Conditional Heteroscedastic Cointegration Analysis with Structural Breaks

A study on the Chinese stock markets

Authors:
Andrea P.G. Kratz
Heli M.K. Raulamo

Supervisor:
Frederik Lundtofte
VT 2011

Abstract

A large number of studies have shown that macroeconomic variables can explain co-movements in stock market returns in developed markets. The purpose of this paper is to investigate whether this relation also holds in China's two stock markets. By doing a heteroscedastic cointegration analysis, the long run relation is investigated. The results show that it is difficult to determine if a cointegrating relationship exists. This could be caused by conditional heteroscedasticity and possible structural break(s) apparent in the sample.

Keywords: cointegration, conditional heteroscedasticity, structural break, China, global financial crisis

Table of contents

Figures and tables

1. Introduction

Modern financial theories state that there is a systematic risk that can explain co-movements in the stock market. Numerous studies performed on developed markets have provided evidence that this systematic risk can be predicted using macroeconomic variables (e.g. Jorion, 1991; Fama and French, 1989; Chen, Roll, and Ross, 1986; Kaul, 1987). Between the years 2001 and 2005 the Chinese stock market's performance went in a different direction than the country's aggregated economy (see figure 1). Even though the annual GDP growth rate was around nine percent, the stock market indices fell 1084 points between those years. This divergent behavior goes against the common belief that the stock market and the real economy move in the same directions (Zeng and Wan, 2007). Given this relation in GDP and stock prices between 2001 and 2005, the question is awakening if macroeconomic factors can explain co-movements in the Chinese emerging stock markets?
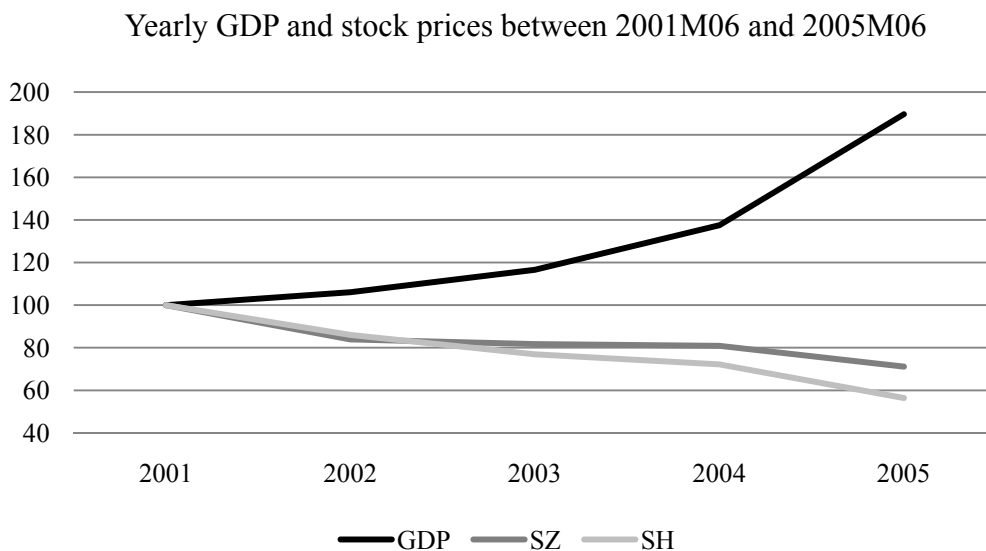
Figure 1 Yearly GDP and stock prices in China, SZ=Shenzhen, SH=Shanghai (index 100=2001-06) Source: NBS (2011)

Some of the macroeconomic variables that explain co-movements in stock market prices have been proven to be rather uneven over time. Some studies[1] have found strong evidence for a specified variable while other studies fail to give any proof of predictability for the same one. Therefore it is difficult to be sure whether macroeconomic variables are useful in predicting the stock market returns. Extensive research has been done of this question in developed economies while less research has been made on the emerging markets. The macroeconomic variables' effect on the stock prices in emerging markets is therefore even more uncertain

---

[1] See for example Chen, Roll and Ross (1986)

than the effect on the developed ones. An even smaller amount of research has been done on the Chinese market over unusual economic periods. As for all stock markets in the world, the global financial crisis which started in 2007 has had an impact on the Chinese markets. It is therefore interesting to investigate if the macroeconomic variables, despite the chaotic period, could predict co-movements in the stock market.

Most studies have looked at the relationship between macroeconomic variables and stock returns in the long run. The rather unexpected behavior of the relationship between stock markets and aggregated economic growth in China during 2001-2005 could just be a temporary move from the long run equilibrium. Another reason might be that the explaining factors for co-movements on the developed markets are not applicable on the emerging markets. Over a longer time period the real economy and the stock market seem to be moving in the same direction.

Yearly GDP and stock prices between 2001M12 and 2010M12
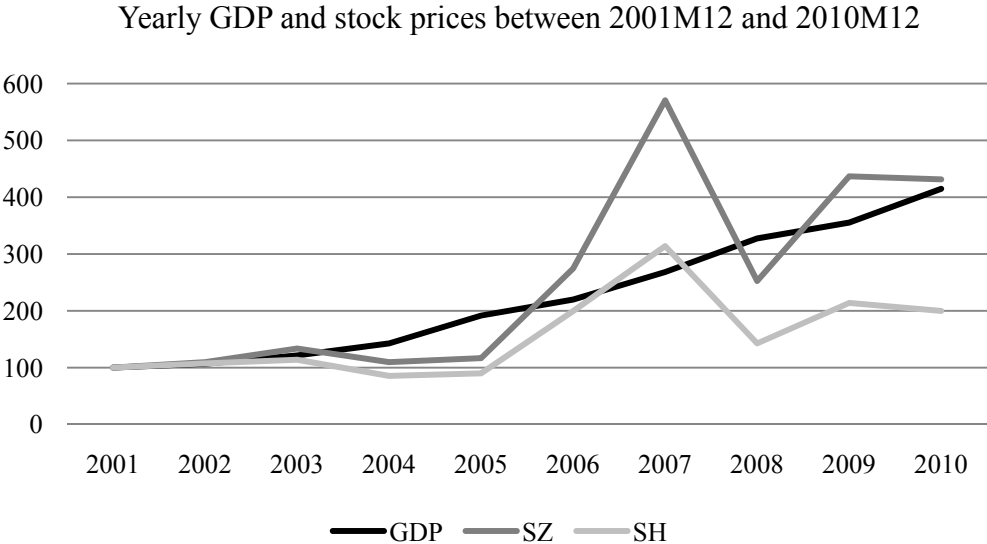
Figure 2 Yearly GDP and stock prices on Shanghai and Shenzhen stock exchange (index 100=2001M12) Source: NBS, 2011

Thus, the aim of this paper is to investigate if underlying macroeconomic variables can explain the long run co-movements in the composite indices of China's two stock markets between 2001 and 2010 by doing a cointegration analysis. The data that will be used are monthly and collected from the National Bureau of Statistics in China (NBS) and Reuter's Datastream. The method that will be used is a replication of a study conducted by Liu and Shrestha (2008) in which they use a heteroscedastic cointegration analysis to look at the long-run relationships between variables. In contrast to Liu and Shrestha's study, this paper will take structural breaks into consideration. The aim of the paper is not to determine what effect

the individual macroeconomic variables have on the market return since extensive studies have already been done on this topic. Therefore, the macroeconomic variables are going to be treated as a bundle to see whether they can explain co-movements.

A problem with using financial and macroeconomic data is that it is often heteroscedastic and autocorrelated. To get reliable results, the data generating process used to find the right regression model is crucial. Unfortunately, several previous research papers have made assumptions on the data set rather than actually testing if they hold which might have led to wrong conclusions. One common assumption is that there are no structural breaks. With this assumption it is reasonable to apply the most frequently used cointegration tests. However, it is a strong assumption to be made during a period which includes an event like the global financial crisis and should therefore not be made. By adding structural breaks a different econometric approach than the most commonly known must be used to test for cointegration. By using an approach developed by Westerlund and Edgerton (2007) this paper will investigate if there is a cointegrating relationship between the Chinese stock markets and a set of macroeconomic variables, when structural breaks are taken into account.

In the next section a short description of the Chinese stock market is given. Section three presents some of the previous research on the co-movements in stock returns. Section four reviews the theory explaining the macroeconomic influences on the stock market as well as the econometric theory used in this paper. Section five describes the empirical study, including a description of the dataset followed by the methodology and results, as well as a discussion. Finally, we present our conclusions.

2. The Chinese stock market

China introduced the Shanghai stock market in November 1990 and the Shenzhen stock market in 1991. There are some differences between the Chinese market and the developed ones such as the New York stock exchange in which most of the previous studies have been preformed (Zeng and Wan, 2007). Almost all listed firms in the Chinese stock market are former State Owned Enterprises (Liu and Shrestha, 2008). The privately tradable part of the stocks in those companies is smaller than ten percent. The reason for the small percentage is that the government wants to prevent private investors from taking over state-owned companies (Zeng and Wan, 2007). These companies hardly ever enter into bankruptcy because the government supports the firms in monetary difficulty. Without a risk of a

takeover and almost no risk of bankruptcy in these firms, corporate governance is usually poorer in Chinese listed firms than for firms in western markets (Liu and Shrestha, 2008).

Interest rates in the bank system are regulated by the government and are often low to stimulate economic growth (Liu and Shrestha, 2008). Alternatives for investment besides the stock market are limited due to the immaturity of other capital markets. As a consequence of the lack of high return investment opportunities for citizens, the Chinese stock market is in contrast to the western markets dominated by private investors (Zeng and Wan, 2007). Approximately 99 percent of all investors in China are of this type. The number of private investors, many without proper financial knowledge, in combination with a low market transparency makes the Chinese market unusually speculative. Many of the stockholders base their investment decisions purely on historical price trends and rumors (Liu and Shrestha, 2008).

The differences between the Chinese market and the more developed ones have led to a discussion on whether the Chinese market is efficient[2] or not. Some studies have shown that the Chinese markets have achieved a weak form of efficiency while others have come to the conclusion that it is not efficient.[3] An example of a mechanism that supports the view on market inefficiency in China is that the possibility of short sale is limited by regulations (Li, 2008). If the stock markets in China suffer from a high degree of inefficiency this could imply there is a possibility that all the information about the economy is not reflected in the asset prices. Since the majority of the investors are private investors, this could also mean that they are not able to collect all the information about business conditions. On top of this, as is commonly known, freedom of the press is limited in China.[4] Important information about business conditions can thus be withheld from investors, and this can affect investment choices.

---

[2] Weak efficiency: Stock prices reflect all information given in the market trading data, e.g. historical prices. Semi-weak efficiency: Stock prices incorporate all information available to the public e.g. publicly published information. Strong efficiency: Stock prices reflect all information, e.g. no arbitrage can be done on insider information (Li, 2008)
[3] For further discussion see Li (2008)
[4] All press in China is state owned (Li, 2008)

3. Previous research

The purpose of this paper is to determine which, if any, macroeconomic variables can be used in predicting stock return on the Chinese markets. Many previous studies have presented variables that can be used as proxies for the business conditions in a country. Some of them will be reviewed here. The first section looks at studies on macroeconomic variables in asset pricing models. The second section will review studies conducted on China's stock markets using the macroeconomic variables that have been found useful in explaining the co-movements on the market. The paper by Liu and Shrestha (2008) is of particular interest because this study employs the same method as theirs and in the Chinese context.

3.1. Stock market and macroeconomic variables

Chen, Roll and Ross (1986) tested if innovations in macroeconomic variables are a part of the risk premium. They argued that the default spread is a measure of business condition. The variables they expected to have an effect on stock returns were industrial production, inflation, risk premium, term structure, market indices, consumption and oil prices. Of these, industrial production, changes in risk premium, term structure, and inflation showed significant results. The researchers therefore concluded that stock market prices reflect the systematic macroeconomic changes.

Fama and French (1989) also tested if expected bond and stock returns are related to business conditions. First, they found that the expected returns of bonds and stocks moved together, so common underlying variables must be influencing both. They then tested if the term and default spread as well as the dividend yield on assets are explanatory variables for excess returns in the United States. They found that the default spread and dividend yields move in a similar way as the long-term business conditions and that the term spread followed the short-term business conditions. They therefore stated that business conditions could be used to forecast bond and stock returns. Their explanation for this relation was that the expected returns on bonds and stocks must be high in order for the investor to be willing to invest when business conditions are bad.

Cheung and Ng (1998) also tested macroeconomic variables relationship with stock returns and looked at the relationship in different time horizons. The study covered several developed economies, and the variables used where oil price, consumption, real money, and real output.

Foreign variables were left out in the study. They found that macroeconomic variables tend to move together in the long run but that divergence occurs in the short run.

Chen (1991) tested if the realized and expected growth rate of GNP can be used to forecast the market risk premium (the market risk premium is the return in excess of the riskless rate). He also looked at how market returns and future GNP growth are related to state variables industrial production, term structure, T-bill rate, default spread, and dividend yield. In his study the NYSE index and US macroeconomic variables were used. His results showed that the market risk premium is negatively correlated to the recent growth and positively correlated with the expected future growth. Another finding was that macroeconomic variables were positively correlated with recent growth as well as positively related to the expected excess market return. In the same way the variables that were negatively correlated with future growth were also negatively correlated with the excess return. He therefore stated that macroeconomic variables can be used to predict future market returns.

The economic theory states that stock return should be positively related to inflation, which is an intuitive notion. Geske and Roll (1983), in looking at the monetary linkage between inflation and stock returns, refer to the findings of e.g. Fama and Schwert (1977) that have shown a negative correlation. They explained this result with that the investors realize that inflation would lead to future reactions on in the macroeconomic climate that will have a negative effect on the investment climate. This would be due to a so called counter-cyclical monetary response by the decision makers in the countries.

Kaul (1987) argues that Geske and Roll (1983) do not analyze the money supply completely. He claims that *"[…] if the central bank follows a pro-cyclical monetary policy, real activity and inflation could be either unrelated or even positively related."* (Kaul, 1987, s. 2). They found that this result occurs depending on which time period or which country that is being investigated since central banks can have either procyclical or countercyclical response policies to economic activity, which lead to different price reactions.

Jensen et al. (1996) added monetary variables as the Federal Reserve's monetary stance to Fama and French's (1989) model. The results showed that the monetary variables explain expected return for bonds in restrictive monetary policy and the expected return on stocks could be explained in expansionary environments.

Jorion (1991) recognized that the exchange rate risk is a big concern for multinational companies and therefore examined if exchange rate risk was priced in the stock market. In the study, Chen, Roll, and Ross (1986) model was used but exchange rate risk was added as a depending variable. When the stock market was divided into industries the results indicated that different industries had different exposures to the exchange rate. His overall results could not state that exchange rate risk is related to expected returns. He therefore left the question on whether exchange rate risk explains stock market movements for further research.

Rapach, Wohar and Rangvid (2005) have done an extensive study to investigate which macroeconomic variables are good in predicting future returns. They investigated stock return predictability in 12 industrialized countries using the variables: inflation rate, money stocks, interest rates, term spread, industrial production, and unemployment rate using both in-of-sample and out-of-sample tests for the predictions. Their results showed that interest rate is the most reliable factor for predicting stock returns and that money stock and term spread also have a predictive power but that the effect is limited to a small number of countries. The results did not give any strong proof for the industrial production and unemployment rate as explanatory variables.

3.2. The Chinese market

Zeng and Wan (2007) looked at a specific change in the Chinese stock market regulations during the period 2001-2007. The reason was to see if the divergence between stock market and the aggregated economy behavior could be explained by temporary market adjustments for the reformation. Price adjusted GDP was the only explanatory variable and only the Shanghai stock exchange was examined. A time-series extrapolation and an out-of-sample forecast based on the vector auto-regression were used to compare how the market would develop without and with the regulation change. The conclusion made was that the unusual behavior probably was due to the changed regulations and that the stock market should move towards the GDP development in the long run.

Yuan and Chen (2010) performed a research on the relationship between inflation rates and the Chinese stock returns between 1991 and 2008. They found that stock market returns cannot be explained by the inflation rate and that this has therefore had a small function in economic forecasting. The response of the stock market on a change in inflation was positive in the first periods but a faint negative correlation appeared in the long run.

Another study conducted by Fen and Tu (2010) covering the Shanghai market between July 2005 and June 2009 and investigated if macro variables influence the market volatility. GDP, exchange rate, money supply and CPI were used as proxies for the aggregated economy. Their results demonstrated that macroeconomic variables did affect the stock market volatility during this time.

Liu and Shrestha (2008) analyzed the long run relationship between macroeconomic variables and stock returns on the Shenzhen and Shanghai stock markets between 1992 and 2001. Even though the economy largely differed from the western markets Liu and Shresthas results were significant. The result showed that macroeconomic variables do affect the stock market in China and that they therefore can be used to forecast the market returns in the long run.

In the next chapter, the theoretical ground that the previous and this research have been based on is going to be presented.

4. Theory

The theoretical background of macroeconomic variables in the stock market and the motivation behind the chosen variables for this paper is going to be described in this chapter. In the second part of the chapter the econometric theories on which the method is going to be based on will be presented.

4.1. Macroeconomic theory

4.1.1. The intertemporal asset pricing model

To explain the market value, one must look at how much investors are willing to pay for assets. Merton (1973) developed a model that takes into account the long run horizon in individual's investment decisions. *The intertemporal asset pricing model* assumes that assets can be traded over a continuous time frame but the trade does only take place at equilibrium prices. The individuals are assumed to be homogenous, rational and they want to maximize their life time utility. Since the model is intertemporal, the consumer-investor is assumed to know the transition probabilities for returns on each asset over the next trading interval as well as the transition probabilities for returns in the future periods, i.e. the market is efficient. Thus, the investor takes into account the relationship between the current returns and the returns available in the future. Another assumption in the model is that the amount of assets

available is constant. Hence, the return on the asset will only be given by the demand function.

The main assumption in the model is that the individual wants to maximize her expected utility of intermediate consumption end of terminal wealth.

$$\max E_t \left[\int_t^T U[C(s), s] \, ds + B[W(T), T] \right] \tag{4.1.1}$$

Where E is an expectation operator conditional on the state of the economy and the current endowment, $C(s)$ is consumption at time s, and U is a utility function that is concave and is increasing. T is the individual's age of death and B is s strictly concave utility-of-terminal wealth function.

Merton assumes that a state variable vector, X can denote the current level of the stock price, P. Therefore the investor makes his decisions between consumption and investments to reach maximum utility by consider her current wealth, the time dimension, and the state variable vector. This means that the optimum consumption and portfolio are a function of the state variables (Merton, 1973). Hence, the maximization equation is subject to these variables. The accumulation equation for an investor is given by:

$$dW = \sum_1^{n+1} w_i W \, dP_i / P_i + (y - c) dt \tag{4.1.2}$$

Where w_i is the fraction of her wealth, W, invested in the ith asset, y is her wage income, and c is the consumption.

The solving of the maximization problem (4.1.1) subject to (4.1.2) and the boundary condition $J(W, X, T) = B(W, T)$, where J is the value function, gives the demand for an asset. A single individual's demand function for asset i is given by:

$$w_i W = A \sum_1^n v_{ij}(\alpha_j - r) + H \sum_1^n \sigma_j v_{ij} \tag{4.1.3}$$

Where α_j is the expected return for asset i, r is the risk free rate, and v_{ij} are the elements of the inverse of the instantaneous variance-covariance matrix of returns, $\Omega = [\sigma_{ij}]$, (where σ_{ij} are the covariances between the returns in the ith and jth assets), $A \equiv -J_w / J_{ww}$, and $H_k \equiv -J_{kw} / J_{ww}$. The subscripts on J denote the partial derivatives. The first term is the mean-variance maximizer given the individuals utility function. The second term is the demand for hedging against unwanted shifts in the investment opportunity set (Merton, 1973).

Another notation for A and H_k are:

$$A = - U_c \Big/ \Big(U_{cc} \frac{\partial c}{\partial W} \Big) > 0 \qquad\qquad (4.1.4)$$

$$H = - \frac{\partial c}{\partial x_k} \Big/ \frac{\partial c}{\partial W} \lesseqgtr 0 \qquad\qquad (4.1.5)$$

Where x_k is the investment opportunity set variable, i.e. the state variables. The risk premium in the asset pricing model is a positive function of the aggregate risk aversion parameter.

All risk-averse individuals will try to hedge against a risk of a lower consumption in the future. A change such that $\frac{\partial c}{\partial x_k} < (>)0$, ceteris paribus, the risk averse individual will demand more of asset i if the asset is positively (negatively) correlated with the change in x_k. Through the increase of asset i the individual will try to minimize the variability of consumption over time to achieve "*consumption smoothing*". If the ex-post opportunity set is less favorable then expected the investor expects to be compensated by a higher level of wealth through the positive correlation of returns. In the same way, the investor will expect a more favorable investment environment if ex-post returns are low (Merton, 1973).

Given this demand function, Merton showed that the equilibrium dynamics for the value on the market is given by:

$$dM = \sum_1^{n+1} D_i dP_i / P_i + \sum_1^{n+1} \big(y^i - c^i \big) \, dt \qquad\qquad (4.1.6)$$

Where D is the aggregate demand function, y is the wage income and c is the consumption and P is the price.

Since the market value is affected by the aggregated demand and all investors are assumed to have the same utility functions and maximizing behavior, the market value will depend on how individuals react to different state variables, this is called individual risk. By combining all assets on the markets, the market portfolio has diversified away all individual risk. The risk left is the risk that is affecting all assets and can therefore not be eliminated by diversification. These are usually macroeconomic factors that affect the economy as a whole. Hence, the total return on the market is the premium for taking on this "macroeconomic" risk, usually called the market risk (Brealey and Myers, 2003).

4.1.2. Implications

The model suggests that asset prices depend on the real markets since the individual's investment behavior depends on the climate of the economy as a whole (Cheung and Ng, 1997). Individuals will save in good times to be able to have money in worse times. If the economy is weak (strong), the individual is less (more) willing to save and invest money (Ely and Salehizadeh, 1999; Fama and French, 1989). Given the above discussion on consumption smoothing and business opportunities *the inter-temporal asset pricing model* gives room to introduce any state variable that are of interest for the investors (Cox, Ingersoll and Ross, 1985).

Also the more general *arbitrage pricing theory* developed by Ross (1976) is open for macroeconomic variables as determinants for the stock prices. The *arbitrage pricing theory* states that several factors may affect stock prices. Different factors are priced depending on how much the investors are willing to pay to avoid a certain risk, i.e. the individual factor risk (Jorion, 1991). Co-movements in the stock markets are therefore reactions of underlying influential macro economic variables since these are risks that affect all assets (Chen, Roll, and Ross, 1986; Cox, Ingersoll and Ross, 1985).

4.1.3. The underlying factors

Inflation, money supply and interest rates are variables that affect the future business opportunities and the individual's willingness to invest in the stock market. These three variables are strongly correlated with each other since a change for example in money supply will affect the inflation and interest rates (Chen, Roll, and Ross, 1986).

Other variables that affect future business opportunities and the individual's investment choices are exchange rates and industrial production. A change in exchange rate affects the aggregated demand and the import and export costs that in turn will affect business opportunities. The change in exchange rate will also affect the investment choices in an open economy since it will be more or less profitable to invest in other countries depending on how the exchange rate changes (Jorion, 1991; He and Ng, 1998).

If the industrial production changes, the future business opportunities will change since there is a change in value of the products that are going to be sold (Chen, Roll, and Ross, 1986). A new level of industrial production will affect the individual's choice between consumption

and investment since it will affect the pay off of investments and consumption costs (Chen, 1991).

Given the above theory and discussion on macroeconomic variables affect on co-movements in the stock market, stock prices, y, can be determined by inflation, CPI, money supply, MS, interest rate, IR, exchange rate, EX and industrial production, IP. This leads to the theoretical model of this paper:

$$y = \alpha_0 + \alpha_1 IF + \alpha_2 MS + \alpha_3 IR + \alpha_4 EX + \alpha_5 IP \qquad (4.1.7)$$

The macroeconomic variables are likely to be depended on each other and also to be affected by common factors in the economy.[5]

4.1.4. Market efficiency

The capital asset pricing models assumes that the market is efficient, thus all current information is available to all investors (Sharpe, 1964). In order for the macroeconomic variables to be able to explain the co-movements in the stock market, the investors must have been given the right information about these variables. If the institutions connected to the market are inefficient in information reporting, the stock market can be mispriced as a consequence. Inefficient information channels also lead to higher transaction costs since it will be more expensive to get the information needed for investment choices (Gilson and Kraakman, 1984).

4.2 Econometric theory

4.2.1. Non-stationarity and unit roots

The majority of economic time series are found to be non-stationary due to their "*random walk*" properties.[6] In a random walk process this period's value depends on the value from the previous period and a random error term. For example, consider the *autoregressive* (*AR*) model of the form:

$$y_t = \rho y_{t-1} + \varepsilon_t \qquad (4.2.1)$$

---

[5] See for example: Lee (1992)

[6] See for example: Nelson and Plosser (1982), Stulz and Wasserfallen (1985), Wasserfallen (1986), Gould and Nelsson (1974), Perron (1988), Samuelsson (1973)

When $|\rho| = 1$ the model (*4.2.1*) is called a random walk and the time series is not stationary (Dickey and Fuller, 1979). The problem with non stationary data is, as noted by Granger and Newbold (1974), that it could lead to a "*spurious regression*". This means that the results falsely indicate a noteworthy relationship between the variables due to a unit root in the error process, giving significant beta estimates, high R^2 and a small Durbin-Watson statistic. One solution to non stationarity is differencing. This means that a series can be made stationary by differencing it:

$$\Delta y_t = y_t - y_{t-1} \qquad (4.2.2)$$

This can be done several times. When the series has become stationary it is said to be *Integrated of order (d)* and denoted *I(d)*.[7]

4.2.2. Cointegration

Granger (1981) introduced the notion of cointegration. He explained that there may be two non stationary series that can wander in the short term but are tied together in the long term. Through an equilibrium relationship they do not drift too far apart. Engle and Granger (1987) define it further, reminding that in general, if two series are both integrated of order *I(d)*, then a linear combination of them is also *I(d)*. However, if there exists a vector β, such that makes their disturbance term, $u_t = y_t - \alpha - \beta x_t$ of lower order *I(d-b)*, then the two series are cointegrated of order *CI(d,b)*. Engle and Granger (1987) (*EG*) tested the cointegration of several macroeconomic variables and found a number of pairs that are cointegrated such as the short and long term interest rates as well as household income and consumption, to name but a few. They also created a procedure to determine if series are cointegrated using the Dickey-Fuller (1979) [8] approach (*DF*) to test for a unit root in the long-run relationship of the series. The null hypothesis equals to no cointegration if the *DF* test implies the existence of a unit root.

The *EG* method detects only a single cointegrating vector and lacks power if there are more than two variables since it does not use all the information available. When there are more than two variables, say *k* amount of variables, there may exist up to *k-1* cointegrating relations (*vectors*). The *EG* test is still appropriate for its purpose, because the null hypothesis tests if there is no cointegration. However, there exist an alternative approach proposed by Johansen (1988) and Johansen and Juselius (1990), here after referred to as the *JJ* method. The *JJ*

---

[7] See for example: Box and Jenkins (1970), Granger, C. W. J. (1981), Engle and Granger (1987)
[8] Based on Dickey (1976) and Fuller (1976)

method uses a *vector autoregressive* (*VAR*) representation and tests through *maximum likelihood* (*ML*) estimation the number of cointegrating relations. The test boils down to likelihood ratio test, but does not have the traditional χ^2 distribution. Instead the *DF* distribution has been extended to a multivariate case. Studies have shown that small sample corrections are needed since the test statistics differ substantially from the asymptotic properties and as a result the tests find cointegration too often.[9] The *JJ* method assumes *Gaussian* errors, i.e. it is not applicable if the model contains structural breaks or if the errors are autocorrelated or heteroscedastic.

4.2.3. Unit root tests

The *DF* approach tests the null hypothesis of a unit root (Dickey-Fuller, 1979). Assuming the series follow an *AR* process one first estimates equation (*4.2.1*) and then tests if $\rho = 1$. In practice the equation is rewritten as:

$$\Delta y_t = (\rho - 1)y_{t-1} + \varepsilon_t \qquad (4.2.3)$$

Where the null becomes equal to testing if $\rho^* = (\rho - 1) = 0$.[10] Under the null hypothesis of non-stationarity the *DF* test does not follow the standard *t*-distribution and therefore the authors have used a distribution computed by Fuller (1976) through a simulation process. If one would use the standard *t*-distribution, it would lead on average to over rejection of the null (Nankervis and Savin, 1985). Using equation (*4.2.3*) to test the unit root requires that one assumes the underlying process for y_t is a simple first-order *AR* with a zero mean and no trend. Since the mean of the series is determent by the initial observation, one is also assuming that y_0 is equal to zero. If it is not know whether the initial value in the process equals zero or not, there is a need to include a constant, μ, into the test equation:

$$\Delta y_t = \mu + \rho^* y_{t-1} + \varepsilon_t \qquad (4.2.4)$$

In the stationary case the constant reflects the non-zero mean of the series, however in the non stationary case it reflects a deterministic trend in y_t. Instead of a unit root the non stationarity can be caused by the deterministic trend in the underlying process. Therefore, if this is suspected one needs to adjust the model to allow for a time trend, γ_t:

$$\Delta y_t = \mu + \rho^* y_{t-1} + \gamma_t + \varepsilon_t \qquad (4.2.5)$$

---

[9] See e.g. Cheun and Lai (1993), Ahn and Reinsel (1990) or Reimers (1992)
[10] Because most computer programs are programmed to test if something equals to zero.

In the stationary case the process is called t*rend stationary*. However, if the null of a unit root holds, then the series contain a *stochastic* trend.

In addition to the previous assumptions, the *DF* also assumes that the errors are white noise. If the underlying process is an *AR* (*p*), then the errors are autocorrelated and this will invalidate the *DF* distribution. Therefore, Said and Dickey (1984) extended the *DF* model, called the Augmented Dickey-Fuller (*ADF*) method, which takes into account the higher order *AR* (*p*) processes. Thus, the test equation becomes:

$$\Delta y_t = \rho^* y_{t-1} + \alpha_1 \Delta y_{t-1} + \cdots + \alpha_p \Delta y_{t-p} + \varepsilon_t \qquad (4.2.6)$$

Where the null of a unit root is equal to testing if $\rho^* = (\rho - 1) = 0$. Their results show that when adding lagged first differences of the dependent variable, letting the number of lags grow with the sample size, the asymptotic distributions of the *DF* statistics will still hold. They also show that the *ADF* tests are valid when the underlying process includes a moving average (*MA*) component. They argue that any *ARMA* model can be represented as an infinite *AR* process. When choosing the number of lags, it is better to go with too many than too few.[11] If too many lags are chosen, the power the test is reduced. However, with too few lags in the test equation, one will end up over-rejecting the null due to the fact that there will still be autocorrelation in the errors and thus the asymptotic distribution is not valid.

An alternative to the *ADF* is the Phillips - Perron (*PP*)[12] Z-test. This is a nonparametric test for a unit root which is also applicable for the more general (unknown) *ARMA* processes. The *PP* test adjusts the original *DF* *t*-statistic through a non-parametric correction to account for the possible autocorrelation in the errors. This test uses the same equations (*4.2.3, 4.2.4, and 4.2.5*) as well as critical values as the *DF*-test.

4.2.4. Structural breaks

Perron (1989) studies the effect of a structural break in a time series on the unit root tests. According to Perron, under the unit root hypothesis shocks have a permanent effect on "*the system*" which implies that economic fluctuations are not transitory. However, Perron finds only two shocks, the Great Crash of 1929 and the oil price shock of 1973, that have had a permanent effect on the variables. Perron means that these shocks are exogenous and not a realization of the underlying *data generating process* (*DGP*) of the series. Thus Perron finds

---

[11] See Banerjee et al (1993) and Harris (1992)
[12] Suggested by Phillips (1987), extended by Perron (1987,1988) and Phillips and Perron (1988) to account for drift as well as trend.

that most macroeconomic time series are not characterized by a single unit root, but rather that variables are *trend stationary*, if one allows a single change in the intercept of the trend function after 1929 and a change in the slope function after 1973. Consequently Perron concludes that fluctuations are indeed transitory. This approach assumes that the time of the break is known. Perron argues in his paper that the usual unit root tests will not be able to reject the null of a unit root if the deterministic trend of the series has a single break.

Perron (1994) develops an approach for the unit root test where a change is allowed in both the null and alternative hypothesis at an unknown time.[13] He adds a dummy variable DU_t that will equal to 1 if $t > T_b$ and 0 otherwise. T_b is the unknown time of a structural break in the series.

$$y_t = \mu_1 + \beta_t + (\mu_2 - \mu_1)DU_t + v_t \qquad (4.2.7)$$

He then chooses T_b so that it minimizes the value of the *t*-statistic for testing $\gamma = 0$, where γ equals to $(\mu_2 - \mu_1)$. Harvey, Leybourne and Newbold (2001) found that using the break point as T_b+1 the test statistic will have more power. Harris and Sollis (2003) adopt the Harvey et al. (2001) procedure and test UK money demand data (1963q1-1989q2), and show that all of the variables except one, the inflation rate, are still non-stationary even after allowing for the possibility of a break in the series.

4.2.5. Structural breaks and cointegration

Gregory and Hansen (1996) (*GH*) propose *ADF* and *Z*-type tests where the null of no cointegration is tested against the alternative of cointegration when there is a break in the intercept or slope coefficients of unknown timing. The model allows for three different forms of structural breaks using a dummy variable in the original *EG* equation thus modifying it. The dummy variable is defined as:

$$\varphi_{tk} = \begin{cases} 0 \; if \; t \le k \\ 1 \; if \; t > k \end{cases} \qquad (4.2.8)$$

A possible structural change is a "*level shift*", where the intercept has changed:

$$y_t = \alpha_1 + \alpha_2 \varphi_{tk} + \beta_1 x_t + u_t \qquad (4.2.9)$$

---

[13] According to Harris and Sollis (2003), pp. 61-63

One can also allow the introduction of a *time trend, t,* into the "*level shift*" model:

$$y_t = \alpha_1 + \alpha_2 \varphi_{tk} + \lambda t + \beta_1 x_t + u_t \qquad (4.2.10)$$

Another possible change permits the slope vector to change as well (without a time trend) and is called a "*Regime shift*":

$$y_t = \alpha_1 + \alpha_2 \varphi_{tk} + \beta_1 x_t + \beta_2 x_t \varphi_{tk} + u_t \qquad (4.2.11)$$

The standard method of testing for cointegration through estimating the cointegrating equation by OLS and then performing a unit root test on the residuals, u_t, will only work if the time of the "*regime shift*" is known *a priori*. Since this is unlikely in practice, *GH* developed a test procedure where the timing information or even the occurrence of a break is not required. For each possible break point in the sample the *ADF* test on u_t is computed for all three models (*4.2.9, 4.2.10 and 4.2.11)*, and then the time period which produces the smallest test statistic is used to test the null of no cointegration.

4.2.6. Conditional heteroscedasticity

Several models and tests on times series assume normally, identically and independently distributed error terms. The residuals are also assumed to be homoscedastic and have a zero mean with a constant variance. Most financial time series have been found to display non-normality and time-varying volatility also known as "*conditional heteroscedasticity*".[14] Time-varying volatility can be described as periods of high volatility following periods of low volatility, i.e. non-constant variance.[15] Non-normality can easily be dealt with by using robust standard errors, however conditional heteroscedasticity may cause problems. Several studies have shown that the *DF* statistics are robust in the occurrence of conditional heteroscedasticity.[16] However, in practice there is an increased probability of spurious regression when used to test cointegration.[17] This is due to the fact that the *DF* test is oversized for the typical sample sizes used in econometrics. This means that that the *DF* test rejects the null of a unit root too often, therefore implying cointegration when there in fact is none. Also the *JJ* method is not applicable when conditional heteroscedasticity is apparent in the errors.

---

[14] See e.g. Engle, R.F., (1982), Bollerslev (1986) or Franses and McAleer (2002)
[15] This volatility clustering was firs noted by Mandelbrot (1963)
[16] See e.g. Pantula (1986,1988) or Phillips & Peron (1988)
[17] See e.g. Kim & Schmidt (1993), Seo (1999) or Boswijk (2001)

4.2.7. *ARCH* and *GARCH*

Engle (1982) introduced the *autoregressive conditional heteroscedastic* (*ARCH*) process. In this process the unconditional variance is constant yet letting the conditional variance change over time "*as a function of past errors*". To allow for a more flexible lag structure and thus a more parsimonious description, Bollerslev (1986) introduced the *generalized ARCH* model (*GARCH*). Bollerslev argued that the extension of the *ARCH* model to *GARCH* resembles the extension of the *AR* model to the *ARMA* model. This means that a higher order *ARCH* model can be represented with a *GARCH* (*p, q*) process. Any *GARCH* (*p, q*) process with low values of *p* and *q* will be a better fit to the data than an *ARCH* (*p*) with a high value of *p*. According to Bollerslev (1986) the *ARCH* (*p*) process is only a linear function of the past sample variances, while the *GARCH* (*p, q*) process permits lagged conditional variances to enter the process as well. He explains this to be "*some sort of adaptive learning mechanism*".

Suppose y_t can be modeled as:

$$y_t = \alpha_0 + \alpha_1 x_{1t} + \cdots + \alpha_k x_{kt} + \varepsilon_t, \qquad \varepsilon_t | \Omega_{t-1} \sim N(0, h_t) \qquad (4.2.12)$$

Then the conditional variance, h_t in a simple *GARCH* (*1,1*) is:

$$h_t = \omega_0 + \omega_1 \varepsilon_{t-1}^2 + \omega_2 h_{t-1} \qquad (4.2.13)$$

If *q=0* in the *GARCH* (*p, q*) process, it reduces to an *ARCH* (*p*) process. Thus, in an *ARCH* (*1*) h_t reduces to:

$$h_t = \omega_0 + \omega_1 \varepsilon_{t-1}^2 \qquad (4.2.14)$$

If both *p* and *q* equal zero the errors, ε_t is white noise. Engle (1982) suggested a Lagrange Multiplier (*LM*) test to pre-test the data with, when *ARCH* is suspected. Bollerslev (1986) proposed a similar test for the *GARCH* model. The first step is to estimate with *OLS* the standard regression model for y_t to obtain the fitted residuals, $\hat{\varepsilon}_t$. Then the squared residuals are regressed on a constant and its lags:

$$\hat{\varepsilon}_t^2 = \omega_0 + \omega_1 \varepsilon_{t-1}^2 + \cdots + \omega_q \varepsilon_{t-q}^2 + v_t \qquad (4.2.15)$$

The null hypothesis is that all the coefficients are zero, i.e. no *ARCH* or *GARCH* effects. In practice the sample size, *T,* is multiplied by R^2 of the $\hat{\varepsilon}_t^2$ regression, which is asymptotically equivalent to the true *LM* statistic. Under the null this has a χ^2 distribution with *q* degrees of freedom.

4.2.8. Conditional heteroscedastic cointegration analysis

Assuming that the error terms follow a *GARCH* process one can test for cointegration by first adjusting the model specification to allow the errors to follow a *GARCH* process and then testing with the *ADF*-test if the errors have a unit root. Due to the robustness of the *DF* statistic one should be able to see if the null of a unit root, i.e. the null of no cointegration can be rejected.[18] However, one must be aware of the increased probability of spurious regression. Seo (1999) also proposed estimating the *AR* and *GARCH* parameters jointly by *ML* method. He then tests the *t*-statistic of the jointly estimated model for a unit root and tabulated critical values that are a mixture of the conventional *DF* distribution and the *standard normal* distribution. Seo (1999) finds that when the *GARCH* effect increases, the power of the *t*-statistic increases relative to the traditional *ADF* test.

4.2.9. Structural breaks and *GARCH*

If a model is miss-specified due to e.g. parameter instability than standard econometric theory will no longer apply. Lamoureux and Lastrapes (1990) showed that the *GARCH* estimates, which are a measure of the persistence of variance, may be biased upwards if possible structural shifts are unaccounted for in the unconditional variance. According to them this can also lead to misinterpretation of the *GARCH* estimates. More recently, other studies, such as Diebold and Inoue (2001), have shown that the occurrence of breaks may be confused with "long memory" (the presence of statistically significant correlations between observation that are a large distance apart).[19] Most standard tests for structural breaks are not applicable in the presence of conditional heteroscedasticity. A change in variance for a typical financial return series with obvious and complex clustering pattern is very difficult to identify. Carrasco and Chen (2002) have shown that most univariate *GARCH* processes are β-mixing[20]. Therefore, the applications of many standard tests for structural breaks are not suitable, since these require uniform mixing conditions.

4.2.10. Cointegration analysis with structural breaks and conditional heteroscedasticity

When a structural break occurs, the cointegrating vector shifts. Thus, there are no linear combinations of the variables that can become stationary (Westerlund and Edgerton, 2007). For that reason, standard cointegration tests are not appropriate and a cointegration test that

---

[18] Process used in Liu and Shrestha (2008)
[19] Definition from Harris and Solis (2003), pp.213
[20] For more detailed description on β-mixing see Carrasco and Chen (2001); Chen, Hansen, and Carrasco, 2010

allows for structural breaks should be used. The *GH* approach of cointegration analysis with structural breaks is very popular but suffers greatly from size distortion by conditional heteroscedasticity (Höglund and Östermark, 2003). Also, it does not allow for more than one break point since it is not computationally practical to extend the model (Westerlund and Edgerton, 2007).

Westerlund and Edgerton (2007) extended the *LM*-based unit-root test.[21] Their test for the null of no cointegration accommodates for heteroscedastic and serially correlated errors, deterministic trends as well as unknown breaks in both the intercept and slope. The model assumes only one cointegrating vector. They find that compared to other existing tests their method performs quite well. They also show that since the asymptotic null distributions of the tests are independent of the nuisance parameters associated with both trend and structural breaks, there is no need to compute different critical values for different breakpoint patterns.

The Westerlund and Edgerton (2007) test is performed in four steps. Initially, the first differenced variables, with dummy variables for the structural breaks, are regressed by OLS to obtain the trend, $\hat{\tau}$ and beta estimates:

$$\Delta Y_t = \hat{\tau} + \Delta x'_t \hat{\beta} + error \qquad (4.2.16)$$

There after the intercept term, α, can be estimated using the coefficients received from the fist regression.

$$\hat{\alpha} = y_1 - \hat{\tau} - x'_1 \hat{\beta} \qquad (4.2.17)$$

The first differenced residual series, $\hat{S}_t$, of the long-run relationship can be written as:

$$\hat{S}_t = y_t - \hat{\alpha} - \hat{\tau}t + x'_t \hat{\beta} \qquad (4.2.18)$$

Finally, the test equation is estimated, but it has to be modified to allow for conditional heteroscedasticity and serial correlation. It is assumed that a standard functional limit theorem applies. Westerlund and Edgerton (2007) argue that it seems reasonable to expect a test that takes explicit account of the structure should work best, as long as the *DGP* can be approximated by a parametric model. They find through simulation that a parametric method indeed works best and apply it to modify the test equation:

$$\Delta \hat{S}_t = constant + \phi \hat{S}_{t-1} + \sum_{j=1}^{P} \delta_j \Delta \hat{S}_{t-j} + error \qquad (4.2.19)$$

---

[21] Developed by Schmidt and Phillips (1992); Ahn (1993); and Amsler and Lee (1995)

According to Westerlund and Edgerton the lag order *P* is chosen so that it whitens the errors. The *LM*-based test statistics can then be obtained using $\hat{\phi}$ and its *t*-ratio. The null of no cointegration is equal to the null that $\phi = 0$. The critical values computed by Schmidt and Phillips (1992) are used.[22]

To allow for conditional heteroscedasticity the test equation can be modified to a *GARCH (p,q)* model:

$$\hat{S}_t = constant + \phi\hat{S}_{t-1} + \varepsilon_t, \qquad \varepsilon_t|\Omega_{t-1}\sim N(0, h_t) \qquad (4.2.20)$$

Where, as a *GARCH (1,1)* model, the conditional variance is:

$$h_t = \omega_0 + \omega_1\varepsilon_{t-1}^2 + \omega_2 h_{t-1} \qquad (4.2.21)$$

Lags may be added to the test equation to *whiten* the errors. The test for the null of no cointegration has changed to test if $\phi = 1$. The asymptotic distribution of the *LM* test statistic should not change due to the modification, and the same critical values computed by Schmidt and Phillips (1992) may be used.

5. Empirical study

In this section the data used is going to be presented. Thereafter the empirical method and results are described. The last part of the section concludes with a discussion of the results.

5.1. Data

The data for this paper was collected from the National Bureau of Statistics of China for all variables except for the interest rate which was obtained from Reuters Datastream. This study uses monthly data from June 2001 to December 2010 which adds up to a sample size of 115 observations. The empirical model has five explanatory variables: Inflation (*CPI)*, Money Supply (*MS*), Interest rate (*LIR*), Exchange rate (*EX*) and Industrial Production (*IP*). The dependent variables are the Shanghai stock exchange (*SH*) and the Shenzhen stock exchange (*SZ*). For the two Stock exchanges the composite index of the respective exchange is used. Inflation is measured as the C*onsumer Price Index (CPI).* To define *MS* we use the data for

---

[22] Schmidt and Phillips (1992), critical values Table 1, $t_n \Leftrightarrow \bar{\tau}$

M2. We use the *long term interest rate* (*LIR*) measured as the 5-year time deposit rate set by the *People's Bank of China* as a proxy for the interest rate variable. The *RMB Yuan* to *USD* exchange rate is used for the variable *EX*. The Gross Output Value of Industry is used as a proxy for the variable *IP*. All of the variables are converted into natural logarithms.

5.2. Methodology and results

5.2.1. The empirical model

The empirical model for Shanghai is:

$$\ln(SH_t) = \alpha_{SH} + \gamma_{SH}t + \beta_{1,SH}\ln(CPI_t) + \beta_{2,SH}\ln(MS_t) + \beta_{3,SH}\ln(LIR_t) +$$

$$+\beta_{4,SH}\ln(EX_t) + \beta_{5,SH}\ln(IP_t) + \varepsilon_{t,SH} \qquad (5.2.1)$$

And for Shenzhen:

$$\ln(SZ_t) = \alpha_{SZ} + \gamma_{SZ}t + \beta_{1,SZ}\ln(CPI_t) + \beta_{2,SZ}\ln(MS_t) + \beta_{3,SZ}\ln(LIR_t) +$$

$$+\beta_{4,SZ}\ln(EX_t) + \beta_{5,SZ}\ln(IP_t) + \varepsilon_{t,SZ} \qquad (5.2.2)$$

Where α_i denotes the intercept, γ_i represent the time trend indicated by a visual inspection of the individual series. $\beta_{i,j}$ is the coefficient for the respective variable *j* and *i* denotes the market (SH and/or SZ). ε_i represents the error term.

According to econometric theory most, if not all, of these series are likely to contain unit roots and a regression with these non-stationary series should be considered spurious if they are not cointegrated. The *EG* method of testing for cointegration assumes that all of the series are *I(1)* and if they are cointegrated, a linear combination of these series will produce *I(0)* residuals. Therefore we first test, with both the usual *ADF*-test and *PP*-test, if all of the series individually contain a single unit root. The results imply with high statistical significance that they all do.

With six variables we may have more than one cointegrating relation, and thus the *EG* method is not the most favorable, since it will only estimate one cointegrating vector. The *JJ* method would be preferable, but it will not be reliable if the series suffer from conditional heteroscedasticity or structural breaks. Therefore, the empirical models are first estimated and tested for *ARCH* and *GARCH* effects. The results are extremely significant (with a *p-value* of 0.000) already at the first lag. Therefore we conclude that the model should be modified to

allow for *ARCH* or *GARCH* specification. Using both the *Akaike's information criterion (AIC)* and the *Schwarz' Bayesian information criterion (BIC)* for model selection we find that the *ARCH (1)* specification has the *best fit* for both models.

5.2.2. *EG* cointegration test

Since the *JJ* method is not reliable to test for cointegration in the presence of conditional heteroscedasticity and the *DF*-statistic has been found to be robust,[23] the *EG* method is chosen. Both the *ADF*- and *PP*-test for the null of a unit root is used to test for cointegrating relation between the variables. The critical values used are the MacKinnon (2010) values for six variables. The tests show with a 5% significance level[24] that the null of no cointegration can be rejected for both models. Since all misspecification in the stationary part of the model is asymptotically dominated by the non stationary part, the cointegrated regression produces "super consistent" estimator of the cointegrating relation, even with incomplete short-run dynamics and autocorrelated errors (Verbeek, 2008).

Assuming no structural breaks have occurred during the test period, the heteroscedastic cointegration analysis suggests that a long run relationship exists between the variables at a 5% significance level. Thus, the results imply that these macroeconomic variables can explain long-run co-movements in the Chinese stock market. However, due to the conditional heteroscedasticity which increases the probability of over rejection of the null, a higher significance level should be used. At the 1% level the null of no cointegration cannot be rejected. This would therefore suggest that the macroeconomic variables cannot be statistically shown to have a long-run relationship with the stock markets.

5.2.3. *GH* cointegration test

What if structural breaks have occurred during the sample period? A visual inspection of the stock indices in figure 3 below suggests a possible level shift sometime between 2006 and 2009.

---

[23] See e.g. Pantula (1986,1988) or Phillips & Peron (1988)
[24] See Table 1 and Table 2 in appendix
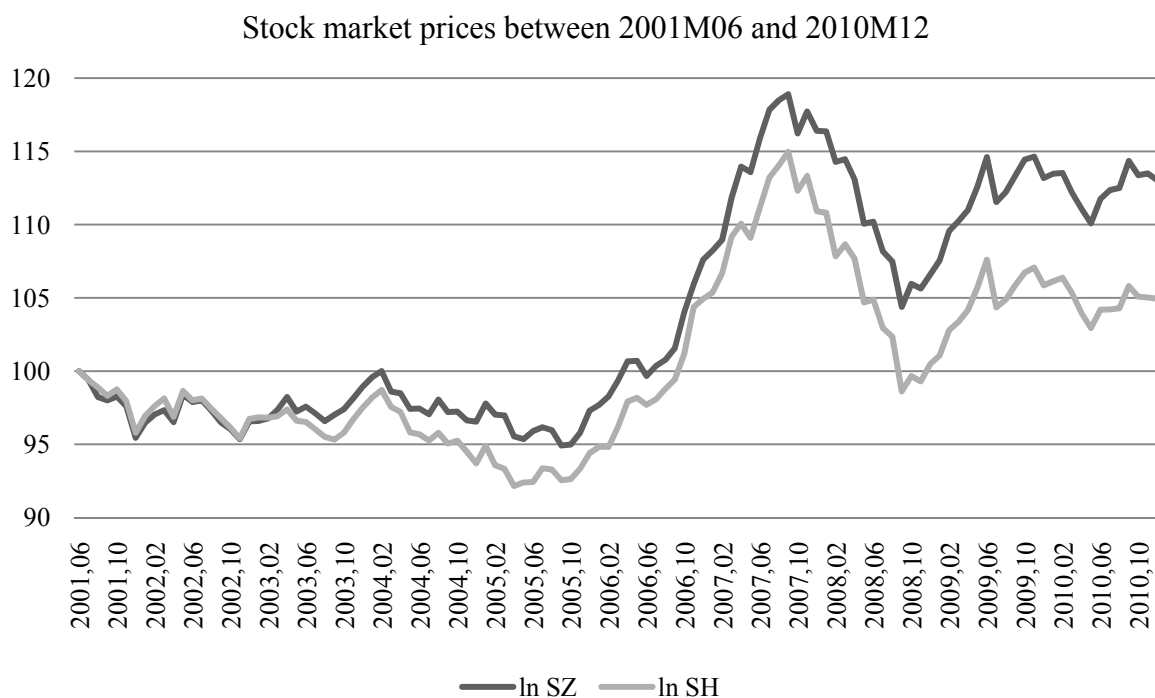
Stock market prices between 2001M06 and 2010M12

Figure 3 Natural logarithms of the nominal stock prices between 2001M06 and 2010M12 (index 100 = 2001M06) Source: NBS, 2011

In the occurrence of a structural shift the usual cointegration tests suffer from loss of power. The *GH* method can be used to test for cointegration with unknown break points. For this test the paper uses the empirical models (*5.2.1*) and (*5.2.2*) and adds dummies for the level shift and/or regime shift. Example for the level shift model:

$$\ln (Y_{i,t}) = \alpha_{1,i} + \alpha_{2,i}\varphi_{tk} + \gamma_i t + \beta_{1,i}\ln (CPI_t) + \beta_{2,i}\ln (MS_t) + \beta_{3,i}\ln (LIR_t) +$$

$$+\beta_{4,i}\ln (EX_t) + \beta_{5,i}\ln (IP_t) + \varepsilon_{t,i} \qquad (5.2.3)$$

Where $Y_{i,t}$ stands for the stock index and *i* equals *SH* and *SZ*. α_2 is the new intercept after the level shift has occurred. φ_{tk} denotes the dummy variable. The break point can be determined by choosing the time period which produces the lowest *ADF*-test *t*-statistic (highest negative value). For this study several (but not all) possible break points[25] were tested and the lowest *t*-statistic for both models was achieved whit a level shift in 2006M10. Using the Gregory and Hansen (1996) critical values for the *ADF*-test with a level shift we cannot reject the null of no cointegration.[26] However, this does not mean that we can accept the null (!). The *GH* method has been shown to suffer greatly from size distortion when the errors display

---

[25] The dates tested are chosen from a visual inspection of figure 3. The break points tested are: 2005M10, 2006M10, 2007M03, 2007M10, 2008M10 which has been tested individually and in different combinations for level shift as well as regime shift.
[26] See Table 3 and Table 4 in appendix

conditional heteroscedasticity. It could therefore be, due to the conditional heteroscedasticity, that the *GH*-test is unable to reject the null of no cointegration.

5.2.4. The Westerlund and Edgerton cointegration test

The final test for cointegration used in this study is the Westerlund and Edgerton (2007) *LM*-type test. Starting without any structural breaks but adjusting the test equation to accommodate for *GARCH* effects and insuring errors are white noise, the test cannot reject the null of no cointegration. This is expected since the omission of existing breaks can lead to substantial loss of power. The test is repeated with several (but not all) possible break points around the supposed break period with both the level and regime shift model. As with the *GH* procedure, the break point and model that produces the lowest *t*-statistic is chosen.

For Shanghai a level shift in 2006M10 and for Shenzhen a level shift in 2007M03 produce the lowest *t*-values. The tests are unable to reject the null of no cointegration.[27] Again, this does not mean that the null can be accepted, only that the test cannot statistically find cointegration. There may very well be a long run cointegrating relation, but this study has been unable to find the right break point. As with other cointegration tests, omission or incorrect placement of the break reduces the power of the test, thus increasing the probability of accepting the null of no cointegration.

The results of this study show that the tests which take into account for structural breaks are unable to reject the null of no cointegration. Taking into account conditional heteroscedasticity but omitting the structural break(s) makes the *EG*-test difficult to interpret. The *EG*-test does not reject the null at the 1% level which would be preferred due to the bias towards over rejecting caused by the heteroscedasticity. However, the *EG*-test does reject the null at the 5% level which might be acceptable due to the lower power of the test when the break(s) are omitted.

5.3. Discussion

5.3.1. The cointegration tests

So why do the result differ when assuming structural break(s), and which model is the most proper one to use? Since the results vary, it is difficult to draw any conclusions on why no cointegrating relationship is found. Is it because of the inefficiency of the Chinese market, or

---

[27] See Table 5 and Table 6 in appendix

because of the difficulty to reject no cointegration when assuming structural breaks, or is it because of the difficulty to find the right break point in the observed time period? All of our methods used contain assumptions and weaknesses that make it hard to determine which method is best in explaining reality.

According to theory, cointegration tests suffer from low power in the presence of structural breaks. This implies that even if the *EG* method cannot reject the null of no cointegration; the hypothesis can be rejected using the *GH*, or Edgerton and Westerlund approach. The result from this study has on the other hand shown the opposite result. The reason for this result might be that we have not been able to find the right structural break.

So, why was the null of no cointegration not rejected in the *GH* and Westerlund and Edgerton test? The *GH* looses power if conditional heteroscedasticity is present, which is likely when dealing with financial data. In the Westerlund and Edgerton test, power is lost when omitting structural break(s) when a structural break is present in the observations. Power is also lost when misplacing the break.

Unfortunately, if a structural break has occurred, this study has not been able to find the right break and the power of the Edgerton and Westerlund and *GH* has thus been too low to be able to reject the null of no cointegration. Another possible reason for rejecting the null is that the right structural break was found, but that the macroeconomic variables are not cointegrated with the stock market. If the right structural break was not found it is partly due to time limitations. It takes a large amount of time to test for cointegration in heteroscedastic models when the structural break is unknown. The reason for this is that each date and possible structure of the break must be tested one by one. The time period for possible break(s) is long; therefore the right break might have been missed since over hundred tests must be performed. It might even be the case that several breaks have occurred in the period which makes the right structure of the break even harder to find.

If it is harder to reject the null of no cointegration in the *EG* method but the null was rejected, why is this result not satisfying? According to the econometric theory, the null of no cointegration is rejected too often due to conditional heteroscedasticity. Even though the test-statistic is asymptotically robust, the null is likely to be rejected too often in these types of sample sizes (as discussed in 4.2.6). Thus, there are two opposite effects that might cancel out. First, the errors contain *ARCH* effects that reject the null hypothesis, and secondly

structural breaks that are likely to be present which make it harder to reject the null. In this study, it is not known which effect is the strongest.

To avoid rejecting no cointegration too easily, the significance level can be lowered. When using a 1% significance level in the *EG* approach the null of no cointegration cannot be rejected. However, if no *ARCH* effects would have been present, the higher significance level could have been acceptable, 10% say. Therefore rejection of the null at a 5 % significance level might be acceptable when both conditional heteroscedasticity and structural break(s) are present under the alternative hypothesis. With the *EG* method it is difficult to determine whether or not the macroeconomic variables are cointegrated with the stock market.

5.3.2. Implication for investment choices

If the 5% significance level is accepted in the *EG* method the results show that macroeconomic variables can explain co-movements in the stock market in the long run. This implies that, when making long term investment decisions in China, the individual should take the macroeconomic variables into account. This is the same result showed in Liu and Shresthas study and is also consistent with the studies performed on the developed markets. This implies that macroeconomic variables can explain movements in both developed markets and the Chinese emerging markets, despite the limited accessibility to information in China. So, despite the limited efficiency in the market, the individual seems to react to changes in business conditions that in turn are affected by macroeconomic variables. This result however has proven to be hard to support if stronger assumptions are made about the observed data.

The theory states that the willingness to pay for an asset is based on the economic business conditions. Imagine that an unexpected event occurs which leads to a new business condition. If the investors think that the new business condition is constant, he will make all his future investment decisions on this new information and not on the historical business opportunities. This will lead to a structural break in the aggregated demand function, and hence the asset prices. When looking at the data series for asset prices in both Shanghai and Shenzhen stock market, it seems that a break has occurred sometime between 2006M03 and 2009M03.

When adjusting the data to the possibility of over rejecting the null in the *EG* approach due to conditional heteroscedasticity, this study is not able to support that macroeconomic variables can explain co-movements in the stock market. The *GH* and Westerlund and Edgerton approaches were also unable to support a cointegrating relation. This implies that, with the

current econometric techniques, without knowledge of where and if a structural break has occurred, it is difficult to conclude that macroeconomic variables should be used when making long term investment decisions in China. Thus, this study cannot give any support to that investors should look at macroeconomic variables when making long term investments.

The global financial crisis has been present in a large part of this study's sample. This could have affected the results. A different result might have been found if this study would have been conducted during a more stable time period on the world markets. During a more stable time period, the cointegration tests have more power. This might be the reason why Liu and Shrestha found stronger evidence for cointegration between macroeconomic variables and the co-movement in the stock markets.

Due to the uncertainty on whether a possible structural break has occurred but not been found, or if it is just that macroeconomic variables cannot explain co-movements in the Chinese markets, it is difficult to draw any conclusions on how the investor should behave. For further research it would be interesting to se if other markets also have problems with rejecting the null hypothesis of no cointegration during the global financial crisis. If not, it might just be the limited information flow, the large amount of private investors (with limited resources to collect and sort information), and other market inefficiencies in China that make it difficult to find evidence of cointegration.

.

6. Conclusions

The purpose of this study was to investigate if macroeconomic variables can be used to explain the co-movements in the stock markets in China, namely those of Shanghai and Shenzhen. A cointegration analysis was performed by using inflation, interest rate, exchange rate, industrial production, and money supply as proxies for the macro economy. Because it is likely that a structural break had occurred in the investigated period, and because the data contained conditional heteroscedasticity, the most commonly used cointegration test approaches made it difficult to determine if a cointegrating relation exists. Therefore, an additional test for cointegration developed by Westerlund and Edgerton was performed. The test could handle both heteroscedasticity and structural breaks but was unable to find cointegration. This study was therefore not able to confirm that macroeconomic variables can be used to explain co-movements in the Chinese stock market. The authors hope for further research in the area of handling conditional heteroscedastic models with structural breaks.

References

Ahn, S.K. and C.G. Reinsel, (1990) "Estimation for Partially Nonstationary Multivariate Autoregressive Models" in: *Journal of the American Statistical Association*, Vol. 85, pp. 813-823, American Statistical Association

Ahn, S. K. (1993) "Some tests for unit roots in autoregressive-integrated-moving average models with deterministic trends" in: *Biometrica*, Vol. 80, pp. 855–868, Oxford University Press

Amsler, C. and J. Lee (1995) "An LM test for a unit root in the presence of a structural break" in: *Journal of Econometric Theory*, Vol. 11, pp. 359–368, Cambridge University Press

Banerjee, A., J.J. Dolado, J.W. Galbraith and D.F. Hendry (1993) *Co-integration, Error-Correction, and the Econometric Analysis of Non-Stationary Data,* Oxford University Press, Oxford

Brealey, R.A. and S.C. Myers (2002), *Principles of coporate finance*, 7th ed. McGraw-Hill/Irwin

Bollerslev, T. (1986) "Generalized autoregressive conditional heteroscedasticity" in: *Journal of Econometrics*, Vol. 31, No. 3, pp. 307-327, Elsevier

Boswijk, H.P. (2001) "Testing for a unit root with near-integrated volatility" *Tinbergen Institute Discussion Paper*, pp. 1-77, Tinbergen Institute, The Netherlands

Box, G.E.P and G.M. Jenkins (1970) *Times Series Analysis, Forecasting and Control*, 1st Ed., Holden Day, San Francisco

Carrasco M. and X. Chen (2002) "Mixing and moment properties of various GARCH and stochastic volatility models" in: *Journal of Econometric Theory*, Vol. 18, No.1, pp. 17-39, Cambridge University Press (CUP)

Chen, N.-F., R. Roll and S.A. Ross (1986) "Economic Forces and the Stock Market" in: *The Journal of Business*, Vol. 59, No. 3, pp. 383-403, The University of Chicago Press

Chen, N.-F. (1991) "Financial Investment Opportunities and the Macroeconomy" in: *The Journal of Finance,* Vol. 46, No. 2, pp. 529-554, Blackwell Publishing

Chen, X., L.P. Hansen and M. Carrasco (2010) "Nonlinearity and temporal dependence" in: *Journal of Econometrics*, Vol.155, No. 2, pp. 155–169, Elsevier

Cheung, Y.–W. and K.S. Lai (1993) "Finite Sample Sizes of Johansen's Likelihood Ratio test for Cointegration" in: *Oxford Bulletin of Economics and Statistics*, Vol. 55, No. 3, pp. 313-332, Blackwell Publishing

Cheung, Y.–W. and L.K. Ng (1997) "International evidence on the stock market and aggregate economic activity" in: *Journal of Empirical Finance*, Vol. 5, No. 3, pp. 281-296, North-Holland

Cox, J.C., J.E. Ingersoll Jr. and S.A. Ross (1985) "An intertemporal General Equilibrium Model of Asset Prices" in: *Econometrica*, 1985, Vol. 53, No. 2, pp. 363-384, The Economy Society

Dickey, D.A. (1976) *Estimation and Hypothesis Testing for Nonstationary Time Series*, Unpublished Ph.D. dissertation, Iowa State University, Ames

Dickey, D.A. and W.A. Fuller (1979) "Distribution of the Estimators for Autoregressive Time Series with a Unit Root" in: *Journal of the American Statistical Association*, Vol. 74, No. 366, pp. 427-431, American Statistical Association

Diebold F.X. and A. Inoue (2001) "Long memory and structural change" in: *Journal of Econometrics*, Vol. 105, No. 1, pp. 131-159, North-Holland

Ely, D. and M. Salehizadeh (1999) "Predicting Stock and Bond Returns: the Impact of Global Variables" in: *International Finance*, Vol. 2 No. 2, pp. 203–226, Blackwell Publishers Ltd.

Engle, R.F. (1982) "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation" in: *Econometrica*, Vol. 50, No. 4, pp. 987-1007, The Econometric Society

Engel, R.F. and C.W.J. Granger (1987) "Cointegration and Error Correction: Representation, Estimation and Testing" in: *Econometrica*, Vol. 55, pp. 251-276, The Econometric Society

Fama, E.F. and K.R. French (1989) "Business Conditions and Expected Returns on Stocks and Bonds" in: *Journal of Financial Economics,* No. 25, pp. 23-49, Elsevier

Fama, E.F. and G.W. Schwert (1977) "Asset returns and inflation" in: *Journal of Financial Economics*, Vol. 5, No. 2, pp. 115-146, Elsevier

Fen, L. and S. Tu (2010) "Empirical Analysis of Factors Influencing China's Stock Market" in: *2010 International Conference on E-Business and E-Government*, pp. 332-336, IEEE

Franses, P.H. and M. McAleer (2002) "Financial Volatility: An introduction" in: *Journal of Applied Econometrics*, Vol. 17, pp. 419-424, John Wiley & Sons

Fuller, W.A. (1976) *Introduction to Statistical Time Series*, 1st Ed., John Wiley & Sons, New York

Geske, R. and R. Roll (1983) "The Fiscal and Monetary Linkage between Stock Returns and Inflation" in: *The Journal of Finance*, Vol. 38, No. 1, pp. 1-30, American Finance Association

Gilson, R.J. and R.H. Kraakman (1984) "The Mechanism of Market Efficiency" in: *Virginia Law Review*, Vol. 70, No. 4, pp. 549-644, Virginia Law Review

Granger, C.W.J. (1981) "Some properties of time series data and their use in econometric model specification" in: *Journal of Econometrics*, Vol. 16, No. 1, pp. 121-130, Elsevier

Granger, C.W.J. and P. Newbold (1974) "Spurious Regressions in Econometrics" in: *Journal of Econometrics*, Vol. 2, No. 2, pp. 111-120, Elsevier

Gould, J.P. and C.R. Nelson (1974) "The Stochastic Structure of the Velocity of Money" in: *American Economic Review*, Vol. 64, No. 3, pp. 405-418, The American Economic Association

Gregory, A.W. and B.E. Hansen (1996) "Residual –based tests for cointegration in models with regime shifts" in: *Journal of Econometrics*, Vol. 70, pp. 99-126, North-Holland

Harris, R.I.D. (1992) "Testing for Unit Roots Using the Augmented Dickey-Fuller Test: Some Issues Relating to the Size, Power and Lag structure of the Test" in: *Economic Letters*, Vol. 38, No. 4, pp. 381-386, Elsevier

Harris, R. and R. Sollis (2003), *Applied time series modelling and forecasting*, 1st ed., Wiley

Harvey, D.I., S.J. Leybourne and P. Newbold (2001) "Innovation outlier unit root tests with an endogenously determined break in level" in: *Oxford Bulletin of Economics and Statistics*, Vol. 63, No.5, pp. 559-575, Blackwell Publishing

He, J. and L.K. Ng (1998) "The Foreign Exchange Exposure of Japanese Multinational Corporations" in: *Journal of Finance*, Vol. 53, No. 2, pp. 733-753, Blackwell Publisher Inc.

Höglund, R. and R.Östermark 2003 "Size and Power of Some Cointegration tests under Structural Breaks and Heteroskedastic Noise" in: *Statistical Papers*, Vol. 44, No. 1, pp. 1-22, Springer

Jensen, G.R., J.M. Mercer and R.R. Johnson (1996) "Business conditions, monetary policy, and expected returns" in: *Journal of Financial Economics*, Vol. 40, No. 2, pp. 213-237, Elsevier

Johansen, S. (1988) "Statistical Analysis of Cointegration Vectors" in: *Journal of Economic Dynamics and Control*, Vol. 12, No. 2-3, pp. 231-254, Elsevier

Johansen, S. and K. Juselius (1990) "Maximum Likelihood Estimation and Inference on Cointegration - With Application to the Demand for Money" in: *Oxford Bulletin of Economics & Statistics*, Vol. 52, No. 2, pp.169-210, Blackwell Publishing

Jorion, P. (1991) "The Pricing of Exchange Rate Risk on the Stock Market" in: *The Journal of Financial and Quantitative Analysis*, Vol. 26, No. 3, pp. 363-376, University of Washington School of Business Administration

Kaul, G. (1987) "Stock Returns and Inflation" in: *Journal of Financial Economics*, Vol. 18, No. 2, pp. 253-276, North-Holland

Kim, K. and P. Schmidt (1993) "Unit root tests with conditional heteroscedasticity" in: *Journal of Econometrics*, Vol. 59, No. 3, pp. 287-300, Elsevier

Lamoureux C.G. and W.D. Lastrapes (1990), "Persistence in variance, structural change and the GARCH model" in: *Journal of Business and Economic Statistics*, Vol. 8, No. 2, pp. 225-234, American Statistical Association

Lee, B.-S. (1992) "Causal Relations among Stock Returns, Interest Rates, Real Activity, and Inflation" in: *The Journal of Finance*, Vol. 47, No. 4, pp. 1591-1603, Blackwell Publishing (American Finance Association)

Li, G. (2008) "Chinas Stock Market: Inefficiencies and Institutional Implications" in: *China & World Economy*, Vol. 16, No. 6, pp.81-96, Blackwell Publishing

Liu, M.-H. and K.M. Shrestha (2008) "Analysis of the long-term relationship between macroeconomic variables and the Chinese stock market using heteroscedastic cointegration" in: *Managerial Finance*, 2008, Vol. 34, No. 11, pp. 744-755, Emerald Group Publishing Limited

MacKinnon, J. G. (2010) *Critical Values for Cointegration Tests*, Queen's Economics Department Working Paper No. 1227, Queen's University

Mandelbrot, B. (1963) "The Variation of Certain Speculative Prices" in: *Journal of Business*, Vol. 36, No. 36, pp. 394-419, University of Chicago Press

Merton, R.C. (1973) "An Intertemporal Capital Asset Pricing Model" in: *Econometrica*, Vol. 41, No. 5, pp. 867-887, The Econometric Society

Nankervis, J.C. and N.E.Savin (1985) "Testing the Autoregressive Parameter with the t-Statistic" in: *Journal of Econometrics*, Vol. 27, No. 2, pp. 143-161, Elsevier

NBS, National Bureau of Statistics (2011), *National Yearbooks*, Monthly statistics

Nelson, C.R. and C.I. Plosser (1982) "Trends and Random Walks in Macroeconomic Time Series" in: *Journal of Monetary Economics*, Vol. 10, pp. 139-162, Elsevier

Pantula, S.G. (1986) "Comment on modeling the persistence of conditional variances" in: *Econometric Reviews*, Vol. 5, No. 1, pp. 71-74, Taylor & Francis

Pantula, S.G. (1988) "Estimation of autoregressive models with ARCH errors" in: *Sankhya: The Indian Journal of Statistics*, Vol. 50, Series B, pp. 461-472, Indian Statistical Institute

Perron, P. (1987) "Test Consistency with Varying Sampling Interval" in: *Cahier de Recherche*, No. 4187, C.R.D.E., Universite de Montreal

Perron, P. (1988) "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach" in: *Journal of Economic Dynamics and Control*, Vol. 12, No. 2-3, pp. 297-332, Elsevier

Perron, P. (1989) "The great crash, the oil price shock, and the unit root hypothesis" in: *Econometrica* Vol. 57, No. 6, pp. 1361-1401, The Econometric Society

Perron, P. (1994) "Trend, Unit Root and Structural Change in Macro Economic Times Series" in B. Bhaskara Rau, in: *Cointegration for the Applied Economist*, Macmillan Press, Basingstoke

Phillips, P.C.B. and P. Perron (1988) "Testing for a Unit Root in Times Series Regression" in: *Biometrika*, Vol. 75, No. 2, pp. 335-346, Biometrika trust

Phillips, P.C.B. (1987) "Time Series Regression with Unit Roots" in: *Econometrica*, Vol. 55, No. 2, pp. 277-302, Wiley

Rapach, D.E., M.E. Woharb and J. Rangvid (2005) "Macro variables and international stock return predictability" in: *International Journal of Forecasting,* No. 21, pp. 137– 166, North-Holland

Reimers, H.-E. (1992) "Comparison of Tests for Multivariate Cointegration" in: *Statistical Papers*, Vol. 33, No. 1, pp. 335-359, Springer

Ross, S. (1976) "The Arbitrage Theory of Capital Asset Pricing" in: *Journal of Econometric Theory,* Vol. 13, No. 3, pp. 341-360, Elsevier

Said, S.E., and D.A. Dickey (1984) "Testing for Unit Roots in Autoregressive - Moving Average Models of Unknown Order" in: *Biometrika*, Vol. 71, No. 3, pp. 599-607, Oxford University Press

Samuleson, P.A. (1973) "Proof that Properly Discounted Present Values of Assets Vibrate Randomly" in: *The Bell Journal of Economics and Management Science*, Vol. 4, No. 2, pp. 369-374, American Telephone and Telegraph Company

Schmidt, P. and P.C.B. Phillips (1992) "LM tests for a unit root in the presence of deterministic trends" in: *Oxford Bulletin of Economics and Statistics*, Vol. 54, No. 3, pp. 257– 287, Blackwell Publishing

Seo, B. (1999) "Distribution theory for unit root tests with conditional heteroscedasticity" in: *Journal of Econometrics*, Vol. 91, No. 1, pp. 113-144, North-Holland

Sharpe, D.W. (1964) "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk" in: *Journal of Finance*, Vol. 19, No. 3, pp. 425-442, American Finance Association

Stulz, R.M. and W. Wasserfall (1985) "Macroeconomic Time Series, Business Cycles and Macroeconomic Policies" in: *Understanding Monetary Regimes*, ed. by Karl Brunner and Allan H. Metzler. Carnegie-Rochester Conference Series on Public Policy, Vol. 22, pp. 9-53, Elsevier

Verbeek, M. (2008) *A guide to modern econometrics*, 3rd Ed., pp. 331, Wiley Publishing

Wasserfallen, W. (1986) "Non-Stationarities in Macro-Economic Time Series -Further Evidence and Implications" in: *Canadian Journal of Economics*, Vol. 19, No. 3, pp. 498-510, University of Toronto Press

Westerlund, J. and D.L. Edgerton (2007) "New Improved Tests for Cointegration with Structural Breaks," in: *Journal of Time Series Analysis*, Vol. 28, No.2, pp. 188-224, Blackwell Publishing

Yuan, L. and L. Chen (2010), "The Relationship of Stock Returns and Inflation in China" in: *2010 International Conference on E-Business and E-Government*, pp. 5319-5322, IEEE

Zeng, Z. and X. Wan (2007) "Understanding the recent performance of China's stock market" in: *China Economic Journal*, Vol. 1, No. 2, pp. 213-225, Routledge

Appendix

Table 1 Cointegration test with *EG* method for Shanghai

Shanghai Cointegration test	
Test	**t-statistic**
ADF	-5.30
PP	-5.44
Test critical values*	
1% level	-5.52
5% level	-4.98
10% level	-4.70

*MacKinnon 2010

Table 2 Cointegration test with *EG* method for Shenzhen

Shenzhen Cointegration test	
Test	**t-statistic**
ADF	-5.19
PP	-5.32
Test critical values*	
1% level	-5.52
5% level	-4.98
10% level	-4.70

*MacKinnon (2010)

Table 3 Cointegration test with *GH* method for Shanghai with constant and trend

Shanghai Cointegration test	
Test	**t-statistic**
ADF(level 2006M10)	-4.08
Test critical values*	
1% level	-6.36
5% level	-5.83
10% level	-5.59

*Gregory and Hansen (1996)

Table 4 Cointegration test with *GH* method for Shenzhen with constant and trend

Shenzhen Cointegration test	
Test	**t-statistic**
ADF(level 2006M10)	-4.45
Test critical values*	
1% level	-6.36
5% level	-5.83
10% level	-5.59

*Gregory and Hansen (1996)

Table 5 Cointegration test with Westerlund & Edgerton method for Shanghai

Shanghai Cointegration test	
Test	**t-statistic**
E&W (no break)	-1.94
E&W (level break 2006M10)	-2.00
Test critical values*	
1% level	-3.63
5% level	-3.06
10% level	-2.77

* Schmidt and Phillips (1992)

Table 6 Cointegration test with Westerlund & Edgerton method for Shenzhen

Shenzhen Cointegration test	
Test	**t-statistic**
E&W (no break)	-0.33
E&W (level break 2007M03)	-1.23
Test critical values*	
1% level	-3.63
5% level	-3.06
10% level	-2.77

* Schmidt and Phillips (1992)