School of Economics and Management
Department of Economics &
Department of Business Administration
Master in Finance Program

Master Thesis

Overconfidence and turnover

-Evidence from the Hong Kong stock market

Authors:

Zeyuan Chen Shuman Zhang

Supervisor:

Frederik Lundtofte

August, 2011

Abstract

High trading volume is a common phenomenon in global financial markets, and it seems to go against the paradigm of classic economic theory based on rational agents and efficient markets. In this thesis, we turn to behavioral finance, as we try to explain this anomaly as a consequence of the so-called overconfidence bias. According to previous research, market return influences the level of overconfidence and overconfidence is the key factor leading to excessive trading. This thesis aims at investigating empirically whether overconfidence exists in the Hong Kong stock market. We construct a market-wide VAR model to investigate the lead-lag relationship between return and turnover, the presence of which can be considered as evidence of overconfidence. Our results suggest that investors are indeed overconfident in the Hong Kong market. To distinguish overconfidence from the disposition effect, we develop an individual security model by using a panel data approach. The empirical findings from the individual security model suggest that the overconfidence found in the Hong Kong market is not just an aggregation of individual disposition effects. Although the disposition effect may also be used as an explanation of the interaction between return and turnover, we find that overconfidence dominates in the Hong Kong stock market.

Key Word: overconfidence, disposition effect, VAR, panel data method

Table of Contents

1. Introduction

1.1 Background

High trading volume has been observed in global financial markets. As one of the most influential financial markets in the world, the New York Stock Exchange (NYSE)'s average monthly turnover in 2010 was approximately 100%[1]. As for the Chinese market, Zhu (2002) reveals that Chinese investors' trading frequency is four times higher than their U.S counterparts. Excessive trading has been considered "the single most embarrassing fact to the standard finance paradigm" (DeBondt and Thaler, 1994, p.392). Since classic models can not explain excessive trading, we resort to behavioral finance theory, which deviates from the assumption of rational agents. Many psychological studies and empirical studies in Finance have found that people are not always rational, and systematic cognitive biases will lead to deviations from inferences drawn by classic theory. In this thesis, we put an emphasis on the overconfidence bias, which is considered the key factor in understanding the trading puzzle in financial markets (DeBondt and Thaler, 1994).

Investors' overconfidence in their security valuation and trading skills has been discussed for years. According to Daniel et al. (1998), investors' overconfidence may be the result of biased self-attribution with regard to past investment outcomes. They argue that overconfidence implies overreaction to private information and underreaction to public signals and thus leads to market mispricing. Gevias and Odean (2001) improve the theory that some investors tend to exaggerate their own ability and ignore the fact that they are in a bull market. Moreover, while previous research emphasizes the return in the capital market, they consider the changes in trading volume as the main testable implication. Later, Statman et al. (2006) conduct empirical research regarding the impact of overconfidence on trading volume in the US market. Given that the level of overconfidence changes with market return, they use market return to measure the degree of overconfidence. They find a significantly positive relationship between market-wide turnover and lagged market returns and

[1] http://www.nyse.com/

view it as evidence of overconfidence. Glaser and Weber (2007) also document that investors with higher degrees of confidence tend to operate more in the German stock market, which is in line with Statman et al.'s (2006) finding.

In contrast to the U.S stock market, there are just a few overconfidence studies conducted on Asia's emerging markets. Allen, Qian & Qian (2005) find that compared to American investors, Chinese investors trade more frequently. By investigating broker accounts, Chen at al (2007) explain this difference with the higher levels of overconfidence among Chinese investors. However, Chinese A stocks are still immature and have a short history. In contrast, the Hong Kong stock market is one of the most mature stock markets in Asia. As it has a longer history, thus enabling us to obtain more precise estimates, we will focus on the Hong Kong stock market.

1.2 Problem discussion

Our paper focuses on the close connection between trading volume and overconfidence. By investigating broker accounts, Chen et al. (2007) have verified the existence of high overconfidence in the Chinese market. Considering data availability, we follow the paper of Statman et al. (2006), which takes turnover as a proxy for the level of overconfidence. Gervias and Odean's (2001) paper provides the basic framework for Statman's model: due to the self-attribution bias, high returns in a bull market will increase investors' overconfidence. On the other hand, the close relation between overconfidence and trading volume has been verified by several studies. Since overconfident investors believe in their abilities and will act based on the information they obtain, trading volume is affected. Hence, if the current trading volume can be explained by the past market return, it can be considered as an evidence of overconfidence. Based on this lead-lag relationship, we will apply a market-wide VAR model to examine the existence of overconfidence.

However, we should also take the disposition effect into consideration, as it also leads to excessive trades. Shefrin and Statman (1985) explain the disposition effect from the aspects of prospect theory, mental accounting, regret aversion and self-control. Investors' disposition to sell winners and ride losers is a consequence of their eagerness to realize the proceeds, and continue to reinvest. In this paper, we will try to distinguish the disposition effect from overconfidence. By connecting individual stocks' turnover with past market returns, we test whether the overconfidence throughout the whole market is driven by the aggregation of individual disposition effects.

1.3 Purpose

Previous empirical studies have verified the existence of investors' overconfidence in many countries. In this paper, we aim at investigating whether the overconfidence effect exists in the Hong Kong stock market by testing the interaction between trading volume and market return. Moreover, we also want to analyze how strong the impact of overconfidence is on market returns, and go further to investigate the reasons behind it. We believe that this may help us gain a better understanding of the Hong Kong stock market.

1.4 Limitation and contribution

Our empirical results, obtained by testing the lead-lag relationship between return and turnover, indicate that overconfidence exists in the Hong Kong stock market. The overconfidence effect seems not as strong as in the U.S. stock market (Statman et al., 2006), but this may well be due to our limited amount of data. On the other hand, we argue that both overconfidence and the disposition effect can explain the lead-lag relation between return and turnover. However, overconfidence has stronger explanatory power than the disposition effect in the Hong Kong stock market.

1.5 Outline

This paper is organized into five parts. Section 1 introduces the background and raises our research question. Section 2 provides a theory review and summarizes related research findings. The third section presents the data and discusses our methodology.

In the empirical section, we analyze our test results. Finally, in the last section, we summarize our main findings.

2. Theory review

The basic paradigm in classic Finance is that of rational agents and efficient markets. In an assumed ideal market, classic theory and models get improved rapidly. However, is traditional theory always valid in the real world? In the 1980s, a number of empirical studies on financial markets found that there were anomalies which could not be explained by classic economic theory. Thus, the initial assumptions were doubted, especially the rationality assumption, since people are not always rational when they make decisions according to psychological research. The application of cognitive psychology to analyze the impact of investors' behavior in financial markets led to the emergence of behavioral finance. In contrast to classic economics, behavioral finance takes subjective cognition of people into consideration, and concentrates on how irrationality influences the market.

In this paper, we will focus on overconfidence bias, an important concept in behavioral finance, which has been discussed by numerous literatures in cognitive psychology. Psychologists point out in earlier literature that people are usually overconfident and overestimate their own abilities. Glaser and Weber (2007) have summarized overconfidence into three forms: miscalibration, underestimation of volatility, and above average effects. We will discuss these three manifestations in turn. Lichtenstein et al. (1982) use the fractile method to analyze assessments of uncertain quantities and reveal that people tend to form too tight probability distributions. In their study people are asked to answer two-alternative-questions and then estimate the correct probability of their question. They find that the estimated probability of correct answer is higher than the assigned probability. Furthermore, people become more overconfident when they answer moderate to extreme difficult questions (Fischhoff et al., 1977). In addition, several questionnaire studies have been performed to investigate investors' estimated confidence interval for return or value of certain index/ stock in the future. Generally they document that estimated intervals are too narrow therefore the volatilities are underestimated (Hilton, 2001). These phenomena are labeled as miscalibration or overconfidence. Particularly, DeBondt (1998) studies 46 individual investors and discovers that their confidence intervals are tighter than the actual price varies. Glaser et al. (2010) obtain a comparable result for students and professional stock traders. Another manifestation of overconfidence is the better than average effect. Based on their survey data, Taylor and Brown (1998) argue that people see themselves as better than average and most individuals judge themselves better than other people view them. One famous example is from Svenson (1981) who shows that 82% of a group of students rank themselves among top 30% in terms of driving safety. It relates to the fact that people often consider themselves have better skills and knowledge compared to others. Glaser and Weber (2007) discover that more than half of investors think their abilities are above average, which make them to trade more. “Overconfidence is characteristic of people, not of markets.”(Odean, 1998a, p.1888). Our interest is that how the characteristic of people affect objective market.

2.1 Market return and overconfidence

The interaction of market return and overconfidence has been studies for years. Daniel et al. (1998) show that overconfidence arising from overreaction to private information and underrection to public signal will result in market mispricing. However, we emphasize on how market return influences overconfidence level. Many scholars such as Miller and Ross (1975) argue that people tend to exaggerate their own responsibility for the success. In other words, it is common to attribute success to their own ability and failure to external matters. People in the financial market are not the exception. As Gervias and Odean (2001) point out, owning to attribution biased, those who make successful investment tend to be more overconfident. They develop a model to measure how traders learn about their ability and how the self-attribution bias results in overconfidence. Their research starts with the assumption that traders don’t know their ability and they learn their ability through experience. In their case the dividend forecast is used to measure success. They reveal that a trader's overconfidence level varies and is associated with previous success and failure. The authors argue that this model could apply to the changing market condition. E.g. in the bull market investors are easier to make profit. Then the traders will become more overconfident as they attribute the success to their own ability and ignore the market circumstance. More specifically, it could be expected that aggregated overconfidence is higher if market gains and otherwise is becoming lower. Given the fact that most stocks have positive return within a bull market, investors tend to be overconfident as they attribute the market success to their own abilities. Similarly, the overconfidence level will decrease in the bear market, though the influence of the market return might be asymmetric as people would like to blame outside factors for the failure.

2.2 Overconfidence and trading volume

As discussed above, overconfidence can be interpreted as a miscalibration of subjective probabilities, underestimation of volatility, or an unrealistic feeling of better than average. Most studies explain the influence of overconfidence from the aspects of miscalibration and better than average effect. Glaser and Weber (2007) use questionnaires and index to evaluate traders' confidence level and find that about 50% of the investors insist that their trading abilities are above average. Moreover, those who think they are better than others trade more. This result accords with Deaves et al.'s (2003) research. Oberlechner and Osler (2003) conduct survey surrounding US currency market professional and figure out that the better than average effect but not the miscalibration explains excess trading volume. Glaser and Weber (2007) also find that there is no significant relationship between miscalibration and trading volume, which is in line with the result from Biais et al. (2005).

While some scholars stick to better than average effect, some others argue that overconfidence also could be explained by miscalibration. Odean (1998a) finds that overconfident investors will lead to increase in market-wide trading volume. He examines the market models that investors are rational in all aspects except the way they treat information. In his models market contains a pricing-taking trader, a strategic-trading insider and a risk-averse market maker. He assumes that these three types of investors are overconfident about the precision of their knowledge than it actually is, thus the posterior are too precise according to calibration literature mentioned above. Overconfidence here also indicates that people tend to overweight the salient information which consistent with their own believes and underweight others. The main contribution of this paper is to provide theoretical support of relation between overconfidence and trading volume. He also examines the case when information is costly and finds the similar result. Odean (1999) tests the hypothesis that overconfident investors trade excessively. He shows that overconfident traders operate even the expected profit could not cover the transaction cost. Based on Odean (1998a) and Gervias and Odean's (2001) arguments that the trading volume changes is the main testable implication of investors' overconfidence, Statman et al. (2006) conduct the empirical research of the US stock market and reveal that the overconfidence could explain the excess trading volume. Following Gervias and Odean (2001), they believe that overconfidence level could be strongly influenced by the market return thus they use market return to represent investors' overconfidence level. They prove that the existence of lead-lag relationship between market return and trading volume and consider it as an evidence of investors' overconfidence which may lead to high trading volume in the U.S stock market. In this paper we will follow their assumption and take market return as an interpreter to measure overconfidence degree and try to establish a relationship between past market return and current trading volume.

2.3 Other explanation of excess trading

The lead-lag relationship between market return and trading volume is also in line with the disposition theory. Shefrin and Statman (1985) show the disposition effect existing in real financial market. Investors tend to sell the winner as they want to realize the profit and hold the loser to avoid regret. However, disposition effect usually refers to investor's attitude towards specific stock but not towards the whole market fluctuation. In contrast, overconfident investors who trade excessively might insist on their belief over the whole market but not limit to the stocks they currently hold. In order to differentiate disposition effect from overconfidence we will test the association from past market return to current trading volume of individual stock. If the relationship presents in the model, then we could infer that the overconfidence effect throughout the whole market is not only the aggregation of the disposition effect in all the individual stocks.

Actually, there is a body of evidences showing the existence of disposition effect in many financial markets. Odean (1998) documents that investors in US market tend to keep losers rather than winners. Baber at al. (2006) show that the disposition effect existing in Taiwan stock market, in which investors are twice likely to sell the winners than the losers. Due to the widespread of disposition effect, we want to differentiate disposition effect from overconfidence.

Besides overconfidence and disposition effect, the literature of differences of opinions could explain the excess trading volume. Differences of opinions could arise because of the difference in prior belief or different understanding of the public signals. Moreover, it is assumed that difference of opinions is common knowledge. Hence, everyone knows others have difference understandings but this fact has no influence on his belief. In other words, investors “agree to disagree”. However, the fact that people know there are other opinions but still insist on their original belief could be considered as overconfidence since people believe they have more precious knowledge. Thus differences of opinions is consistent with better than average effect (Shiller, 1999). Several empirical studies have been performed to test the link between differences of opinions and trading volume. For instance, Antweiler and Frank (2004) examine the influence of the messages posted on Yahoo!Finance and Raging Bull and then find that the disagreements among the posted information result in soared trading volume. Glaser and Weber (2009) shows that both past market return and past portfolio return impact individual investor’s trading behavior, while the past market return has greater power. One explanation is that high market return increase difference opinions. Graham and Harvey (2001) make surveys of CFO stock return expectation and discover that after market wins the different ideas increase.

3. Methodology

3.1 Data

The data of Hong Kong transactions in this paper comes from Thomson Datastream. Since it's difficult to find daily based historical transaction records of individual stocks of whole market, we use the data of constituent stocks in Hang Seng Index instead. Hang Seng Index is a market-weighted index of 33 stocks representing approximately 70% of the market capitalization of all stocks traded on the Stock Exchange of Hong Kong[2], which is good indicator for a total market. For simplicity, in this paper we treat Hang Seng Index as a whole market and constituent stocks inside Hang Seng Index as individual stocks in the presumed market.

The data of the Hang Seng index began from 1964 with a base of 100. However, due to the lack of information about trading volume for early years, our sample period is defined from April 1989 to April 2011. The data should be monthly based, and daily based data is also needed to calculate monthly volatility. We have approximately 6000 daily observations of trading volume, price and market value for each constituent stock of Hang Seng Index. It's supposed to keep 33 constituent stocks inside the index, but actually 45 stocks are included during our sample period due to the limitation of the database. The constitution of Hang Seng Index is a rolling procedure, there may be the case that old constituent stocks which losing its market impact are kicked out and new one comes in. However, the database keeps them both in records (some are kept default and some still have data), and it's difficult for us to find the changes of the stock because of lack of information disclosure. After computing the number of constituent stocks for each month (eliminating stocks with default data), we find that

[2] *Wall Street Words: An A to Z Guide to Investment Terms for Today's Investor* by David L. Scott, 2010, Houghton Mifflin Harcourt Publishing Company.

there are at most 36 stocks for one month, which is acceptable. We think that the additional 3 stocks which may be kicked out of the index could not have such big influence on total market shares.

According to previous literatures, trading volume (in shares) and turnover ratio are both commonly used indicators to measure trading activities. In this paper, considering growing trend of trading volume in long sample period, we prefer turnover ratio since turnover ratio is a relative measure which eliminates the influence of growth. Since we only have data of trading volume (in shares) for individual constituent stocks, we have to compute the turnover ratio for each stock. Moreover, since we have no trading volume data for whole market, the market turnover ratio is calculated by value-weighted averaging of individual turnover.

Lo and Wang (2000) provide detailed calculation formulas for both share turnover and value-weighted turnover. Suppose X_i represents shares traded monthly for individual stock i, N_i is the outstanding shares of stock i. Hence, the individual turnover is calculated by $t_i=X_i/N_i$; The weights w_i for each stock is differed with its own market value divided by total market capitalization. By applying different weights to turnover ratio for each stock, we have market turnover as $t_{vw} = \sum_{i=1}^{n} w_i t_i$.

By repeating the calculation for each stock during the whole sample period, we obtain our market turnover series. Figure 1 is the plotted graph of monthly market turnover. There are two big volatility clusters over the whole sample period, which reflects the impact of financial crisis in 1998 and 2008. The market turnover plunged to the bottom and had a steady growth tendency until it reach the second highest peak in 2008, which indicates the series may be accompanied with a time trend. We apply Augmented Dickey-Fuller (ADF) unit root test to verify our guess, and we find it fails to reject the null hypothesis of existence of a unit root at 5% confidence level[3], that suggests market turnover is non-stationary with no intercept and trend.

In this paper, we plan to test overconfidence effect through interaction between market return and market turnover by applying vector autoregression (VAR) model. Hence, the market turnover series is required to be stationary to ensure the model estimation non-biased and valid. Since taking log of market turnover series still cannot solve the stationarity problem, following the Statman et al. (2006) then we introduce Hodrick-Prescott (1997) filter further to logged maket turnover to get a detrended stationary series. The basic framework of Hodrick-Prescott filter is decomposition of original series to a growth component and a cyclical component. Figure 2 demonstrates that the original turnover (blue line) is separated to a smooth trend (red line) and a stationary series (green line) keeping the same pattern of volatility as before. The stationary series called cyclical part is what we need, and will be taken into our VAR model. For simplicity, this detrended logged turnover series is named as *mturn* in this thesis. It's worth noting that Statman et al. (2006) also mention that using detrended series will lead to bias in capturing response of long-term trading activity to market return. The long- term trend of market turnover induced by business cycle (bull and bear) and financial crisis may be explained as overconfidence effect. However, to ensure valid model estimation, the detrended logged market turnover series *mturn* is finally chosen.

There are two ways to compute the market return series. One is using Hang Seng Index to calculate the monthly market return directly, and another is using value-weighted average of individual stock returns. Since we use the value-weighted average market turnover before, here we choose the value-weighted average as well for consistency and accuracy. For monthly market return, we should first calculate return for all stocks within the index for each month. Then we apply corresponding

[3] The P-value of ADF test is 0.4617, we can not reject H_0: The series has a unit root, indicating that it is non-stationary.

market value weights to each stock return for one month to get one market return observation. Repeating the process for all months during the sample period, we generate the market return series, *mret*. As another important variable, market return passes the stationary test (ADF unit root test) at 1% significance level with its p-value less than 0.0001. Similar to market turnover series, market return also has two manifest volatility clusters during the sample period. (See Figure 3) Combined with graph of market turnover (Figure 1), we think these two clusters may result from financial crisis happened in 1998 and 2008 respectively. To eliminate the impact of financial crisis completely, we design two dummy variables, *dum98* takes value of 1 from September 1997 to November 1998(other months keep 0), while *dum08* is from May 2008 to June 2009.

We have made a statistical description of our variables for full sample period and two sub-periods (predefined financial crisis period in 1998 and 2008) in Table 1. Compared with full sample period, the sub-period of financial crisis in 1998 have the negative mean return (-0.01357), and we find the maximum and minimum return for whole period is also realized in this period. The volatility of returns for this sub-period doubled with standard deviation of 0.156 (standard deviation of the full sample is 0.075), while the standard deviation of sub-period of 2008 is 0.124. These statistical figures may provide some support to our introduction of dummy variables.

Beside the dummy variables designed for financial crisis period with high volatility, market volatility is also employed as a control variable. The volume-volatility relationship is our motivation to introduce market volatility into the model. The Mixture of Distribution Hypothesis (MDH) initiated by Clark (1973) is theoretical foundation supporting a positive relationship between price changes and volume. Focusing on Asia-Pacific stock markets, Deo et al. (2008) find that there is a positive relationship between return variance and lagged trading volume for most of the selected markets. There are also varieties of similar empirical studies for Austrian market (Mestel, 2003), Brazilian market (OR. De Medeiros, 2006), Pakistani market

(Mubarik, 2009), etc. We follow Statman et al.'s(2006) specification of the monthly volatility provided by French et al. (1987), which is computed by adding squared daily returns with twice the sum of the products of adjacent returns.[4]

The motivation to introduce the second control variable *dispersion* comes from working paper of Campbell and Lettau (1999), which measures volatility not only in market level, but also considers the industry-specific and firm-specific influence. In order to capture the idiosyncratic risk for individual firms, we employ dispersion, which is the cross-sectional volatility of individual firms within Hang Seng Index on monthly basis. Following Campbell and Lettau (1999), we first compute squared deviation from mean return for each stock, and then multiply corresponding market-capitalization weights to generate *disp* series.[5] Table 1 also provides statistical description of two control variables, mean dispersion over the full sample period is 6.7%, which is a bit lower compared to dispersion (7.1%) in US market reported by Statman et al. (2006).

In order to distinguish overconfidence effect from disposition effect, we also define variables for constructing individual security model. In our individual model, individual turnover is introduced as endogenous variables, while market return, stock return and volatility are exogenous variables. Since this model is built up for each stock individually, the dispersion variables representing cross-sectional effect on stock portfolios can be eliminated. The sample period is the same to market-wide

---

[4] Assume that there are *N* days for the month *t*, monthly volatility should be calculated as $misg_t = \sum_{i=1}^{N} r_i^2 + 2\sum_{i=1}^{N\text{-}1} r_i r_{i+1}$, where r_i is the return for the *i*th trading day in the month.

[5] Suppose we have N stocks in the market for month *t*, the return of stock *i* is r_i, and $\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i$, w_i is the market capitalization weight for stock *i*, the monthly dispersion should be: $disp_t = \sqrt{\sum_{i=1}^{N} w_i (r_i - \bar{r})^2}$

model, which is from April 1989 to April 2011, including 265 months in total. As mentioned above, there are 45 constituent stocks over the whole period while default data is also included.

In terms of stock turnover, we first use the market value of the firm and the stock price to calculate number of outstanding shares, and then the turnover ratio for individual stock can be computed by monthly trading volume divided by the number of outstanding shares. We employ the same method (HP filter) in previous part to deal with non-stationarity of the turnover ratio. Detailed statistical figures (minimum, maximum, standard deviation) about *turn* series for each stock can be found in Table 2. Detrended logged turnover series (*ret*) is generated for each stock in Hang Seng Index. Return for individual stock (named by *ret*) is easily computed with trading price on monthly basis. Daily return for individual stock is needed to obtain the monthly realized volatility, *sig*. The methodology is the same to calculation of monthly market volatility, which has been described in the footnote 4.

3.2 Vector Autoregressive Model (VAR)

Vector autoregressive model (VAR) is constructed to investigate whether there are lead-lag relationships among variables. Compared to univariate time series model, standard VAR model estimates several equations simultaneously without specifying which variables are exogenous or endogenous. Besides the AR terms in the model, the lag-term of other variables may help to capture more features of the data, and also is the key to explain interactions of the variables.

In this thesis, we use the VARX model, the advanced version of VAR with control variables (exogenous variables which values are decided outside the model). The basic VARX model is specified as following:

$$Y_t = a + \sum_{i=1}^{N} A_i Y_{t-i} + \sum_{j=0}^{M} B_j X_{t-j} + e_t$$

where Y_t is a $n \times 1$ vector for n variables with t observations each, and A is a $n \times n$

matrix measuring the coefficients of lagged terms of Y_t itself. With two control variables, X_t is a 2×1 vector of exogenous variables, while B is a $n\times2$ matrix of coefficients. N and M represents the optimal lag length for Y_t and X_t, respectively.

To find the optimal lag length, we compare likelihood ratio (LR) and information criteria(IC) from the 1 lag to 8 lags. Likelihood ratio (LR) test is based on constructing a restricted model with last *k* lags having jointly zero coefficients, while according to information criteria the optimal lag length is chosen by minimum value related to variance-covariance matrix of residuals and number of regressors and observations.[6] Eviews computes the likelihood ratio and different version of information criteria until the 8th lags which is shown in Table 3. Both likelihood ratio (LR) and Akaike information criterion (AIC) indicate that two lags of endogenous variables are the most appropriate. With fixed two lags of endogenous variables, we build up models by adding lags of exogenous variables, and compare the information criteria of these models. It turns out that the optimal lag length for the exogenous variables is one lag.

Our first model is constructed to investigate the lead-lag relationship between market return and market turnover, which is specified as following:

$$\begin{bmatrix} mturn_t \\ mret_t \end{bmatrix} = \begin{bmatrix} \alpha_{1,t} \\ \alpha_{2,t} \end{bmatrix} + \sum_{i=1}^{2} A_i \begin{bmatrix} mturn_{t-i} \\ mret_{t-i} \end{bmatrix} + \sum_{j=0}^{1} B_j \begin{bmatrix} msig_{t-j} \\ disp_{t-j} \end{bmatrix} + \gamma_1 dum98 + \gamma_2 dum08 + \begin{bmatrix} e_{1,t} \\ e_{2,t} \end{bmatrix} \quad (1)$$

In addition to this model, Granger causality test will be performed to verify the correlation between market return and market turnover. Specifically, in equation (1) if the lag terms of market return are significant in explaining market turnover, we can say that market return Granger cause market turnover. If this relationship holds, it can be explained as overconfidence effect in the market.

[6]Three difference IC specified as following:

$$AIC = \log\left|\hat{\Sigma}\right| + \frac{2k}{T}; \qquad SBIC = \log\left|\hat{\Sigma}\right| + \frac{k}{T}\log(T); \qquad HQIC = \log\left|\hat{\Sigma}\right| + \frac{2k}{T}\log(\log(T))$$

where Σ is variance-covariance matrix of residuals, k is number of regeressors while T is number of observations.

Although Granger causality test can examine the existence of impact of one variable on future value of other variables, it fails to tell us the sign of the relationship and the duration the impact. To compensate the drawbacks of Granger causality test, we introduce impulse response functions, which capture the reaction of the endogenous variables in our VARX model to external shocks. Assume we impose one unit shock on the error term $e_{1,t}$, and the impact of the $e_{1,t}$ on market return and market turnover can be observed by tracing back of VARX model. Meanwhile, we can also see how the market return and market turnover will react to one unit shock to $e_{2,t}$. Thus, the impulse response function is able to depict a dynamic picture about how and how long one variable will affect another.

3.3 Individual security model estimation with panel data approach

Following Statman et al. (2006), individual security model is built up to distinguish the overconfidence effect from disposition effect. However, our individual model only focuses on the equation in which turnover is endogenous variable. This equation is the crucial for both overconfidence and disposition effect. This simple regression model involves three main variables: individual stock return (*ret*), individual stock turnover (*turn*) and market return (*mret*). Stock volatility (*sig*) is also considered while the dispersion over stocks returns is not needed since this model will applied to each stock individually. The lag length also keeps the same as market-wide model.

$$turn_t = \alpha_0 + \sum_{i=1}^{2}\alpha_{1i} turn_{t-i} + \sum_{i=1}^{2}\alpha_{2i} ret_{t-i} + \sum_{i=1}^{2}\alpha_{3i} mret_{t-i} + \sum_{j=0}^{1}\beta_j sig_{t-j} + \varepsilon_t \quad (2)$$

To differentiate overconfidence effect from sample aggregations of disposition effect, it's necessary to discuss the influence of individual stock return and market return on individual stock turnover. If the parameter of individual stock return is significant in explaining individual turnover while market return's parameters is not, it can be explained that the disposition effect dominates overconfidence, and vice versa. There may be the case both disposition effect and overconfidence effect existing simultaneously, which indicates overconfidence effect can explain some part which can not be captured by disposition effect.

The most difficult problem for individual stock model is combination of the 45 stocks into one aggregated model. Simply averaging coefficient of 45 individual models may help to get the mean coefficients, but this will lead to biased estimation of standard error for coefficients. Statman et al. (2006) use the bootstrap procedure to simulate the precise standard error for coefficients while in this paper we apply panel data approach instead.

One advantage of panel data approach is that it allows cross-sectional heterogeneity. Generally speaking, a panel data set contains a number of the repeated units (individuals, firms, countries) over a chosen period. Naturally, a panel data set covers both time and individual dimension thus it has higher accuracy. In particular, if one is interested in the changes over time, a panel data set is more efficient estimators than a series of cross-sections (Verbeek, 2004). Compared to time series or cross-sectional data set, panel data yields better identifications of parameters (Verbeek, 2004). Practically, it eases the problem of omitted variables by introducing individual dummy and time dummy, where individual dummy is to capture the unobserved effect for certain units that does not vary over the sample period and time dummy is to capture the unobserved effect for all units in the certain time point. Rewrite the individual security model in panel data approach as

$$turn_{k,t} = \alpha_0 + \sum_{i=1}^{2}\alpha_{1i}turn_{k,t-i} + \sum_{i=1}^{2}\alpha_{2i}ret_{k,t-i} + \sum_{i=1}^{2}\alpha_{3i}mret_{k,t-i} + \sum_{j=0}^{1}\beta_j sig_{k,t-j} + f_k + \varepsilon_{k,t} \qquad (3)$$

Where k refers to the individual security (k=1,…,45). Following Holtz-Eakin et al.'s (1988) dynamic panel approach, the individual fixed effect f_k is included to index the individual heterogeneity. For example, f_k measure the unobserved fix effect of k security which is constant over the time. In static model, fixed effect model estimated with OLS is consistent and efficient (Brooks, 2008). However, in dynamic model the lag term will depend on the fixed effect dummy (Verbeek, 2004). For example, $turn_{k,t-i}$ is correlated to f_k thus OLS estimation is not consistent any more. Nickell (1981) takes the first difference and then estimates the function by using generalized instrument variables. Holtz-Eakin et al. (1988) proposes GMM method to estimate the coefficients. Both methods require the absence of autocorrelation in residual to guarantee the validity of the moment condition (Verbeek, 2004). We will use the GMM method to estimate the security model in Eviews.

In order to avoid the spurious regression, we perform the unit root test to examine the stationarity of our variables. Levin and Lin (2002) (LL), introduce a model for panel series set. The null hypothesis is that all *k* series in the panel set have a unit root (common root), while the alternative is the all individual series are stationary. The shortcoming of LL model is that it assumes all series in the panel has the same unit root structure, i.e they are either I(1) or I(0). Im, Peseran and Shin (2003), hereafter IPS, develop another model to test the individual unit root process. Different from LL, they relax the homogeneity constraint and allow different lags for *k* unit. Their null hypothesis is that each series in the panel contains a unit root for all *k*, and the alternative is that at least one of the individual series in the panel is stationary. In particular, they take the average of the ADF test statistics. IPS is a generalization of LL test and more powerful than LL (Harris and Sollis, 2003). However, this model requires large T as the distribution of panel test statistics depends on T. Therefore, it doesn’t fit the unbalanced panel. Moreover, those two models suffer from low power when models include fixed effects. Maddala and Wu (1999) improve IPS model by contain Fisher-type test to combine the p-value obtained from the ADF tests for all *k* series. It does not require the balanced panel, thus they argue that Fisher-type test is better than IPS test. Since our data is an unbalance panel set, ADF-Fisher unit root test is the most proper model to test the stationary. The test results in Table 4 suggest that all the variables in the security model are stationary.

3.4 Validity and Reliability

Validity refers to the issue that whether a variable used to capture a concept could really reflect the concept (Breman, 2004). In other words, it considers whether conclusion generated from the empirical results could report the reality. In our paper we try to verify the existence of overconfidence in Hong Kong Stock market by identifying the lead-lag relationship between market return and market turnover. First the market-wide model is applied. Following Statman et al. (2006), the past market return is selected to measure the overconfidence level. The logic is that people's overconfidence levels change according to their previous outcome. Especially in the bull market people tend to be more overconfident and ignore the fact the whole market return increases because of the self-attribution bias. The close linkage from market return to overconfident level has been identified by Gervias and Odean (2001). On the other hand, many scholars (Odean, 1998a; etc) suggest that overconfident investors trade more. In this paper we use the turnover ratio as a proxy of trading activities, which is a common method in academia. To distinguish the disposition effect from overconfidence, the individual model is introduced. The individual turnover is the dependent variable. Furthermore, compared to the market-wide model, individual return is included to measure the disposition effect. Here the market return is also used as a representative of overconfidence. If the individual return is affected by the past market turnover, then it suggests overconfidence exists other than disposition effect. In a word, the selected variables and method could ensure us to reflect the reality in the Hong Kong Stock market.

Reliability concerns the consistency of measuring a concept (Breman, 2004). Moreover, it considers in what degree the result could be replicated, i.e, if other scholars could obtain the same result by repeating the same method to the same data. In order to assure the sufficient result, both the data and method should be examined prudently. Regarding the data, we gather it from Thomson DataStream which is widely used in academia. In particular, due to the limited access to the information of all the individual stocks, the 33 constituent stocks of Hang Seng index is viewed as an indicator of the whole market. As Hang Seng Index could represent approximately 70% of the market capitalization of the whole Hong Kong market, we consider it as a reasonable proxy. As mentioned in section 3.1, given the limitation of the data series, our sample includes at most 36 stocks in a specific month. However, we view the additional 3 stocks will not lead to great error in the final result. Furthermore, all the model and regression are conducted in Eviews which is commonly used in econometric study. We perform VAR in the market-wide model, and draw the impulse response function to identify the relationship between market return and market turnover. To make our model more reliable, we also design dummy variables to eliminate the effect of financial crisis period. Panel GMM regression, which is proved to be the consistent and efficient in the panel data set, is applied in the individual security model. In the process of building the models, we carefully use the diagnostic tests and control the residual behaviors. To sum up, all these data and methods could suggest the reliable results.

4. Empirical Results

4.1 Market VAR estimation and impulse response function

Table 5 outlines the estimation of market VAR system which contains endogenous variables: detrended logged market turnover, *mturn*, and market return, *mret*. The control variables are market volatility, *misg*, and market dispersion, *disp*. The table is organized by columns for dependent variables and rows for lag terms and control variables. For each coefficient, we report the estimated value and standard error. The p-values are also given to show whether the estimated coefficient is different from zero. Generally we consider the p-value smaller than 10% as significant, and less than 5% or 1% as highly significant. Note that the dummy variable refers to 2008 financial cirisis, *dum08* is not statistically significant in the VAR system thus we exclude it from our regression.

According to Table 5, market turnover is autocorrelated. The coefficients of the first lagged and second lagged market turnover are highly significant, with the estimated parameters of 0.1307 and 0.1440. It suggests that the market turnover is affected by its own behavior in past two periods. Moreover, market turnover also depends on the first lagged market return. Given that the market turnover is the dependent variables, the estimated coefficient for the first lagged market return is 0.3330 and it is significant at 10% confidence level. The influence of the market return to the market turnover only exists in the first lag, since the second lag of market return is not significant. The positive impact of the lagged market return on the market turnover fits our overconfidence hypothesis, although the affect is not as strong as we expected. We will discuss more in the impulse response section.

Considering the relationship between volatility and trading volume, our estimate result indicates that influence of volatility, *misg* is large. More specifically, the coefficient of current volatility is 2.2410 and the first lag volatility is -1.5874 with highly significant. The positive affect of the *misg* is in line with Mixture of Distribution Hypothesis (MDH) (Clark ,1973). But the sign turns to be negative in the first lagged volatility. Here we have the same situation with Statman et al. (2006). Since the volatility is autocorrelated, there might be multicollinearity problem between *misg* and its lag, and this will affect the estimation output. The parameter of current dispersion is also significant. The inclusion of these two exogenous variables is to control other explanations of the trading volume behavior. The figure implies that both the inter-temporal market return changes (*misg*) and cross-sectional volatility of individual firms (*disp*) have strong influence on the market turnover. In particular, the former variables have stronger impact. Following Statman et al. (2006), the information events in the market might be the factor behind this phenomenon.

The dummy variable for 1998 financial crisis (*dum 98*) is significant when the market turnover is the dependent variable. We could infer that the *dum 98* influences the trading volume. However, the positive sign (0.1366) shows that people tend to trade more during the crisis. The possible explanation comes from the dummy design. The *dum98* contains the observations from September 1997 to November 1998 when both market return and market return fluctuates dramatically. It includes the period that *mturn* and *mret* climb up to the peak and drop to the bottom. According to Table 1, the mean of *mturn1* during this dummy period is remarkable larger than the whole sample period. Moreover, the maximum value over the twenty years appears in late 1997. In our consideration, the model may not be able to capture the trading behavior during the extreme period. That's the motivation we initiate this dummy variable, and the significant coefficient shows that the people indeed act differently in this period.

Table 5 also demonstrates that market return has positive autoregression term in lag one. The estimated value of first lagged *mret* is 0.1053 with the standard error -0.0632. The first lagged market turnover is significant in 10% level and the sign is positive. This finding is consistent with Gervais et al. (2001) that trading volume could be used to predict the return. However, the impact of market turnover decays in the first lag, as the second lagged *mturn* is not statistically significant. Moreover, *misg* and *disp* have positive impact on the market return contemporarily. The coefficient of dummy variable for 1998 crisis (*dum98*) has negative sign but is not significant.

The Granger causality test result is associate with the overconfidence hypothesis (See Table 6). Given the null hypothesis "MRET does not Granger Cause MTURN", the p-value is 0.0903 that we reject the H_0 in 10% significant level. Along with the figures in Table 6, we infer that past market return has positive impact on current trading volume, although the affect is weak and doesn't last long. However, this relationship doesn't hold in the opposite way. The p-value is greater than 10% when the dependent variable is market return that we can't reject the null hypothesis. Thus we could not see the influence of past trading volume on the market return in the Granger causality test.

4.2 Market impulse response function

Impulse response function relies on the VAR estimation to check the impact of one standard deviation shock from the residual. It measures the response of $Y_{j,t+s}$ to an response in Y_{1t}, given all other variables are constant (Verbeek, 2004). Figure 3 shows the four possible scenarios of our VAR system. The blue line in Panel A and Panel B outlines the response of *mturn* to a shock from *mturn* and *mret*, while the red line refers to two standard error band. Since we use detrended logged market turnover series, the vertical axis in Panel A and B measures the percentage increase in market turnover compare to situation with no shock. Moreover, the impulse function becomes zero in the long term as the *mturn* is detrended.

According to Panel B, the response of market turnover to shock of market return exists until the fourth lag. More specifically, in lag one the response is not evident, but turns to large and positive in lag two. The shock of market return at second lag brings 2% increase in market turnover. Interestingly, the response function is not monotonically. The impulse becomes negative in the third lag and dies out after the fourth lag. However, the upper and lower bands of the two standard error are not in the same quadrants from the beginning to the end, so that the signs of the response are not significant. This graph accords with our model estimation and Granger causality test. This result indicates weak response of market turnover to market return in Hong Kong stock market, compare to the large and persistent response in US market. Staman et al. (2006) show that one standard deviation in market return results in 8.6% increase in trading volume. We could infer that the Hong Kong stock has weak form overconfidence. Culture difference could be one factor. Hong Kong used to be an important link between Mainland China and the western world, thus its stock market involves the investors from difference background. Those investors might act in a contradicted way as some of them are more overconfidence while others tend to be more cautious. It will result in the weak overconfidence pattern. Moreover, we need to explain that from the data aspect. As discuss in the methodology section, we use the stocks in Hang Seng index to capture the behaviors in the whole stock market. Due to the fact that the involved stocks changes over time (some new stocks are included while some others are excluded), our data for individual stock might have null value in some period. We regard it as a disadvantage of the data and think it might lead to biased result, i.e, the estimated overconfidence might be smaller than the actual one. Panel A indicates that market turnover is highly autocorrelated. In particular, one standard deviation shock in its first lag leads to 20% increase in trading volume. The response drops to 4% in the second lag and diminishes until the fifth lag.

Panel C and D illustrates the response of market return to market return and market turnover, respectively. Again the red line is the two standard error band. Panel D shows that *mret* is autocorrelated. One standard deviation shock of first lagged market return results in 7% increase in market return. The impulse response function declines and dies out after the forth lag. Note that the response of *mret* is only significant in the first lag. This strong autoregreesive behavior could be covered by the momentum theory that positive return will follow gains in the short time horizon (Rouwenhorst, 1998). Panel C reveals that response of market return to the shock from the market turnover is weak and small as it only exists in the second lag. One unit shock in second lagged trading volume will lead to 1% increase in market return. The impulse response in other lags is nearly zero. The non-significance of the response in Panel C is in line with the result of the Granger Causality test (see Table 6).

4.3 Individual stock regression

Market-wide model has verified the lead-lag relationship between market return and market turnover, which may provide powerful evidence for presence of overconfidence effect. However, as we mentioned in theoretical section, there are also other explanation for this lead-lag relationship. Disposition effect is one alternative theory, which indicates investors' tendency to sell winners too early and hold losers too long. Shefrin and Statman (1985) prove the existence of deposition effect empirically, and provide several interpretations for disposition effect from the aspect of behavioral finance theory. However, disposition effect is always explained as investor's behavioral bias to individual stocks while overconfidence level is believed to have close relationship with whole market fluctuation. Gervias and Odean (2001) believe that market return can measure the overconfidence level. High market returns will enhance the level of overconfidence, since the investors tend to attribute the success of generating more profits to their own abilities. To differentiate disposition effect from overconfidence, we build up a new model to test the relationship between past market return and current trading volume of individual stock. More specifically, the response of security turnover to security return is viewed as a proxy of disposition effect. If the test results reveal that the market-wide return has distinct impact on the security turnover, then we could infer that the overconfidence is not only the aggregation of disposition effects of the all the individual securities.

The individual security model is estimated as equation (3) in methodology part. In the individual security model, we consider that how the security turnover (*turn*) responses to both the security return (*ret*) and market-wide return (*mret*). Beside individual security volatility (*sig*), a fixed effect factor is employed to capture the individual heterogeneity of different stocks. In order to be consistent with the market-wide model, we also use same lag structure (*turn*, *ret*, *mret* keep 2 lags, and *sig* for 1 lag). We have 45 individual securities with 265 monthly observations for each variable. Repeating the model for 45 individual stocks will produce 45 regression equations. However, what we need is only one aggregated model of 45 individual stocks. In order to get precise estimation of coefficients and corresponding standard errors, Statman et al. (2006) resort to bootstrap program while we use GMM to estimate the aggregated model directly by panel data approach provided by Eviews.

Empirical results (See Table 7, Model 1) reveal that individual stock turnover is highly autocorrelated since both AR (1) and AR (2) terms are significant at 1% confidence level. Moreover, as a highly significant explanatory variable in market-wide model, the volatility (*sig*) is also correlated with individual stock turnover with p-value of 0.0448. Lagged market return is able to explain the individual securities turnover with significant coefficient of 0.2797 (p-value of 0.0835). In contrast, coefficients of the lagged term of individual stock return fail to pass the significance test which suggests weak disposition effect compared to overconfidence. This result is also consistent with the research of Statman et al. (2006), in which the mean coefficients of individual security return remains non-significant for 10 lags.

To further investigate the explanatory power of disposition effect and overconfidence, based on original model, two additional models are constructed with default of market return and individual security return respectively (See Table 7). Excluding lagged market return, Model 2 is to model the disposition effect, and it finds that the first lagged individual return becomes significant with p-value of 0.0594. On the other hand, the result of Model 3 implies overconfidence effect turns to be stronger without including individual security return (*ret*), since the first and second lag of market return are both significant at confidence level of 5%. To sum up, both disposition effect and overconfidence may explain the lead-lag relationship between return and turnover. Our empirical results verified the existence of overconfidence and disposition effect in the Hong Kong stock market. However, the disposition effect has weaker explanatory power compared to overconfidence, and overconfidence dominates the disposition effect in Hong Kong stock market.

5. Conclusion

The effects of investors' overconfidence have been investigated by many scholars. Gervias and Odean (2001) show that, due to the self-attribution bias, the level of overconfidence is positively related to market return. Glaser and Weber (2007) demonstrate that overconfident investors tend to trade more. These previous studies provide a framework for our study. Focusing on the U.S. stock market, Statman et al. (2006) apply a VAR model to examine the lead-lag relationship between return and turnover, which can be interpreted in terms of overconfidence.

Following Statman et al. (2006), we first build a market-wide VAR model to examine the interaction between turnover and return in the Hong Kong stock market. The results show that previous market return is positively related to market turnover. To further analyze the sign and duration of the relationship, impulse response functions are also employed. Both market turnover and market return are found to be autocorrelated. The response of market turnover to shocks in market return is stronger than the response in the opposite direction, which is consistent with overconfidence. However, this overconfidence effect in the Hong Kong market is not as strong as in the U.S. market. One possible explanation is that Hong Kong investors are not as overconfident as American investors. It may also be that the Hang Seng Index (70% of the market capitalization) is not able to capture investors' behavior in the market as a whole.

To distinguish the disposition effect from overconfidence, an individual security model is constructed. The basic idea of this model is to use the response of individual turnover to market return and individual stock return as proxies for overconfidence and the disposition effect, respectively. We find that both overconfidence and the disposition effect may explain excess trading in the Hong Kong market. Nevertheless, the disposition effect has relatively weak explanatory power compared to overconfidence. In the case of putting both market return and individual return into regression, we find that market return is statistically significant in explaining individual return while individual return is insignificant, which means that overconfidence may dominate the disposition effect in the Hong Kong stock market.

Our main findings provide new empirical evidence to overconfidence theory. Due to data limitation and availability, we use the Hang Seng index to represent the Hong Kong stock market. Future research using more extensive data might result in better estimations. Moreover, we suggest a deeper investigation of individual investors, e.g. through broker accounts.

Appendix

Figure 1 Monthly market turnover of the Hong Kong market (based on Hang Seng Index)

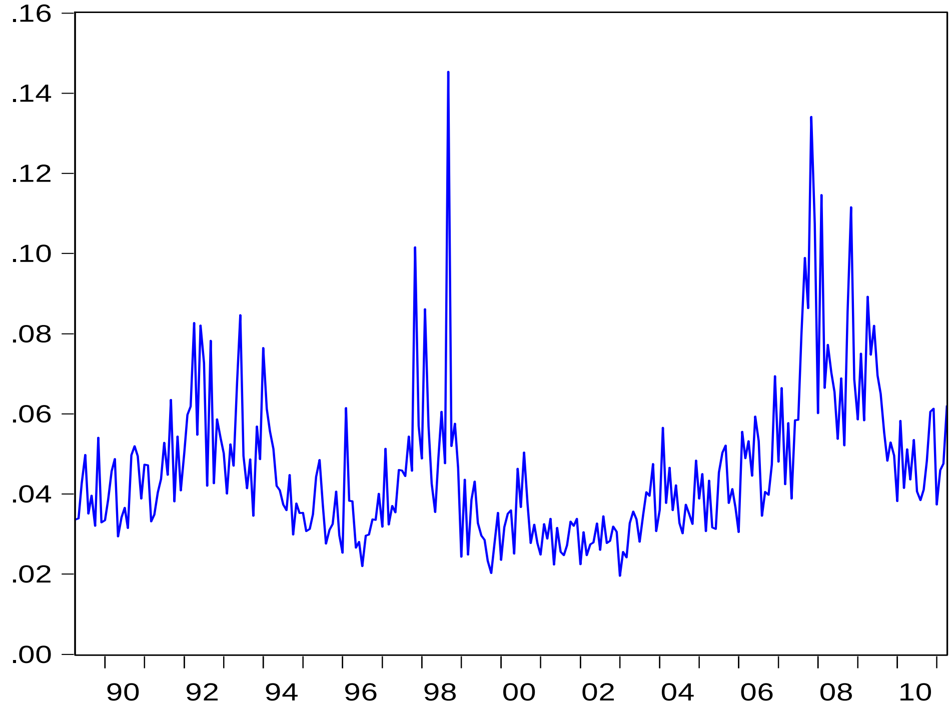

Figure 2 Detrended logged market turnover by HP filter

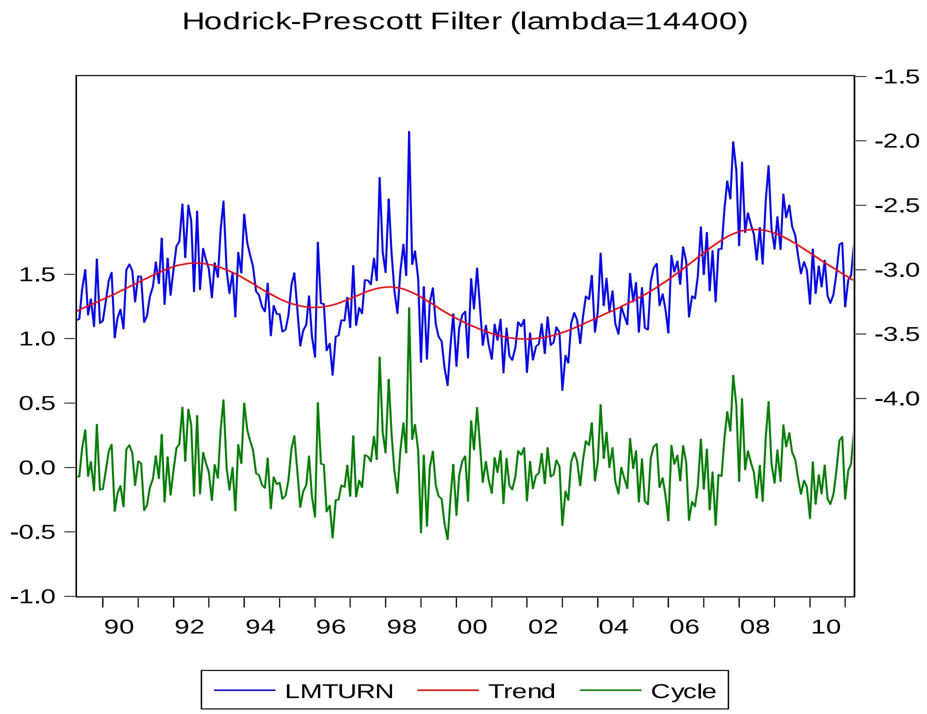

Figure 3 Monthly market return of the Hong Kong market (based on Hang Seng Index)

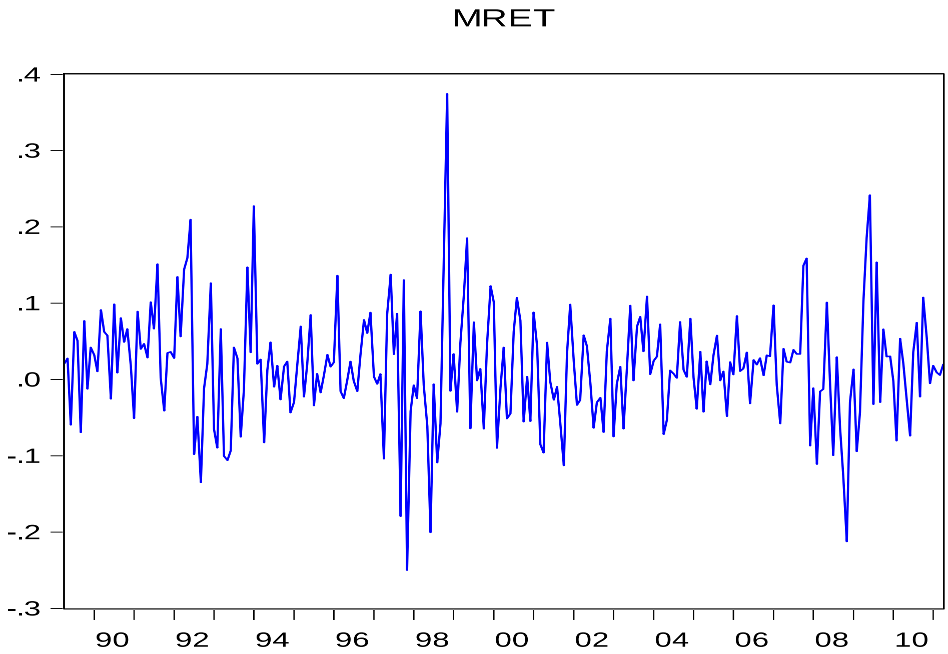

Figure 4 Impulse response function of market-wide model

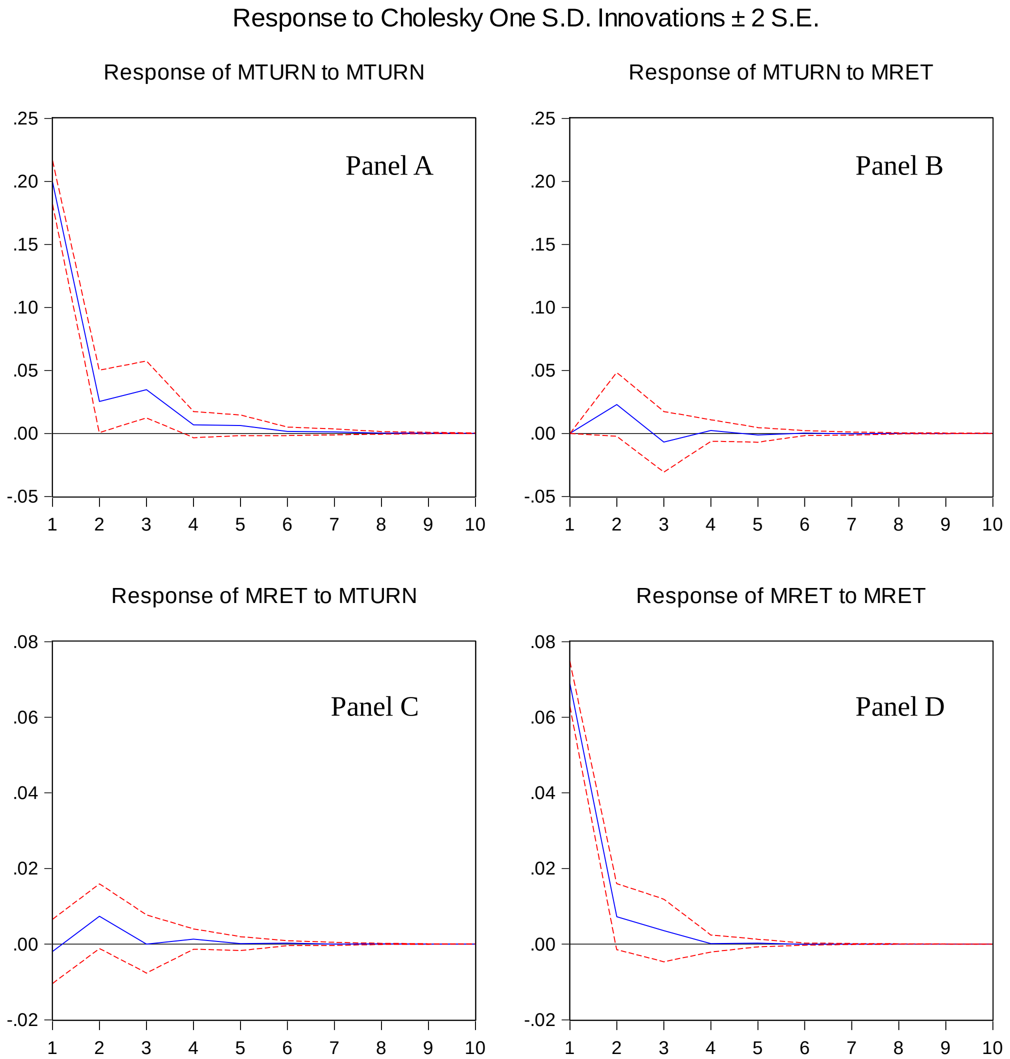

Table 1 Statistical description of variables in market-wide model

1989m4-2011m4					
	MTURN1	Mturn	Mret	Misg	Disp
Mean	0,045866	-2,67E-03	0,016258	0,068456	0,067681
Median	0,041406	-0,00044	0,017938	0,057262	0,059897
Maximum	0,145399	1,238486	0,374521	0,282245	0,226139
Minimum	0,019533	-0,55867	-0,25008	0,020199	0,019822
Std. Dev.	0,018682	0,244925	0,074714	0,03722	0,033248
Skewness	1,891562	0,844499	0,305518	1,946881	1,748192
Kurtosis	8,369107	5,351125	5,530319	8,475638	7,224694

1997m9-1998m11					
	MTURN1	MTURN	MRET	MISG	DISP
Mean	0,062711	0,310531	-0,01357	0,134904	0,094594
Median	0,054329	0,239801	-0,02457	0,135261	0,079366
Maximum	0,145399	1,238486	0,374521	0,201535	0,226139
Minimum	0,035499	-0,1982	-0,25008	0,055059	0,037236
Std. Dev.	0,028309	0,363755	0,156053	0,045692	0,05301
Skewness	1,931614	1,238538	0,804878	-0,2683	1,484402
Kurtosis	5,990713	4,087149	3,756665	1,992456	4,311332

2008m5-2009m6					
	MTURN1	MTURN	MRET	MISG	DISP
Mean	0,07243	0,068238	7,46E-05	0,115692	0,0973
Median	0,069555	0,031582	-0,01749	0,108838	0,089197
Maximum	0,111545	0,509022	0,241775	0,282245	0,18648
Minimum	0,052098	-0,26167	-0,21235	0,055069	0,067551
Std. Dev.	0,015984	0,219918	0,124443	0,060249	0,032636
Skewness	0,916852	0,298566	0,368528	1,474107	1,565092
Kurtosis	3,592513	2,394048	2,522478	5,137107	4,936947

Note: *Mturn1* is original market turnover, and *mturn* represents detrended logged market turnover.

Table 2 Statistical description for individual stock turnover

Code	Number of Obs.	Maximum	Minimum	Mean	Std
ALUM	112	0,46769	-0,37279	1,77E-13	0,15393
BOCH	59	0,574202	-0,31792	1,41E-14	0,174917
BOCC	70	0,365602	-0,30331	5,68E-13	0,139482
BEAA	265	0,594165	-0,55227	-0,00223	0,183491
BIHL	47	0,476831	-0,3872	1,19E-13	0,15352
BOC	105	0,725975	-0,56833	5,28E-13	0,197503
CATH	265	0,559692	-0,42556	-0,001	0,165134
CHGK	265	0,508367	-0,345	0,001399	0,139503
CCEC	52	0,347276	-0,32772	4,98E-13	0,13699
CCBN	66	0,632411	-0,2218	5,57E-13	0,143234
CLS	80	0,34476	-0,30466	-0,00806	0,139131
CHT	80	0,410126	-0,23953	-0,0038	0,113526
HAIH	80	0,280321	-0,23999	-0,00498	0,109792
COLI	80	0,408596	-0,27522	-0,00348	0,114792
CHPE	80	0,427023	-0,30364	-0,0073	0,137639
CHRE	80	0,390178	-0,29001	-0,00709	0,131573
CREP	80	0,519021	-0,33576	-0,01068	0,155542
CHBE	80	0,447066	-0,72756	0,00391	0,189232
CSHE	70	0,410152	-0,30891	3,83E-13	0,144536
UNIC	130	0,382359	-0,28917	1,81E-13	0,127618
CTIP	265	0,865574	-0,68684	-0,00158	0,236799
CLIG	265	0,479202	-0,3393	0,000565	0,135528
CNOO	122	0,407308	-0,26473	2,82E-13	0,118965
FLGR	196	0,59016	-0,7735	2,15E-14	0,190483
ESPR	208	1,215248	-0,63306	-1,9E-14	0,228504
FOXC	74	0,485223	-0,2861	4,99E-13	0,158111
AMOY	265	0,782614	-0,4902	0,00222	0,179452
HSBA	265	0,582728	-0,4493	1,4E-06	0,147227
HELD	265	0,527023	-0,36497	0,002358	0,147663
HKCG	265	0,546414	-0,56779	0,000366	0,147317
HKEX	130	0,526445	-0,38171	2,41E-13	0,162655
HSBC	265	0,672651	-0,3358	0,001007	0,145723
HUTI	265	0,576174	-0,39441	0,001925	0,148571
ICBC	54	0,303432	-0,23596	-1,2E-13	0,135473
FUNG	226	0,951885	-0,63865	-4E-13	0,19788
MTRC	126	0,680291	-0,33313	2,15E-13	0,186446
NWDV	265	0,570535	-0,4816	0,002769	0,171143
PECH	132	0,696045	-0,37735	2,52E-13	0,165174
PING	82	0,408731	-0,34109	3,55E-13	0,162606
HKEL	265	0,489175	-0,3011	0,002019	0,130024
SILC	265	0,930901	-0,59271	-0,00015	0,239666
SHKP	265	0,614991	-0,31956	0,001613	0,148085

SWPA	265	0,465574	-0,36957	0,002871	0,131552
TCNT	82	0,664776	-0,34211	2,08E-13	0,163208
HKWH	265	0,436244	-0,38207	0,001481	0,139279

Table 3 Lag structure criteria for endogenous variables in market-wide VARX model

Lag	LogL	LR	FPE	AIC	SC	HQ
0	377,1758	NA	0,000215	-2,771138	-2.636054*	-2,716863
1	385,8834	16,95529	0,000207	-2,806667	-2,61755	-2.730683*
2	391,8537	11.53506*	0.000204*	-2.821538*	-2,578386	-2,723843
3	394,6325	5,326934	0,000206	-2,812321	-2,515136	-2,692917
4	395,5323	1,711154	0,000211	-2,788923	-2,437704	-2,647808
5	395,9327	0,75564	0,000217	-2,761757	-2,356504	-2,598932
6	397,3952	2,737243	0,000221	-2,742605	-2,283319	-2,558071
7	398,3741	1,817477	0,000226	-2,719805	-2,206485	-2,513561
8	399,9311	2,867209	0,00023	-2,701367	-2,134013	-2,473413

Value with star(*) is chosen by specific criterion.

Table 4 Results of different unit root tests of variables (p-value)

Variables	LL	IPS	ADF-Fisher
turn	0.0000	0.0000	0.0000
ret	0.0000	0.0000	0.0000
mret	0.0000	0.0000	0.0000
sig	0.3909	0.0000	0.0000

Table 5 Market VAR estimation

		MRET	MTURN
MRET(-1)	coefficient	0.1053	0.3330
	standard error	-0.0632	-0.1828
	P-value	0.0971	0.0697
MRET(-2)	coefficient	0.0281	-0.1766
	standard error	-0.0595	-0.1720
	P-value	0.6369	0.3056
MTURN(-1)	coefficient	0.0380	0.1307
	standard error	-0.0212	-0.0614
	P-value	0.0745	0.0342
MTURN(-2)	coefficient	-0.0084	0.1440
	standard error	-0.0188	-0.0545
	P-value	0.6545	0.0087
MISG	coefficient	-0.5600	2.2410
	standard error	-0.1675	-0.4844
	P-value	0.0010	0.0000
MISG(-1)	coefficient	0.0209	-1.5874
	standard error	-0.1709	-0.4942
	P-value	0.9029	0.0015
DISP	coefficient	1.0303	1.8986
	standard error	-0.1589	-0.4595
	P-value	0.0000	0.0000
DISP(-1)	coefficient	-0.2873	0.2098
	standard error	-0.1784	-0.5160
	P-value	0.1086	0.6847
C	coefficient	0.0020	-0.1989
	standard error	-0.0135	-0.0390
	P-value	0.8849	0.0000
DUM98	coefficient	-0.0197	0.1366
	standard error	-0.0221	-0.0640
	P-value	0.3727	0.0336

Table 6 Granger causurity test: mret and mturn

Null Hypothesis:	Obs	F-Statistic	Prob.
MRET does not Granger Cause MTURN	265	2.42737	0.0903
MTURN does not Granger Cause MRET		0.94012	0.3919

Table 7 Individual security model estimation

Dependent Variable: TURN Independent variables		Model 1	Model 2	Model 3
mret(-1)	Coeficient	0,2797	null	0,3218
	standard error	0,1616		0,1424
	p-value	0,0835		0,0239
mret(-2)	Coeficient	-0,0653	null	-0,0661
	standard error	0,0408		0,0285
	p-value	0,1095		0,0203
ret(-1)	Coeficient	0,0312	0,1715	null
	standard error	0,1012	0,0910	
	p-value	0,7580	0,0594	
ret(-2)	Coeficient	-0,0022	-0,0173	null
	standard error	0,0272	0,0190	
	p-value	0,9352	0,3622	
AR(1)	Coeficient	0,2303	0,2226	0,2316
	standard error	0,0166	0,0174	0,0147
	p-value	0,0000	0,0000	0,0000
AR(2)	Coeficient	0,1332	0,1387	0,1317
	standard error	0,0122	0,0121	0,0117
	p-value	0,0000	0,0000	0,0000
sig	Coeficient	1,8532	1,2115	1,8398
	standard error	0,9236	0,9888	0,9078
	p-value	0,0448	0,2205	0,0427
sig(-1)	Coeficient	-0,2175	-0,1104	-0,2049
	standard error	0,3612	0,3880	0,3550
	p-value	0,5471	0,7761	0,5638
C	Coeficient	-0,0399	-0,0271	-0,0400
	standard error	0,0160	0,0157	0,0159
	p-value	0,0125	0,0841	0,0118
R-squared		0,1821	0,1520	0,1804
Durbin-Watson stat		1,9930	1,9929	1,9932
cross-sectional effect		fixed	fixed	fixed

Reference

Allen, F., Qian, F., Qian, M. (2005). “Law, finance, and economic growth in China.” Journal of Financial Economics **77**(1): 57-116.

Antweiler, W. and Frank,M.Z., (2004). “Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards.” Journal of Finance **59**(3):1259–1294.

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005), “Judgemental Overconfidence, Self-Monitoring, and Trading Performance in an Experimental Financial Market.” Review of Economic Studies **72**(2): 287-312.

Brooks, C. (2008). Introductory Econometrics for Finance, Cambridge University Press.

Bryman, A. (2004). Social research methods, Oxford University Press.

Chen, G., Kim, K. A., Nofsinger, John R. and Rui, Oliver M. (2007). “Trading performance, disposition effect, overconfidence, representativeness bias, and experience of emerging market investors.” Journal of Behavioral Decision Making **20**(4): 425-451.

Campbell, J. Y. and Lettau, M. (1999). “Dispersion and Volatility in Stock Returns: an Empirical Investigation.” SSRN eLibrary.

Clark, P. K. (1973). “A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices.” Econometrica **41**(1): 135-155.

Daniel, K., Hirshleifer, D. and Subrahmanyam, A. (1998). “Investor Psychology and Security Market under- and Overreactions.” Journal of Finance **53**(6): 1839-1885.

De Bondt, W. F. M, Thaler, R. H..(1994). Financial decision-making in markets and firms: a behavioral perspective, National Bureau of Economic Research.

De Bondt, W. F. M. (1998). “A portrait of the individual investor - Heuristics and biases.” European Economic Review **42**(3-5): 831-844.

Deaves, R., Lüders, E. and Luo, G.Y. (2009). “An Experimental Test of the Impact of Overconfidence and Gender on Trading Activity.” Review of Finance **13**(3): 555-575.

De Medeiros, Otavio R. and Van Doornik, B. F. N. (2008). “The Empirical Relationship between Stock Returns, Return Volatility and Trading Volume in the Brazilian Stock Market.” Brazilian Business Review (English Edition) **5**(1): 1-17.

Deo, Malabika, Srinivasan, K. Devanadhen, K. (2008). “The Empirical Relationship between Stock Returns, Trading Volume and Volatility: Evidence from Select Asia-Pacific Stock Market.” European Journal of Economics, Finance and Administrative Sciences **12**: 58-68.

Fischhoff, B. Slovic, P and Lichtenstein, S. (1977). Knowing with Certainty: The Appropriateness of Extreme Confidence, Journal of Experimental Psychology: Human Perception and Performance.

French, K. R., Schwert, G. W. and Stambaugh, R.F. (1987). “Expected Stock Returns and Volatility.” Journal of Financial Economics **19**(1): 3-29.

Gervais, S. and Odean, T. (2001). “Learning to be overconfident.” Review of Financial Studies **14**(1): 1-27.

Gervais, S., Kaniel, R. and Mingelgrin, D. (2001). “The High-Volume Return Premium.” Journal of Finance **56**(3): 877-919.

Glaser, M. and Weber, M. (2009). “Which past returns affect trading volume?” Journal of Financial Markets **12**(1): 1-31.

Glaser, M. and Weber, M. (2007). “Overconfidence and trading volume.” The Geneva Risk And Insurance Review **32**(1): 1-36.

Glaser, M., Weber, M. and Langer, T. (2010). “Overconfidence of Professionals and Lay People: Individual Differences within and between Tasks?”,working paper.

Graham, J. R. and Harvey, C. R. (2001). “Expectations of Equity Risk Premia, Volatility and Asymmetry from a Corporate Finance Perspective.”, working paper.

Harris, R. and Sollis, R. (2003). Applied Time Series Modelling and Forecasting. Wileys

Hilton, D. J. (2001). “The Psychology of Financial Decision-Making: Applications to Trading, Dealing, and Investment Analysis.” Journal of Behavioral Finance **2**(1): 37-53.

Hodrick, R.J. and Prescott, Edward C.(1997) “Postwar U.S. Business Cycles: An Empirical Investigation.” Journal of Money, Credit, and Banking **29**(1): 1-16.

Holtz-Eakin, D., Whitney, N. and Rosen, H. S. (1988). “Estimating Vector Autoregressions with Panel Data.” Econometrica **56**(6): 1371-1395.

Im, K. S., Pesaran, M. H. and Shin, Y. (2003). “Testing for unit roots in heterogeneous panels.” Journal of Econometrics **115**(1): 53-74.

Levin, A., Lin, C. F. and Chu, C. J. (2002). “Unit root tests in panel data: asymptotic and finite-sample properties.” Journal of Econometrics **108**(1): 1-24.

Lichtenstein, S., Fischhoff, B., and Phillips, L.D. (1982). “Calibration of Probabilities: The State of the Art to 1980.” in Judgment Under Uncertainty: Heuristics and Biases, Kahneman, D., Slovic P. and Tversky, A. (Eds.), Cambridge: Cambridge University Press, pp. 306–334.

Lo, A.W. and Wang, J. (2000). “Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory.” Review of Financial Studies **13**(2): 257-300.

Maddala, G. S. and Wu, S. (1999). “A Comparative Study of Unit Root Tests With Panel Data and a New Simple Test.” Oxford Bulletin of Economics & Statistics **61**(4): 631.

Mestel, R., Gurgul, H. and Majdosz, P. (2003), “The Empirical Relationship between Stock Returns, Trading Volume and Volatility on the Austrian stock market.”, working paper

Miller, D. and Ross, M. (1975) “Self-serving Biases in Attribution of Causality: Fact or Fiction?.” Psychological Bulletin **82**:213–225.

Mubarik, F. and Javi, A. Y. (2009). “Relationship Between Stock Return, Trading Volume And Volatility: Evidence From Pakistani Stock Market.” Asia Pacific Journal of Finance & Banking Research **3**(3): 1-17.

Nickell, S. (1981). “Biases In Dynamic Models With Fixed Effects.” Econometrica **49**(6): 1417-1426.

Oberlechner, T. and Osler, C. L. (2008). “Overconfidence in Currency Markets.” working paper.

Odean, T. (1998). “Are Investors Reluctant to Realize Their Losses?” Journal of Finance **53**(5): 1775-1798.

Odean, T. (1998a). “Volume, Volatility, Price, and Profit When All Traders Are above Average.” Journal of Finance **53**(6): 1887-1934.

Odean, T. (1999). “Do Investors Trade Too Much?” American Economic Review

89(5): 1279-1298.

Rouwenhorst, K. G. (1998). "International Momentum Strategies." Journal of Finance **53**(1): 267-284.

Shefrin, H. and Statman, M.(1985). "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence." Journal of Finance **40**(3): 777-790.

Shiller, R. J. (1999). "Human Behavior and the Efficiency of the Financial System." In Handbook of Macroeconomics, J. Taylor and M. Woodford (Eds.), Amsterdam: Elsevier Science, pp. 1305–1340.

Svenson, O. (1981). "Are we all less risky and more skillful than our fellow drivers?" Acta Psychological **47**(2): 143-148.

Statman, M., Thorley, S., Vorkink, K.(2006). "Investor Overconfidence and Trading Volume." Review of Financial Studies **19**(4): 1531-1565.

Taylor, S. E. and Brown, J. D. (1988). "Illusion and well-being: a social psychological perspective on mental health." Psychological Bulletin **103**(2): 193-210.

Verbeek, M. (2004).A guide to modern econometrics. Wiley.

Zhu, N. (2002). "The Local Bias of Individual Investors." working paper.