The Intraday Option Information and Stock Return Predictability

I FAN CHIANG

11 April 2019

Abstract

The Intraday IVS contains certain imformation about intraday stock return.

1 Introduction

Probabilistic inference has become a core technology in AI, largely due to developments in graph-theoretic methods for the representation and manipulation of complex probability distributions. Whether in their guise as directed graphs (Bayesian networks) or as undirected graphs (Markov random fields), *probabilistic graphical models* have a number of virtues as representations of uncertainty and as inference engines. Graphical models allow a separation between qualitative, structural aspects of uncertain knowledge and the quantitative, parametric aspects of uncertainty...

2 Empirical Methodology

2.1 Deviation From Put-Call Parity

The put-call parity relations derived from **stoll1969relationship** is a classical options pricing concepts in finance. It characterized the relationship that must exist between European put and call options with the identical underlying asset, expiration and strike prices. The equation must hold for European options on no-dividends paying underlying in a perfect market.

$$C - P = S - PV(K) \tag{1}$$

Where C and P represent call prices and put prices, and S is Stock price. With same maturity and exercise price K, the arbitrage opportunity would exist if the equation is not hold. The **black1973pricing** formula satisfies the put-call parity for any assumed value of the volatility parameter σ, therefore,

$$C^{BS}(\sigma) + PV(K) = P^{BS}(\sigma) + S \tag{2}$$

where $C^{BS}(\sigma)$ and $P^{BS}(\sigma)$ indicate Black-Scholes call and put prices, respectively. Combine the above equation, we can derive the equation

$$C^{BS}(\sigma) - C = P^{BS}(\sigma) - P \tag{3}$$

which implies that the implied volatility of call option and put option should be the same if all equation holds.

$$IV^{call} = IV^{put} \tag{4}$$

Of course the equation may not be hold once the option is American-style. However, our primary studies on SPX option is European style. Therefore, we do not need to consider

the dividend payment or early exercise case in our further research.

Clearly, the larger implied volatilities are the higher the call or put option prices claim. Following **amin2004index**, we refer to the difference between call and put implied volatilities as the call-put implied volatility spread(CPIV). It is suggested that a positive(negative) CPIV could be viewed as a bullish(bearish) signal regarding the underlying stock.

The aggregate Intraday CPIV are construced as following steps:

1. We first divided a single day into 14 of 5-minutes interval. Each interval contains the tick data from 2.5 minutes ahead and behind. For example, the 9 a.m. invertal, we collect valid data from 08:47:30 to 09:32:30 to represent this interval. As for open(close) interval, we choose to accumulate the full 5 minutes data behind(ahead)[1]

2. Similar to **xing2010does**, in each interval, we eliminate an option from the sample if its time to expiration is less than 10 days or more than a year, if its open interest is negative, if its moneyness[2] is smaller than 0.9 or more than 1.1.Furthermore, the option quotes must not violate basic no-arbitrage relations.

3. Then, in each time interval, there must be several valid option pairs with identical maturity(T) and exercise price(K). For each option pair we choose only one pair to be the representitive. For quote data, we choose the mean of best bid(β^*) best offer(α^*) as the chosen call and put price. For trade data, we capture the transaction that is closet to the centering time.

4. After collecting several time interval valid option pairs. we calculated the CPIV by applying,

$$CPIV_t = IV_t^{call} - IV_t^{put} = \sum_{j=1}^{N_t} w_{j,t}(IV_{j,t}^{call} - IV_{j,t}^{put}) \tag{5}$$

$CPIV_t$ denotes the implied volatility spread on interval t; $IV_{j,t}$ describe the B-S implied volatility, where the j refers to valid pairs of put and call options; $w_{j,t}$ are weights, there are N_t valid pairs of option on interval t.

In our study, the amount of intraday CPIV should be $38668(14 Intervals * 2762 Days)$. However, most of the quotes are short date contract(less than 10 days) in the middle of month so that we have mutiple missing values by this approach.

2.2 Data

In our analysis, the primary quote and trade intraday data for SPX option originates from CBOE MDR. The sample period studied is from January 2007 to December 2017. The option data includes trade date, trade time, expiration date, put-call code, exercise price, maturities, bid price, ask price, underlying price. The size of the sample data is about 1-TB around and the amount is about 1 billion. After we exclude the tick data fall outside the 5-minutes intervals, it remains about 40 million. Furthermore, we follow the approach from **ofek2004limited** to exclude the invalid option pairs. Finally, we have 1,692,542 valid volatility spreads for SPX option from January 2007 to December 2017.

---

[1]We collect the whole transaction data in 5 minutes for trade data. However, the size of quote data is extremely unbalanced in different interval, we restricted 1000 to 2000 quotes as maximum for call and put in collecting quote data.

[2]Moneyness is defined as the ratio of the strike price to the stock price.

Following the prior studies(**bollerslev2009expected**), several macroeconomic variable are suggested to be crucial and informative with regard to future returns. Specifically, we collect data of the default spread(between Moody's BAA and AAA corporate bond spreads), the term spread(between the 10-year T-Bond and 3-month T-bill yields) [3] as control variables in our regression analysis. The daily price of S& P 500 index is obtained from Bloomberg. The zero-coupon bond (ZCB) rate represent risk-free rate in B-S formula are collected from WRDS with different duration.

Table 1 contains descriptive statistics on the volatility spread.

2.3 Figure

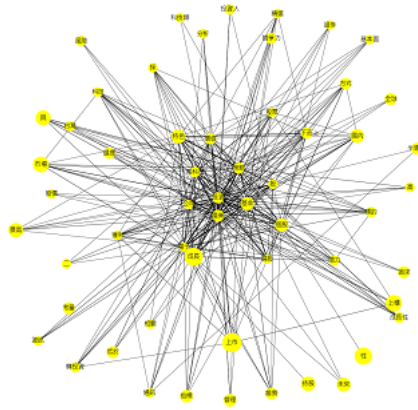1stock trading volume, and stock returns data are taken from CRSP for the construction of control measures.

Figure 1: Network ot words

2.4 Variable Definitions

All measures using implied volatility are calculated using all options with 90 or fewer days to expiration. Following **cremers2010deviations**, CPIV is the open interest-weighted call implied volatility less open interest weighted put implied volatility. CPIV STD is the standard deviation of CPIV over the past 20 days. IV is the open interest-weighted implied volatility. ME is stock price multiplied by the number of shares outstanding at the end of the measurement period and is reported in billions. Return is the daily, weekly, or monthly return on the day, week, or month, respectively, following CPIV measurement. Reversal is return in the calendar month prior to CPIV measurement. Momentum is the cumulative return in calendar months in brackets relative to the date of CPIV measurement. Turnover is monthly volume divided by the number of shares outstanding over calendar months prior to CPIV measurement with months designated in brackets. Illiquidity is the illiquidity measure of Amihud (2002), the absolute value of the return divided by the dollar 1trading volume averaged over calendar days prior to CPIV measurement, with days designated in brackets.

---

[3]The daily data are collected from the public website of the Federal Reserve Bank of St. Louis.

2.5 Descriptive Statistics

This table presents descriptive statistics for xxxx.

Table 1: Multivariate Regressions of SPX Return on IVS and Controls -2

	12:00	12:30	13:00	13:30	14:00	14:30	15:00
Intercept	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	-0.19	(-0.33)	-0.84	-0.50	-0.16	-0.35	-0.09
IVS	-0.03	-0.05	0.03	0.01	-0.01	0.00	-0.04
	(-1.08)	(-1.97)***	-0.87	-0.19	(-0.32)	(-0.02)	(-0.81)
DDEF	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(-0.82)	(-0.97)	(-0.55)	(-0.86)	(-0.68)	(-0.74)	(-1.01)
DTERM	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	-0.54	-0.96	-0.99	-1.21	-1.02	-1.10	-1.14
lag_return	-0.10	-0.10	-0.10	-0.10	-0.10	-0.11	-0.10
	(-3.13)***	(-3.04)***	(-3.07)***	(-3.01)***	(-2.89)***	(-3.26)***	(-3.23)***
Adj Rsquare(%)	1.24	1.99	1.25	1.14	1.02	1.27	1.36