators contain export turnover, the output of light industry and sales of goods, and the real estate investment index.

The price levels for real estate vary greatly from city to city, where one for example have that, according to data from SouFun-CREIS which have reported average housing prices for one hundred cities in China since 2010, the average per square meter price of residential housing
in Shanghai, a tier 1 city, was 45,847 \textit{yuan} in November of 2016, and in the same period the per square meter price in Chengdu, a tier 2 city, was 8,234 \textit{yuan}. To curb rocketing housing prices in Shanghai, tighter restrictions on mortgage down payment was introduced in march in 2016 which required no less than 50\% down payment on normal secondary residential housing, and 70\% down payment on secondary residential housing which exceeded 144 square meters (Shanghai Municipal People's Government, 2016). Similar restrictions were imposed in Beijing in march this year, which raised the required mortgage down payment for normal secondary residential housing from 50\% to 60\%, and for housing above 144 square meters were raised from 70\% to 80\% (Beijing Municipal Commission of Housing and Urban-Rural Development, 2017). Currently, 18 of major cities in China have implemented similar measures to Shanghai and Beijing and placed restrictions on home purchases which include stricter mortgage requirements, limiting the number of housing locals can own and requiring non-locals to have to pay local taxes for several years before being allowed to buy housing (Zheng, 2017). Whether or not these measures will work as intended to calm down the rising prices remains to be seen.

After this relatively brief introduction to the two markets, I now move on to the analysis of the housing markets on the SSE, beginning with the methodology that is used.

3. Methodology

In this section, I will go through the main methodology used in this paper to analyse the possible effects the housing price in different city tiers in China has on the SSE. I will be employing a general-to-specific approach to examine the dynamics between the two markets in this paper, meaning that I start out with a general model which then is simplified through econometric procedures until I am left with congruent and well-specified models. The analysis begins by examining the properties of the variables used in the analysis: first, the attributes and distributions of the variables are briefly discussed through a table of descriptive statistics and a QQ-plot. Further, the variables are tested for stationarity through an adjusted Dickey-Fuller unit root test. The model used in this paper to analyse the problem is a straightforward autoregressive distributed-lag model (ADL) with finite lags using Ordinary Least Squares (OLS), in which the Shanghai Stock Exchange Composite Index is the dependent variable, and the different city tiers are the independent variables. After the ADL model is estimated and insignificant lags removed, weak exogeneity of the independent variables are tested using the Durbin-Wu-Hausman test. Further dynamic analysis is then conducted using a
reparameterization of the ADL model called the error-correction model (ECM). Through the ECM, short-run and long-run effects of the housing price indexes from the different city tiers can be identified and discussed. Finally, the constancy of parameters and the variance of the coefficients is evaluated using graphs plotted through recursive estimation. The methodology of these steps and tests are described below.

3.1 The model

The unrestricted ADL model with m finite lags:

$$SHCOMP_t = \alpha + \sum_{i=1}^{m} \gamma_i SHCOMP_{t-i} + \sum_{i=0}^{m} \beta_i Tier1_{t-i} + \sum_{i=0}^{m} \mu_i Tier2_{t-i}$$

$$+ \sum_{i=0}^{m} \eta_i Tier3_{t-i} + \sum_{i=0}^{m} \lambda_i Tier45_{t-i} + \epsilon_t$$  \hspace{1cm} (1)

Where SHCOMP is the month by month growth rate of the Shanghai Exchange Composite Index in period t. \(\alpha\) is the constant term. \(\gamma, \beta, \mu, \eta\) and \(\lambda\) are the coefficients of the variables Tier1, Tier2, Tier3 and Tier45 respectively. The model contains m lags. \(\epsilon_t\) is the error term. The analysis will be done in the following way: multiple time series will be run where zero-restrictions are put on all but one city category for every regression. This is done with the intention of finding out which city category’s housing price index has the most (if any) effect on the stock market index. The ADL model was selected for its versatile use in dynamic analysis with the possibility of a reparameterization into the ECM model which is further described in a sub-section below.

3.2 Unit root

Before beginning my analysis of the model, the nature of the variables must first be established. The variables of interest are first tested to see if they are unit root processes, i.e. being non-stationary, or if they are stationary. If the variables are unit root processes, then OLS cannot be used as it can provide invalid estimates and standard asymptotic inference theory cannot be applied as the danger of spurious regression will be present (Granger and Newbold, 1974, p. 117-119). However, if the variables are stationary, then standard inference
theory applies, which makes the job of analysing much easier. To determine if the variables have unit roots, Augmented Dickey-Fuller (ADF) tests will be applied. The ADF-test tests the null hypothesis of a unit root in a time series while the alternative hypothesis is that of stationarity. The ADF is called the augmented Dickey-Fuller test because of the addition of dynamics in the model when lags are added. The addition of lags is done to reducing autocorrelation of the residuals in the time series. The ADF(m) test with a constant term but no trend on SHCOMP with m lags is shown below:

\[
\Delta SHCOMP = \alpha + (\beta - 1)SHCOMP_{t-1} + \sum_{i=1}^{m} y_i \Delta SHCOMP_{t-i} \tag{2}
\]

More formally, the null hypothesis, \( H_0: \beta = 1, \) or \( \delta = 0 \) if we rename \( (\beta - 1) \) into \( \delta \). Whereas the alternative hypothesis is \( H_1: \delta \neq 0 \). The more negative the t-statistic is, the stronger the rejection of the null hypothesis of unit root is. The critical values for this test follow the Dickey-Fuller distribution. The DF-distribution is dependent on both the data generating process (DGP) and of the model that is estimated. This means that the distribution of the t-statistic, \( \hat{\delta} \), that I estimate depends on whether a constant term and/or a trend term is added. “A rule of thumb is to add deterministic trends that are of a higher order than what one thinks is the case in the DGP” (Bårdsen and Nymoen, 2014, p. 290).

### 3.3 Model specification and goodness-of-fit

The modeling approach in this paper is general-to-specific, therefore the importance of properly reducing the model is paramount. Reducing the model by simply removing insignificant variables may result in a badly specified model unless one takes into account information loss because of reduction as well as how reliable the results from such a model might be. Information criteria and various misspecification are used to make sure the restricted model is well specified. The misspecification tests include tests for error autocorrelation, normality, heteroscedasticity and functional form, and are included in the software package “PcGive” as a part of the econometrics software “Oxmetrics” developed by Jurgen A. Doornik. The misspecification tests employed here are described briefly below, whereas further information about these misspecification tests can be found in volume 1 of the manual for “PcGive 14".
3.3.1 Information Criteria

To determine the appropriate number of lags in the ADL there exists several different information criteria that can help us decide on lag length. According to Liew (2004), the Hannan-Quinn Criterion (HQC) identifies the correct lag-length better than other lag selection criteria when the number of observations exceeds 120. However, the study also found that Aikaike Information Criterion (AIC) and Final Prediction Error (FPE) had the least probability of under-estimation among the criteria in the study. The lower the values of the AIC and HQC information criteria the model has, the lower the information loss the model has compared to models with higher values of these information criteria. The formulas for the AIC and HQC are shown below:

AIC: \( \ln(\sigma^2) + \frac{2k}{T} \)

HQC: \( \ln(\sigma^2) + \frac{2k(\ln(\ln(T))))}{T} \)

\( \sigma^2 \) is the variance estimator, k is the number of parameters and T is the number of observations. As the dataset used in this paper contains 131 observations for every variable, both the AIC and HQC together with t-values, F-tests and misspecification tests are consulted when removing insignificant terms and lags to get congruent and well-specified models.

3.3.2 Error Autocorrelation test

The error autocorrelation test is based on Harvey (1990, p. 275), and reported in F-form. The null hypothesis of this test indicates that there is no autocorrelation (that errors are white noise), whereas a rejection of the null hypothesis suggests the presence of autocorrelation. If one rejects the null hypothesis of no autocorrelation, then that means the error terms may be autocorrelated. Autocorrelation of the error terms could arise because a relevant independent variable is omitted, which may bias the coefficient estimates of the other variables. The greatest issue of autocorrelation is that the true value of the variance may be larger than the estimates OLS gives, resulting in lower standard errors and thus higher t-statistics, which may result in falsely rejecting the null hypothesis of variables in the regression. In the worst case, the results of a model with autocorrelated error terms are useless.
3.3.3 Normality test

The normality test is based on Doornik and Hansen (1994, p. 2), where the null hypothesis is that of normal distribution using critical values derived from the Chi distribution, and the alternative hypothesis rejects normality. The test statistic is formally represented in (x):

\[ t_{\text{norm}} = z_1^2 + z_2^2 \overset{\text{approx}}{\sim} \chi^2(2) \]  

Where \( z_1 \) and \( z_2 \) are transformed skewness and kurtosis based on D’Agostino (1970, p. 680), and are also detailed in the appendix of Doornik and Hansen (1994, p. 7). If the null hypothesis of normal distribution is rejected, then the dependent variable may follow a different distribution than the normal distribution. This may not necessarily be a problem when using OLS regression due to the Gauss-Markov Theorem. According to this theorem, OLS is the best linear unbiased estimator (BLUE) as long as the errors have a zero mean, are uncorrelated, and have constant variance.

3.3.4 Heteroscedasticity test

The heteroscedasticity test used in this paper is based on White (1980, p. 821), and is done with an auxiliary regression of the dependent variable squared, which is then regressed on the original variables and their squares. The null hypothesis of this test is unconditional homoscedasticity, and the alternative hypothesis is that the variance of the dependent variable depends on the independent variables and their squares, i.e. that the variance is not constant over time. When the variance is not constant over time, the true variance, and thus standard errors, are underestimated resulting in the possible erroneous rejection of the null hypothesis.

3.3.5 Functional form test

A test for functional form, the Regression Specification Error Test (RESET23), is employed which is based on Ramsey (1969, p. 361). This test tests the null hypothesis of correct specification of the model against the alternative hypothesis: that powers of the dependent variable have been omitted. The RESET23 test tests for square and cubic powers of the dependent variable, whereas the RESET test in Ramsey’s paper only tests for cubic powers. Rejection of the null hypothesis would suggest that the functional form of the model should be revised for a better goodness-of-fit.
3.4 Weak exogeneity

To do single-equation modelling efficiently, one must first test if the explainable variables are weakly exogenous. This is important because if the variables are not at least weakly exogenous, then one cannot simply use the conditional model but also have to estimate the marginal model in the system because the marginal model contains information that could increase the efficiency of the estimation of the parameters in the conditional model (Bårdsen and Nymoen, 2014, p. 273). The conditional model can be interpreted as a regression model representation of the vector autoregressive model (VAR), and the second row in the VAR represents the marginal model. Below is the conditional model, (3), and the marginal model, (4), with zero-restrictions in place on the other city tier variables.

\[
SHCOMP_t = \alpha + \sum_{i=1}^{m} \gamma_i SHCOMP_{t-i} + \sum_{i=0}^{m} \beta_i Tier1_{t-i} + \epsilon_t \tag{3}
\]

\[
Tier1_t = \alpha + \sum_{i=0}^{m} \gamma_i SHCOMP_{t-i} + \sum_{i=1}^{m} \beta_i Tier1_{t-i} + \epsilon_t \tag{4}
\]

To test for weak exogeneity, the Durbin-Wu-Hausman (DWH) test will be applied. The DWH-test will be employed with the artificial regression used in Davidson and MacKinnon (1993, p. 239). This is done by first regressing the marginal model, (4), then saving the residuals (or fitted values) and then testing the significance of the residuals from the marginal model in the conditional model, (3), with a t-test. The null hypothesis of this test implies that the marginal model does not contain any information that could increase efficiency of the conditional model and that the explainable variable tested is weakly exogenous, while the alternative hypothesis rejects weak exogeneity.

3.5 Error Correction Model

The Error Correction Model (ECM) is a reparameterization of the ADL model specified above. Since it is only a reparameterization of the ADL model, the statistical properties of the ADL model are unchanged. Alogoskoufis and Smith (1991, p. 104), clarified the concept of ECM by dividing it into three main directions of applications: The terminology of ECM was first introduced by Philips (1954, 1957) when analysing feedback control mechanisms for
stabilisation policy. The usage and form of the ECM evolved further through Sargan (1964) and Hendry & Mizon (1978), and later by Engle and Granger (1987) with the context of error-correction and cointegration. The ECM formulation that will be used in this paper is mainly the one proposed by Hendry. Hendry emphasized the ECM as a reparameterization of the ADL that can give estimates of short-run dynamics, error-correction and long-run relationships. Though the ECM is most commonly used for cointegration purposes when the data is non-stationary (Engle and Granger 1987), it can be used with stationary data as well (Keele, 2004). The motivation behind using the model is to model both short-run and long-run effects and see whether there are any long-run effects of the city tiers housing price growth rate on the Shanghai Composite Index.

The ADL specified in (1) can be reparametrized into the following multivariate single-equation ECM model with helpful notation from Bårdsen and Nymoen (2014, p. 233):

\[
\Delta \text{SHCOMP}_t = \alpha + \sum_{i=0}^{m-1} \gamma_i^{\Delta} \Delta \text{SHCOMP}_{t-i} + \sum_{i=0}^{m-1} \beta_i^{\Delta} \Delta \text{Tier1}_{t-i} + \sum_{i=0}^{m-1} \mu_i^{\Delta} \Delta \text{Tier2}_{t-i} \tag{5}
\]

\[
+ \sum_{i=0}^{m-1} \eta_i^{\Delta} \Delta \text{Tier3}_{t-i} + \sum_{i=0}^{m-1} \lambda_i^{\Delta} \Delta \text{Tier45}_{t-i} + \left( \sum_{j=1}^{m} \gamma_j - 1 \right) \text{SHCOMP}_{t-1} + \sum_{j=0}^{m} \beta_j \text{Tier1}_{t-1} + \sum_{j=0}^{m} \mu_j \text{Tier2}_{t-1} + \sum_{j=0}^{m} \eta_j \text{Tier3}_{t-1} + \sum_{j=0}^{m} \lambda_j \text{Tier45}_{t-1} + \varepsilon_t
\]

Which can also be written as:

\[
\Delta \text{SHCOMP}_t = \sum_{i=0}^{m-1} \gamma_i^{\Delta} \Delta \text{SHCOMP}_{t-i} + \sum_{i=0}^{m-1} \beta_i^{\Delta} \Delta \text{Tier1}_{t-i} + \sum_{i=0}^{m-1} \mu_i^{\Delta} \Delta \text{Tier2}_{t-i} \\
+ \sum_{i=0}^{m-1} \eta_i^{\Delta} \Delta \text{Tier3}_{t-i} + \sum_{i=0}^{m-1} \lambda_i^{\Delta} \Delta \text{Tier45}_{t-i} \\
- \left( 1 - \sum_{j=1}^{m} \gamma_j \right) \left[ \text{SHCOMP}_{t-1} - \frac{\alpha}{1 - \sum_{j=1}^{m} \gamma_j} - \frac{\sum_{j=0}^{m} \beta_j}{1 - \sum_{j=1}^{m} \gamma_j} \text{Tier1}_{t-1} - \frac{\sum_{j=0}^{m} \mu_j}{1 - \sum_{j=1}^{m} \gamma_j} \text{Tier2}_{t-1} \\
+ \frac{\sum_{j=0}^{m} \eta_j \text{Tier3}_{t-1}}{1 - \sum_{j=1}^{m} \gamma_j} - \frac{\sum_{j=0}^{m} \lambda_j \text{Tier45}_{t-1}}{1 - \sum_{j=1}^{m} \gamma_j} \right] + \varepsilon_t
\]

And more compactly by renaming the long-run coefficients:
\[
\Delta SHCOMP_t = \sum_{i=0}^{m-1} \gamma_i^+ \Delta SHCOMP_{t-i} + \sum_{i=0}^{m-1} \beta_i^+ \Delta Tier1_{t-i} + \sum_{i=0}^{m-1} \mu_i^+ \Delta Tier2_{t-i} + \sum_{i=0}^{m-1} \eta_i^+ \Delta Tier3_{t-i} + \sum_{i=0}^{m-1} \lambda_i^+ \Delta Tier45_{t-i} \\
- \tau [SHCOMP_{t-1} - \alpha^* - \beta^* Tier1_{t-1} - \mu^* Tier2_{t-1} + \eta^* Tier3_{t-1} - \lambda^* Tier45_{t-1}] + \epsilon_t
\]

Where

\[
\gamma_i^+ = - \sum_{j=1}^{m} \gamma_j, \quad i = [1, m-1], \\
\beta_i^+ = - \sum_{j=1}^{m} \beta_j, \quad i = [1, m-1], \\
\mu_i^+ = - \sum_{j=1}^{m} \mu_j, \quad i = [1, m-1], \\
\eta_i^+ = - \sum_{j=1}^{m} \eta_j, \quad i = [1, m-1], \\
\lambda_i^+ = - \sum_{j=1}^{m} \lambda_j, \quad i = [1, m-1]
\]

\[
\beta_0^+ = \beta_0, \quad \mu_0^+ = \mu_0, \quad \eta_0^+ = \eta_0, \quad \lambda_0^+ = \lambda_0
\]

The first differences are represented by \(\Delta\); for example, the first difference of \(SHCOMP_t\) would be \(\Delta SHCOMP_t = SHCOMP_t - SHCOMP_{t-1}\). Notice that this again is merely a reparameterization of the ADL-model, and as such the statistical properties of the model itself are unchanged, i.e. the properties of \(\epsilon_t\) is unchanged. The equation inside the clamps in (6) is the deviation of \(SHCOMP_{t-1}\) from its long-run equilibrium value. Thus, if there is a long-run solution, then \(\tau\) is significantly different from zero and tells us the speed of how the SHCOMP adjusts back to the equilibrium value. \(\beta^*\) is then the long-run causal effect of Tier1 on SHCOMP and since I examine growth rates in this paper, \(\beta^*\) would represent the long-run elasticity of Tier1 on SHCOMP. The first differences of the lags are the impact effects, or the short-run effects. The model is quite flexible so the distributed lag lengths do not need to be the same for different variables, the levels of SHCOMP and Tier1 can be on different lags instead of the first lag as in the model above, and the differences can also be adjusted over more than one period if need be. The long-run effects are invariant to different formulations of the ECM, but the lags of the first differences in the example above are not. The long-run solution will also include a Wald-test which has a null hypothesis of all long-run coefficients except the constant are zero.
3.6 Invariance of the variables and constancy of parameters

For the estimated models to be reliable, it is important that any possible structural breaks are taken into account. If there exist structural breaks and they are not accounted for, then the estimated coefficients for in the model may become unreliable because of structural changes might have taken place that fundamentally may change the way an economic variable acts. Recursive estimation will be used to examine whether or not the estimated coefficients are invariant or if there exist structural breaks and will be initialized with 10 observations. Should structural breaks be detected in the data, then dummy variables or step dummies might be included in the models to account for the regime shifts. The recursive graphics estimated will include recursively estimated plots of the coefficients +/- 2 SE (95% confidence interval), 1-step residuals +/- 2 SE, and three variants of Chow: one-step Chow test, break-point Chow test and forecast Chow test. Requirements for the Chow tests to be suitable for use are that the time series must be stationary and that regressors must be exogenous, which will be examined through the sections on unit root and weak exogeneity mentioned above. The one-step Chow test statistic has the form:

\[
\frac{(RSS_t - RSS_{t-1})/(t - k - 1)}{RSS_{t-1}} \sim F(1, t - k - 1)
\]

under the null hypothesis of constant parameters if the distribution of the dependent variable is normal. RSS is the residual sum of squares and k is the number of parameters estimated from the sample size t. The one-step Chow test used in this paper is based on the prediction interval test in Chow (1960) but is estimated recursively over the sample and in effect tests one step ahead forecast failure at each time step and was introduced in PCgive 4.0 as a part of tests for misspecification together with the breakpoint and forecast Chow tests (Nielsen and Whitby, 2015, p. 1). The breakpoint Chow (also called Ndown Chow) test are sequences of Chow tests where the number of forecasts goes down from \( N = T - M + 1 \) to 1. It has the form:

\[
\frac{(RSS_T - RSS_{t-1})/(t - k - 1)}{RSS_{t-1}(T - t + 1)} \sim F(T - t + 1, t - k - 1)
\]

For \( t = M, \ldots, T \).

The forecast Chow (also called Nup Chow) test, like the breakpoint Chow test, is a sequential Chow test. However, in the forecast Chow test, the forecast horizon increases from M to T.
This test tests the model over 1 to $M - 1$ against an alternative model which allows any form of change over M to T. Forecast Chow test has the form:

$$\frac{(RSS_t - RSS_{M-1})/(M - k - 1)}{RSS_{M-1}(t - M + 1)} \sim F(t - M + 1, M - k - 1)$$

For $t = M, ..., T$.

Persistent breaks in these Chow tests would indicate structural breaks.

4. Empirical Analysis

This part contains data description, empirical results and a discussion of the results. In the first sub-section, data description, a short description of the dataset is presented for the reader to understand how the dataset used was constructed. In the second sub-section, results from the ADL and ECM models are presented as well as results from misspecification tests to test the robustness of the models. In the third and last sub-section, I discuss the results obtained in the second sub-section. The software used for the analysis is Oxmetrics 6.01 with PcGive 13.

4.1 Data description

The SSE Composite Index will be used as the dependable variable in this analysis. The Shanghai Stock Exchange Composite Index is a capitalization-weighted index. The index tracks the daily price performance of all A-shares and B-shares listed on the Shanghai Stock Exchange. The index was developed on December 19, 1990 with a base value of 100. Index trade volume on Q is scaled down by a factor of 1000 (Bloomberg). The month by month growth rates used in this paper are calculated by taking the closing value of the composite index on the last day of each month since January 2006. These growth rates are then made centered around 0. The closing value of the Composite Index was chosen because the closing value is a good measure of market sentiments from period to period.

The independent variables in this paper are made using the housing price index developed by the National Bureau of Statistics in China (NBS). NBS has since July 2005 released monthly housing price indices on 70 large- and medium-sized cities in China; the so-called NBS 70-city index. I will briefly describe the method used by NBS to develop the NBS 70-city index here. Commercially traded residences were separated into three size categories: below 90 square meter housing, 90-144 square meter housing and above 144 square meter housing categories. Samples of reported transaction prices for housing complexes were gathered by NBS every month and then averaged. The averaged reported transaction price for the current
month was compared with the previous month’s transaction price in the same housing complex. Then in the last step, city level housing price changes were calculated as an average, weighted by transaction volume, of all complexes’ price change. There exist however critiques against this dataset. One critique being that the NBS adjusted its computational strategies in this series in 2011, which makes the data before and after this date not strictly comparable. Critique has also been raised at the data for not matching the reality of the housing price markets on average, because of the price growth being reported to be much lower than that noted by observers (Wu et al. 2016, p. 93). Fang et al. (2016, p.21) compared the official housing price index against one they made, which consisted of mortgage data from 120 major cities in China. They found that their dataset showed synchronized co-movements with the official dataset, but that the official dataset was very smooth and did not show much real price housing growth during the last decade. This critique is kept in mind when I analyse the empirical results in the next section of the paper.

To analyse how housing price growth in different city tiers may affect the stock markets, I sort cities into their corresponding city tiers, and then the mean monthly price growth of each city tier is computed using these cities’ month by month price growth. There are no official lists that divide the cities of China into city tiers, however regardless of this, the city classification system is widely used to describe the size and importance of cities in China. In this paper, I will use the classification system proposed by the consulting firm Nexus-Pacific, which classifies the cities into tiers ranging from 1 to 5: Tier 1 cities are the main metropolitan centres with large populations. Tier 2 cities are often provincial capitals and other developed cities of size and importance. Tier 3 cities are smaller and not as widely known as tier 2 cities but are still attributed with high economic growth and development. Tier 4 and 5 cities are cities of smaller size and importance, with varying economic growth. Which tier a city belongs to is open for debate at all times, and cities may and will move between the city tiers as the local economy changes. Table 1 shows which cities in this dataset belong to which category as well as some description on the city tier which is taken from Nexus-Pacific.
<table>
<thead>
<tr>
<th>City category</th>
<th>Variable name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier 1</td>
<td>Tier1</td>
<td>Large densely populated urban metropolises with huge economic, cultural and political influence in China. Tier 1 cities attract great attention from foreign enterprises due to income levels much higher than the national average, and larger middle class representation and increasing consumption habits. Cities that fall within this category represent China’s most developed markets in terms of consumer behaviour. <strong>This list includes:</strong> Shanghai, Beijing, Guangzhou, Shenzhen, Tianjin, Chongqing</td>
</tr>
<tr>
<td>Tier 2</td>
<td>Tier2</td>
<td>These are generally made up of Provincial capitals, sub-provincial cities, Special Economic Zones (SEZ), and other more developed cities with cultural and economic influence. Over the past decade, Tier 2 cities have received increased attention and investment from foreign companies due to lower labour costs, less competition, lower operating costs for retailers, and rapidly increasing consumer spending habits. <strong>This list includes:</strong> Changchun, Changsha, Chengdu, Dalian, Fuzhou, Guiyang, Haikou, Hangzhou, Harbin, Hefei, Hohhot, Jinan, Kunming, Lanzhou, Nanchang, Nanjing, Nanning, Ningbo, Qingdao, Sanya, Shenyang, Taiyuan, Wuhan, Wuxi, Xiamen, Xian, Zhengzhou</td>
</tr>
<tr>
<td>Tier 3</td>
<td>Tier3</td>
<td>These are generally made up by open coastal cities, high income cities, and cities with significant economic development which are less known. <strong>This list includes:</strong> Beihai, Guilin, Huizhou, Jinhua, Qinhuangdao, Quanzhou, Shijiazhuang, Tangshan, Wenzhou, Xining, Xuzhou, Yantai, Yinchuan, Zhangjiagang</td>
</tr>
<tr>
<td>Tier 4 &amp; 5</td>
<td>Tier45</td>
<td>Cities of smaller importance and lower economic development than tier 3 <strong>This list includes:</strong> Baotou, Jilin, Luoyang, Yangzhou, Xiangfan, Yichang, Yueyang</td>
</tr>
</tbody>
</table>

Source: Nexus-pacific.com

The city tier list proposed by Nexus-Pacific contains more cities than what is shown in table 1, however only the cities in the city tier list that were also in the NBS-70 list have been included in the dataset used in this paper. Tier 4 and tier 5 were merged into one variable in the dataset used in this paper, as the differences between the cities of these tiers are not that great. It must be noted that the city tier system is not close to being a perfect way of classifying cities, as it does not take into account the different natures of the cities it includes. A better system of city classification could be one that accounted for the differences in the cities such as the one proposed by The Demand Institute, in which cities are classified into...
categories such as Super, Affluent, Satellite, Integrated industrial, Tourism and so on according to the type of city and the main economic drivers of that city. However, the conventional tier classification system was chosen for this paper regardless, as it is more commonly referred to and used both in literature and in the news.

The dataset made for this paper has a total of 131 monthly observations, and begins from January 2006 until November 2016.

4.2 Empirical results

In this section, the results of tests and methodology briefly explained in the previous section will be reported and discussed. The order of this section is as follows: first some statistical properties of the variables in the analysis will be discussed, then tests for unit root and exogeneity properties will be employed which are necessary for efficient ADL analysis. The results from the ADL analysis are then presented and discussed followed by results from the ECM reparameterization in order to analyse long-run effects. Finally, invariance of coefficients and constancy of parameters is examined at the end of this section.

First, I examine the descriptive statistics table that is tabulated below:

<table>
<thead>
<tr>
<th>Variables</th>
<th>SHCOMP</th>
<th>Tier1</th>
<th>Tier2</th>
<th>Tier3</th>
<th>Tier45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
<td>131</td>
</tr>
<tr>
<td>Mean</td>
<td>1,10</td>
<td>0,53</td>
<td>0,32</td>
<td>0,20</td>
<td>0,22</td>
</tr>
<tr>
<td>Std.Devn.</td>
<td>8,9412</td>
<td>0,8263</td>
<td>0,5320</td>
<td>0,4514</td>
<td>0,4993</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0,27</td>
<td>1,01</td>
<td>0,45</td>
<td>-0,11</td>
<td>0,30</td>
</tr>
<tr>
<td>Excess Kurtosis</td>
<td>0,86</td>
<td>2,06</td>
<td>1,07</td>
<td>0,15</td>
<td>1,51</td>
</tr>
<tr>
<td>Minimum</td>
<td>-24,63</td>
<td>-1,08</td>
<td>-0,96</td>
<td>-0,96</td>
<td>-1,20</td>
</tr>
<tr>
<td>Maximum</td>
<td>27,45</td>
<td>3,90</td>
<td>2,16</td>
<td>1,31</td>
<td>1,87</td>
</tr>
<tr>
<td>Asymptotic test:Chi^2(2)</td>
<td>5.6618 [0.059]</td>
<td>45,316 [0.000]**</td>
<td>10,697 [0.005]**</td>
<td>0,4111 [0.814]</td>
<td>14,363 [0.001]**</td>
</tr>
<tr>
<td>Normality test:Chi^2(2)</td>
<td>5.8348 [0.054]</td>
<td>18,278 [0.000]**</td>
<td>7,4761 [0.024]*</td>
<td>0,9289 [0.629]</td>
<td>12,395 [0.002]**</td>
</tr>
</tbody>
</table>

* = 5% significance, ** = 1% significance

The distributions of Tier1, Tier2 and Tier45 seem to have heavy tails indicated by excess kurtosis, while Tier1 also has substantial skewness in table 2. If excess kurtosis exceeding 1 or is below -1, then that suggests the distribution is too peaked or too flat respectively.

Skewness refers to asymmetry in the distribution with respect to the mean. If skewness exceeds 1 or is below -1 then the skewness suggests that the distribution is substantially right-skewed (long right tails) or left-skewed (long left tails) respectively. Tier1 seems substantially
right-skewed with a skewness of 1.01. *Tier1*, *Tier2* and *Tier45* have substantial excess kurtosis, pointing in the direction of extreme outliers that makes the tails of their distributions thick and the peaks sharper than if they were normally distributed, i.e. lower probability of achieving values near the mean and high probability of extreme values compared to normally distributed variables. The results of the normality tests based on Doornik and Hansen (1994), which is a function of skewness and kurtosis, suggests that it is unlikely that *Tier1*, *Tier2* and *Tier45* have normal distributions, and especially so for *Tier1*. QQ-plots are presented in the appendix to tell us more about the possible distribution of the variables.

By examining the QQ-plots in A1 in the appendix, one can see that most of the deviations from the QQ-line in *Tier1*, *Tier2* and *Tier45* are at the tails and especially at the upper quantile. For *Tier1*, *Tier2* and *Tier45*, the bottom quantile residuals that deviate from the QQ-plot are mostly within the red 95% CI bands, however at the top quantile several of the deviations go beyond the CI. *Tier1* shows skewness in its somewhat flat U-shape but it does not seem to be very serious, however it looks like it might follow a gamma distribution by the shape of the QQ-plot and by looking at a histogram. However, aside from these deviations the distribution of *Tier1* does not look too much different from the normal distribution, so the inefficiency of using OLS in our analysis should not be too great, and thus p-values should still prove to be decent estimates. As long as the assumptions of the error term having a zero mean, being uncorrelated, and having constant variance are not broken, then OLS is still BLUE.

I proceed by testing for unit root with the use of the ADF-test. The ADF test was applied with a constant term and 1 lagged first difference. The results are tabulated below in table 3:

**Table 3: ADF tests (T=127, Constant; 5%=−2.88 1%=−3.48)**

<table>
<thead>
<tr>
<th>D-lag</th>
<th>t-adf</th>
<th>beta Y_1</th>
<th>sigma</th>
<th>t-DY_lag</th>
<th>t-prob</th>
<th>AIC</th>
<th>F-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-6.652**</td>
<td>0.23422</td>
<td>8.921</td>
<td>-1.217</td>
<td>0.2259</td>
<td>4.4</td>
<td>4.396</td>
</tr>
<tr>
<td>0</td>
<td>-9.743**</td>
<td>0.14353</td>
<td>8.939</td>
<td>4.4</td>
<td>0.2259</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tier1:**

<table>
<thead>
<tr>
<th>D-lag</th>
<th>t-adf</th>
<th>beta Y_1</th>
<th>sigma</th>
<th>t-DY_lag</th>
<th>t-prob</th>
<th>AIC</th>
<th>F-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4.074**</td>
<td>0.76348</td>
<td>0.5186</td>
<td>1.268</td>
<td>0.2071</td>
<td>0.2259</td>
<td>4.396</td>
</tr>
<tr>
<td>0</td>
<td>-3.862**</td>
<td>0.78733</td>
<td>0.5199</td>
<td>0.2071</td>
<td>0.2259</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tier2:**

<table>
<thead>
<tr>
<th>D-lag</th>
<th>t-adf</th>
<th>beta Y_1</th>
<th>sigma</th>
<th>t-DY_lag</th>
<th>t-prob</th>
<th>AIC</th>
<th>F-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4.443**</td>
<td>0.74913</td>
<td>0.3216</td>
<td>1.801</td>
<td>0.0741</td>
<td>0.2246</td>
<td>0.0741</td>
</tr>
<tr>
<td>0</td>
<td>-4.041**</td>
<td>0.78264</td>
<td>0.3244</td>
<td>0.2071</td>
<td>0.2246</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tier3:**

<table>
<thead>
<tr>
<th>D-lag</th>
<th>t-adf</th>
<th>beta Y_1</th>
<th>sigma</th>
<th>t-DY_lag</th>
<th>t-prob</th>
<th>AIC</th>
<th>F-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-3.356*</td>
<td>0.79745</td>
<td>0.2912</td>
<td>-1.443</td>
<td>0.1516</td>
<td>0.2444</td>
<td>0.1516</td>
</tr>
<tr>
<td>0</td>
<td>-4.043**</td>
<td>0.76869</td>
<td>0.2924</td>
<td>0.2071</td>
<td>0.2444</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tier45:**

<table>
<thead>
<tr>
<th>D-lag</th>
<th>t-adf</th>
<th>beta Y_1</th>
<th>sigma</th>
<th>t-DY_lag</th>
<th>t-prob</th>
<th>AIC</th>
<th>F-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-4.426**</td>
<td>0.6807</td>
<td>0.3729</td>
<td>-0.1441</td>
<td>0.8857</td>
<td>0.1949</td>
<td>0.8857</td>
</tr>
<tr>
<td>0</td>
<td>-4.902**</td>
<td>0.6758</td>
<td>0.3715</td>
<td>-1.443</td>
<td>0.1949</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*= 5% significance, **= 1% significance
The null hypothesis of unit root in the ADF tests is rejected for all the variables, thus I find that all the variables are stationary at a 1% significance level. Having testing that the variables are indeed stationary, I eliminate the risk of erroneous conclusions later due to the possibility of spurious causality as is the case with non-stationary processes. With these results at hand, I can use standard inference theory to analyse the time series variables in our model.

To be able to estimate an efficient ADL model with OLS, weak stationarity of the explainable variables is a requirement unless the error term is white noise. I employ the DWH-test to see if the explainable variables are weakly exogenous; if the explainable variables are weakly exogenous, then I can continue to analyse our conditional model without having to estimate the marginal model in order to not lose efficiency. Below I have tabulated the results of the ADL model and the DWH-test to test for weak exogeneity of the city tier variables on SHCOMP. First, ADL models with zero-restrictions on all but one of the city tiers were run with 8 lags on each variable. The models were then reduced by the removal of insignificant lags while minding tests for misspecification and trying to minimize the AIC and HQH information criteria. Misspecification tests had priority when restricting the model, and then HQH over AIC.

### Table 4, ADL results with DWH-test

**SHCOMP by OLS**

<table>
<thead>
<tr>
<th>Tier1 model</th>
<th>Constant</th>
<th>SHCOMP_1</th>
<th>SHCOMP_4</th>
<th>Tier1_2</th>
<th>Tier1_4</th>
<th>residuals Tier1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.968</td>
<td>0.124</td>
<td>0.279**</td>
<td>-2.724**</td>
<td>1.931</td>
<td>-</td>
</tr>
<tr>
<td>SE</td>
<td>(0.959)</td>
<td>(0.968)</td>
<td>(0.087)</td>
<td>(1.225)</td>
<td>(1.301)</td>
<td>-</td>
</tr>
<tr>
<td>F-test (4,118)</td>
<td>3.766 [0.006]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>0.968</td>
<td>0.124</td>
<td>0.279**</td>
<td>-2.724**</td>
<td>1.931</td>
<td>-0.479</td>
</tr>
<tr>
<td>SE</td>
<td>(0.962)</td>
<td>(0.088)</td>
<td>(0.09)</td>
<td>(1.23)</td>
<td>(1.31)</td>
<td>(1.176)</td>
</tr>
<tr>
<td>F-test (5,117)</td>
<td>3.025 [0.013]*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tier2 model</th>
<th>Constant</th>
<th>SHCOMP_4</th>
<th>SHCOMP_5</th>
<th>Tier2_3</th>
<th>Tier2_5</th>
<th>residuals Tier2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.135</td>
<td>0.279**</td>
<td>0.139</td>
<td>-6.087**</td>
<td>4.099*</td>
<td>-</td>
</tr>
<tr>
<td>SE</td>
<td>(0.903)</td>
<td>(0.085)</td>
<td>(0.086)</td>
<td>(1.817)</td>
<td>(1.880)</td>
<td>-</td>
</tr>
<tr>
<td>F-test (3,119)</td>
<td>6.616 [0.000]**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coefficient</td>
<td>1.182</td>
<td>0.289**</td>
<td>0.145</td>
<td>-6.994**</td>
<td>4.499*</td>
<td>-0.01675</td>
</tr>
<tr>
<td>SE</td>
<td>(0.904)</td>
<td>(0.085)</td>
<td>(0.086)</td>
<td>(1.931)</td>
<td>(1.903)</td>
<td>(1.639)</td>
</tr>
</tbody>
</table>

31
<table>
<thead>
<tr>
<th>F-test (4,122)</th>
<th>5.213 [0.001]**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier3 model</td>
<td>Constant SHCOMP_2 SHCOMP_3 SHCOMP_4 Tier3_1 residuals Tier3</td>
</tr>
<tr>
<td>Coefficient</td>
<td>1.281 0.114 0.041 0.259** -4.205* -</td>
</tr>
<tr>
<td></td>
<td>(0.833) (0.087) (0.087) (0.087) (1.725) -</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F-test (2,120)</th>
<th>6.375 [0.002]**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.270 0.113 0.040 0.260** -4.229* 1.739</td>
</tr>
<tr>
<td></td>
<td>(0.835) (0.086) (0.087) (0.087) (1.729) (2.655)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F-test (3,119)</th>
<th>4.43 [0.005]**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tier45 model</td>
<td>Constant SHCOMP_4 Tier45_3 - - residuals Tier45</td>
</tr>
<tr>
<td>Coefficient</td>
<td>1.295 0.242** -3.131* - - -</td>
</tr>
<tr>
<td></td>
<td>(0.834) (0.085) (1.522) - - -</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F-test (2,120)</th>
<th>5.997 [0.003]**</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.2752 0.243** -3.051* - - -0.714</td>
</tr>
<tr>
<td></td>
<td>(0.840) (0.085) (1.536) - - (2.058)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F-test (3,119)</th>
<th>4.43 [0.005]**</th>
</tr>
</thead>
</table>

Notes: * = 5% significance, ** = 1% significance. The numbers in parentheses are standard errors.

Explanation of table 4: the reduced ADL models with zero-restrictions on all but one city category has all been regressed through OLS. The results of the DWH-test for the Tier 1 model is represented by the variable named residual tier1 and similarly for the other models. As explained in the methodology section, if the residuals are significant then one can reject the null of weak exogeneity, whereas insignificant values support non-rejection of the null-hypothesis. The test results from misspecification tests from the results in table 4 are tabulated in table 5.
From the results of the DWH-test in table 4, I find that the residuals of the marginal models of all the city tiers are not significant in the conditional models, thus I cannot reject the hypothesis of weak exogeneity of Tier1, Tier2, Tier3 and Tier45 at a 5% significance level. Interestingly, while the conditional models did not exhibit statistically significant misspecification tests at the 5% level, the marginal models did however show clear signs of misspecification in rejection of normality test, heteroscedasticity test and RESET23 test. This implies that the system that decides SHCOMP growth rate and the city tiers housing price growth rate is more complex than the conditional model expressed here. However, since I have not rejected weak exogeneity for all the independent variables, the regressions do not lose information by not estimating the marginal models.

Since I have established that the variables I want to analyse rejected non-stationarity and did not reject weak exogeneity, I can proceed with discussing the results of the ADL-model in table 4. In all the models considered, I have indeed found significant lags of the independent city tier variables. The fourth lag of SHCOMP seems to be consistently significant on all the models reported in table 4. Looking at the Tier1 model, I found that the second lag of Tier1 has a significant effect on SHCOMP, and that the coefficient is negative. The fourth lag has a
positive coefficient, but is not significant even at a 10% significance level. zero-restricting the
fourth lag lead to a higher AIC value and rejection of the normality test and the RESET test
for functional form at 5% significance, and is therefore not zero-restricted in the model even
though it is not significantly different from zero at conventional significance levels. The tests
for misspecification are not rejected at a 5% significance level, which suggests that the model
fits well, and that lags of Tier1 does a good job of explaining SHCOMP. The significant
coefficients may be interpreted as elasticities, where for example a 1% increase in the housing
price in first tier cities leads to a lagged -2.724% change in the SSE composite index.
In the Tier2 model, both the third and fifth lag of Tier2 are significant, where the third lag is
negative and the fifth lag positive. The negative coefficient of the third lag of Tier2 is
significant at a 0.1% significance level, whereas the positive coefficient of the fifth lag is
significant at a 5% significance level. The fifth lag of SHCOMP is not significant, but it is
included because the rejection of the ARCH test for residual heteroscedasticity at a 5%
significance level when it was zero-restricted implied residual heteroscedasticity and also that
the removing the fifth lag lead to a higher AIC value.
In the Tier3 model, the first lag of Tier3 is significant and negative, with the second and third
lag of SHCOMP being insignificant. The insignificant lags of SHCOMP are kept since zero-
restricting them resulted in a rejection of the ARCH test at a 5% significance level.
Finally, in the Tier45 model, I found that the third lag of Tier45 is significant at a 5%
significance level, with a negative coefficient. None of the misspecification tests were
rejected with the model specified, and the HQH value is lowest with this specification even
though the AIC value is a bit higher compared to a competing specification of the model with
more lags.
From the results of the ADL model, there seems to be a negative relationship between
SHCOMP and the housing price indexes of at least Tier1, Tier3 and Tier45. The negative
relationship suggests that increases in housing prices in these cities has a somewhat lagged
negative impact on the stock market prices. The dynamic relationship between SHCOMP and
Tier2 is somewhat more complicated, since I have both significant positive and negative
coefficients of lags. The negative coefficients have higher absolute value than the positive
coefficients, suggesting that a long-run relationship could be negative. To examine the notion
of more persistent dynamic effects further, I will examine the long-run solution through an
ECM reparameterization. Reparameterization of the ADL model used previously into the
ECM lets us see the long-run solution of the city tier housing price index on the SSE. The
results of the ECM are tabulated below in table 6:
Table 6, ECM

<table>
<thead>
<tr>
<th>DSHCOMP by OLS:</th>
<th>Constant</th>
<th>D3SHCOMP 1</th>
<th>D2Tier1 2</th>
<th>SHCOMP 1</th>
<th>Tier1 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>0.953</td>
<td>-0.282**</td>
<td>-1.942</td>
<td>-0.594**</td>
<td>-0.79</td>
</tr>
<tr>
<td></td>
<td>(0.945)</td>
<td>(0.087)</td>
<td>(1.284)</td>
<td>(0.118)</td>
<td>1.057</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long run solution:</th>
<th>Constant</th>
<th>Tier1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.607</td>
<td>-1.375</td>
</tr>
<tr>
<td></td>
<td>(1.574)</td>
<td>(1.766)</td>
</tr>
</tbody>
</table>

WALD test: Chi^2(1) = 0.606132 [0.4362]

<table>
<thead>
<tr>
<th>D4SHCOMP by OLS</th>
<th>Constant</th>
<th>DSHCOMP 4</th>
<th>D2Tier2 3</th>
<th>SHCOMP 4</th>
<th>Tier2 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.135</td>
<td>-0.139</td>
<td>4.1*</td>
<td>-0.583**</td>
<td>-1.988</td>
</tr>
<tr>
<td></td>
<td>(0.903)</td>
<td>(0.086)</td>
<td>(1.88)</td>
<td>(0.114)</td>
<td>(1.636)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long run solution:</th>
<th>Constant</th>
<th>Tier2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.949</td>
<td>-3.413</td>
</tr>
<tr>
<td></td>
<td>(1.562)</td>
<td>(2.931)</td>
</tr>
</tbody>
</table>

WALD test: Chi^2(1) = 1.3558 [0.2443]

<table>
<thead>
<tr>
<th>D2SHCOMP by OLS</th>
<th>Constant</th>
<th>D2SHCOMP 2</th>
<th>DSHCOMP 3</th>
<th>SHCOMP 2</th>
<th>Tier3 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.281</td>
<td>-0.299*</td>
<td>0.041</td>
<td>-0.587**</td>
<td>-4.205*</td>
</tr>
<tr>
<td></td>
<td>(0.833)</td>
<td>(0.115)</td>
<td>(0.087)</td>
<td>(0.132)</td>
<td>(1.725)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long run solution:</th>
<th>Constant</th>
<th>Tier3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>2.183</td>
<td>-7.162*</td>
</tr>
<tr>
<td></td>
<td>(1.462)</td>
<td>(3.639)</td>
</tr>
</tbody>
</table>

WALD test: Chi^2(1) = 3.8742 [0.0490] *

<table>
<thead>
<tr>
<th>D4SHCOMP by OLS</th>
<th>Constant</th>
<th>SHCOMP 4</th>
<th>Tier45 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.295</td>
<td>-0.758**</td>
<td>-3.131*</td>
</tr>
<tr>
<td></td>
<td>(0.834)</td>
<td>(0.085)</td>
<td>1.522</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long run solution:</th>
<th>Constant</th>
<th>Tier45</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>1.710</td>
<td>-4.133*</td>
</tr>
<tr>
<td></td>
<td>(1.095)</td>
<td>(2.054)</td>
</tr>
</tbody>
</table>

WALD test: Chi^2(1) = 4.05018 [0.0442] *

Notes: * = 5% significance, ** = 1% significance. The numbers in parentheses are standard errors.
Explanation of table 6: $D2SHCOMP$ is the first difference of $SHCOMP$ over 2 lags, while $D3SHCOMP$ is the first difference of $SHCOMP$ over 3 lags and so on. The long-run solution is the equation inside the clamps of equation (6). The WALD-test tests the null hypothesis of all the long-term coefficients being equal to zero.

From table 6, one can see that the reparametrized Tier1 and Tier2 models do not have significant long-run effects on $SHCOMP$, with $t$-probabilities of 0.44 and 0.25 respectively. This implies that Tier1 and Tier2 only have “contemporaneous” effects on $SHCOMP$ (over 2 periods for Tier1 and 4 for Tier2), which does not persist over time. Tier3 and Tier45 however, does have significant long-run effects on $SHCOMP$ at a 5% significance level, and the Wald tests also confirm these results. The significant long-run effects of Tier3 and Tier45 have both negative coefficients which suggest a negative long-run equilibrium relationship between the stock markets and the housing price indexes of these city tiers. In other words, the long-run elasticity of Tier3 and Tier45 on $SHCOMP$ are -7.162 and -4.133 respectively. However, this long-run elasticity is not adjusted immediately. The effects of Tier3 and Tier45 on $SHCOMP$ occurs over several time periods and returns to equilibrium long-run solution at rates 0.587 and 0.758 respectively.

A final examination of the coefficients and parameters must be made for the results above to have meaning, namely to test for invariance of the coefficients and constancy of parameters in the ADL. If the coefficients are not invariant over time, then that would imply that there exist structural breaks in the data. Structural breaks may be caused by shifts in the time series due to economic shocks and could make the estimated coefficients unreliable. The financial crisis of 2007/08 and the stock bubble burst in 2015 are likely candidates for structural breaks in $SHCOMP$. To examine for structural breaks, I estimate the ADL model recursively, and then I take a look at the coefficients +/- 2 SE to see how the coefficients change when estimated recursively. This can tell us whether or not the estimated coefficients have varied over time. The three variants of Chow-tests that are estimated and plotted for every model and are shown in the appendix.

I begin by examining the Tier1 model for structural breaks through the recursive graphics in figure A2. By examining the plots, no lasting structural breaks in the plots that represents the coefficient +/- 2 SE are found. The initial part of the recursive coefficients appears to be unstable but that could also be due to the nature of recursive estimation and not necessarily signs of structural breaks. The one-step Chow test, or the 1up Chow test as it is called in the figure, breaks through the 1% significance level in July 2009, and through the 5%
significance level in October 2007 which corresponds to the stock market crash in China as a result of the subprime crisis originating in the USA. June and August in 2015 are significant at a 10% level while December in 2015 is significant at 5% which corresponds to the stock market crash during the summer of 2015 and aftershocks lasting until February 2016 (Bloomberg). The Ndn Chow test (Break point Chow test) and the Nup Chow test (Forecast Chow test) do not indicate the presence of structural breaks in during the financial crisis nor during the stock market bubble burst in 2015. The graphic plots of the Tier2 model are shown in figure A3 and show a little more initial instability of the lags of SHCOMP and Tier2 than in the Tier1 model. The one-step Chow test breaks through the 1% significance level in July 2009, and through the 5% significance level in September and October 2007 which corresponds to the subprime crisis from the USA. This Chow test is also broken through at a 5% significance level in June, July and December in 2015 which corresponds to the stock market crash during the summer of 2015 as in the Tier1 model. Both the Ndn Chow test and the Nup Chow test do not break through the 1% significance level, which indicates that there are no clear or lasting structural breaks. The Tier3 model shows similar results as the Tier1 model and the Tier2 model in figure A4, though the coefficients appear more stable through the recursive estimation. The Tier45 model in figure A5 appears to be very stable when looking at the plots of the recursively estimated coefficients, except for the initialization. The Nup Chow test is very close to the 1% in 2009 but does not break through. All in all, the models do not appear to have any large and persistent structural breaks that could render the coefficients unreliable.

4.3 Discussion

Aside from during macroeconomic shocks like the financial crisis, the stock market and the real estate market appear to have a negative linear relationship, where the negative effects of real estate growth are lagged effects. The effects being lagged makes sense, as changes in the housing prices are not immediately known. Price growth in the residential real estate markets may attract investors away from the stock market so that demand for stocks decrease, which can lead to lower stock prices and thus lower return for investors on the stock market. The results from the ADL models shows that housing price in tier 1 cities appears not to have a very large lagged negative impact on the stock market as compared to the other city tiers. Tier 1 and 2 cities also did not appear to have any long-run effects on the Shanghai Composite Index. The results are a bit surprising, as one would imagine housing in these types of cities to
be more liquid than that of the following tiers, because of the high price and demand. A more liquid housing market could respond more to shocks in the stock markets, as was argued by Lin and Fuerst (2014). Tier 3 and combined Tier 4 and 5 cities were found to have negative lagged relationship with the SSE in addition to more persistent negative long-run relationships, through the use of the ECM. This result implies that not only do they have lagged effects, but also long-run equilibriums to which the SSE Composite Index readjusts over time after an initial shock in housing prices. The negative relation between the two markets was found to be stronger in the Tier 3 and Tier 4 & 5 models than the first two city tiers.

I theorize that one reason for these results may be that the prices of housing in these lower tier cities are not as high as in the tier 1 and tier 2 cities, which could be a contributing factor to the negative relationship between the stock markets and these housing markets. High housing prices puts the entrance bar into the housing market high, such that investors that cannot afford to enter the housing market in these cities might rather invest in housing in lower tier cities. Investors that cannot afford to enter the housing markets at all would have little alternatives except the capital markets to gain returns higher than the deposit interest rates in the bank, and thus may not react that much to increases in the housing prices. CHFS reported that households in tier 1 and tier 2 cities have lower percentages of uninhabited household-owned (vacant, for brevity) housing than in tier 3 cities. Reasons for owning vacant housing were among others: high risk taking behaviour, doubling of price rental ratios and double return on the first housing. That the vacant housing rate in tier 3 cities is higher than that in tier 1 and tier 2 cities could support the results here, since the results show that an increase in housing prices in tier 3 cities leads to a decrease in the SSE Composite Index which may prompt investors to invest in housing with the focus on expectation of future price increase, rather than cash stream from renting out. Since an increasing number of higher tier cities have begun to introduce stiffer mortgage down payment requirements, especially for secondary housing, investors may be further incentivized to invest in housing in the lower tier cities.

5. Conclusion

Through the use of an Autoregressive Distributed-model (ADL) and a reparameterization of this model into ECM form, I have in this paper analysed a linear relationship between the Shanghai Stock Exchange through the Shanghai Composite Index, and housing price growth rates in 4 different city categories based on a city categorization system commonly referred as
the “tier system”. The motivation for this paper was to examine how the housing price growth in different types of cities in China may affect the capital markets. Since Chinese retail investors have few options for investment, examining how the real estate market and the capital market move in relation to one another is very important, not only for investors themselves but also for policy makers. If the two markets are related, then policy makers must be wary when implementing new policies in one of the markets as it may affect the other market as well. I found negative relationships between the housing prices of all the city tiers, where the negative relation was found to be strongest between the tier 3 and tier 4 & 5 cities, and weaker between tier 1 and 2 cities. The results imply that investors could include them both in their portfolio without being exposed to correlated risk, as they do not move together in the same direction. The stronger effect from the lower tier cities may be due to the higher prices on housing in tier 1 and 2 cities compared to tier 3 and 4 & 5 combined with greater mortgage down payment requirements, especially on secondary housing.

A weakness in this paper that may have influenced the results, is the accuracy of the NBS 70-city Index data. Since critique against the data not truly representing price development observed in the real estate market exist, the results in this paper may subsequently be less accurate. It would be interesting to examine this subject further with other datasets and compare the results. In addition, using a more suitable city categorization system (such as the one proposed by the Demand Institute) could give a more accurate picture of the relationship between housing prices and the capital market. Lastly, the approach in this paper was a linear one, however several papers on the subject has also examined the relationship between the stock market and the real estate market through a non-linear approach and achieved conflicting results. The relationship between the stock market and real estate market is thus not clear-cut and should warrant more attention, especially when it comes to economies like China, that practice strict capital controls.
Reference list


Schell, Orville. (2015) “Why China’s stock market bubble was always bound to burst”. The Guardian. Link: https://www.theguardian.com/world/2015/jul/16/why-chinas-stock-market-bubble-was-always-bound-burst Accessed 06.05.2017
Appendix

Figure A1. Residual QQ plot against N(0,1)

- SHCOMP × normal
- Tier1 × normal
- Tier2 × normal
- Tier3 × normal
- Tier45 × normal

Figure A2. Graphs of recursive estimated coefficients and Chow tests for tier 1 model

SHCOMP_1 × ±/−2SE
SHCOMP_4 × ±/−2SE
Constant × ±/−2SE
Tier1_2 × ±/−2SE
Tier1_4 × ±/−2SE
RealStep
1up CHOWs 1%
Ndx CHOWs 1%
Nap CHOWs 1%
Figure A3, Graphs of recursive estimated coefficients and Chow tests for tier 2 model

Figure A4, Graphs of recursive estimated coefficients and Chow tests for tier 3 model
Figure A5, Graphs of recursive estimated coefficients and Chow tests for tier 4 & 5 model