# 國立交通大學

## 財務金融研究所

## 博 士 論 文

選擇權評價和選擇權交易量

Essays on Option Pricing and Option Trading Volume

研 究 生：王志瑋

指導教授：周幼珍 博士

李漢星 博士

中 華 民 國 一百零二 年 六 月

# 選擇權評價和選擇權交易量

# Essays on Option Pricing and Option Trading Volume

研 究 生：王志瑋　　Student：Chih-Wei Wang

指導教授：周幼珍 博士　　Advisors：Dr. Yow-Jen Jou

李漢星 博士　　Dr. Han-Hsing Lee

國 立 交 通 大 學

財 務 金 融 研 究 所

博 士 論 文

A Thesis

Submitted to Graduate Institute of Finance

College of Management

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Finance

June 2013

Hsinchu, Taiwan, Republic of China

中華民國 一百零二 年 六 月

# 選擇權評價和選擇權交易量

學生：王志瑋　　　　　　　　　　　　　　　　　　指導教授：周幼珍 博士

李漢星 博士

國立交通大學財務金融研究所博士班

## 摘要

本文研究主要包含二個不同議題。第一個議題是，我們擴展 Log-HAR 波動度模型在選擇權定價上，這是比其他的已實現波動率選擇權定價模型更簡單且更方便。此外，我們用 S＆P 500 指數選擇權市場資料去比較 HAR 模型和 NGARCH 選擇權定價模型，其中 NGARCH 已被證明為最好的模型在 GARCH 模型裡面。我們的實證分析是在最近的金融危機的選擇權交易，從 2007 年 7 月 3 日 2008 年 12 月 31 日。我們發現，HAR 型模型成功地預測了樣本外選擇權價格 ，可能原因是因為已實現波動率比較接近波動率指數（VIX）在金融市場的波動。HARG 和 Log-HAR 選擇權定價模型存在 mixed 結果，因為在動盪時期，Log-HAR 選擇權定價模型比 HARG 選擇權定價模型好，而在相對不穩定的時期 HARG 選擇權定價模型 Log-HAR 選擇權定價模型好。第二個議題是相關 informed 投資者可以預測未來指數報酬率在台灣等新興市場。不同於以往的實證研究結果，我們發現，在更近的時期，國內機構投資者表現出顯著的預測能力在 Daily TAIEX 報酬率，除了在 2008 年的金融危機之外。相反的，外國機構投資者只有微弱的可預測性在金融危機之前。我們進一步探討指數報酬率和不同的交易者類型之間的 lead-lag 關係。我們的研究結果顯示，2008 年金融危機之前，只有國內機構投資者對 intraday TAIEX 報酬率具有預測能力。在 2008 年的金融危機，intraday TAIEX 報酬率顯著領先國外和國內機構投資者的選擇權交易量，這表明雖然機構投資者密切關注市場的波動做出反應，卻無法事先預測市場的波動。

**關鍵字：**已實現波動率、Log-HAR、HARG、NGARCH、S&P500 指數選擇權、國外投資人、國內投資人、Intraday lead-lag 分析

# Essays on Option Pricing and Option Trading Volume

Student：Chih-Wei Wang　　Advisors：Dr. Yow-Jen Jou
Dr. Han-Hsing Lee

Graduate Institute of Finance, National Chiao Tung University

## Abstract

This dissertation consists of two separate issues. The first issue is we extend Log-HAR option pricing model, which is more convenient compared to other option pricing models associated with realized volatility in the way of simpler estimation procedure. In addition, we test the empirical implications of HAR-type models in the S&P 500 index options market with the comparison of the NGARCH option pricing model that has been documented as the best model in pricing options among GARCH-type models. Our empirical analysis is based on options traded from July 3, 2007 to December 31, 2008 covering the recent financial crisis, where has never been discussed in existing literature. Overall, we find that the HAR-type models successfully predict out-of-sample option prices probably because they are based on realized volatilities, which are closer to expected volatility (VIX) in financial markets. However, it seems to exist the mixed result between the Log-HAR and the HARG models in pricing options since the Log-HAR is better than the HARG in times of turmoil, while it is worse during the rather unstable period. The second issue is related to informed investors can predict future index returns in emerging markets like Taiwan. Unlike previous empirical results, we find that in more recent periods, the put-call ratio of domestic institutional investors show significant predictive power for daily TAIEX returns, except during the 2008 financial crisis. In contrast, the put-call ratio of foreign institutional investors only has weak predictability for the TAIEX returns prior to the severe global market downturn in late 2008. We further explore the intraday lead-lag relationship among index returns and put-call ratios of different trader types. Our results show that only the trading of domestic institutional investors possesses predictive capability for intraday TAIEX returns prior to the 2008 financial crisis. During the 2008 financial crisis, intraday TAIEX returns significantly lead option trades of foreign and domestic institutional investors, suggesting that although institutional investors closely watch and react to market fluctuations, they are unable to predict market movement beforehand.

**Keywords**: Realized volatility, Log-HAR, HARG, NGARCH, S&P500 index option, Foreign investors, Domestic investors, Intraday lead-lag analysis

# 誌謝

從交大財金所碩士班到博士班這段時間，能夠順利完成論文的作業，首先我最要感謝的是指導教授周幼珍老師及李漢星老師的辛勤教導及殷殷教誨。其中周幼珍老師更是支持我出國念書的最大動力，多年的學習路途一路走來， 周老師總是不厭其煩地提醒我在學術研究上應該具備的正確態度及習慣，老師除了學習上的教導也時常關心我的日常生活，並且幾次在我陷入困境時，他是將我拉出困境的重要推手； 在和李漢星老師做論文的期間，除了學習到很多程式能力和如何尋找論文題目，他也隨時提供設備上的幫助。對我而言，周幼珍老師就像是我的媽媽一樣；而李漢星老師就像是個亦師亦友的角色。在此同時，我感謝交大財金所諸位老師辛勤的教導，特別感謝李正福老師、鍾惠民老師和黃星華老師，在百忙之中仍撥空擔任我的論文口試委員。亦感謝台灣大學石百達老師及成功大學林軒竹老師擔任本篇論文的口試委員並給予我許多寶貴的建議及方向。

感謝在國外一起念書的婉茜，她是我從交大碩班一起做論文到現在的最佳戰友，雖然每次在討論論文時會因為不同意見而爭吵，而我們卻可以因為彼此的意見而做出更好的論文，這也成為我這幾年研究生活中不可或缺的前進動力。

最後我要感謝宜穎，一個我生命中最重要的人，一直很相信我也支持我的決定，謝謝你相信我可以從統計轉到財金領域，謝謝你相信我可以在交大完成博士學業和繼續出國念書。而後我會更用心在研究領域上，希望可以做出更多好品質的論文。

志瑋
謹誌於 新竹交大
民國一百零二年六月

# Contents

# List of Tables

# List of Figures

## Chapter1. Introduction

There is a growing demand for theoretical and empirical knowledge of the financial volatility. It is well-known that financial volatility has played such a central role in derivative pricing, asset allocation, and risk management. The behavior of dynamic volatility process has been widely studied since the generalized autoregressive conditional heteroskedastic (GARCH) asset return process is proposed by Bollerslev (1986), which considers volatility as an unobservable variable and therefore uses a specified conditional mean and variance model to estimate the latent volatility. For the recent decade, an alternative approach is to construct an observable proxy for the latent volatility with high frequency intraday data for estimating lower frequency volatility, called realized volatility. Besides, TAIEX options in 2008 were ranked the fifth most frequently traded index options in the world, according to the annual Futures Industry Association's survey (Burghardt and Acworth, 2009). The structure of options trades in Taiwan is different in that while there is high institutional investor participation in the U.S. market, in Taiwan over 40% of trades in derivatives and stocks are done by individual investors. Based on above description, there are two parts in this essay. They present two independent papers, respectively.

In the first part of this dissertation, we continue the existing issue of option valuation with the specification of the HAR-type volatility model, but, different from the HARG model, we only consider the case when the risk neutral and physical dynamics of realized volatilities are the same, and use the Log-HAR model to capture the dynamic volatility. Our setting gives us less parameters to estimate in a more simple way than HARG. Although HARG expresses a solid foundation from theoretical viewpoint, we wonder whether a model with simple procedure to implement but with poor theory to back up can perform better through empirical viewpoint.

In the second part of this dissertation, foreign investors are generally believed to possess superior private information to domestic traders in an emerging market. Barber et al. (2009) documented that domestic individual investors suffer systemic and economically large losses in the Taiwan stock market while foreign institutional investors are the main winners. Therefore, experienced domestic investors should recognize this fact and closely watch trades of foreign institutions. In order to confirm this evidence, we use the more recent data on TAIEX options, including the period before and during the 2008 global financial tsunami, to re-examine the predictive ability of private information from different classes of investors for future index returns. This is important, because over the years domestic investors have acknowledged that foreign investors may possess superior information and foreign capital flows can create price pressure for the host-country stock market. The discrepancy between our analysis and those in the past is worth exploring, because similar patterns can also appear in other emerging markets with high individual participation, such as China. Moreover, we conduct daily and intraday VAR (Vector Autoregression) analyses to examine the relationship between options trading activities of various classes of investors and their relation to TAIEX returns.

To sum up, the dissertation provides some insights into the issues of realized volatility for option pricing and index option volume for predicting future index return. With these points in mind, the research results will provide us with the empirical evidences to comprehend the occasion of some distinctive phenomena in financial markets.

# Chapter2. Is the Realized Volatility Good for Option Pricing during the Recent Financial Crisis?

## 2.1 Introduction

With the fact that options are heavily and frequently traded for reasons of speculation or hedge, it is crucial to find the method that can not only forecast option more accurately but also spend less computing time. Although Black-Scholes model (called BS model, 1973) provides a simple framework to price options, some well-known pricing biases of this model have been widely documented.[1] The unreasonable assumptions of BS model that the volatility is constant should be responsible to those pricing biases, and a large proportion of literature introduces time-varying volatility to the BS model.[2] The key to the success of any option pricing model is determined by the capability of return process consistent with the distributional and time series properties of the underlying asset (Bates, 1995). That is, embedding the dynamic volatility process in option pricing model is necessary in order to evaluate options more accurately.

The behavior of dynamic volatility process has been widely studied since the generalized autoregressive conditional heteroskedastic (GARCH) asset return process is proposed by Bollerslev (1986), which considers volatility as an unobservable variable and therefore uses a specified conditional mean and variance model to estimate the latent volatility. For the recent decade, an alternative approach is to construct an observable proxy for the latent volatility with high frequency intraday data for estimating lower frequency volatility, called realized volatility. It is a non-parametric method of volatility measurement by calculating the summation of square intraday asset returns, formalized by Andersen et al. (2001) (Henceforth ABDL). In order to model realized volatilities, Corsi (2009) establishes a conditional volatility model, called Heterogeneous Autoregressive model of the Realized Volatility (HAR henceforth). The advantage of this volatility model is in the way of rather easy to be estimated with multivariate settings in contrast to

[1] Such as, underpricing of out-of-the-money options (Gultekin et al. 1982), underpricing of options on low-volatility securities (Black and Scholes 1972; Gultekin et al. 1982; Whaley 1982), and underpricing of short-maturity option (Black 1975; Whaley 1982).

[2] See for example, Eisenberg and Jarrow (1994), Lee at al. (2004), and Lee at al. (2005).

GARCH-type models. The author's empirical findings also support that HAR model successfully captures the main empirical features of financial data such as long memory and fat tail. In theory, although the HAR model does not formally possess long-memory, the mixture of relatively few volatility components makes it capable of reproducing remarkable slow volatility autocorrelation decay in a simple and parsimonious way.[3] On the basis of the HAR model, Andersen, Bollerslev, and Diebold (2007) state that the logarithmic HAR model (Log-HAR henceforth) are more amenable to the use of standard time series procedures, since the logarithmic daily realized volatilities are approximate unconditionally normal distributed. The advantage of HAR (or Log-HAR) in modeling volatilities over the GARCH-type model motivates us to investigate whether its superiority is able to extend to the realm of pricing options.

As far as we know, only three papers embed realized volatility models into the application of pricing options. Stentoft (2008) has shown that realized volatility following Inverse Gaussian process can reduce pricing errors on three individual stock options. Feunou and Meddahi (2009) extend the class of affine model to non-Markovian dynamics to build an option pricing model using realized volatility measures. However, Corsi et al. (2013) argue that neither the former model nor the latter model provides a formal change of risk neutral measure for the realized volatility process; the latter model even does not consider leverage effect. They fill this gap by combining a flexible discrete time option pricing framework (Bertholon et al., 2008) and a simple approximate long-memory model (Corsi, 2009), to make a formal change of measure for the realized volatility process, and then proposed the option pricing model, called the Heterogeneous Auto-Regressive Gamma (HARG) model. Their empirical findings on the S&P 500 index option (henceforth SPX) show HARG outperforms two GARCH-type option pricing models: the GARCH model proposed by Heston & Nandi (2000), and the component GARCH model, which gives rise to a GARCH process with a more persistent dynamics than the standard one (Christoffersen et al., 2008).

This paper continues the existing issue of option valuation with the specification of the HAR-type volatility model, but, different from the HARG model, we only consider the case when the risk neutral and physical dynamics of realized volatilities are the same, and use the Log-HAR model to capture the dynamic volatility. Our setting gives us less parameters to estimate in a more simple way than HARG. Although HARG expresses a solid foundation from theoretical viewpoint, we wonder whether a model with simple procedure to implement but with poor theory to back up can perform better through empirical viewpoint.

To the extent which GARCH-type options, it is firstly developed by Duan (1995). In the following, the closed-form GARCH option pricing model is proposed and its corresponding

[3] Ding, Granger, and Engle (1993), Baillie, Bollerslev, and Mikkelsen (1996), and Bollerslev and Mikkelsen (1999) suggest that the autocorrelation of volatility decreases with a hyperbolic rate and it could be displayed by long-memory processes to be a more adequate representation for the conditional variance of returns.

empirical findings on SPX also support its superiority over *ad hoc* Black-Scholes model of Dumas et al. (1998) (see Heston and Nandi, 2000). In the comparison of various GARCH-type option valuation models, a comprehensive study on SPX documents that the non-linear asymmetric GARCH (NGARCH) option model is superior in removing biases from pricing residuals for all moneyness and maturity categories especially for out-the-money contracts (Hsieh and Ritchken, 2005).[4] Overall, the existing empirical evidence tends to support that the NGARCH option pricing model is the best to the use of evaluating options. It thus motivates us to provide empirical analysis to the comparison of out-of-sample option pricing performance between HAR-type option valuation models and NGARCH, which is never discussed before.

In sum, the contribution of this paper is in three-folds. First, we propose Log-HAR option pricing model, which is more convenient compared to other option pricing models associated with realized volatility in the way of simpler estimation procedure.[5] Second, although the NGARCH option pricing model has been documented as the best model among GARCH-type models in pricing options, the comparison between NGARCH and HAR-type option valuation models is still missing. This paper fills this gap by comparing the empirical results of the out-of-sample valuation errors on four models: (1) our Log-HAR option pricing model, (2) the HARG proposed by Corsi et al. (2013), (3) the NGARCH option pricing model, and (4) the Black-Scholes model as the benchmark model. Third, as our Figure 2.1 shows that the realized volatility becomes dramatically volatile and higher since Lehman Brothers filed for bankruptcy on September 15, 2008. However, as far as we know, no existing literature provides empirical investigation on SPX during the recent financial crisis. It is important to have knowledge which option pricing model can forecast option prices more accurately since options are widely used as financial instruments to hedge in the times of turmoil.

Our empirical analysis on SPX (from July 3, 2007 to December 31, 2008) shows that the out-of-sample valuation errors from the HAR-type models are lower than those from other models, including the NGARCH option pricing model that has been documented as the best model in pricing options among GARCH-type models, and the Black-Scholes model that has been regarded as the traditional benchmark model for many literatures. The HAR-type models successfully predict out-of-sample option prices probably because they are based on realized volatilities, which are closer to expected volatility (VIX) in financial markets as shows in Figure 2.1. Moreover, in times

[4] More empirical evidence of NGARCH option pricing models are provided as follows. First, in terms of Hang-Seng index options, the empirical performance of NGARCH option pricing model outperforms two ad hoc versions of the BS model (Duan and Zhang, 2001). Comparing stochastic volatility model on the FTSE 100 index, Lehar, et al. (2002) conclude the empirical results to out-of-sample option pricing performance of FTSE 100 index option show that NGARCH dominates both stochastic volatility and the benchmark BS model.
[5] For example, HARG includes non-linear optimization procedure, while our Log-HAR option price model only requires the Ordinary-Least-Square method.

of extremely turmoil (it defines the period from September 15 to December 31, 2008 in our paper), the Log-HAR performs better than the HARG model in mid-term and long-term contracts, while worse than the HARG model in short-term puts. As a result, it seems to support that the model constructed based on realized volatilities and with simpler framework could value option prices more accurately in the fluctuant period as contracts with longer than 46 days to expiration. During the rather unstable period (it defines the period from July 3, 2007 to September 14, 2008), it again holds up the superiority of the HAR-type models over other models, and the HARG model performs better than the Log-HAR model. Overall, it seems to exist the mixed result between the Log-HAR and the HARG models in pricing options during very turmoil or rather unstable periods.

The remainder of this paper is organized as follows. Section 2 describes the data, and Section 3 provides the literature review of realized volatility and HAR-RV model. Then, Section 4 demonstrates option pricing models. Subsequently, Section 5 describes the analysis of out-of-sample pricing performance. The conclusion is finally drawn in Section 6, along with recommendations for future research.

## 2.2 Data

### 2.2.1 Option Data

This paper investigates SPX because it is heavily traded in U.S. and has been paid much attention in many empirical literatures on options pricing.[6] Our option data is collected from TICK DATA database,[7] and covers the recent financial crisis period from July 3, 2007 to December 31, 2008. Based on some concerns in existing literature, the following rules are applied to filter the raw tick option data to be the daily option data set for the later empirical test.

1) For a given day, although an option with the same strike, maturity, and type (call or put) may be quoted many times, the same option is represented only once. That is, this procedure changes tick-by-tick option data set into daily option data set.
2) Due to the different trading rule between SPX and S&P 500 index (the former ceases trading at 4:15 PM Eastern Standard Time, while the later ceases trading at 4:00 PM Eastern Standard Time), researchers face a serious nonsynchroneity problem when using the OptionMetrics database to evaluate other option pricing models (Battalio and Schultz, 2006).[8] To avoid the nonsynchroneity problem, for each type of option, we only select the option data that is before and most close to 4:00 p.m. and match corresponding S&P 500 index level at the moment

[6] See, for example, Bakshi et al. (1997), Heston and Nandi (2000), and Corsi et al. (2010).
[7] Tick Data database sources intraday option data from the Chicago Board Options Exchange (CBOE), and then provides historical tick-by-tick options data for all listed U.S. equity and index options contracts reported by the Options Price Reporting Authority (OPRA) back to July 2, 2004. All data is recorded in a timestamp to the second prior to July 1, 2008, and to the millisecond after July 1, 2008.
[8] In corsi et al. (2010), the authors gather option data from OptionMetrics, and thus ignore the nonsynchroneity problem, which could generate bias empirical analysis in evaluating options.

when the selected option price is recorded.

3) To mitigate the impact of price discreteness, on option valuation, options with values smaller than $3/8 are excluded. As options with less than 6 days to expiration may induce liquidity-related biases, they are excluded from the sample.[9] We also eliminate an option as its trading volume is less than five.
4) Following what earlier literature normally did,[10] keeps only options with moneyness between 0.9 and 1.1, where moneyness defined as $M = \left(S/K\right)$, $S$ and $K$ are the S&P 500 index level price and the strike price, respectively.
5) As in Bakshi et al. (1997), we take out the option as its price violates the arbitrage restriction. That is, the call price has to be greater than or equal to (1) the underlying asset spot price minus the strike price, and (2) the underlying asset spot price minus the present value of the remaining dividends minus the discounted strike price. Similar, the put price has to be greater than or equal to (1) the strike price minus the underlying asset spot price, and (2) the present value of the remaining dividends plus the discounted strike price minus the underlying asset spot price.

In order to evaluate option, ones needs to adjust daily index level price by taking into account dividends paid in many of the stocks in the S&P 500 index. We collect a time series data of daily dividend yield form Datastream.[11] For each option contract, the dividend-adjusted corresponding index level price is computed by subtracting the present value of the future realized dividends until the maturity of the option from the current index level. Later, we use the adjusted index level to option pricing models as the initial underlying asset price.

We consider only out-of-the-money (OTM) options and divide the moneyness into six groups including: (<0.94), (0.94–0.97), (0.97–1.00), (1.00–1.03), (1.03–1.06), and (>1.06). Since this paper only investigates out-of-the-money options, the former three groups contain only call options while the later three groups contain put options. Following Hsieh and Ritchken (2005), the days-to-maturity is divided into 3 groups: short-term contracts have maturities between 10 and 45 days; mid-term contracts have maturities between 46 and 90 days; and long-term contracts have maturities between 91 and 180 days. Finally, our option data set has been split up into 18 categories and remains a total number of 12,408 observations, where call and put options separately include 6690 and 5718 observations. Table 2.1 demonstrates the summary statistics of our option data set. We can observe that the average option prices increases as days-to-maturity increases in both call or put options. Also, the average call price for each category (see panel A) increases with the increase

[9] See, for example, Bakshi et al. (1997) and Yung and Zhang (2003).
[10] See, for example, Henton and Nandi (2000).
[11] See Santa-Clara and Yan (2010).

of moneyness; on the contrary, the average put price (see panel B) decreases as moneyness increases. All of these results are consistent with common knowledge about options.

< Table 2.1 is inserted about here >

### 2.2.2 Index Data

In order to price SPX, one needs time series data of S&P 500 index, which is the underlying asset of SPX, and should contain a period of data prior to the target option date for obtain estimated parameters. We use 1000 observations before the date of evaluated option to estimate parameters for alternative option valuation methods.[12] Since our first evaluated option date is on July 3, 2007, we select S&P 500 index begins with July 2, 2003, and it thus leads our S&P 500 index covering from July 2, 2003 to December 31, 2008. The index is collected from TICK DATA database, which takes an advantage of providing different interval data series, including tick-by-tick, five-minute, ten-minute, and daily interval. This paper uses daily data in the Black-Scholes model and NGARCH model, while uses five-minute interval data to calculate realized volatilities in the application of both Log-HAR and HARG model. The choice of five-minute horizon is to mitigate market microstructure friction effects, including price discreteness, infrequent trading, and bid-ask bounce effects (see ABDL 2001).

## 2.3 Literature Review of Realized Volatility and HAR Model

### 2.3.1 Realized Volatility

The main framework of realized volatility was proposed by ABDL (2001). Assuming that the logarithmic asset price follows a continuous-time process

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) \tag{1}$$

where $p(t)$ is the logarithm of instantaneous price, $\mu(t)$ is a continuous, finite variation drift process, $dW(t)$ is the standard Brownian motion, and $\sigma(t)$ is a stochastic process independent of $dW(t)$. For this diffusion process, the integrated volatility at day $t$, is the integral of the instantaneous volatility over the one day interval $(t-1;t)$,

$$\sigma_t = \left(\int_{t-1}^{t} \sigma^2(\omega)\,d\omega\right)^{1/2} \tag{2}$$

ABDL (2001), applying the quadratic variation theory, suggest that the sum of intraday squared returns converges (as the maximal length of returns go to zero) to the integrated volatility of the

[12] In line with Corsi (2009), for every time of implementing estimation, he selects a data set contains 1000 observations in estimating parameters.

prices. This nonparametric estimator is called realized variance.[13] The definition of the realized variance over a time interval of one day is

$$RV_t = \sum_{j=0}^{M-1} r_{t-j\Delta}^2 \tag{3}$$

where $\Delta = 1/M$ is discretely sampled period and $r_{t-j\Delta} = p(t-j\Delta) - p(t-(j+1)\Delta)$ defines continuously compounded returns. Under these assumptions, the ex-post realized volatility is an unbiased volatility estimator. Moreover, as the sampling frequency is increased, the realized volatility provides a consistent nonparametric measure of the integrated volatility over the fixed time interval.

$$p\lim_{M\to\infty} RV_t = \sigma_t^2 . \tag{4}$$

### 2.3.2 Empirical Properties of Realized Volatility

In measuring daily realized volatilities, we use five-minute interval S&P 500 index return series in the period of July 2, 2003–Dec. 31, 2008, total 1375 trading days. Figure 2.1 demonstrates annualized realized volatilities (see the solid and black line) during this period. The realized volatility has started fluctuating since the second half year of 2007. In particular, it has a dramatically increase on September 15, 2008 (Lehman Brothers filed for bankruptcy), and arrives its peak on October 10, 2008. Overall, after July 2007, it is obvious to see that the realized volatilities are higher and more unstable than those before June 2007. As we only take look at realized volatilities after July 2007, it is able to observe two parts representing one rather unstable period of July 2007–Sep. 14, 2008, and the other very fluctuant period of Sep. 15, 2008–Dec. 31, 2008. We thus partition our data series into three partitions and define (1) the period before July 3, 2007 as the very stable period, (2) the period between July 3, 2007 and Sep .14, 2008 as a rather unstable period, and (3) the period after Sep. 15, 2008 as very fluctuant period. Note that even the most recent empirical investigation on SPX only deals with data over the very stable period.[14] We wonder if results of existing literature still hold in times of turmoil as in the very stable period.

< Figure 2.1 is inserted about here >

Furthermore, Figure 2.1 also plots the volatility captured by NGARCH model (see the dotted line and henceforward we call it NGARCH volatility) and VIX (see the solid and gray line). There

[13] To be clear, this paper denotes the square root of the realized variance as realized volatility.
[14] Corsi et al. (2010) study SPX from January 5, 2000 to December 31, 2004.

are some remarkable difference between the realized volatility and NGARCH volatility. First, NGARCH volatility expresses an extremely stable trend in contrast to the realized volatility through the whole period. Second, while NGARCH volatility also reveals an increasing trend after Sep. 15, 2008, the scale seems too small, and the movement seems too stable compared to the realized volatility. Third, in contrast to NGARCH, realized volatility represents a more similar dynamic activity as VIX does. At least, we believe that realized volatility is a better measure of capturing market's volatility than NGARCH volatility since VIX is a popular measure of the implied volatility of SPX, and often represents one measure of the market's expectation of stock market volatility.

To the extent which the distribution of realized volatilities itself and of standardized returns adjusted by realized volatilities, the Panel A of Figure 2.2 demonstrates that the unconditional distribution of realized daily volatilities seems to be highly non-normal and skew to the right; in contrast, the Panel B of Figure 2.2 displays the distribution of logarithmic realized volatilities is closer to a normal distribution. Our findings are consistent with earlier evidence that the distribution of logarithmic realized volatilities is generally much closer to Gaussian distribution than it of the raw realized volatility series (ABDL, 2001). Furthermore, although the unconditional distribution of daily returns is leptokurtic, the distribution of daily returns normalized by the realized volatility is close to normal (see Panel C of Figure 2.2). For a more detail report, the right side of Table 2.2 also indicates that the distribution of daily standardized returns is approximate unconditionally normal distribution because of the skewness close to zero as well as the kurtosis close to three.

< Figure 2.2 is inserted about here >

< Table 2.2 is inserted about here >

### 2.3.3 HAR and Log-HAR Model

The Heterogeneous Autoregressive of the Realized Volatility (HAR) firstly proposed by Corsi (2009), can directly model and forecast the time series behavior of realized volatilities by a conditional volatility model which is able to reproduce the memory persistence observed in the data. At the same time, this approach remains an easy way to estimate. The simulation results in Corsi (2009) seem to confirm that the HAR-RV model successfully fits the main empirical features of financial data (long memory and fat tail) in a simple and parsimonious way. For example, empirical results on USD/CHD data by applying HAR-RV model steadily and substantially outperform other previous models (the standard GARCH model and the stochastic volatility model).

The HAR-RV model includes past volatilities aggregated over different time horizons as explanatory variables. Regarding to realized variance over different time horizons longer than one day, these multi-period variables are normalized sums of the one-period realized variances (i.e. a simple average of the daily quantities). More specifically, the multi-period variation measures by the sum of the corresponding daily measures:

$$RV_{t-k,:t-1}=\frac{1}{k}\sum_{j=1}^{k}RV_{t-j} \tag{5}$$

In Corsi et al. (2009), they take into account $k=5$ and $k=22$ approximately corresponding to one week and one month, respectively. Furthermore, they express the one-day ahead volatility as a linear combination of past one day, average five days, and average twenty-two days realized volatilities. The equation (6) is labeled as HAR-RV model

$$\left(RV_t\right)^{1/2}=\beta_0+\beta^{(d)}\left(RV_{t-1}\right)^{1/2}+\beta^{(W)}\left(RV_{t-5,t-1}\right)^{1/2}+\beta^{(M)}\left(RV_{t-22,t-1}\right)^{1/2}+\varepsilon_t \tag{6}$$

where $t=1,2,\ldots,T$

To the extent that documented in Table 2.2 and by ABDL (2003), the distribution of logarithmic realized volatility approximates to a normal distribution, in turn allowing for the use of standard normal distribution theory and related mixture models. Therefore, the logarithmic realized volatility is more suitable than the realized volatility for the prospective of modeling (ABD, 2007). The Log-HAR model can be represented as follows:

$$\log\left(RV_t\right)^{1/2}=\beta_0+\beta^{(d)}\log\left(RV_{t-1}\right)^{1/2}+\beta^{(W)}\log\left(RV_{t-5,t-1}\right)^{1/2}+\beta^{(M)}\log\left(RV_{t-22,t-1}\right)^{1/2}+\varepsilon_t \tag{7}$$

### 2.4 Option Pricing Models

The aim of this paper is to provide the empirical investigation of price forecasting ability on SPX based on two alternative option pricing model: option models with the application of realized volatility (RV-type option pricing models) and GARCH-type option pricing models. The former, emerging in the rather recent literature, is the use of realized volatility measured together with HAR-RV model; the later mainly treats GARCH-type as variance process in capturing returns process. Furthermore, the concept of realized volatility is to take volatility as observable, while GARCH-type models consider volatility as latent.

In the realm of RV-type option pricing models, Corsi et al. (2013) uses the HAR multi-component model for the conditional mean of the RV, specifies the conditional distribution of the HAR is from non-centered gamma for transiting density, and results in the model with an Auto-Regressive Gamma process, called HARG. This paper develops the model also together with RV-type option model. However we include the Log-HAR, instead of HAR, to obtain forecasted volatilities used in returns process, because it has been regarded as an improved model than the HAR due to the distribution of logarithmic realized volatility following normality. This paper labels the model as the Log-HAR option pricing model. Different from HARG, we only consider the case when the risk neutral and physical dynamics of realized volatilities are the same. That gives us a more simple way to estimate parameters used in pricing models than HARG. Although HARG expresses a solid foundation from theoretical viewpoint, we wonder whether a model with simple procedure to implement but with poor theory to back up can perform better through empirical viewpoint.

In terms of GARCH-type option models, the NGARCH has been documented its superior to the model proposed by Heston and Nandi (2000), which was widely used as the best model in pricing European options before (see Hsieh and Ritchken, 2005). It motivates us to compare the HARG, the Log-HAR, and the NGARCH option pricing models. The more detail description of key option pricing models is represented as below.

### 2.4.1 HARG Option Pricing Model

Corsi et al. (2013) propose the HARG model as the volatility process of returns. The advantage of this volatility model is: firstly, this model captures a crucial feature of return volatility, the long memory, which implies the volatility autocorrelation decreases hyperbolically; secondly, it is a discrete-time process, combined with an exponential affine stochastic discount factor, giving rise to the risk-neutral dynamics in pricing options. Within HARG, the dynamics for the log-return is described as follows:

$$\ln\left(\frac{S_{t+1}}{S_t}\right) = y_{t+1} = r + \left(\tilde{\gamma} - \frac{1}{2}\right)V_{t+1} + \sqrt{V_{t+1}}\varepsilon_{t+1} \tag{8}$$

where $\varepsilon_{t+1}|V_{t+1} \sim N(0,1)$ and $V_{t+1}$ is the corresponding variance for $y_{t+1}$.

Following Corsi (2009), the dynamics of the variance is set up as an HAR process, in which the variance is denoted as $RV_{t+1}$, measured by using the method of Two Scales Realized Volatility (see Zhang et al., 2005). For the purpose of considering another well-known feature, the asymmetric changes of volatilities with respective to different directions of stock price changes, the extended HAR model includes $L_t$ to be a factor of the leverage effect. In order to execute the risk-neutral valuation to evaluate options, they assume RV following an Auto Regressive Gamma process (see Gourieroux and Jasiak, 2006), with shape and scale parameters $\delta$ and $c$, respectively, and location parameter $\beta'(\mathbf{RV_t}, L_t)$. It denotes $RV_t = (RV_t, RV_{t-1}, \ldots, RV_{t-22})$ and $L_t = I_{(y_t<0)}RV_t$ **(**$I$ is an indicator function**)**. The formal process of realized volatility is shown in the following framework:

$$RV_{t+1}|F_t \sim \Gamma\left(\delta, \beta'(\mathbf{RV_t}, L_t), c\right) \tag{9}$$

where $\beta'(\mathbf{RV_t}, L_t) = \beta_1 RV_t + \frac{\beta_2}{4}\left(\sum_{i=1}^{4} RV_{t-i}\right) + \frac{\beta_3}{17}\left(\sum_{i=5}^{21} RV_{t-i}\right) + \beta_4 L_t$, and $L_t = I_{(y_t<0)}RV_t$

To estimate parameters under $\mathbb{P}$, the Pseudo-maximum Likelihood method based on Gaussian pseudo-family is used. Following Gourieroux and Jasiak (2006), the expectation and variance of $RV_{t+1}|RV_t$ can be represented as

$$\left(\hat{c},\delta,\beta\right)' = \arg\max_{c,\delta,\beta} \sum_{t=22}^{T} \left\{ -\frac{1}{2}\log V\left(RV_{t+1}\middle|RV_t\right) - \frac{1}{2}\frac{\left(RV_{t+1} - E\left(RV_{t+1}\middle|RV_t\right)\right)^2}{V\left(RV_{t+1}\middle|RV_t\right)} \right\} \tag{10}$$

where $E\left(RV_{t+1}\middle|RV_t\right) = c\delta + c\beta' RV_t$ and $V\left(RV_{t+1}\middle|RV_t\right) = c^2\delta + 2c^2\beta' RV_t$

Furthermore, with the combination of the affine discrete time HARG (3) model and a discrete time Stochastic Discount Factor (SDF), the dynamics system of log-return and volatilities is derived under the martingale measure $\mathbb{Q}$, and only $\nu_1$ remains a free parameter.[15] With this SDF specification, some parameters transitions are required for the change of probability measure. It includes

$$\beta^* = \beta/(1+c\lambda),\ \ \delta^* = \delta,\ \text{and}\ \ c^* = c/(1+c\lambda) \tag{11}$$

where $\lambda = \nu_1 + \gamma^2/2 - 1/8$

Finally, they stated that, under $\mathbb{Q}$ measure, the log-return follow a discrete-time stochastic volatility, with risk premium $\gamma^* = -1/2$.

Overall, they suggest that option prices can be measured by executing the following five steps: first, the estimated parameters are obtained under the physical probability $\mathbb{P}$ by applying for the Pseudo-Maximum Likelihood method. Second, $\nu_1$ is calibrated by matching the unconditional mean of at-the-money implied volatility. Third, the estimated parameters under $\mathbb{P}$ are changed in line with Equation (11). Fourth, we simulate both the volatility and log-return to obtain simulated asset prices under measure $\mathbb{Q}$ by using Monte Carlo Simulation method. Finally, options can be evaluated by putting the average of all terminal simulated assets prices (denoted as $S_T$) into the following formulas:

$$C_t = \exp\left(-r(T-t)\right)\times\max\left(S_T - K, 0\right)\ \text{and}\ \ P_t = \exp\left(-r(T-t)\right)\times\max\left(K - S_T, 0\right) \tag{12}$$

### 2.4.2 Log-HAR Option Pricing Model

We propose the Log-HAR option pricing model in the way of positing that the risk neutral and physical dynamics of realized volatilities are the same. Under the risk-neutral world, assuming that the logarithmic asset price follows a continuous-time process:

$$dp_t = \mu_t dt + \sigma_t dW_t \tag{13}$$

By Ito's lemma, the above equation can be represented

$$d\ln S_t = \left(r - \sigma_t^2/2\right)dt + \sigma_t dW_t \tag{14}$$

where $S_t$ is the underlying asset price. When considering discrete time process, the process becomes:

[15] See Appendix.

$$\ln S_{t+\Delta t} - \ln S_t = \left(r - \sigma_t^2/2\right)\Delta t + \sigma_t \varepsilon \sqrt{\Delta t} \tag{15}$$

where $r$ is risk-free interest rate, $\Delta t$ is time interval, $\varepsilon \overset{i.i.d.}{\sim} N(0,1)$, and the variance process is captured by Log-HAR as in Equation (7). In this setting, all the unknown parameters are included in the variance equation, and are estimated by implementing the Ordinary Least Square (OLS) method.[16] We thus are able to forecast the volatility for the next day at time *t*:[17]

$$\log\left(RV_{t+1,for}\right)^{1/2} = \hat{\beta}_0 + \hat{\beta}^{(d)} \log\left(RV_t\right)^{1/2} + \hat{\beta}^{(W)} \log\left(RV_{t-4,t}\right)^{1/2} + \hat{\beta}^{(M)} \log\left(RV_{t-21,t}\right)^{1/2} \tag{16}$$

where $\log\left(RV_{t-4,t}\right)^{1/2}$ and $\log\left(RV_{t-21,t}\right)^{1/2}$ are the average of the combination of past four and twenty-one realized volatilities and a new one-day forecasting realized volatility respectively. Let $\sigma_t = \left(RV_{t+1,for}\right)^{1/2}$, $\Delta t = 1$, and it generates the asset price for the next day $S_{t+1} = S_t \exp\left(r - RV_{t+1,for}/2 + \varepsilon\sqrt{RV_{t+1,for}}\right)$. Repeating this procedure until achieving the expiration date (denoted as $T$) of options, and we redo this simulation in *N* times. It results in *N* number of estimated $S_T$ for evaluating options.

By definition, the discounted average value of $Max\left\{\left(S(T) - X\right), 0\right\}$ and the discounted average value of $Max\left\{\left(X - S(T)\right), 0\right\}$ are predicted European call and put option prices, respectively.

### 2.4.3 NGARCH Option Pricing Model

We briefly review the Duan's NGARCH model. Let $S_t$ be the asset price and $h_t$ is the conditional variance of the logarithmic return. Thus, under the risk-neutral measure, we have,

$$\ln\left(\frac{S_{t+1}}{S_t}\right) = r - \frac{1}{2}h_{t+1} + \sqrt{h_{t+1}}v_{t+1} \tag{17}$$

where $v_{t+1} \sim N(0,1)$. The variance equation is

$$h_{t+1} = \beta_0 + \beta_1 h_t + \alpha_1 h_t (v_t - w) \tag{18}$$

where $w = \gamma + \lambda$. Here, $\lambda$ is the unit risk premium for the asset, and $\gamma$ is a nonnegative parameter that captures the negative correlation between return and volatility innovations.

## 2.5 Empirical Results

We test the empirical implications of HAR-type models (including HARG and our Log-HAR model) in the S&P 500 index options market. As benchmark models, we choose the NGARCH option pricing model that has been documented as the best model in pricing options among

---

[16] The OLS method is also used in Corsi (2009), ABD (2007).
[17] The word "for" means forecast.

GARCH-type models, and the BS model that has been regarded as the traditional benchmark model for many literatures. The primary goal of our paper is to compare models in terms of out-of-sample performance due to two reasons. First, it is able to provide the forecasting pricing errors, which is more important to the fitting performance. Second, the model with more parameters is likely to produce better pricing results for in-sample approach, while it may be not due to its good structure, but more parameters. The presence of more parameters may be penalized if the extra parameters do not improve its structural fit in out-of-sample forecasting.

Table 2.3 reports the out-of-sample valuation errors for the various models during the period of July 3, 2007 to December 31, 2008, which covers the most recent financial crisis. Our numerical results are demonstrated based on the root-mean-squared valuation error in dollars (RMSE henceforth). By construction, the smaller reported number in RMSE implies the better pricing performance. Table 2.3 also reports the valuation errors by different option moneyness and maturity categories for calls (Panel A) and puts (Panel B) respectively. It is found that HARG option pricing model demonstrates consistently smaller valuation errors than other three models aggregated across all moneyness, and over different maturities (short-term, mid-term, and long-term). In terms of calls (Panel A), we observe that the Log-HAR performs better in the deep-out-of-the money calls (with moneyness <0.94) over all three maturities except for the short-term contacts, and calls with moneyness (0.94–0.97) short-term maturity. In contrast, HARG option pricing model consistently dominates other models in remaining moneyness and maturity, except BS model performs better in short-term the maturity with moneyness (<0.94). In general, the HARG tends to have lower valuation errors in calls. Looking at valuation errors on puts (see Panel B), we find that the HARG performs best for all moneyness and maturities. Overall, the HAR-type models, especially the HARG model, successfully predict out-of-sample option prices probably because it is based on realized volatilities, which are closer to expected volatility (VIX) in financial markets as shows in Figure 2.1.

< Table 2.3 is inserted about here >

Furthermore, even if one certain model could outperform other models in pricing options in times of tranquility, we concern whether this superiority still hold during a very fluctuant period. In the section 3.2, we partition the full option data set into two different groups based on volatility of S&P 500 index, including (1) the rather unstable period of July 3, 2007–September 14, 2008, and (2) the very fluctuant period of September 15, 2008–December 31, 2008. Table 2.4 reports the out-of-sample RMSE for each out-of-sample period. First, during a very fluctuant period, we find that the Log-HAR model outperforms other three models in aggregated moneyness and over all three maturities. For calls (the third to fifth columns in Panel A), the Log-HAR again has the smallest forecasted error, except that the NGARCH performs better in short-term moneyness (<0.94). For puts (the third to fifth columns in Panel B), we also observe that the Log-HAR model outperforms other models in mid-term and long-term contracts except for deep-out-of-money (>1.06) puts, where HARG has the smallest valuation errors. Furthermore, HARG model dominates other models in the short-term puts. Overall, in times of turmoil, the HAR-type models have smaller valuation errors than our two benchmarks models (the NGARCH and the BS model), and the Log-HAR performs better than the HARG model in mid-term and long-term contracts, while worse than the HARG model in short-term puts. As a result, it seems to support that the model constructed based on realized volatilities and with simpler framework could value option prices more accurately in the fluctuant period as contracts with longer than 46 days to expiration.

We now move on to analyze our empirical results in the rather unstable period compared to our very fluctuant period (see the sixth column to eighth column in Table 2.4). The HARG model exhibits smaller valuation errors for five moneyness categories in terms of call and put option, except in the deep-out-of-the-money calls (with moneyness <0.94). The Log-HAR, BS, and NGARCH model outperforms than HARG model only in moneyness (<0.94). Overall, the HAR-type models still dominates our benchmark models and especially the HARG performs better than the Log-HAR model during the rather unstable period. It again holds up the superiority of the HAR-type models than other models, but it exists mixed results between the Log-HAR and the HARG models during very turmoil or rather unstable periods.

< Table 2.4 is inserted about here >

## 2.6 Conclusion

In sum, the contribution of this paper is in three-folds. First, we propose Log-HAR option pricing model, which is more convenient compared to other option pricing models associated with realized volatility in the way of simpler estimation procedure. Second, this paper compares the empirical results of the out-of-sample valuation errors on four models: (1) our Log-HAR option pricing model, (2) the HARG proposed by Corsi et al. (2013), (3) the NGARCH option pricing model that has been documented as the best model in pricing options among GARCH-type models, and (4) the Black-Scholes model that has been regarded as the traditional benchmark model for many literatures. Third, as far as we know, no existing literature provides empirical investigation on S&P 500 index options during the recent financial crisis (from July 3, 2007 to December 31, 2008). We fill this gap since it is important to have knowledge which option pricing model can forecast option prices more accurately since options are widely used as financial instruments to hedge in the times of turmoil.

Overall, we find that the HAR-type models successfully predict out-of-sample option prices probably because they are based on realized volatilities, which are closer to expected volatility (VIX) in financial markets. However, it seems to exist the mixed result between the Log-HAR and the HARG models in pricing options since the Log-HAR is better than the HARG in times of turmoil, while it is worse during the rather unstable period.

For further researches, Value-at-Risk (VaR) is a standard risk measure linked to holding a portfolio and is often used to be criterion in determining the capital requirement for financial institutions. Therefore, one could extend to compare the prediction of VaR on options by using the model proposed in this paper.

**Appendix**

Assuming $r_t = 0$ (for computational convenience), the SDF satisfies the following framework:

$$M_{t,t+1} = \exp\left(-\nu_0 - \nu_1 RV_{t+1} - \nu_2 \left(RV\right)_t - \nu_3 \left(\sum_{i=1}^{4} RV_{t-i}\right) - \nu_4 \left(\sum_{i=1}^{21} RV_{t-i}\right) - \nu_5 y_{t+1} - \nu_6 L_t\right)$$

(A1)

It complies with the no arbitrage conditions if the following implicit parameter-restrictions are satisfied:

$$\begin{aligned}
\nu_0 &= -\delta \log\left(1 + c\left(\frac{1}{2}\gamma^2 + \nu_1 - \frac{1}{8}\right)\right) \\
\nu_2 &= -c\beta_1 \frac{\frac{1}{2}\gamma^2 - \frac{1}{8} + \nu_1}{c\left(\gamma^2/2 - 1/8 - \nu_1\right) + 1} \\
\nu_3 &= -c\beta_2 \frac{\frac{1}{2}\gamma^2 - \frac{1}{8} + \nu_1}{c\left(\gamma^2/2 - 1/8 - \nu_1\right) + 1} \\
\nu_4 &= -c\beta_3 \frac{\frac{1}{2}\gamma^2 - \frac{1}{8} + \nu_1}{c\left(\gamma^2/2 - 1/8 - \nu_1\right) + 1} \\
\nu_5 &= \gamma + \frac{1}{2} \\
\nu_6 &= -c\beta_4 \frac{\frac{1}{2}\gamma^2 - \frac{1}{8} + \nu_1}{c\left(\gamma^2/2 - 1/8 - \nu_1\right) + 1}
\end{aligned} \tag{A2}$$

where $\nu_1$ remains a free parameter.

**Table 2.1 The summary statistics of SPX during the second-half of 2007 and 2008**

The table reports the summary statistics of out-of-the-money S&P500 index option from July 03, 2007 to December 31, 2008, total 12,408 number of observations. All option are categorized into 18 groups in line with Moneyness (S/K, the index level divided by the strike price) and Days-to-Maturity, where the former is separated into six groups, including: <0.94, 0.94–0.97, 0.97–1.00, 1.00–1.03, 1.03–1.06, and >1.06, and the later is divided into three different period, short-term (10-45 days), medium-term (46-90 days), and long-term (>90 days). The amount in parentheses stands for standard deviation and the number of observation is reported in the braces.

| Moneyness | Days-to-Maturity | | | |
|---|---|---|---|---|
| S/K | 10 to 45 | 46 to 90 | 91-180 | Subtotal |
| Panel A: out-of-the-money call options (6690 observations) | | | | |
| <0.94 | $4.39 | $7.28 | $18.14 | |
| | (5.78) | (8.62) | (15.37) | |
| | {905} | {1247} | {656} | {2808} |
| 0.94–0.97 | $7.49 | $23.79 | $43.72 | |
| | (7.33) | (13.58) | (13.66) | |
| | {1089} | {499} | {227} | {1815} |
| 0.97–1.00 | $21.12 | $47.40 | $66.49 | |
| | (13.06) | (12.98) | (12.97) | |
| | {932} | {766} | {369} | {2067} |
| Panel B: out-of-the-money put options (5718 observations) | | | | |
| 1.00–1.03 | $23.70 | $41.94 | $60.65 | |
| | (11.96) | (12.82) | (13.67) | |
| | {802} | {652} | {327} | {1781} |
| 1.03–1.06 | $13.05 | $30.26 | $45.11 | |
| | (9.55) | (11.69) | (11.06) | |
| | {609} | {363} | {204} | {1176} |
| >1.06 | $5.56 | $12.47 | $22.78 | |
| | (6.13) | (10.65) | (15.61) | |
| | {1326} | {958} | {477} | {2761} |
| Subtotal | {5663} | {4485} | {2260} | {12408} |

**Table 2.2 The unconditional daily logarithmic realized volatilities and standardized returns**

This table shows that the summary statistics of the logarithmic realized volatility and daily standardized return $r_t / \upsilon_t$ from July 2, 2003 to Dec. 31, 2008. The realized volatilities are calculated from five-minute intraday returns of S&P500 index.

| | Series: Logarithmic Realized Volatility<br>Sample: 7/02/2003–12/31/2008<br>Observation: 1375 | Series: Standardized Returns<br>Sample: 7/02/2003–12/31/2008<br>Observation: 1374 |
|---|---|---|
| Mean | -5.0609 | 0.1095 |
| Median | -5.1696 | 0.1403 |
| Maximum | -2.5764 | 4.5853 |
| Minimum | -6.3907 | -4.0787 |
| Std. Dev. | 0.5279 | 1.2704 |
| Skewness | 1.3293 | 0.1101 |
| Kurtosis | 5.5437 | 3.0144 |

**Table 2.3 Out-of-sample valuation errors for out-of-the-money SPX**

Option data is from July 3, 2007 to December 31, 2008, total 12,408 contracts and 379 trading days. The valuation error is calculated by comparing the market and forecasted prices. The reported numbers are the root-mean-squared valuation error in dollars (RMSE). Panel A and B report valuation errors in out-of-the-money call and put options respectively.

| Moneyness | Model | Maturity 10–45 | Maturity 46–90 | Maturity 91–180 |
|---|---|---|---|---|
| All | Log-HAR | 7.62 | 13.69 | 22.11 |
| | HARG | 6.68 | 12.56 | 20.63 |
| | NGARCH | 9.48 | 15.51 | 21.24 |
| | BS | 9.41 | 15.42 | 21.57 |
| | | Panel A: out-of-the-money call options | | |
| <0.94 | Log-HAR | 6.17 | 6.02 | 13.00 |
| | HARG | 8.77 | 13.54 | 25.63 |
| | NGARCH | 6.06 | 7.83 | 13.75 |
| | BS | 6.05 | 7.74 | 14.00 |
| 0.94–0.97 | Log-HAR | 5.24 | 12.63 | 22.1 |
| | HARG | 5.41 | 12.15 | 16.67 |
| | NGARCH | 6.43 | 14.92 | 18.09 |
| | BS | 6.21 | 14.57 | 19.44 |
| 0.97–1.00 | Log-HAR | 8.57 | 18.02 | 27.10 |
| | HARG | 7.05 | 13.76 | 18.19 |
| | NGARCH | 11.32 | 19.12 | 23.41 |
| | BS | 11.06 | 19.46 | 25.73 |
| | | Panel B: out-of-the-money put options | | |
| 1.00–1.03 | Log-HAR | 11.19 | 18.69 | 28.32 |
| | HARG | 8.40 | 14.36 | 21.37 |
| | NGARCH | 14.00 | 21.29 | 29.15 |
| | BS | 13.74 | 20.38 | 27.47 |
| 1.03–1.06 | Log-HAR | 9.46 | 18.24 | 28.74 |
| | HARG | 6.77 | 13.04 | 21.63 |
| | NGARCH | 11.67 | 20.24 | 25.77 |
| | BS | 11.96 | 20.41 | 25.62 |
| >1.06 | Log-HAR | 5.57 | 11.12 | 19.51 |
| | HARG | 3.91 | 8.18 | 14.55 |
| | NGARCH | 7.33 | 13.23 | 20.78 |
| | BS | 7.45 | 13.48 | 21.09 |

**Table 2.4 Out-of-sample valuation errors for out-of-the-money SPX in times of turmoil**

Option data of the fluctuant period is from Sep. 15, 2008 to Dec. 31, 2008, total 2121 contracts and 76 trading days (the third to the fifth columns). Option data within the rather unstable period is from July 3, 2007 to Sep. 12, 2008, total 10,287 contracts and 303 trading days (the sixth to the eighth columns). The valuation error is calculated by comparing the market and forecasted prices. The reported numbers are the root-mean-squared valuation error in dollars (RMSE). Panel A and B report valuation errors in out-of-the-money call and put options respectively.

| | | Data Period | | | | | |
|---|---|---|---|---|---|---|---|
| | | Sep. 15, 2008 to Dec. 31, 2008 (fluctuant period) | | | July 3, 2007 to Sep. 12, 2008 (rather unstable period) | | |
| | | Maturity | | | Maturity | | |
| Moneyness | Model | 10–45 | 46–90 | 91–180 | 10–45 | 46–90 | 91–180 |
| All | Log-HAR | 11.53 | 15.72 | 20.09 | 6.70 | 13.22 | 22.64 |
| | HARG | 12.50 | 24.11 | 36.76 | 4.99 | 8.27 | 12.99 |
| | NGARCH | 16.79 | 25.53 | 29.43 | 7.48 | 12.41 | 18.33 |
| | BS | 17.00 | 25.50 | 29.83 | 7.29 | 12.29 | 18.64 |
| | | Panel A: out-of-the-money call options | | | | | |
| <0.94 | Log-HAR | 9.78 | 10.29 | 16.26 | 1.84 | 3.76 | 11.57 |
| | HARG | 13.73 | 26.03 | 46.87 | 3.12 | 5.16 | 9.38 |
| | NGARCH | 9.54 | 14.57 | 21.85 | 2.00 | 3.68 | 9.08 |
| | BS | 9.57 | 14.61 | 22.27 | 1.85 | 3.35 | 9.22 |
| 0.94–0.97 | Log-HAR | 13.72 | 22.47 | 25.18 | 3.97 | 10.77 | 21.76 |
| | HARG | 15.22 | 28.73 | 38.32 | 3.81 | 7.87 | 12.51 |
| | NGARCH | 17.88 | 34.28 | 35.07 | 4.59 | 10.11 | 15.34 |
| | BS | 17.50 | 33.65 | 35.22 | 4.36 | 9.8 | 17.02 |
| 0.97–1.00 | Log-HAR | 14.63 | 23.31 | 23.61 | 7.95 | 17.24 | 27.75 |
| | HARG | 15.84 | 28.22 | 29.57 | 5.90 | 10.62 | 14.86 |
| | NGARCH | 27.07 | 38.54 | 36.78 | 9.14 | 14.96 | 19.64 |
| | BS | 26.90 | 38.20 | 38.97 | 8.83 | 15.56 | 22.11 |
| | | Panel B: out-of-the-money put options | | | | | |
| 1.00–1.03 | Log-HAR | 16.97 | 23.13 | 24.72 | 10.35 | 17.98 | 29.03 |
| | HARG | 15.09 | 31.15 | 37.88 | 7.29 | 9.97 | 15.82 |
| | NGARCH | 27.54 | 37.62 | 43.18 | 11.57 | 17.85 | 25.21 |
| | BS | 27.70 | 36.91 | 41.18 | 11.17 | 16.82 | 23.58 |
| 1.03–1.06 | Log-HAR | 17.07 | 17.39 | 30.66 | 8.46 | 18.34 | 28.55 |
| | HARG | 13.77 | 26.72 | 51.06 | 5.74 | 10.32 | 16.15 |
| | NGARCH | 26.22 | 36.53 | 37.17 | 9.33 | 17.36 | 24.39 |
| | BS | 27.95 | 37.74 | 38.99 | 9.27 | 17.29 | 23.93 |
| >1.06 | Log-HAR | 7.94 | 12.02 | 18.22 | 4.88 | 10.84 | 20.11 |
| | HARG | 5.66 | 11.90 | 19.77 | 3.40 | 6.68 | 11.16 |
| | NGARCH | 13.34 | 19.96 | 25.33 | 5.10 | 10.4 | 18.17 |
| | BS | 13.62 | 20.37 | 25.91 | 5.15 | 10.57 | 18.30 |

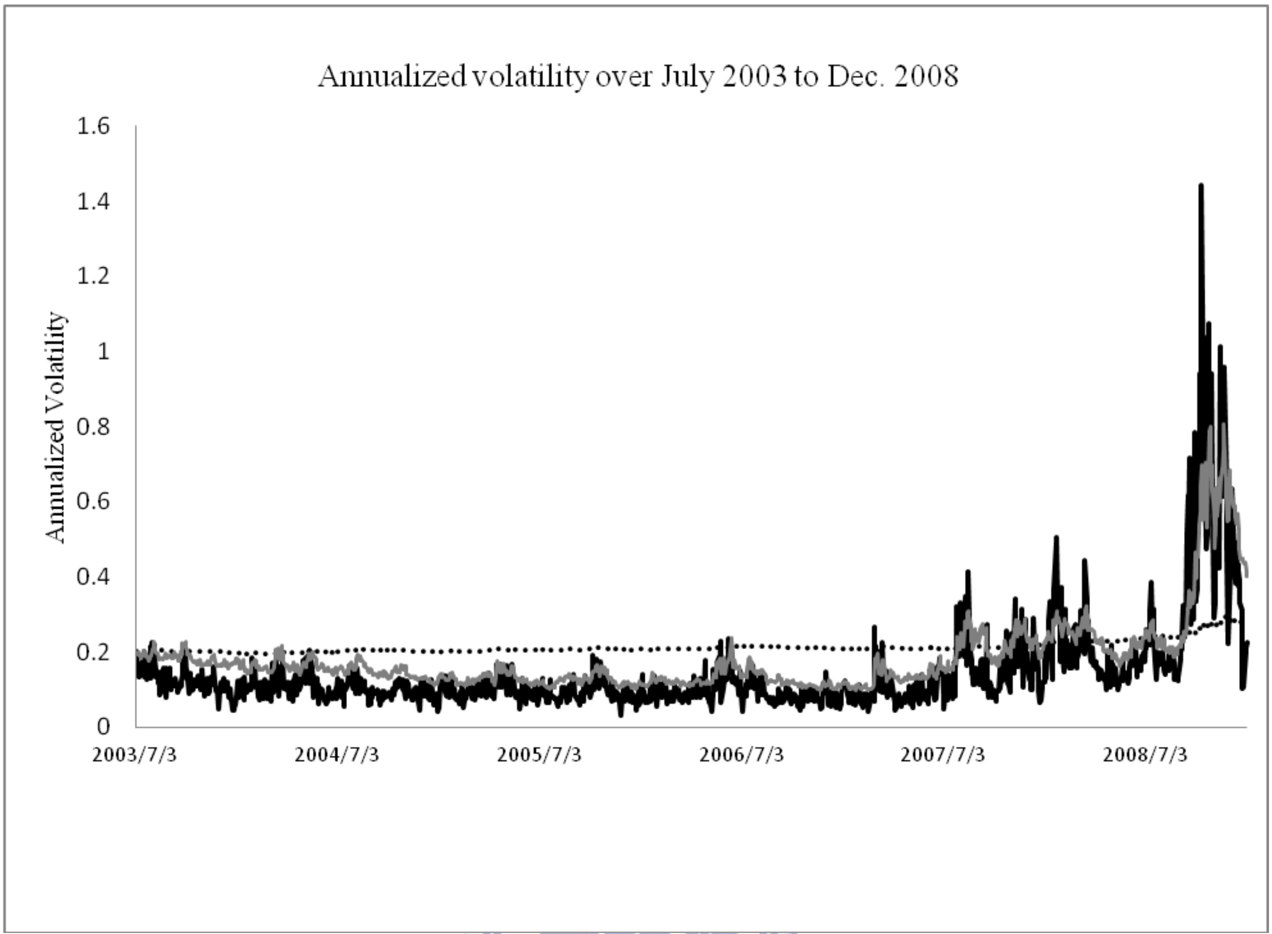
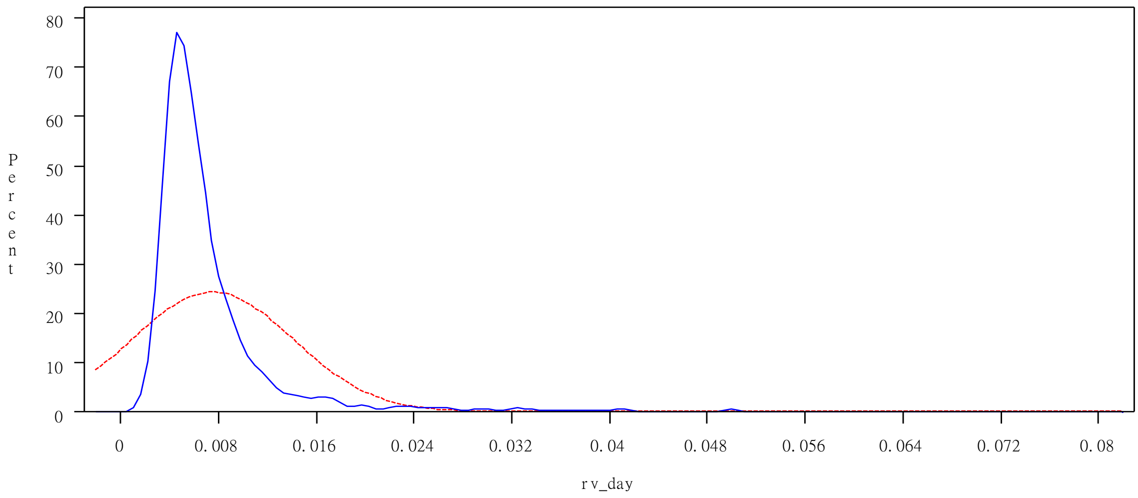
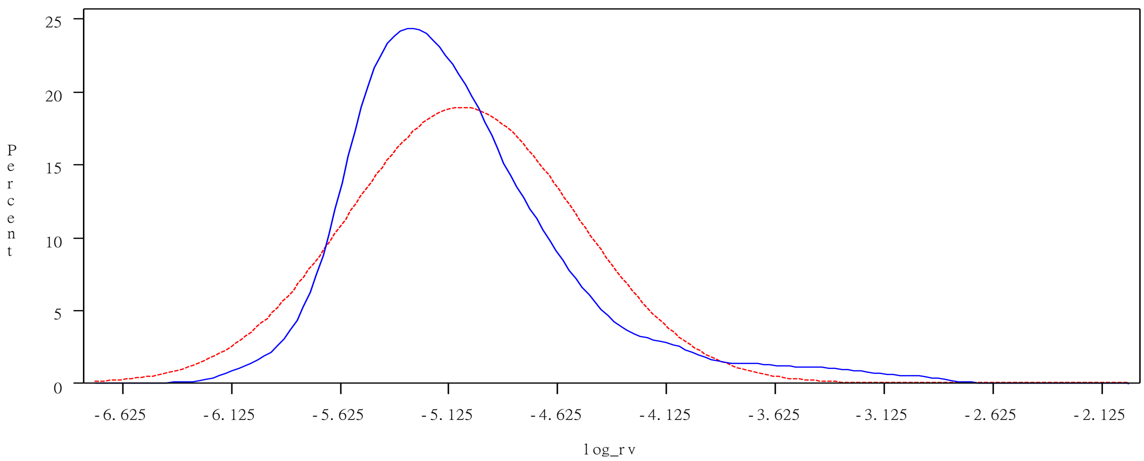


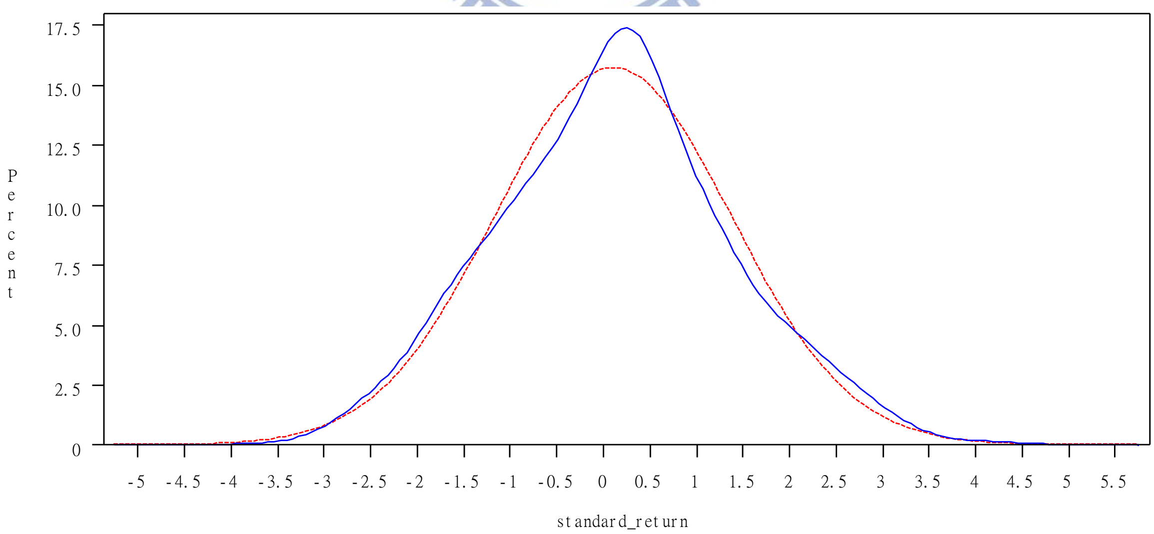

**Figure 2.1 The annualized volatility over July 2003 to Dec. 2008**
This graph plots volatilities obtained from realized volatility measure, NGARCH model, and VIX in the period of July 2003–December 2008. The solid and black line stands for realized volatility, the solid and grey line stands for VIX, and the dotted line stands for NGARCH volatility.




Panel A: The unconditional distribution of daily realized volatilities




Panel B: The unconditional distribution of daily logarithmic realized volatilities




Panel C: The unconditional distribution of daily standardized returns

**Figure 2.2 The unconditional distribution of daily RV, logarithmic RV, and standardized return**

The realized volatilities are calculated from five-minute intraday returns of S&P500 index, from July 2, 2003 to December 31, 2008, for a total of 3146 observations. The dotted line refers to the normal density. The top graph demonstrates that the unconditional distribution of realized volatilities, the middle graph shows the unconditional distribution of daily logarithmic realized volatility, and the bottom graph stands for standardized daily returns, $r_t / \upsilon_t$, where $\upsilon_t$ is the realized volatility at time *t*.

# Chapter3. Have domestic institutional investors become as market savvy as foreign investors? Evidence from the Taiwan options market

## 3.1 Introduction

The impact of foreign investors trading in host countries has recently drawn considerable attention from academics and practitioners alike. Griffin et al. (2004) and Richards (2005) described the relationship between foreign capital flows and stock returns in emerging markets, with both finding that foreign capital flows have a positive correlation with host-country stock returns. Richards (2005) argued that informed traders may benefit from access to capital flow information since foreign capital flows are so large (as a proportion of the host market's capitalization) in emerging markets that they can create price pressure.[18] Chang et al. (2009) examined the relationship between the information content of foreign investors' options trading and the host-country returns by using data of the Taiwan market. Trading in options instead of equities provides many advantages due to higher leverage, lower transaction cost, and even volatility trading. Their results indicate that only foreign institutional investors have significant power to predict returns of the Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX).

Chang et al. (2009) investigated the predictive power of the put and call positions of different types of traders in Taiwan for the period December 2001 to December 2005, whereas we choose the more recent period, from 2 January 2007 to 31 December 2008, which covers the period when markets around the globe plunged following one of the most shocking and largest bankruptcy in the history of world markets in September 2008 – that of Lehman Brothers. Most countries around the globe suffered a serious economic downturn in late 2008, during the worldwide financial tsunami. Figure 3.1 depicts movement of the TAIEX index, as well as trading volumes of four main types of investor classes, during the period 2007 to 2008.

< Figure 3.1 is inserted about here >

Data on the TAIEX options from the Taiwan Futures Exchange (TAIFEX) are quite unique since they comprise detailed tick-by-tick transaction records.[19] Unlike datasets of previous empirical studies, detailed information of transactions by types of traders (trader identification code), as well as trading directions (buy/sell), are also available in our dataset. The distinguished

[18] For example, Chang et al. (2009) indicated that a foreign trader who knows that a large mutual fund tracking the emerging markets will increase the weightings of some specific countries, will take advantage of this information to build long call positions. Conversely, the informed foreign trader will build long put positions when knowing that the weightings of some countries will decrease.

[19] The TAIEX is an order-driven market, and TAIEX options (TXO) are European style and cash settled. The expiration months of the contracts include spot month, the next two calendar months, and the next two quarterly months in a March cycle. Trading hours are from 08:45a.m. to 1:45p.m., Monday through Friday on regular Taiwan Stock Exchange business days, and 08:45a.m. to 1:30p.m. on the last trading day for the delivery month contract.

features of the dataset enable us to analyze the influence of options volumes on future index returns realized by different investor classes. Furthermore, since intraday trading volumes of different classes of investors are also observed, the unique characteristics of the TAIEX options transaction dataset allow us to explore the relationships between the trading of different investor groups.

TAIEX options in 2008 were ranked the fifth most frequently traded index options in the world, according to the annual Futures Industry Association's survey (Burghardt and Acworth, 2009). The structure of options trades in Taiwan is different in that while there is high institutional investor participation in the U.S. market, in Taiwan over 40% of trades in derivatives and stocks are done by individual investors. Although there has been a downward trend in Taiwanese individual participation over the years, individual investors still play an important role in both derivatives and stock transactions. Figure 3.1 shows that individual participation continues to dominate domestic and foreign institutional investors in terms of trading volumes in our data period, from 2007/1/2 to 2008/12/31. Note that even during the global financial crisis in late 2008, individual participation remained high, in contrast to the trading volumes of market makers. This indicates that characteristics of Taiwan's options market are very different from the U.S. index options market.

Foreign investors are generally believed to possess superior private information to domestic traders in an emerging market. Barber et al. (2009) documented that domestic individual investors suffer systemic and economically large losses in the Taiwan stock market while foreign institutional investors are the main winners.[20] Therefore, experienced domestic investors should recognize this fact and closely watch trades of foreign institutions. In the Taiwan stock market, most investors, including individual investors, have acknowledged that foreign investors usually have superior information on the domestic market. Even if this is not true, the net buys of foreign investors are often large enough to cause price pressure on the Taiwan stock market. Therefore, transactions of foreign investors have been closely watched by most stock analysts and media in Taiwan. Furthermore, the domestic exchanges also provide market information regarding trading activities of foreign investors. The Taiwan Stock Exchange (TWSE) realizes the importance of foreign investors' trades and every day after the market reports market information on "Total Buy", "Total Sell", and the net difference between buys and sells – "Net Buy". The TAIFEX reports the long, short, and net amount of trades by foreign institutional investors (FINIs).

The contributions of this paper are two-fold. First, we use the more recent data on TAIEX options, including the period before and during the 2008 global financial tsunami, to re-examine the predictive ability of private information from different classes of investors for future index returns.

[20] Barber et al. (2009) demonstrated that the aggregate portfolio of individual investors suffers losses of 3.8 percentage points annually, which are equivalent to 2.2% of Taiwan's gross domestic product or 2.8% of total personal income. By contrast, institutions enjoy an annual performance boost of 1.5 percentage points, and foreign institutions reap nearly half of these institutional profits.

This is important, because over the years domestic investors have acknowledged that foreign investors may possess superior information and foreign capital flows can create price pressure for the host-country stock market. The discrepancy between our analysis and those in the past is worth exploring, because similar patterns can also appear in other emerging markets with high individual participation, such as China. Second, we conduct daily and intraday VAR (Vector Autoregression) analyses to examine the relationship between options trading activities of various classes of investors and their relation to TAIEX returns. We are quite thankful for the unique characteristics of our dataset from the Taiwan Futures Exchange. To the best of our knowledge, the VAR analysis of daily and intraday trading activities for different classes of investors is the first empirical study of its kind. Moreover, the 2008 global financial crisis provides a testing ground to investigate investors' trading behaviors during such a rarely seen economic condition. Thus, we conduct analyses by further dividing our data into two sub-periods: the pre-crisis period (2007/01~2008/06) and the period of the global financial crisis (2008/07~2008/12).

Our empirical evidence from daily data frequency suggests that unlike the findings of Chang et al. (2009), the put-call ratios of foreign institutional investors only have marginally significant predictive power for next-day TAIEX returns prior to the 2008 global financial tsunami and do not have any useful information content for future index returns during 2008 financial crisis. In contrast, domestic institutional investors have stronger prediction capability for next-day TAIEX returns before the global financial crisis. These results are robust even after controlling for some variables related to TAIEX returns.

The intraday VAR analyses reveal that the put-call ratios of domestic institutional and foreign institutional investors are positively correlated during the pre-crisis period, suggesting that these two groups of investors adopt similar options trading strategies. During the 2008 global financial crisis, no trader type existed that was able to predict TAIEX returns. The TAIEX returns lead the put-call ratios of domestic institutional and foreign institutional investors for 30 to 60 minutes, indicating that these two classes of traders closely watch and react to the changes in the market index. Finally, individual investors tend to adopt contrarian option strategies and do not exhibit any predictive ability for future index returns. Overall, it is apparent that domestic institutional investors are becoming more market savvy in predicting future index returns during the more recent period of 2007 to 2008. Our empirical results suggest that domestic institutional investors are becoming increasingly better informed as the host-country security market matures over time.

### 3.2 Data and Descriptive Statistics

Our TAIEX options dataset consists of detailed tick-by-tick transaction records. The advantage of tick-by-tick data is that we can retrieve more information from every transaction. In addition to the general information on options contracts, such as strike price and time-to-maturity, information of the trader type identification codes and trading directions (buy/sell) is also available. This enables us to classify transaction records into four main classes of investors: domestic institutional investors, foreign institutional investors, individual investors, and market makers.[21] Furthermore, trading directions in our dataset are known for certain, and we do not have to speculate on the signs by using the Lee and Ready (1991) algorithm. These unique features differentiate our analysis from studies of the U.S. or European options markets.

Some prior studies have used these special characteristics of the TAIEX options dataset to provide interesting and unique perspectives. Chang et al. (2009) examined the predictive power of the call and put positions of different traders in the TAIEX options market and found that only foreign institutional investors can predict future index returns. Chang, Hsieh and Wang (2010) investigated whether volatility information exists in the Taiwan options market, noting that different categories of traders use different trading strategies to realize their volatility information. Shiu et al. (2010) looked at if net buying pressure affects the implied volatility function of TAIEX options in a market with high individual participation, finding that implied volatility changes of TAIEX options are dominated by buying pressure for index calls, rather than index puts, which is the case for S&P 500 options.

Following Pan and Poteshman (2006), we further classify our data into four trade types: open-buy, open-sell, close-buy, and close-sell positions. This classification also demonstrates the advantages of our unique dataset over previous studies. We summarize the descriptive statistics in Table 3.1. Table 3.1 reports trading volumes by different classes of positions: open-buy, open-sell, close-buy, and close-sell. The average trading volume is calculated as the average of the daily volume during the sample period from 2007 to 2008. One can easily observe that individual investors account for around one half of all kinds of contracts during the sample period. Since market makers are often regarded as liquidity providers, if they are excluded, then individual participation makes up for almost 70% of the volume in the TAIEX options market. In terms of average trades, open-buy-call positions are much greater than other positions as documented by the prior studies of Chan et al. (2009) and Chang et al. (2009). Note that this is very different from the market structure of the U.S. where open-buy call positions are the smallest portion of open-buy and open-sell positions taken together (Pan and Poteshman, 2006).

< Table 3.1 is inserted about here >

### 3.3 Empirical Methodology: Testing the Predictability of Private Information on Index Returns and the Lead-lag Relationship for Different Trader Types

[21] Transactions by foreign individual investors are rare in our dataset, and thus foreign individual investor is not classified as a major investor class in this paper.

This paper first investigates whether different categories of investors can have predictive power during the data period. Instead of the information-based model by Easley et al. (1998) and Chang et al. (2009), we adopt the multivariate VAR (Vector autoregression) model by Hasbrouck (1991) and extend the VAR model to multiple classes of investors so that we can explore the predictability of the information content of the options market on stock markets as well as the relationship among options trading of different classes of investors.

$$\begin{aligned} R_t &= a_1 R_{t-1} + \cdots + a_p R_{t-p} + b_1 X_{t-1} + \cdots + b_p X_{t-p} + \varepsilon_{1,t} \\ X_t &= c_1 R_{t-1} + \cdots + c_p R_{t-p} + d_1 X_{t-1} + \cdots + d_p X_{t-p} + \varepsilon_{2,t} \end{aligned} \tag{1}$$

where $X_t = [X_{1t}, X_{2t}, \cdots, X_{It}]$. Here, $R_t$ represents TAIEX index returns, and $X_{it}$ is the information content variables for the *ith* investor class.

We follow Pan and Poteshman (2006) and use the put-call ratio to measure the information content. The put-call ratio is defined as:

$$X_{it} = P_{it} / (P_{it} + C_{it})$$

where $P_{it}$ and $C_{it}$ are the trading volumes of put and call options contracts, respectively.[22] In our paper we mainly use open-buy put and open-buy call options to examine whether investors possess superior private information. As argued by previous studies, close-buy positions have complex trading purposes and may not be suitable for clear interpretation of the results. Lakonishok et al. (2004) pointed out that many of the sell-call positions are actually part of covered call strategies. As opposed to the intuitive interpretation that traders expect a bull market in the future, a covered call is induced by relatively much more complicated information.

We next add some control variables to examine if these groups of investors still have superior prediction ability for next-day index returns. The regression equation is:

$$R_{t+1} = \alpha_i + \beta_i X_{i,t} + \gamma_i control_{i,t} + \varepsilon_{i,t+1} \tag{2}$$

where $R_{t+1}$ denotes next-day TAIEX return, $X_{i,t}$ denotes the current (or lagged) open-buy put-call ratio of a class of investors, and $control_{i,t}$ are the control variables that may affect TAIEX index returns.

We choose seven control variables that may influence next-day index returns. The first control variable is an interaction term of a dummy variable (*D_*$_{Mat}$), which is one if the next trading date is the maturity date, and the put-call ratio calculated by options matured in the next day. The second is logarithmic daily trading volume (*ln_Vol*), serving as the liquidity control variable. The

[22] The put/call ratio ignores the actual quantities traded. Thus, we also use net put minus call volume as the measure of information content. The results are qualitatively similar, but with weaker statistical significance.

third is the past one-day index return ($R_{-1}$) to control for the reversal effect of Lo and Mackinlay (1990). These three control variables are from Chang et al. (2009). The fourth variable is the one-day lagged Hang Seng index return ($R_{-1}^{HSI}$), which is the most influential stock market in the Greater China area and has been recognized as being highly correlated with TAIEX returns.

We also add three other variables that may be associated with capital flows, including the level of the 3-Month U.S. Treasury bill rate ($r$), the percentage change in the TWD/USD exchange rate ($\Delta ex$), and a proxy variable for the capital flows (*Sln_CF*) incorporating the natural logarithm of the daily foreign exchange trading volume between New Taiwan dollar (TWD) and US dollar (USD) (*ln_VolFX*).[23] If $\Delta ex$ is positive, then *Sln_CF* is equal to negative one times *ln_VolFX*; otherwise, *Sln_CF* is set to *ln_VolFX*. By construction, *Sln_CF* represents capital flows to the host country, since a positive *Sln_CF* represents capital inflows while a negative *Sln_CF* represents capital outflows.

In addition to the daily VAR analyses, in this paper we also conduct intraday VAR analyses to investigate the intraday predictability of private information on index returns as well as the relationship between trading of different classes of investors. In the spirit of Copeland (1976), we choose the put-call ratio constituted by trading volumes during the specific period of time as a measure of the amount of information that flows into the market. Many studies in the literature have already shown that a lead-lag relationship exists between stocks and stock options. Manaster & Rendleman (1982), Bhattacharya (1987), and Anthony (1988) found that stock options trading leads stock trading. Stephan and Whaley (1990) and Chan, Chung and Johnson (1993) noted the opposite results - that is, stocks lead stock options. To the best of our knowledge, intraday market lead-lag analyses have rarely been conducted. A relatively close but somewhat different intraday market analysis has been done by Kurov and Lasser (2004) who investigated the price dynamics for trades initiated by exchange locals and off-exchange customers.

The intraday VAR analyses are similar to the daily counterparts in Equation (1), except that we use the 30-minute index return and put-call ratio instead of daily frequency. The TAIEX options market opens at 8:45a.m and closes at 1:45p.m. Our analysis is limited to regular trading hours, and we pool trading activity data series for each 30-minute time interval.[24] To examine whether intra-market trading activities of institutional investors and individual investor are related, one needs to test if there exists any significant intraday lead-lag relationship between the put-call ratios of different classes of investors.

---

[23] The 3-Month Treasury bill rate and the exchange rate are from the Federal Reserve Bank of St. Louis website. Daily foreign exchange trading volume between TWD and USD is computed as the sum of the trading volumes of Taipei Foreign Exchange and YuanTai Corporation. These two companies are the only two chartered foreign exchange dealers in Taiwan. The data of trading volume are from Taiwan Economic Journal (TEJ).

[24] Following the literature, we eliminate the overnight index returns in intraday analyses.

### 3.4 Empirical Results

### 3.4.1 Prediction Ability and Options Trading of Different Investor Classes – Daily VAR Analyses

To investigate whether different trader types have different predicting powers on index returns, we consider three classes of investors, including domestic institutional, foreign institutional, and individual investors (excluding market makers since they are often regarded as liquidity providers), and use the open-buy put-call ratios of each group as the information content variables in Equation (1). Here, VAR (2) is our best model based on the selection criteria, AIC and SBC.

In Panel A of Table 3.2, we find that only the lagged two-day put-call ratio of domestic institutional investors *PC_r*(*d*) has predictive power for the current index return in our full sample period. It appears that foreign institutional and individual investors have no predictive ability for next-day index returns. These results are distinctively different from the findings of Chang et al. (2009), who suggested that only foreign institutional investors have private information and are able to reap profits in the Taiwan market. Moreover, we observe that the put-call ratio of individual investors, *PC_r*(*i*), has a significantly positive relationship with the past-day TIAEX return and has a slightly negatively significant relationship with the lagged 1-day put-call ratio of domestic institutional investors. It is apparent that individuals employ a contrarian strategy for the market, i.e., buying puts in a bull market and buying calls in a bear market, and individual investors adopt an opposite strategy to those of the domestic institutional investors on the previous day.

< Table 3.2 is inserted about here >

Our finding is of interest since our sample covers the recent global financial crisis in 2008. To further analyze the significance and magnitude of prediction abilities among the three kinds of investors, we divide our data into two periods: the pre-crisis period (2007/01/02~2008/06/30) and the period of the global financial crisis (2008/07/01~2008/12/31). The subprime mortgage crisis broke out in July 2007 (CSI: Credit Crunch; 2007), but the severe decline in equity markets did not occur until August 2008. In the U.S., the daily closing of the Dow Jones Industrial Average dropped from 11,517 at the beginning of September 2008 to 7,552 by mid-October (approximately 34%). Taiwan was no exception as credit tightened and international trade declined. The TAIEX plunged from 6,813 at the beginning of September to 4,090 by mid-October (approximately 40%). We set the sub-period from July 2008 to December 2008 in this paper to investigate whether the global financial crisis may have affected options trading of different investor classes.

Before reporting sub-period VAR results, we provide some descriptive statistics of the trading volumes of domestic and foreign institutional investors. Figure 3.2 (open-buy) and Figure 3.3 (open-sell) report the time charts of the trading volumes of domestic and foreign institutional investors, as well as the TAIEX index levels. One can observe from these figures that foreign institutional investors' trading volumes are approximately 4 to 5 times the volumes of domestic institutional investors in the case of open-buy and 2 to 3 times in the case of open-sell positions. In addition, both foreign and domestic institutional investors buy more puts than calls while they sell more calls than puts. In contrast, from Figure 3.4 it is apparent that individual investors buy more calls than puts, even during the global financial crisis. This is in line with the results in Panel A of Table 3.2, whereby individuals in general employ different options trading strategies from those of institutional investors.

< Figure 3.2 is inserted about here >

< Figure 3.3 is inserted about here >

< Figure 3.4 is inserted about here >

In Panel B of Table 3.2, not only does the lagged 2-day put-call ratio of domestic institutional investors still have explanatory power for the current index return, but the lagged 1-day put-call ratio of foreign institutional investors can predict index returns. The result of the pre-crisis period (2007/01/02~2008/06/30) regarding the predictive power of foreign institutional investors on next-day index returns is in line with Chang et al. (2009). We also find a significantly positive relationship between the lagged 1-day put-call ratio of domestic institutional investors and current put-call ratio of the foreign institutional investors, which may imply that domestic institutional investors lead transactions made by foreign investors. The remaining results in Panel B are similar to those of the full sample period.

It appears that all three classes of investors have no forecasting ability for next-day index returns during the period of the global financial crisis (2008/07/01~2008/12/31), as shown in Panel C of Table 3.2. These results are not surprising since the stock market was extremely volatile and unpredictable during the financial crisis. The results here motivate us to conjecture whether foreign institutional investors might not be capable of adopting appropriate trading strategies just like other classes of investors during the global financial crisis period. In order to examine our conjecture, we later perform intraday VAR analyses to closely investigate the relationship of options trading among these three classes of investors and TAIEX returns.

We also conduct VAR analyses for other types of option trades across three different types of investors and different time periods. To conserve space, we do not provide VAR(2) results of open-sell, close-buy, and close-sell put-call ratios since there is virtually no predictive ability in the information content of these put-call ratios on daily TAIEX returns.[25] Our empirical results are

[25] The results for the other types of option trades are available upon request.

consistent with the previous empirical findings such as Pan and Poteshman (2006) in the literature, whereby the information within open-buy volume is clearly the most informative and parsimoniously captures the information of the put and call volumes. A potential explanation suggested by Pan and Poteshman (2006) is that traders can only use information to close positions if they happen to have the appropriate positions open at the time they become informed. Thus, the information content from closing trades may be lower compared with opening trades. Since informed traders are very likely to open new positions in cases when they possess private information on the underlying asset, we adopt the open-buy trading volume as the main variable in this study to examine the information content of trading volume.

### 3.4.2 Predictive Abilities of Domestic and Foreign Institutional Investors when Adding Control Variables

In this subsection, we specifically examine whether domestic and foreign institutional investors still maintain predictive abilities as presented in the previous section by taking into account control variables associated with next-day index returns. Based on daily VAR results, it is apparent that the lagged 2-day put-call ratio of domestic institutional investors displays a stronger forecasting power for the current index return. Since the lagged 1-day put-call ratio of domestic institutional investors, even controlling for other variables, can predict future index returns, it is reasonable to conjecture that the predictive power of the put-call ratio of domestic institutional investors could be enhanced when incorporating information of the lagged 2-day put-call ratio. Therefore, we compute $PC\_r(d)_{-1,-2AVG}$ as the average of the lagged 1- and 2-day put-call ratios of domestic institutional investors and perform regression tests including the same set of control variables.

Table 3.3 reports the correlation matrix of TAIEX returns, lagged put-call ratios of domestic and foreign institutional investors, and the other control variables. From the first columns of all panels, the lagged Hang Seng index returns, $R_{-1}^{HSI}$, have the highest correlation with TAIEX returns, especially during the sub-period of the subprime crisis. Prior to the global financial crisis, the percentage change in the exchange rate, $\Delta$ex, has the highest correlation (-0.123) with TAIEX returns, followed by $PC\_r(d)_{-1,-2AVG}$ (-0.096), $PC\_r(f)_{-1}$ (-0.091), $PC\_r(d)_{-1}$ (-0.075), and Hang Seng index returns (0.071), in terms of the absolute values of correlation. During the global financial crisis sub-period, $R_{-1}^{HSI}$ becomes the variable with the highest correlation (0.200) with TAIEX returns, followed by $PC\_r(d)_{-1,-2AVG}$ (-0.156), $PC\_r(d)_{-1}$ (-0.125), and $ln\_Vol$ (0.119). The absolute correlation coefficients of $\Delta ex$ and $PC\_r(f)_{-1}$ turn much lower than those in the pre-crisis period.

< Table 3.3 is inserted about here >

Tables 3.4 and 3.5 summarize results of predictive regressions for the TAIEX returns using

information on the put-call ratios of domestic institutional investors up to the past day and the past two days, respectively. Using only the lagged 1-day put-call ratio of domestic institutional investors, Panel A of Table 3.4 shows that all coefficients of open-buy put-call ratios are robustly negatively significant. This suggests that domestic institutional investors may possess private information and are able to trade TAIEX options profitably. The lagged 1-day TAIEX return ($R_{-1}$) and lagged 1-day Hang Seng index return ($R_{-1}^{HSI}$) are statistically significant when simultaneously included. Other control variables are not significant in explaining the next-day index return in the full sample regression tests.

< Table 3.4 is inserted about here >

Panel B of Table 3.4 shows that during the pre-crisis period the put-call ratio of domestic institutional investors still has a negative relation to the next-day TAIEX return, but its statistical significance becomes lower. Excluding the data for the global financial crisis, the lagged 1-day index return and the Hang Seng index return also lose their forecasting power for the future index return. In contrast, the percentage change of exchange rate ($\Delta ex$) is significantly negatively related to the next-day TAIEX return, which is consistent to the capital flow argument that appreciation (depreciation) of a domestic currency in general will accompany capital inflows (outflows) and boost (lower) the stock market of the host country.[26]

Another proxy variable of capital flows, *Sln_CF*, exhibits a positive relation between capital flows and the next-day TAIEX returns, confirming that the effect of capital flows on the future index returns is in line with the capital flows argument. Note that the statistically insignificant relation between *Sln_CF* and TAIEX returns merely shows that the lagged one-day *Sln_CF* is not able to capture the effect of capital flows on index returns, but the other functional form or a longer period of information for *Sln_CF* may significantly explain the next-day TAIEX returns.[27]

Panel C of Table 3.4 shows that during the subprime crisis sub-period, the put-call ratio of domestic institutional investors does not explain the next-day TAIEX returns, although its correlation coefficient (-0.125) in Table 3.3 is higher than those in the full period and the pre-crisis sub-period. The change in the exchange rate also plays no role in forecasting index returns during the financial crisis. In contrast, $R_{-1}^{HSI}$ is statistically positively significant in explaining TAIEX returns, indicating the strong co-movement and positive correlation among worldwide financial markets during the 2008 global financial tsunami.

[26] Note that a positive $\Delta ex$ means depreciation of the New Taiwan dollar and appreciation of the US dollar, while a negative $\Delta ex$ means appreciation of the New Taiwan dollar and depreciation of the US dollar. The negative coefficient of $\Delta ex$ means that in general TWD depreciation (positive $\Delta ex$) is associated with a down stock market, while TWD appreciation (negative $\Delta ex$) is associated with an up stock market.

[27] The regression model assumes that foreign capital flows will affect the host-country stock market the next day. An unreported result shows that the exchange rate dummy variable $D_{ex}$, which is one if $\Delta ex$ is negative and zero otherwise, is positively (but insignificantly) related to the next-day index return.

< Table 3.5 is inserted about here >

In Table 3.5, $PC\_r(d)_{-1,-2AVG}$, the average 2-day put-call ratio of domestic institutional investors, exhibits stronger predictive ability than the lagged 1-day put-call ratio. In the pre-crisis sub-period (Panel B of Table 3.5), the put-call ratio of domestic institutional investors now presents some prediction power for next-day index returns at the 10% statistical significance, while the lagged 1-day put-call ratio has no predictive power for next-day index returns (Panel B of Table 3.4). The incorporation of both lagged 1-day and 2-day options volume proves to be a better information content variable for domestic institutional investors. Unlike the results of Chang et al. (2009) in which only the coefficients of the open-buy put-call ratios of foreign institutional investors show significant predictive power for future index returns while there is no prediction abilities among other classes of investors, our results regarding domestic institutional investors are substantially different.

< Table 3.6 is inserted about here >

We next present the results of predictive regression for foreign institutional investors and compare them with those by Chang et al. (2009). Surprisingly, Table 3.6 shows that the put-call ratio of foreign institutional investors only has marginally significant predictive power for next-day TAIEX returns during the pre-crisis period (Panel B) and has no prediction capability in the full sample period (Panel A) controlling for other variables. Our results are very different from those of Chang et al. (2009), in which the lagged 1-day put-call ratio of foreign institutional investors could significantly forecast next-day index return. The differences could not be only explained by the inclusion of additional control variables such as interest rates or exchange rates. From the sub-period analyses in Panel C of Table 3.6, the loss of explanatory power for $PC\_r(f)_{-1}$ is partly due to the complete loss of predictability during the global financial crisis period (2008/7~2008/12). Overall, based on results above, domestic institutional investors exhibit a stronger ability than foreign institutional investors to forecast future market returns during our sample period from 2007 to 2008.

### 3.4.3 Prediction Ability and Options Trading of Different Investor Classes – Intraday VAR Analyses

To explore the interaction among trading behaviors of domestic, foreign institutional investors, and individual investors, we extend regression analyses of Equation (1) from daily to intraday frequency so as to investigate the lead-lag relationships among TAIEX returns and options trading of the three classes of investors. We choose 30 minutes as a time interval for intraday frequency, and VAR(2) is our best model based on model selection criteria, AIC and SBC. We present in Panel A of Table 3.7 the full-sample intraday VAR analysis of three different classes of investors. The lagged put-call ratio of domestic institutional investors is negatively significantly related to the current 30-minute index returns, while the put-call ratios of foreign institutional and individual investors exhibit no statistical significance on next period index returns. Given the significant prediction ability of domestic institutional investors for the next 30-minute TAIEX returns, domestic institutional investors can be regarded as groups of better informed investors. For the other two groups of investors, there are no lead-lag relationships between their put-call ratios and TAIEX returns, implying that foreign institutional and individual investors have no private information for the next period index returns.

< Table 3.7 is inserted about here >

Regarding the relationships among the three types of investors, it appears that groups of domestic institutional investors and individual investors use opposite trading strategies, because the Lag1 put-call ratio (0~30 minutes) of domestic institutional investors is negatively significant to that for individuals. One can also observe that the options trading of foreign institutional investors seem to lead domestic institutional investors for 30 minutes, because of the positively significant result in the Lag1 put-call ratio of foreign individual investors to that of domestic institutional investors. However, it is apparent that only the put-call ratios of domestic institutional investors can forecast future TAIEX returns. This motivates us to further explore the possible intraday lead-lag relationship among these three classes of investors in the two sub-periods.

From Panel B of Table 3.7, during the pre-crisis period (2007/1/1 to 2008/6/30) the put-call ratio of domestic institutional investors still negatively significantly leads 30-minute TAIEX returns. Regarding the lead-lag relation between foreign institutional and domestic institutional investors, one can observe that both Lag1 of PC_r(f) to PC_r(d) (t-value=2.5002) and Lag2 of PC_r(d) to PC_r(f) (t-value=2.0186) are statistically positively significant, which implies these two investors seem to adopt similar option trading strategies within 60 minutes. On the contrary, individual investors use the opposite option strategy to that of domestic institutional investors within 30 minutes. From results above, a potential explanation is that domestic institutional investors may possess more information on impending stock market events and thus can better predict market movement in the host country. Therefore, domestic institutional investors will utilize a highly leveraged financial instrument like options when they expect stock market fluctuations. Foreign institutional investors are relatively less informed and are not able to trade options to predict the intraday stock index returns during the pre-crisis period.

We next investigate the lead-lag relations during the global financial tsunami in Panel C of Table 3.7. It appears no put-call ratios of any class of investors exist that can predict intraday TAIEX return, and there is no significant lead-lag relationship among domestic institutional, foreign institutional, and individual investors. The evidence during the global financial crisis sub-period does not support the common wisdom that foreign institutional investors can trade options for hedging ahead of the other classes of investors, especially as foreign institutional investors might have possessed more private information than domestic investors about capital flows from global asset management during this severe downturn.

During the crisis period, it is apparent that TAIEX returns lead the options trading of institutional investors. The Lag1 index return is strongly significant at the 1% level to explain the put-call ratio of domestic institutional investors. For foreign institutional investors, the Lag2 index return is significant at the 1% level and the Lag1 index return is slightly significant at the 10% level in explaining put-call ratios. These results suggest that domestic and foreign institutional investors could not predict TAIEX returns ahead of time, but did take actions after observing changes in the market index. This might be explained by the hedging transactions of their portfolios when triggered by the stock market's downturn. Compared with the pre-crisis period results in Panel B, domestic institutional investors also lost their forecasting ability or private information during this unusual global financial tsunami.

Overall, unlike the evidence of previous literature that foreign institutional investors possess more private information than other types of investors, our study suggests that domestic institutional investors may also exhibit similarly superior predictability for future index returns. We provide some possible explanations for our empirical results in this section. First, domestic institutional investors may employ option strategies based on market conditions or signals. Therefore, their trades are correlated with these market events/signals, and thus one can observe a high correlation in their options trading with foreign institutional investors. Second, institutional investors may obtain trading information of those better informed investors through some channels and thus follow their trades in the options market. Nonetheless, we should note that the Taiwan Futures Exchange publishes institutional trading information only after trading hours. The only real-time information during trading hours is of prices and aggregate volumes. Accordingly, we cannot provide direct evidence of this conjecture.[28] Finally, we cannot rule out the possibility that the correlation may just be a coincidence for our analyzing period. No matter what the root cause is for the positive correlation between the trades of domestic and foreign institutional investors, the conclusion emerging from this section is that domestic institutional investors are becoming increasingly better informed at predicting index returns in our sample period.

**3.5 Conclusion**

In order to capture the dynamics among TAIEX returns and options trading of different trader

[28] However, as far as we know, it is not impossible for traders to obtain some information on some foreign institutional investors through brokers.

types, we perform daily and intraday VAR analyses using index returns and the put-call ratios of the three trader classes: domestic institutional, foreign institutional, and individual investors. We separately conduct analyses of the whole data period (2007/1~2008/12) and the two sub-periods, including the pre-crisis period (2007/1~2008/6) and the most impacted sub-period of the financial tsunami (2008/7~2008/12). Following the literature, we also investigate predictability of different trader types through predictive regression, controlling for lagged variables related to index returns.

In the daily analyses for the whole data period, the put-call ratios of domestic institutional investors are negatively related to future index returns, while those of foreign institutional and individual investors exhibit no significant relation to next period index returns. We also find that domestic institutional and foreign institutional investors seem to use similar option strategies, while individual investors tend to use contrarian option strategies in the market that are opposite to the trades of institutional investors. In the sub-period analyses, during the pre-crisis sub-period, domestic institutional investors still maintain their predictive ability for future index returns, and foreign institutional investors show marginal prediction capability for future index returns. During the 2008 global financial crisis period, no trader type seems to possess private information in forecasting future index returns. In the intraday analyses, most of the results are similar to those for daily frequency regarding predictability of different trader types. The major difference is that intraday TAIEX returns strongly lead the options trading of both domestic and foreign institutional investors during the period of the global financial tsunami, indicating that institutional investors do not possess superior private information over individuals during this unusual period. However, institutional investors indeed closely watch the market and trade options accordingly.

The analyses of predictive regression incorporate a set of lagged control variables related to index returns, such as TAIEX returns, and Hang Seng index returns, the change in the exchange rate, and the proxy for capital flows. The results of predictive regression still support the daily VAR analysis that domestic institutional investors can predict index returns in the full period and the pre-crisis sub-periods, while foreign institutional investors only exhibit marginal predictability in the pre-crisis sub-period. There exists no significant predictability for any trader type in the financial crisis sub-period. Our empirical findings reveal different evidence in contrast to the previous literature that foreign institutional investors possess superior information over domestic investors.

In sum, our empirical results suggest that domestic institutional investors are becoming as market savvy as foreign institutional investors in predicting future index returns, which can shed some light on the evolvement of interaction among various trader types, especially when the progressive development of emerging financial markets has been getting vast attention from academics, businesses, and professionals. In addition to the contribution to the literature, our findings of daily predictive regression are also useful for investors in practice, because all the variables we incorporate in our analyses are public information. Nonetheless, the root cause for the better performance by domestic institutional investors has not been uncovered in this study and can be a potentially fruitful research in the future.

**Table 3.1 Daily options trading volume by different trade types and different classes of investors**

| Trading volume (contracts) | Open | | | | Close | | | |
|---|---|---|---|---|---|---|---|---|
| | Buy | | Sell | | Buy | | Sell | |
| | Call | Put | Call | Put | Call | Put | Call | Put |
| Average (no.) | 83,014 | 61,641 | 52,970 | 46,511 | 50,371 | 44,224 | 75,475 | 57,362 |
| Domestic institutional investors (%) | 2.47 | 5.06 | 6.08 | 5.13 | 6.20 | 5.04 | 2.78 | 4.47 |
| Foreign institutional investors (%) | 12.69 | 17.13 | 10.63 | 10.40 | 2.67 | 2.66 | 2.60 | 4.91 |
| Individual investors (%) | 62.83 | 52.59 | 46.23 | 49.87 | 44.81 | 46.22 | 62.62 | 55.44 |
| Market makers (%) | 22.01 | 25.22 | 37.05 | 34.62 | 46.51 | 46.39 | 32.20 | 35.40 |

Notes: Sample period is from 2 January 2007 to 31 December 2008. The average trading volume is calculated by the time-series average of the aggregated volume with different strike prices and the time-to-expiration at each trading day; the proportion of trading volume by different classes of investors (expressed as a percentage) is calculated by the time-series average of each trade type by different classes of investors divided by the total volume of the corresponding trade types.

**Table 3.2 The VAR estimation results of daily index return and open-buy put-call ratios of different classes of investors**

| | Explanatory variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Lagged R* | | *Lagged PC_r(d)* | | *Lagged PC_r(f)* | | *Lagged PC_r(i)* | |
| Dependent variable | Lag 1 | Lag 2 | Lag 1 | Lag 2 | Lag 1 | Lag 2 | Lag 1 | Lag 2 |
| Panel A: Full Period (2007/01-2008/12) | | (484 observations) | | | | | | |
| *R* | 0.002 | 0.078 | -0.006 | -0.011 | -0.003 | 0.006 | -0.010 | 0.002 |
| | *0.03* | *1.55* | *-0.99* | ***-1.87***$^{*}$ | *-0.56* | *1.34* | *-0.68* | *0.16* |
| *PC_r(d)* | -0.622 | 0.157 | 0.248 | 0.190 | -0.011 | -0.018 | 0.079 | -0.141 |
| | *-1.57* | *0.39* | *5.24*$^{***}$ | *4.02*$^{***}$ | *-0.30* | *-0.51* | *0.69* | *-1.25* |
| *PC_r(f)* | 0.399 | 0.513 | 0.083 | 0.066 | 0.283 | 0.111 | -0.063 | 0.186 |
| | *0.75* | *0.96* | *1.30* | *1.05* | *5.90*$^{***}$ | *2.30*$^{**}$ | *-0.41* | *1.22* |
| *PC_r(i)* | 0.365 | 0.072 | -0.034 | -0.027 | 0.020 | 0.022 | 0.528 | 0.153 |
| | *2.31*$^{**}$ | *0.45* | *-1.83*$^{*}$ | *-1.43* | *1.42* | *1.57* | *11.55*$^{***}$ | *3.39*$^{***}$ |
| Panel B: Pre-Crisis Period (2007/01-2008/06) | | | (359 observations) | | | | | |
| *R* | -0.015 | -0.067 | -0.003 | -0.010 | -0.008 | 0.004 | -0.007 | 0.016 |
| | *-0.25* | *-1.15* | *-0.53* | ***-1.80***$^{*}$ | ***-1.77***$^{*}$ | *1.00* | *-0.51* | *1.09* |
| *PC_r(d)* | -0.412 | 0.210 | 0.206 | 0.150 | 0.003 | -0.001 | 0.092 | -0.180 |
| | *-0.74* | *0.38* | *3.74*$^{***}$ | *2.69*$^{***}$ | *0.08* | *-0.03* | *0.68* | *-1.32* |
| *PC_r(f)* | 0.254 | 0.738 | 0.154 | 0.101 | 0.275 | 0.084 | -0.052 | 0.040 |
| | *0.33* | *0.96* | ***2.02***$^{**}$ | *1.31* | *4.92*$^{***}$ | *1.51* | *-0.27* | *0.21* |
| *PC_r(i)* | -0.163 | -0.241 | -0.042 | -0.013 | 0.008 | 0.022 | 0.523 | 0.217 |
| | *-0.76* | *-1.12* | *-1.96*$^{**}$ | *-0.59* | *0.49* | *1.42* | *9.90*$^{***}$ | *4.12*$^{***}$ |
| Panel C: Financial Crisis Period (2008/07-2008/12) | | | | (125 observations) | | | | |
| *R* | 0.011 | 0.225 | -0.010 | -0.012 | 0.009 | 0.006 | -0.043 | -0.017 |
| | *0.10* | *2.05* | *-0.63* | *-0.73* | *0.64* | *0.41* | *-1.11* | *-0.47* |
| *PC_r(d)* | -0.555 | 0.309 | 0.342 | 0.256 | -0.016 | -0.053 | 0.054 | -0.023 |
| | *-0.89* | *0.47* | *3.56*$^{***}$ | *2.68*$^{***}$ | *-0.19* | *-0.61* | *0.23* | *-0.11* |
| *PC_r(f)* | -0.276 | -0.429 | -0.072 | 0.049 | 0.136 | 0.052 | -0.034 | 0.296 |
| | *-0.37* | *-0.55* | *-0.63* | *0.44* | *1.33* | *0.51* | *-0.12* | *1.14* |
| *PC_r(i)* | 0.905 | 0.306 | -0.015 | -0.060 | 0.048 | 0.004 | 0.505 | 0.047 |
| | *3.60*$^{***}$ | *1.16* | *-0.40* | *-1.58* | *1.38* | *0.13* | *5.44*$^{***}$ | *0.54* |

Notes: This table consists of three panels each corresponding to the results of the full sample period (Panel A) and two sub-periods, pre-crisis period from 2007/01/02 to 2008/06/30 (Panel B) and financial crisis period from 2008/07/01 to 2008/12/31. *R* denotes the daily TAIEX return. *PC_r*(*d*), *PC_r*(*f*), and *PC_r*(*i*) are the put-call ratios of domestic institutional, foreign institutional, and individual investors, respectively. VAR (2) is our best model based on the selection criteria, AIC and SBC. The regression coefficients for the two lags and t-statistics are reported below the coefficient. One, two, and three asterisks (*) indicate t-values are significant at the 0.1, 0.05 and 0.01 levels, respectively.

**Table 3.3 Correlation matrix**

Full period (484 observations)

| | *R* | $PC\_r(d)_{-1}$ | $PC\_r(d)_{-1,-2AVG}$ | $PC\_r(f)_{-1}$ | $D\_{}_{Mat(d)}$ | $D\_{}_{Mat(f)}$ | *ln_Vol* | $R_{-1}$ | $R_{-1}^{HSI}$ | *r* | *Δex* | *Sln_CF* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *R* | 1.000 | | | | | | | | | | | |
| $PC\_r(d)_{-1}$ | -0.100 | 1.000 | | | | | | | | | | |
| $PC\_r(d)_{-1,-2AVG}$ | -0.127 | 0.818 | 1.000 | | | | | | | | | |
| $PC\_r(f)_{-1}$ | -0.026 | 0.139 | 0.144 | 1.000 | | | | | | | | |
| $D\_{}_{Mat(d)}$ | 0.020 | -0.039 | -0.048 | 0.011 | 1.000 | | | | | | | |
| $D\_{}_{Mat(f)}$ | -0.006 | -0.094 | -0.105 | 0.059 | 0.921 | 1.000 | | | | | | |
| *ln_Vol* | 0.051 | -0.116 | -0.141 | 0.057 | -0.011 | 0.005 | 1.000 | | | | | |
| $R_{-1}$ | 0.031 | -0.310 | -0.250 | -0.303 | -0.048 | -0.009 | 0.134 | 1.000 | | | | |
| $R_{-1}^{HSI}$ | 0.144 | -0.170 | -0.121 | -0.177 | -0.069 | -0.034 | 0.076 | 0.591 | 1.000 | | | |
| *r* | 0.059 | -0.069 | -0.087 | 0.283 | -0.004 | 0.034 | 0.187 | 0.072 | 0.062 | 1.000 | | |
| *Δex* | -0.099 | 0.089 | 0.083 | 0.080 | -0.003 | -0.010 | -0.133 | -0.280 | -0.193 | -0.001 | 1.000 | |
| *Sln_CF* | 0.063 | -0.047 | -0.084 | -0.132 | -0.001 | 0.005 | 0.104 | 0.229 | 0.149 | -0.086 | -0.591 | 1.000 |

Pre-crisis Period: 2007/1-2008/6 (360 observations)

| | *R* | $PC\_r(d)_{-1}$ | $PC\_r(d)_{-1,-2AVG}$ | $PC\_r(f)_{-1}$ | $D\_{}_{Mat(d)}$ | $D\_{}_{Mat(f)}$ | *ln_Vol* | $R_{-1}$ | $R_{-1}^{HSI}{}_{1}$ | *r* | *Δex* | *Sln_CF* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *R* | 1.000 | | | | | | | | | | | |
| $PC\_r(d)_{-1}$ | -0.075 | 1.000 | | | | | | | | | | |
| $PC\_r(d)_{-1,-2AVG}$ | -0.096 | 0.798 | 1.000 | | | | | | | | | |
| $PC\_r(f)_{-1}$ | -0.091 | 0.160 | 0.198 | 1.000 | | | | | | | | |
| $D\_{}_{Mat(d)}$ | 0.012 | -0.075 | -0.116 | 0.030 | 1.000 | | | | | | | |
| $D\_{}_{Mat(f)}$ | -0.029 | -0.092 | -0.122 | 0.067 | 0.956 | 1.000 | | | | | | |
| *ln_Vol* | -0.034 | -0.039 | -0.043 | -0.042 | -0.007 | -0.006 | 1.000 | | | | | |
| $R_{-1}$ | 0.024 | -0.283 | -0.224 | -0.316 | 0.025 | 0.020 | 0.021 | 1.000 | | | | |
| $R_{-1}^{HSI}$ | 0.071 | -0.156 | -0.077 | -0.165 | -0.006 | -0.012 | -0.022 | 0.530 | 1.000 | | | |
| *r* | 0.022 | -0.041 | -0.047 | 0.198 | -0.002 | 0.027 | -0.190 | 0.035 | 0.051 | 1.000 | | |
| *Δex* | -0.123 | 0.023 | 0.066 | 0.082 | -0.018 | 0.008 | -0.098 | -0.217 | -0.111 | 0.106 | 1.000 | |
| *Sln_CF* | 0.068 | 0.040 | -0.024 | -0.132 | 0.027 | 0.000 | 0.096 | 0.188 | 0.093 | -0.162 | -0.686 | 1.000 |

Financial Crisis Period: 2008/7-2008/12 (124 observations)

| | *R* | $PC\_r(d)_{-1}$ | $PC\_r(d)_{-1,-2AVG}$ | $PC\_r(f)_{-1}$ | $D\_{}_{Mat(d)}$ | $D\_{}_{Mat(f)}$ | *ln_Vol* | $R_{-1}$ | $R_{-1}^{HSI}$ | *r* | *Δex* | *Sln_CF* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *R* | 1.000 | | | | | | | | | | | |
| $PC\_r(d)_{-1}$ | -0.125 | 1.000 | | | | | | | | | | |
| $PC\_r(d)_{-1,-2AVG}$ | -0.156 | 0.855 | 1.000 | | | | | | | | | |
| $PC\_r(f)_{-1}$ | 0.045 | 0.168 | 0.104 | 1.000 | | | | | | | | |
| $D\_{}_{Mat(d)}$ | 0.033 | 0.041 | 0.094 | -0.042 | 1.000 | | | | | | | |
| $D\_{}_{Mat(f)}$ | 0.037 | -0.100 | -0.050 | -0.002 | 0.865 | 1.000 | | | | | | |
| *ln_Vol* | 0.119 | -0.234 | -0.277 | 0.018 | -0.022 | -0.011 | 1.000 | | | | | |
| $R_{-1}$ | 0.023 | -0.353 | -0.276 | -0.412 | -0.161 | -0.089 | 0.290 | 1.000 | | | | |
| $R_{-1}^{HSI}$ | 0.200 | -0.192 | -0.168 | -0.292 | -0.157 | -0.098 | 0.196 | 0.643 | 1.000 | | | |
| *r* | -0.061 | 0.116 | 0.117 | 0.237 | -0.012 | -0.033 | 0.166 | -0.032 | -0.020 | 1.000 | | |
| *Δex* | -0.069 | 0.161 | 0.088 | 0.160 | 0.013 | -0.038 | -0.144 | -0.324 | -0.241 | 0.091 | 1.000 | |
| *Sln_CF* | 0.050 | -0.255 | -0.212 | -0.187 | -0.075 | 0.017 | 0.093 | 0.312 | 0.245 | -0.171 | -0.571 | 1.000 |

Notes: This table consists of three panels and each reports the correlation matrix of full sample period (Panel A) and two sub-periods, pre-crisis period from 2007/01/02 to 2008/06/30 (Panel B) and global financial crisis period from 2008/07/01 to 2008/12/31 (Panel C). *R* denotes the daily TAIEX return. $PC\_r(d)_{-1}$, $PC\_r(d)_{-1,-2AVG}$, and $PC\_r(f)_{-1}$ are the lagged 1-day put-call ratios of domestic institutional investors, the average of lagged 1-day and 2-day put-call ratios of domestic institutional investors, and the lagged 1-day put-call ratios of foreign institutional investors, respectively. An interaction term of a dummy variable ($D\_{}_{Mat}$), which is one if the next trading date is the maturity date, and the put-call ratio calculated by options matured in the next-day ($PC\_r_{Near\ Maturity}$). The logarithm of the daily closing index trading volume (*ln_Vol*), the lagged one-day index return ($R_{-1}$), the lag one-day Hang Seng index return ($R_{-1}^{HSI}$), interest rate (*r*), and exchange rate (Δ*ex*). The proxy variable for the capital flows (*Sln_CF*) is constructed as *ln_VolFX* if Δ*ex* is positive and -1 times *ln_VolFX* otherwise, where *ln_VolFX* is the natural logarithm of daily foreign exchange trading volume between the New Taiwan dollar and US dollar.

**Table 3.4 Predictive regressions with control variables – Predictability from the put-call ratio of domestic institutional investors**

| | Int. | $PC\_r(d)_{-1}$ | $D\_{Mat(d)} \times PC\_r_{Near\ Maturity}$ | $ln\_Vol$ | $R_{-1}$ | $R_{-1}^{HSI}$ | $r$ | $\Delta ex$ | $Sln\_CF$ | $R^2$ | Obs. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: Full Period (2007/01-2008/12) | | (484 observations) | | | | | | | | | |
| Model 1 | 0.0060 * | -0.0117 ** | | | | | | | | 0.0099 | 484 |
| | *1.85* | *-2.18* | | | | | | | | | |
| Model 2 | -0.0193 | -0.0113 ** | 0.0036 | 0.0015 | -0.1428 ** | 0.1346 ** | 0.0503 | -0.4975 | | 0.0461 | 484 |
| | *-0.39* | *-1.97* | *0.63* | *0.47* | *-1.96* | *2.54* | *1.01* | *-1.59* | | | |
| Model 3 | -0.033 | -0.0115 ** | 0.0039 | 0.0025 | -0.1197 * | 0.1384 *** | | | | 0.0368 | 484 |
| | *-0.64* | *-1.99* | *0.68* | *0.75* | *-1.67* | *2.6* | | | | | |
| Model 4 | -0.0277 | -0.011 * | 0.0026 | 0.0021 | -0.005 | | 0.051 | | | 0.0139 | 484 |
| | *-0.51* | *-1.88* | *0.46* | *0.59* | *-0.09* | | *1* | | | | |
| Model 5 | -0.0236 | -0.0112 * | 0.0023 | 0.0019 | -0.0285 | | | -0.5206 | | 0.0197 | 484 |
| | *-0.44* | *-1.93* | *0.41* | *0.54* | *-0.52* | | | *-1.57* | | | |
| Model 6 | -0.0286 | -0.0113 * | 0.0039 | 0.0021 | -0.1207 * | 0.1375 ** | 0.0466 | | | 0.0387 | 484 |
| | *-0.56* | *-1.95* | *0.67* | *0.65* | *-1.66* | *2.57* | *0.93* | | | | |
| Model 7 | -0.0245 | -0.0115 ** | 0.0036 | 0.002 | -0.1412 ** | 0.1357 ** | | -0.4874 | | 0.0438 | 484 |
| | *-0.49* | *-2* | *0.63* | *0.6* | *-1.97* | *2.58* | | *-1.56* | | | |
| Model 8 | -0.0178 | -0.011 * | 0.0024 | 0.0014 | -0.0311 | | 0.0549 | -0.5314 | | 0.0233 | 484 |
| | *-0.34* | *-1.89* | *0.41* | *0.41* | *-0.56* | | *1.09* | *-1.61* | | | |
| Model 9 | -0.0226 | -0.0115 ** | 0.0038 | 0.0017 | -0.1343 * | 0.1364 ** | 0.0545 | | 0.0001 | 0.0423 | 484 |
| | *-0.45* | *-1.99* | *0.64* | *0.53* | *-1.83* | *2.56* | *1.08* | | *1.38* | | |
| Panel B: Pre-Crisis Period (2007/01-2008/06) | | (360 observations) | | | | | | | | | |
| Model 1 | 0.0040 | -0.0075 | | | | | | | | 0.0056 | 360 |
| | *1.32* | *-1.38* | | | | | | | | | |
| Model 2 | 0.0398 | -0.0079 | 0.0008 | -0.0024 | -0.0691 | 0.0594 | 0.025 | -0.8582 ** | | 0.0282 | 360 |
| | *0.8* | *-1.4* | *0.11* | *-0.75* | *-0.96* | *0.92* | *0.42* | *-2.21* | | | |
| Model 3 | 0.0333 | -0.0074 | 0.0011 | -0.002 | -0.0385 | 0.06 | | | | 0.0115 | 360 |
| | *0.73* | *-1.3* | *0.15* | *-0.65* | *-0.52* | *0.9* | | | | | |
| Model 4 | 0.0332 | -0.0074 | 0.0009 | -0.0019 | 0.003 | | 0.013 | | | 0.0071 | 360 |
| | *0.66* | *-1.3* | *0.12* | *-0.61* | *0.05* | | *0.22* | | | | |
| Model 5 | 0.0467 | -0.0081 | 0.0006 | -0.0028 | -0.0265 | | | -0.8435 ** | | 0.0232 | 360 |
| | *1.01* | *-1.42* | *0.08* | *-0.92* | *-0.45* | | | *-2.18* | | | |
| Model 6 | 0.0313 | -0.0073 | 0.0011 | -0.0019 | -0.0386 | 0.0597 | 0.0108 | | | 0.0116 | 360 |
| | *0.63* | *-1.29* | *0.15* | *-0.57* | *-0.52* | *0.9* | *0.18* | | | | |
| Model 7 | 0.0444 | -0.008 | 0.0008 | -0.0026 | -0.0684 | 0.06 | | -0.8435 ** | | 0.0277 | 360 |
| | *0.97* | *-1.42* | *0.11* | *-0.88* | *-0.96* | *0.93* | | *-2.16* | | | |
| Model 8 | 0.0417 | -0.008 | 0.0006 | -0.0025 | -0.0278 | | 0.0273 | -0.86 ** | | 0.0238 | 360 |
| | *0.84* | *-1.4* | *0.08* | *-0.79* | *-0.47* | | *0.46* | *-2.22* | | | |
| Model 9 | 0.0359 | -0.0082 | 0.0008 | -0.0021 | -0.0564 | 0.06 | 0.0243 | | 0.0001 | 0.0177 | 360 |
| | *0.73* | *-1.43* | *0.1* | *-0.68* | *-0.76* | *0.9* | *0.41* | | *1.5* | | |
| Panel C: Financial Crisis Period (2008/07-2008/12) | | (124 observations) | | | | | | | | | |
| Model 1 | 0.0074 | -0.0187 | | | | | | | | 0.0157 | 124 |
| | *1.01* | *-1.54* | | | | | | | | | |
| Model 2 | -0.1700 | -0.0179 | 0.0098 | 0.0118 | -0.2652 ** | 0.2127 ** | -0.2248 | -0.1810 | | 0.0947 | 124 |
| | *-0.97* | *-1.24* | *1.02* | *1.01* | *-1.98* | *2.44* | *-0.76* | *-0.46* | | | |
| Model 3 | -0.1517 | -0.0196 | 0.0104 | 0.0105 | -0.2533 * | 0.2147 ** | | | | 0.0893 | 124 |
| | *-0.87* | *-1.39* | *1.16* | *0.92* | *-1.92* | *2.47* | | | | | |
| Model 4 | -0.1854 | -0.0160 | 0.0065 | 0.0128 | -0.0460 | | -0.2453 | | | 0.0317 | 124 |
| | *-0.94* | *-1.09* | *0.70* | *0.98* | *-0.45* | | *-0.82* | | | | |
| Model 5 | -0.1584 | -0.0172 | 0.0064 | 0.0109 | -0.0597 | | | -0.2754 | | 0.03 | 124 |
| | *-0.82* | *-1.17* | *0.71* | *0.85* | *-0.60* | | | *-0.64* | | | |
| Model 6 | -0.1737 | -0.0181 | 0.0101 | 0.0121 | -0.2557 * | 0.2141 ** | -0.2361 | | | 0.0935 | 124 |
| | *-0.99* | *-1.26* | *1.07* | *1.03* | *-1.95* | *2.45* | *-0.80* | | | | |
| Model 7 | -0.1486 | -0.0193 | 0.0101 | 0.0103 | -0.2644 ** | 0.2130 ** | | -0.2102 | | 0.0909 | 124 |
| | *-0.86* | *-1.35* | *1.09* | *0.91* | *-1.96* | *2.46* | | *-0.54* | | | |
| Model 8 | -0.1802 | -0.0157 | 0.0061 | 0.0124 | -0.0608 | | -0.2298 | -0.2455 | | 0.0339 | 124 |
| | *-0.92* | *-1.06* | *0.65* | *0.96* | *-0.62* | | *-0.77* | *-0.57* | | | |
| Model 9 | -0.1737 | -0.0181 | 0.0101 | 0.0121 | -0.2560 * | 0.2141 ** | -0.2354 | | 0.0000 | 0.0935 | 124 |
| | *-0.98* | *-1.23* | *1.06* | *1.03* | *-1.90* | *2.47* | *-0.78* | | *0.01* | | |

Notes: The table describes the results of time-series regressions of the next-day TAIEX spot index returns on the open-buy put-call ratio constructed from domestic institutional investors ($PC\_r(d)_{-1}$) and a set of control variables as the following. An interaction term of a dummy variable ($D\_{Mat}$), which is one if the next trading date is the maturity date, and the put-call ratio calculated by options matured in the next-day ($PC\_r_{Near\ Maturity}$). The logarithm of the daily closing index trading volume ($ln\_Vol$), the lagged one-day index return ($R_{-1}$), the lag one-day Hang Seng index return ($R_{-1}^{HSI}$), interest rate ($r$), and exchange rate ($\Delta ex$). The proxy variable for the capital flows ($Sln\_CF$) is constructed as $ln\_VolFX$ if $\Delta ex$ is positive and -1 times $ln\_VolFX$ otherwise, where $ln\_VolFX$ is the natural logarithm of daily foreign exchange trading volume between the New Taiwan dollar and US dollar. All results reported are corrected by the Newey-West method. One, two, and three asterisks (*) indicate t-values are significant at the 0.1, 0.05 and 0.01 levels, respectively.

**Table 3.5 Predictive regressions with control variables – Predictability from the two-day average put-call ratio of domestic institutional investors**

| | Int. | $PC\_r(d)_{-1,-2AVG}$ | $D\_{Mat(d)}$ × $PC\_r_{Near\ Maturity}$ | $ln\_Vol$ | $R_{-1}$ | $R_{-1}^{HSI}$ | $r$ | $\Delta ex$ | $Sln\_CF$ | $R^2$ | Obs. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: Full Period (2007/01-2008/12) | | (484 observations) | | | | | | | | | |
| Model 1 | 0.0094 *** | -0.0183 *** | | | | | | | | 0.0162 | 484 |
| | *2.58* | *-2.84* | | | | | | | | | |
| Model 2 | -0.0098 | -0.018 *** | 0.0033 | 0.0012 | -0.1448 ** | 0.1369 *** | 0.0467 | -0.492 | | 0.0522 | 484 |
| | *-0.2* | *-2.6* | *0.57* | *0.36* | *-2.03* | *2.6* | *0.93* | *-1.61* | | | |
| Model 3 | -0.0228 | -0.0184 *** | 0.0035 | 0.0021 | -0.1225 * | 0.1406 *** | | | | 0.0434 | 484 |
| | *-0.44* | *-2.68* | *0.61* | *0.63* | *-1.74* | *2.66* | | | | | |
| Model 4 | -0.0185 | -0.0173 ** | 0.0023 | 0.0017 | -0.0052 | | 0.0478 | | | 0.0194 | 484 |
| | *-0.34* | *-2.45* | *0.4* | *0.48* | *-0.09* | | *0.93* | | | | |
| Model 5 | -0.0142 | -0.0177 ** | 0.002 | 0.0015 | -0.0286 | | | -0.5164 | | 0.0254 | 484 |
| | *-0.26* | *-2.51* | *0.35* | *0.43* | *-0.53* | | | *-1.59* | | | |
| Model 6 | -0.0188 | -0.0181 *** | 0.0035 | 0.0017 | -0.1233 ** | 0.1397 *** | 0.0431 | | | 0.045 | 484 |
| | *-0.37* | *-2.61* | *0.61* | *0.53* | *-1.73* | *2.62* | *0.85* | | | | |
| Model 7 | -0.0145 | -0.0184 *** | 0.0032 | 0.0016 | -0.1436 ** | 0.1379 *** | | -0.4825 | | 0.0503 | 484 |
| | *-0.29* | *-2.67* | *0.57* | *0.48* | *-2.04* | *2.64* | | *-1.58* | | | |
| Model 8 | -0.0089 | -0.0173 ** | 0.002 | 0.0011 | -0.0308 | | 0.0516 | -0.5266 | | 0.0277 | 484 |
| | *-0.17* | *-2.44* | *0.36* | *0.31* | *-0.57* | | *1.01* | *-1.63* | | | |
| Model 9 | -0.014 | -0.0179 *** | 0.0035 | 0.0014 | -0.1346 ** | 0.1387 *** | 0.0504 | | 0.0001 | 0.0479 | 484 |
| | *-0.27* | *-2.58* | *0.59* | *0.43* | *-1.88* | *2.62* | *0.99* | | *1.25* | | |
| P Panel B: Pre-Crisis Period (2007/01-2008/06) | | | (360 observations) | | | | | | | | |
| Model 1 | 0.0066 * | -0.0120 * | | | | | | | | 0.0091 | 360 |
| | *1,86* | *-1.90* | | | | | | | | | |
| Model 2 | 0.0431 | -0.0122 * | 0.0001 | -0.0024 | -0.0704 | 0.0641 | 0.0335 | -0.8254 ** | | 0.0313 | 360 |
| | *0.86* | *-1.81* | *0.02* | *-0.77* | *-0.97* | *0.98* | *0.37* | *-2.1* | | | |
| Model 3 | 0.0373 | -0.0125 * | 0.0003 | -0.0019 | -0.0431 | 0.0647 | | | | 0.0158 | 360 |
| | *0.81* | *-1.88* | *0.04* | *-0.67* | *-0.58* | *0.96* | | | | | |
| Model 4 | 0.0373 | -0.012 * | 0.0001 | -0.002 | 0.0028 | | 0.0111 | | | 0.0107 | 360 |
| | *0.74* | *-1.81* | *0.02* | *-0.63* | *0.05* | | *0.18* | | | | |
| Model 5 | 0.0494 | -0.0119 * | -0.0001 | -0.0028 | -0.0237 | | | -0.8119 ** | | 0.0257 | 360 |
| | *1.06* | *-1.78* | *-0.01* | *-0.93* | *-0.4* | | | *-2.07* | | | |
| Model 6 | 0.0357 | -0.0124 * | 0.0003 | -0.0019 | -0.0413 | 0.0647 | 0.0084 | | | 0.0159 | 360 |
| | *0.71* | *-1.86* | *0.04* | *-0.6* | *-0.58* | *0.96* | *0.14* | | | | |
| Model 7 | 0.0473 | -0.0123 * | 0.0001 | -0.0027 | -0.0698 | 0.0647 | | -0.8119 ** | | 0.0309 | 360 |
| | *1.02* | *-1.84* | *0.01* | *-0.89* | *-0.98* | *0.99* | | *-2.05* | | | |
| Model 8 | 0.0447 | -0.0117 * | -0.0001 | -0.0026 | -0.0248 | | 0.0252 | -0.8271 ** | | 0.0262 | 360 |
| | *0.88* | *-1.74* | *-0.01* | *-0.8* | *-0.42* | | *0.42* | *-2.12* | | | |
| Model 9 | 0.0394 | -0.0126 * | 0.0001 | -0.0022 | -0.0578 | 0.0644 | 0.0212 | | 0.0001 | 0.0211 | 360 |
| | *0.79* | *-1.88* | *0.01* | *-0.69* | *-0.79* | *0.96* | *0.35* | | *1.39* | | |
| Panel C: Financial Crisis Period (2008/07-2008/12) | | | | (124 observations) | | | | | | | |
| Model 1 | 0.0125 | -0.0273 * | | | | | | | | 0.0244 | 124 |
| | *1.41* | *-1.88* | | | | | | | | | |
| Model 2 | -0.1417 | -0.0247 | 0.0117 | 0.0102 | -0.2557 ** | 0.2099 ** | -0.2027 | -0.2188 | | 0.0999 | 124 |
| | *-0.78* | *-1.46* | *1.20* | *0.85* | *-1.98* | *2.45* | *-0.68* | *-0.57* | | | |
| Model 3 | -0.1254 | -0.0263 | 0.0125 | 0.0091 | -0.2403 * | 0.2119 ** | | | | 0.0946 | 124 |
| | *-0.70* | *-1.58* | *1.35* | *0.78* | *-1.91* | *2.48* | | | | | |
| Model 4 | -0.1567 | -0.0235 | 0.0083 | 0.0112 | -0.0395 | | -0.2233 | | | 0.0377 | 124 |
| | *-0.76* | *-1.33* | *0.88* | *0.82* | *-0.42* | | *-0.74* | | | | |
| Model 5 | -0.1283 | -0.0255 | 0.0083 | 0.0092 | -0.0557 | | | -0.3085 | | 0.0374 | 124 |
| | *-0.64* | *-1.47* | *0.89* | *0.70* | *-0.59* | | | *-0.74* | | | |
| Model 6 | -0.1473 | -0.0245 | 0.0120 | 0.0106 | -0.2436 * | 0.2116 ** | -0.2178 | | | 0.0981 | 124 |
| | *-0.80* | *-1.45* | *1.26* | *0.88* | *-1.94* | *2.46* | *-0.73* | | | | |
| Model 7 | -0.1208 | -0.0264 | 0.0120 | 0.0088 | -0.2543 ** | 0.2099 ** | | -0.2475 | | 0.0968 | 124 |
| | *-0.68* | *-1.58* | *1.28* | *0.76* | *-1.96* | *2.47* | | *-0.65* | | | |
| Model 8 | -0.1494 | -0.0238 | 0.0079 | 0.0107 | -0.0571 | | -0.2040 | -0.2796 | | 0.0405 | 124 |
| | *-0.73* | *-1.35* | *0.82* | *0.79* | *-0.61* | | *-0.68* | *-0.67* | | | |
| Model 9 | -0.1474 | -0.0244 | 0.0121 | 0.0106 | -0.2441 * | 0.2115 ** | -0.2164 | | 0.0000 | 0.0981 | 124 |
| | *-0.80* | *-1.42* | *1.25* | *0.88* | *-1.88* | *2.47* | *-0.71* | | *0.03* | | |

**Table 3.6 Predictive regressions with control variables – Predictability from the put-call ratio of foreign institutional investors**

| | Int. | $PC\_r(f)_{-1}$ | $D\_{Mat(f)} \times PC\_r_{Near\ Maturity}$ | $ln\_Vol$ | $R_{-1}$ | $R_{-1}^{HSI}$ | $r$ | $\Delta ex$ | $Sln\_CF$ | $R^2$ | Obs. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Panel A: Full Period (2007/01-2008/12) | | | (484 observations) | | | | | | | | |
| Model 1 | 0.0001 | -0.0023 | | | | | | | | 0.0007 | 484 |
| | *0.05* | *-0.57* | | | | | | | | | |
| Model 2 | -0.0326 | -0.0038 | -0.0003 | 0.0021 | -0.1273 * | 0.132 ** | 0.0679 | -0.5002 | | 0.0385 | 484 |
| | *-0.66* | *-0.97* | *-0.04* | *0.65* | *-1.72* | *2.49* | *1.34* | *-1.63* | | | |
| Model 3 | -0.0479 | -0.0021 | 0.0001 | 0.0031 | -0.097 | 0.136 ** | | | | 0.0279 | 484 |
| | *-0.94* | *-0.52* | *0.01* | *0.95* | *-1.37* | *2.54* | | | | | |
| Model 4 | -0.0407 | -0.0038 | -0.0009 | 0.0026 | 0.0078 | | 0.0686 | | | 0.0072 | 484 |
| | *-0.77* | *-0.95* | *-0.15* | *0.76* | *0.14* | | *1.33* | | | | |
| Model 5 | -0.0383 | -0.002 | -0.001 | 0.0025 | -0.0074 | | | -0.5192 | | 0.0116 | 484 |
| | *-0.73* | *-0.49* | *-0.16* | *0.73* | *-0.14* | | | *-1.61* | | | |
| Model 6 | -0.0418 | -0.0038 | -0.0001 | 0.0027 | -0.1051 | 0.1348 ** | 0.0639 | | | 0.0311 | 484 |
| | *-0.83* | *-0.95* | *-0.01* | *0.82* | *-1.45* | *2.51* | *1.25* | | | | |
| Model 7 | -0.0394 | -0.002 | -0.0001 | 0.0026 | -0.1181 | 0.1334 ** | | -0.4869 | | 0.0349 | 484 |
| | *-0.79* | *-0.51* | *-0.02* | *0.8* | *-1.64* | *2.52* | | *-1.59* | | | |
| Model 8 | -0.0309 | -0.0039 | -0.0011 | 0.0019 | -0.0184 | | 0.0727 | -0.5332 | | 0.0157 | 484 |
| | *-0.59* | *-0.98* | *-0.18* | *0.58* | *-0.33* | | *1.42* | *-1.65* | | | |
| Model 9 | -0.0365 | -0.0036 | -0.0002 | 0.0023 | -0.1164 | 0.1329 ** | 0.0707 | | 0.0001 | 0.0342 | 484 |
| | *-0.73* | *-0.9* | *-0.03* | *0.71* | *-1.59* | *2.5* | *1.37* | | *1.29* | | |
| P Panel B: Pre-Crisis Period (2007/01-2008/06) | | | (360 observations) | | | | | | | | |
| Model 1 | 0.0033 * | -0.0065 * | | | | | | | | 0.0083 | 360 |
| | *1.67* | *-1.82* | | | | | | | | | |
| Model 2 | 0.0352 | -0.0073 * | -0.0024 | -0.0022 | -0.0785 | 0.0589 | 0.0521 | -0.8439 ** | | 0.032 | 360 |
| | *0.71* | *-1.85* | *-0.35* | *-0.68* | *-1.05* | *0.89* | *0.84* | *-2.15* | | | |
| Model 3 | 0.0337 | -0.0066 * | -0.0025 | -0.002 | -0.0463 | 0.0601 | | | | 0.0148 | 360 |
| | *0.73* | *-1.71* | *-0.37* | *-0.66* | *-0.63* | *0.89* | | | | | |
| Model 4 | 0.0291 | -0.0072 * | -0.0027 | -0.0017 | -0.0081 | | 0.0398 | | | 0.0115 | 360 |
| | *0.58* | *-1.79* | *-0.41* | *-0.55* | *-0.13* | | *0.63* | | | | |
| Model 5 | 0.046 | -0.0065 * | -0.0025 | -0.0028 | -0.0309 | | | -0.813 ** | | 0.0254 | 360 |
| | *0.99* | *-1.73* | *-0.38* | *-0.92* | *-0.5* | | | *-2.08* | | | |
| Model 6 | 0.0273 | -0.0072 * | -0.0025 | -0.0016 | -0.0494 | 0.0591 | 0.0375 | | | 0.0159 | 360 |
| | *0.55* | *-1.78* | *-0.38* | *-0.51* | *-0.65* | *0.87* | *0.59* | | | | |
| Model 7 | 0.0438 | -0.0065 * | -0.0023 | -0.0026 | -0.0732 | 0.0602 | | -0.8134 ** | | 0.0299 | 360 |
| | *0.95* | *-1.73* | *-0.35* | *-0.88* | *-1.01* | *0.92* | | *-2.06* | | | |
| Model 8 | 0.037 | -0.0073 * | -0.0026 | -0.0023 | -0.0374 | | 0.0545 | -0.8449 ** | | 0.0277 | 360 |
| | *0.74* | *-1.87* | *-0.39* | *-0.72* | *-0.6* | | *0.88* | *-2.17* | | | |
| Model 9 | 0.0305 | -0.007 * | -0.0026 | -0.0019 | -0.0618 | 0.0591 | 0.0487 | | 0.0001 | 0.0204 | 360 |
| | *0.62* | *-1.74* | *-0.38* | *-0.59* | *-0.81* | *0.88* | *0.77* | | *1.3* | | |
| Panel C: Financial Crisis Period (2008/07-2008/12) | | | (124 observations) | | | | | | | | |
| Model 1 | -0.0065 | 0.0065 | | | | | | | | 0.002 | 124 |
| | *-1* | *0.53* | | | | | | | | | |
| Model 2 | -0.2025 | 0.0123 | 0.0130 | 0.0130 | -0.1937 | 0.2096 ** | -0.3396 | -0.2010 | | 0.088 | 124 |
| | *-1.17* | *1.01* | *1.08* | *1.12* | *-1.40* | *2.27* | *-1.16* | *-0.53* | | | |
| Model 3 | -0.1837 | 0.0088 | 0.0145 | 0.0117 | -0.1845 | 0.2106 ** | | | | 0.0774 | 124 |
| | *-1.06* | *0.70* | *1.34* | *1.01* | *-1.37* | *2.30* | | | | | |
| Model 4 | -0.2149 | 0.0104 | 0.0100 | 0.0139 | 0.0159 | | -0.3460 | | | 0.0263 | 124 |
| | *-1.13* | *0.81* | *0.91* | *1.10* | *0.17* | | *-1.20* | | | | |
| Model 5 | -0.1859 | 0.0074 | 0.0097 | 0.0119 | -0.0044 | | | -0.3039 | | 0.0209 | 124 |
| | *-0.99* | *0.56* | *0.88* | *0.95* | *-0.05* | | | *-0.75* | | | |
| Model 6 | -0.2072 | 0.0122 | 0.0137 | 0.0133 | -0.1830 | 0.2112 ** | -0.3521 | | | 0.0865 | 124 |
| | *-1.19* | *1.00* | *1.19* | *1.15* | *-1.37* | *2.29* | *-1.22* | | | | |
| Model 7 | -0.1790 | 0.0090 | 0.0136 | 0.0114 | -0.1973 | 0.2087 | | -0.2438 | | 0.0796 | 124 |
| | *-1.04* | *0.72* | *1.21* | *0.99* | *-1.41* | *2.28* | | *-0.66* | | | |
| Model 8 | -0.2087 | 0.0106 | 0.0091 | 0.0135 | 0.0000 | | -0.3296 | -0.2627 | | 0.0288 | 124 |
| | *-1.10* | *0.82* | *0.79* | *1.07* | *0.00* | | *-1.13* | *-0.64* | | | |
| Model 9 | -0.2063 | 0.0123 | 0.0135 | 0.0133 | -0.1871 | 0.2103 ** | -0.3423 | | 0.0000 | 0.0868 | 124 |
| | *-1.18* | *1.01* | *1.14* | *1.13* | *-1.33* | *2.29* | *-1.15* | | *0.20* | | |

**Table 3.7 The VAR estimation results of intraday index return and open-buy put-call ratios of different classes of investors**

| | Explanatory variables | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | *Lagged R* | | *Lagged PC_r(d)* | | *Lagged PC_r(f)* | | *Lagged PC_r(i)* | |
| Dependent variable | Lag 1 | Lag 2 | Lag 1 | Lag 2 | Lag 1 | Lag 2 | Lag 1 | Lag 2 |
| Panel A: Full Period (2007/01-2008/12) | | (4354 observations) | | | | | | |
| *R* | 0.0006 | 0.0034 | -0.0136 ** | 0.0064 | -0.0026 | 0.0048 | 0.0068 | -0.0031 |
| | *0.0379* | *0.2245* | *-2.1436* | *1.0150* | *-0.4278* | *0.7701* | *0.3603* | *-0.1646* |
| *PC_r(d)* | 0.0535 | 0.0199 | 0.3114 *** | 0.1321 *** | 0.0371 ** | 0.0009 | 0.0179 | -0.0820 * |
| | *1.4831* | *0.5525* | *20.6539* | *8.7617* | *2.5224* | *0.0645* | *0.3974* | *-1.8169* |
| *PC_r(f)* | 0.0034 | 0.0075 | 0.0142 | 0.0248 | 0.4438 *** | 0.1359 ** | -0.0196 | -0.0105 |
| | *0.0930* | *0.2030* | *0.9185* | *1.6061* | *29.4967* | *9.0340* | *-0.4243* | *-0.2269* |
| *PC_r(i)* | -0.0017 | 0.0023 | -0.0115 ** | 0.0038 | -0.0015 | 0.0023 | 0.4850 *** | 0.2224 *** |
| | *-0.1421* | *0.1922* | *-2.3270* | *0.7616* | *-0.3186* | *0.4820* | *32.7232* | *14.9880* |
| Panel B: Pre-Crisis Period (2007/01-2008/06) | | (3238 observations) | | | | | | |
| *R* | 0.0008 | 0.0036 | -0.0173 ** | 0.0082 | -0.0029 | 0.0056 | 0.0099 | -0.0018 |
| | *0.0469* | *0.2029* | *-2.0657* | *0.9713* | *-0.3654* | *0.7034* | *0.3868* | *-0.0686* |
| *PC_r(d)* | 0.0594 | 0.0259 | 0.2877 *** | 0.1152 ** | 0.0419 ** | -0.0100 | 0.0580 | -0.0634 |
| | *1.6143* | *0.7044* | *16.4130* | *6.5723* | *2.5002* | *-0.5944* | *1.0889* | *-1.1899* |
| *PC_r(f)* | 0.0062 | 0.0099 | 0.0213 | 0.0368 ** | 0.4358 *** | 0.1348 *** | -0.0165 | -0.0297 |
| | *0.1612* | *0.2576* | *1.1679* | *2.0186* | *25.0006* | *7.7311* | *-0.2970* | *-0.5351* |
| *PC_r(i)* | -0.0019 | 0.0020 | -0.0118 ** | 0.0107 * | -0.0058 | 0.0019 | 0.4566 *** | 0.2433 *** |
| | *-0.1569* | *0.1687* | *-2.0880* | *1.8930* | *-1.0838* | *0.3533* | *26.6924* | *14.2148* |
| Panel C: Financial Crisis Period (2008/07-2008/12) | | | (1116 observations) | | | | | |
| *R* | 0.0337 | 0.0187 | -0.0008 | -0.0009 | -0.0020 | 0.0016 | -0.0009 | -0.0047 |
| | *1.0969* | *0.6098* | *-0.7139* | *-0.7702* | *-1.6071* | *1.2765* | *-0.2803* | *-1.3972* |
| *PC_r(d)* | -3.0304 *** | -1.2417 | 0.3538 *** | 0.1665 *** | 0.0086 | 0.0602 * | -0.0417 | -0.1150 |
| | *-3.9259* | *-1.6080* | *11.8909* | *5.6660* | *0.2716* | *1.9119* | *-0.4984* | *-1.3678* |
| *PC_r(f)* | -1.3355 * | -2.0706 *** | -0.0033 | 0.0009 | 0.4199 *** | 0.0933 *** | -0.0590 | -0.0291 |
| | *-1.8038* | *-2.7956* | *-0.1161* | *0.0327* | *13.8123* | *3.0868* | *-0.7354* | *-0.3613* |
| *PC_r(i)* | 0.3738 | -0.0879 | -0.0058 | -0.0169 | 0.0130 | -0.0026 | 0.5599 *** | 0.1531 *** |
| | *1.3582* | *-0.3193* | *-0.5475* | *-1.6139* | *1.1527* | *-0.2346* | *18.7651* | *5.1072* |

Notes: This table consists of three panels with each corresponding to the results of full sample period (Panel A) and two sub-periods, pre-crisis period from 2007/01/02 to 2008/06/30 (Panel B) and financial crisis period from 2008/07/01 to 2008/12/31. *R* denotes the daily TAIEX return. PC_r(d), PC_r(f), and PC_r(i) are the put-call ratios of domestic institutional, foreign institutional, and individual investors, respectively. VAR (2) is our best model based on the selection criteria, AIC and SBC. The regression coefficients for the two lags and t-statistics are reported below the coefficient. One, two, and three asterisks (*) indicate t-values are significant at the 0.1, 0.05 and 0.01 levels, respective.

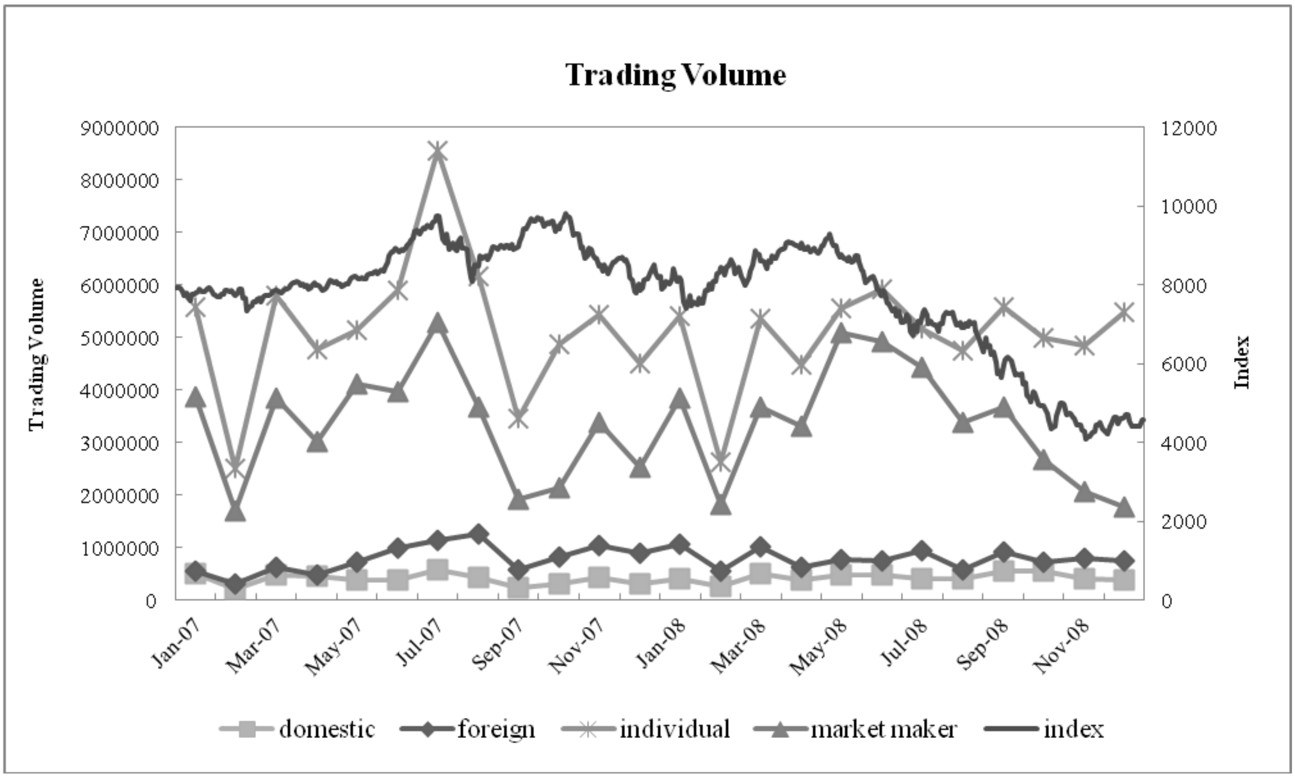


**Figure 3.1 Overall monthly trading volumes for four different classes of investors and the TWSE weighted stock index from 2007/1 to 2008/12**

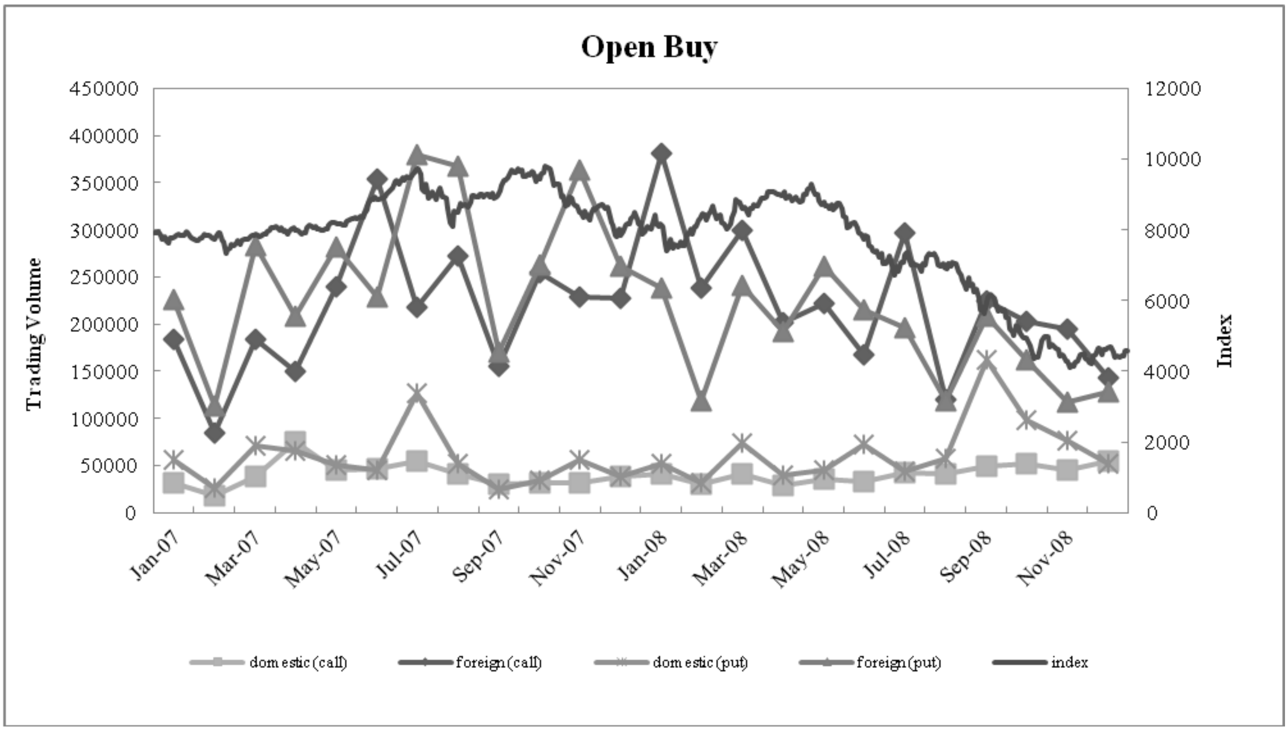



**Figure 3.2 Overall trading volumes (monthly) for open-buy options contracts of domestic and foreign institutional investors and the TWSE weighted stock index from 2007/1 to 2008/12**

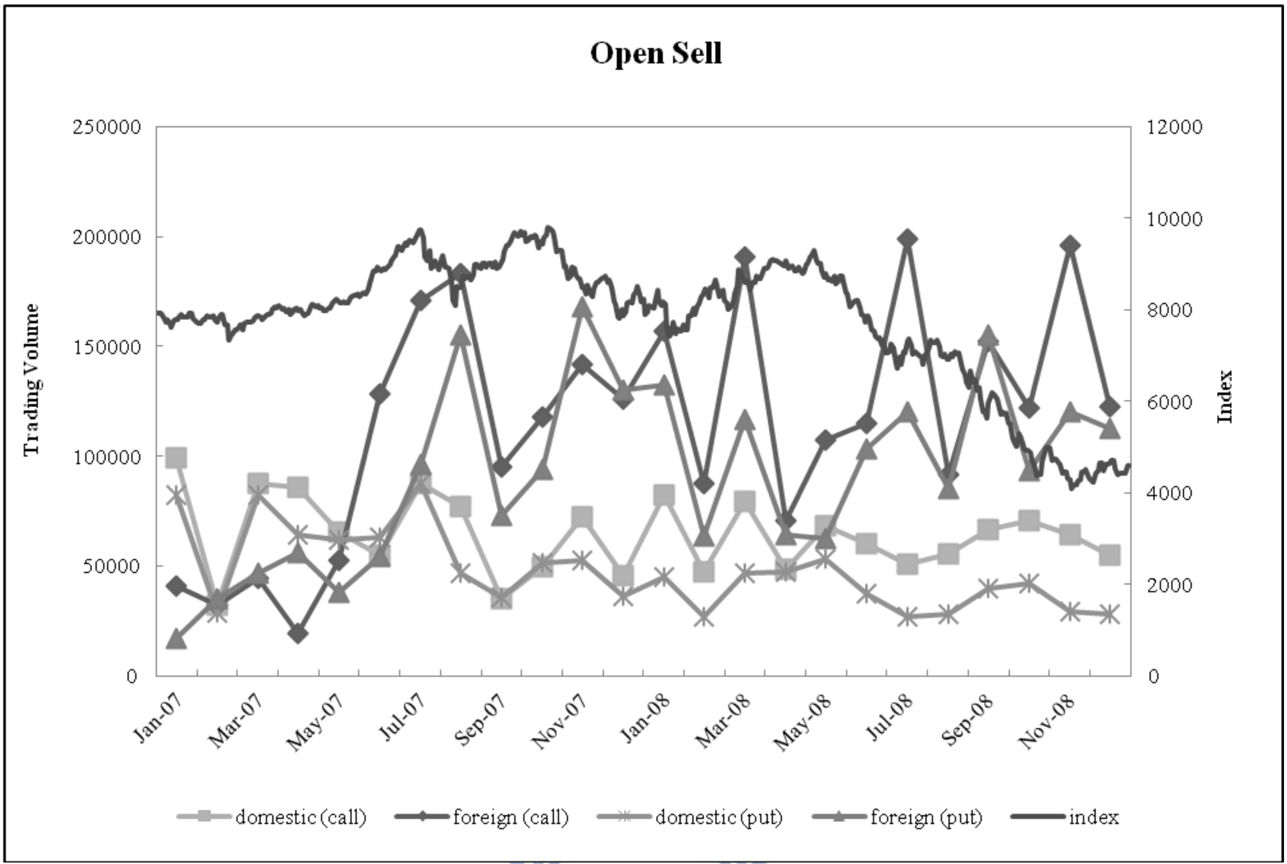



**Figure 3.3 Overall trading volumes (monthly) for open-sell options contracts of domestic and foreign institutional investors and the TWSE weighted stock index from 2007/1 to 2008/12**

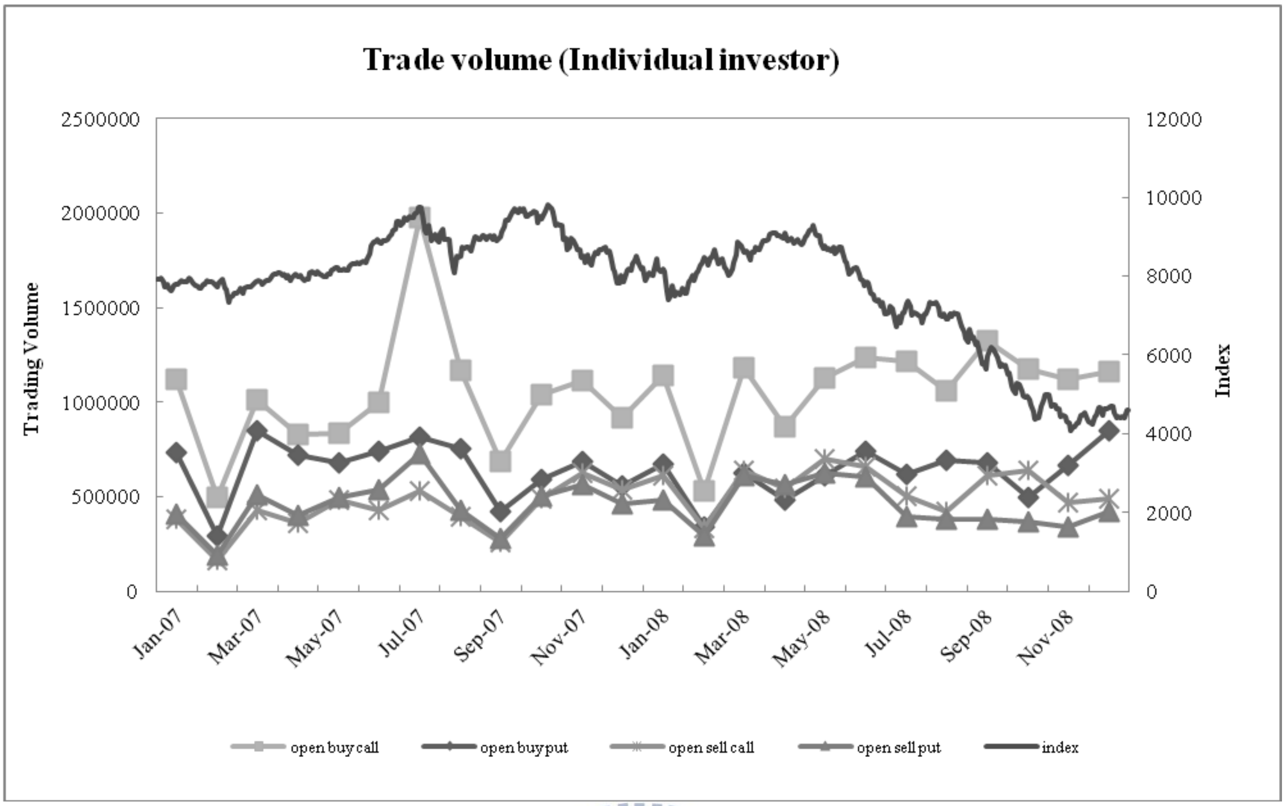


**Figure 3.4 Overall trading volumes (monthly) for open-buy and open-sell options contracts of individual investors and the TWSE weighted stock index from 2007/1 to 2008/12**

# Chapter4. Conclusions

Realized volatility plays an important role in recent application of finance. Also, the option trading volume in Taiwan ranked the fifth most frequently traded index options in the world. Therefore, this dissertation investigates two important issues in the financial market, including realized volatility for option pricing and index option volume for predicting future index return.

The first issue in this dissertation is to analyze SPX from July 3, 2007 to December 31, 2008 and show that the out-of-sample valuation errors from the HAR-type models are lower than those from other models, including the NGARCH option pricing model that has been documented as the best model in pricing options among GARCH-type models, and the Black-Scholes model that has been regarded as the traditional benchmark model for many literatures. The HAR-type models successfully predict out-of-sample option prices probably because they are based on realized volatilities, which are closer to expected volatility (VIX) in financial markets. Moreover, in times of extremely turmoil (it defines the period from September 15 to December 31, 2008 in our paper), the Log-HAR performs better than the HARG model in mid-term and long-term contracts, while worse than the HARG model in short-term puts. As a result, it seems to support that the model constructed based on realized volatilities and with simpler framework could value option prices more accurately in the fluctuant period as contracts with longer than 46 days to expiration. During the rather unstable period (it defines the period from July 3, 2007 to September 14, 2008), it again holds up the superiority of the HAR-type models over other models, and the HARG model performs better than the Log-HAR model. Overall, it seems to exist the mixed result between the Log-HAR and the HARG models in pricing options during very turmoil or rather unstable periods.

The second issue in this dissertation is that empirical evidence from daily data frequency suggests that unlike the findings of Chang et al. (2009), the put-call ratios of foreign institutional investors only have marginally significant predictive power for next-day TAIEX returns prior to the 2008 global financial tsunami and do not have any useful information content for future index returns during 2008 financial crisis. In contrast, domestic institutional investors have stronger prediction capability for next-day TAIEX returns before the global financial crisis. These results are robust even after controlling for some variables related to TAIEX returns.

The intraday VAR analyses reveal that the put-call ratios of domestic institutional and foreign institutional investors are positively correlated during the pre-crisis period, suggesting that these two groups of investors adopt similar options trading strategies. During the 2008 global financial crisis, no trader type existed that was able to predict TAIEX returns. The TAIEX returns lead the put-call ratios of domestic institutional and foreign institutional investors for 30 to 60 minutes, indicating that these two classes of traders closely watch and react to the changes in the market index. Finally, individual investors tend to adopt contrarian option strategies and do not exhibit any predictive ability for future index returns. Overall, it is apparent that domestic institutional investors are becoming more market savvy in predicting future index returns during the more recent period of 2007 to 2008. Our empirical results suggest that domestic institutional investors are becoming increasingly better informed as the host-country security market matures over time.

To conclude, this dissertation provides detailed insights into the issues in realized volatility for option pricing and index option volume for predicting future index return. Also, institutional managers and investors could utilize the results of this paper to set up the trading strategies.

# Reference


Andersen, T.G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized return volatility. *Journal of Financial Economics*, 61, 43–76.

Andersen, T.G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96, 42–55.

Andersen, T.G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica*, 71, 579–625.

Andersen, T.G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *The Review of Economics and Statistics*, 89, 701–720.

Anthony, J.H., (1998). The interrelation of stock and options market trading-volume data. *Journal of Finance* 43, 949-964.

Baillie, R., T. Bollerslev, and H. Mikkelsen (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74, 3–30.

Bakshi, G., C. Cao, and Z. Chen (1997). Empirical performance of alternative option pricing models. *Journal of Finance*, 52, 2003–2049.

Barber, B.M., Y.T. Lee, Y.J. Liu, T. Odean (2009). Just How Much Do Individual Investors Lose by Trading? *Review of Financial Studies*, 22, 151-186.

Bates, D.S. (1995). Testing option pricing models. The Wharton School of the University of Pennsylvania, Unpublished Manuscript.

Battalio, R. and P. Schultz (2006). Options and the bubble. *Journal of Finance*, 61, 2071–2102.

Bertholon, H., A. Monfor, and F. Pegorara (2008). Econometric asset pricing modeling. *Journal of Financial Econometrics*, 73, 151–184.

Bhattacharya, M. (1987). Price changes and related securities: The case of call options and stocks. *Journal of Financial and Quantitative Analysis*, 22, 1-15.

Black, F., and M. Scholes (1972). The valuation of option contracts and a test of market efficiency. *Journal of Finance*, 27, 399–417.

Black, F. and M. Scholes (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81, 637–653.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31, 307–327.

Bollerslev, T. and H. Mikkelsen (1999). Long-term equity anticipation securities and stock market

volatility dynamics. *Journal of Econometrics*, 92, 75–99.
Burghardt, G. and W. Acworth (2009). Annual Volume Survey: 2008 A Wild Ride. Futures Industry, (March 2009), 16-29.

Chan, K., Y. Chang, and P. Lung (2009). Informed trading under different market conditions and moneyness: Evidence from TXO options. *Pacific-Basin Finance Journal*, 17, 189-208.

Chan, K.C., Y.P. Chang, and H. Johnson (1993). Why option prices lag stock prices: A trading based explanation. *Journal of Finance*, 48, 1957-1967.

Chang, C.C., P.F. Hsieh, and Y.H. Wang (2010). Information content of options trading volume for future volatility: Evidence from the Taiwan options market. *Journal of Banking and Finance*, 34, 174-183.

Chang, C.C., P.F. Hsieh, and H.N. Lai (2009). Do informed option investors predict stock returns? Evidence from the Taiwan stock exchange. *Journal of Banking and Finance*, 33, 757-764.

Copeland, T.E. (1976). A model of asset trading under the assumption of sequential information arrival. *Journal of Finance*, 31, 1149-1168.

Corsi, F. (2009). A simple long memory model of realized volatility. *Journal of Financial Econometrics*, 7, 174–196.

Corsi, F., N. Fusari, and D. La Vecchia (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics*, 107, 284–304.

Christoffersen, P., K. Jacobs, C. Ornthanalai, and Y. Wang (2008b). Option valuation with long-run and short-run volatility components. *Journal Financial Economic*, 90, 272–297.

CSI: Credit Crunch (2007). *The Economist*, 4-8.

Ding, Z., C. Granger, and R. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, 83–106.

Duan, J. C. (1995).The GARCH option pricing model. *Mathematical Finance*, 5, 13–32.

Duan, J. C. and H. Zhang (2001). Hang Seng index options around the Asian financial crisis: A GARCH approach. *Journal of Banking and Finance*, 25, 1989–2014.

Dumas, B., J. Fleming, and R. Whaley (1998). Implied volatility functions: Empirical tests. *Journal of Finance*, 53, 2059–2106.
Easley, D., M. O'Hara, P. Srinivas (1998). Option volume and stock prices: Evidence on where informed traders trade. *Journal of Finance*, 53, 431-465.

Eisenberg, L. and R. Jarrow (1994). Option pricing with random volatilities in complete markets. *Review of Quantitative Finance and Accounting*, 4, 5–17.

Feunou, B. and N. Meddahi (2009). Generalized Affine Models. Available at SSRN: http://ssrn.com/abstract=1367033.

Gourieroux, C. and J. Jasiak (2006). Autoregressive Gamma Process. *Journal of Forecasting*, 25, 129–152.

Griffin, J.M., F. Nardari, and R.M. Stulz (2004). Are daily cross-border equity flows pushed or pulled? *Review of Economics and Statistics*, 86, 641-657.

Gultekin, B., R. Rogalski, and S. Tinic (1982). Option pricing model estimates: Some empirical results. *Financial Management*, 11, 58–69.

Heston, S. and S. Nandi (2000). A closed-form GARCH option valuation model. *Review of Financial Studies*, 13, 585–625.

Hsieh, K. and P. Ritchken (2005). An empirical comparison of GARCH option pricing models. *Review of Derivatives Research*, 8, 129–150.

Kurova, A. and D. Lasser (2004). Price Dynamics in the Regular and E-Mini Futures Markets. *Journal of Financial and Quantitative Analysis*, 39, 2, 365-384.

Lakonishok, J., I. Lee, and A.M. Poteshman (2004). Behavioral finance and the option market. *Working Paper*. Department of Finance, University of Illinois at Urbana-Champaign.

Lee, C. F., T.-P. Wu, and R.-R. Chen (2004). The Constant Elasticity of Variance Models: New Evidence from S&P 500 Index Options." *Review of Pacific Basin Financial Markets and Policies*, 7, 173–190.

Lee, C. F., G.-H. Tzeng, and S.-Y. Wang (2005). A fuzzy set approach for generalized CRR model: An empirical analysis of S&P 500 index options. *Review of Quantitative Finance and Accounting*, 25, 255–275.

Lee, C.M.C. and M.J. Ready (1991). Inferring Trade Direction from Intraday Data. *Journal of Finance*, 46, 733-746.

Lehar, A., M. Scheicher, and C. Schittenkopf (2002). GARCH vs. stochastic volatility: Option pricing and risk management. *Journal of Banking and Finance*, 26, 323–345.

Lo, A. and C. Mackinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3, 175-205.

Manaster, S. and R. Rendleman (1982). Option prices as predictors of equilibrium stock prices. *Journal of Finance*, 37, 1043-1058.

Pan, J. and A.M. Poteshman (2006). The information in option volume for future stock prices. *Review of Financial Studies*, 19, 871-908.

Richards, A. (2005). Big fish in small ponds: The trading behavior and price impact of foreign investors in Asian emerging equity markets. *Journal of Financial and Quantitative Analysis*, 40, 1-27.

Santa-Clara, P. and S. Yan (2010). Crashes, volatility, and the equity premium. *Review of*

*Economics and Statistics*, 92, 435–451.

Shiu, Y.M., G.G. Pan, S.H. Lin, and T.C. Wu (2010). Impact of Net Buying Pressure on Changes in Implied Volatility: Before and After the Onset of the Subprime Crisis. *Journal of Derivatives*, 17, 4, 54-66.

Stentoft, L. (2008). Option pricing using realized volatility. *Working Paper* at CREATES, University of Copenhagen.

Stephan, J. and R. Whaley (1990). Intraday price change and trading volume relations in the stock and stock option markets. *Journal of Finance*, 45, 191-220.

Whaley, R. (1982). Valuation of American call options on dividend-paying stocks. *Journal of Financial Economics*, 10, 29–58.

Yuan, H., M. and H. Zhang (2003). An empirical investigation of the GARCH option pricing model: Hedging performance. *The Journal of Futures Markets*, 23, 1191–1207.

Zhang, L., Y. Ait-Sahalia, and P. A. Mykland (2005). A tale of two times scales: Determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association*, 100, 1394–1411.