AVIX: An Improved VIX Based on Stochastic Interest Rates and an Adaptive Screening Mechanism

Zhenlong Zheng, Zhengyun Jiang, and Rong Chen *

An improved model-free implied variation index (AVIX) is proposed in this article. The AVIX is developed under a generalized semi-martingale process with stochastic interest rates. An adaptive option screening mechanism is proposed to accommodate different market conditions. The effect of dividend protection is also considered. An empirical study of the China 50 ETF option market suggests that the AVIX is a better barometer of aggregate implied variation and investor sentiment than the traditional VIX. It reacts to market changes more rapidly and more sensitively. The AVIX also contains more information about future volatility and provides a more efficient forecast of future realized volatility. © 2016 Wiley Periodicals, Inc. Jrl Fut Mark 37:374–410, 2017

1. INTRODUCTION

Since it was introduced in 2003 by the Chicago Board Options Exchange (CBOE), the new model-free CBOE Volatility Index (VIX) has received increasing attention. Referred to as the "fear gauge," the VIX has been considered by many people to be a good barometer of aggregate implied volatility and investor sentiment in real time. Because of its broad use in the construction of trading and hedging strategies, and even in the creation of public policy, many other option markets have launched similar volatility indexes, such as the VSTOXX, the Nikkei 225 VI, the VKOSPI, the VHSI, and the TXO VIX. In June 2015, the Shanghai Stock Exchange (SSE) launched a model-free volatility index, the iVIX, for the China 50 ETF option market, with a computational procedure that is similar to that of the CBOE VIX.

However, it is observed that the iVIX often exhibits significant differences from most Black-Scholes (B-S) implied volatilities in the market, which may mean that its implicit information is biased. This phenomenon leads us to examine the theoretical underpinning and the calculation procedure of the traditional VIX with the goal of making it more general and adaptable.

The idea of a model-free VIX stems from the academic research in finance. The pioneering work of Breeden and Litzenberger (1978) laid the foundation for the subsequent

---

Zhenlong Zheng, Zhengyun Jiang, and Rong Chen are at the Department of Finance, School of Economics, Xiamen University. We thank the National Natural Science Foundation of China (Grant No. 71101121, 71371161, and 71471155) for its financial support. In particular, we thank Xiaoquan Liu for the helpful comments provided at the 2016 China Derivatives Markets Conference.

JEL Classification: G12, G13

*Correspondence author, Department of Finance, Xiamen University, 422 Simingnan Road 361005, Xiamen, Fujian, P.R.C. Tel: +8613860126618, Fax: 86-592-2186633, e-mail: aronge@xmu.edu.cn

Received August 2016; Accepted September 2016

The Journal of Futures Markets, Vol. 37, No. 4, 374–410 (2017)

Published online 9 December 2016 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/fut.21824

research regarding extracting risk-neutral distributions and model-free variance from option prices. In 1997, the working paper of Carr and Madan (2001) demonstrates that any payoff function with bounded expectation can be spanned by a continuum of out-of-the-money (OTM) European calls and puts. Demeterfi, Derman, Kamal, and Zou (1999) explore the pricing of variance swaps under a diffusion process with a constant interest rate. Later, the concept of the fair delivery value of future variance, which is developed in this article, becomes the basis for the CBOE VIX. Britten-Jones and Neuberger (2000) derive the risk-neutral expectation of future variance under a diffusion process with a zero interest rate. They call it the "model-free" implied variance because its derivation is not based on a particular option-pricing model. Jiang and Tian (2007) show that the model-free implied variances that were developed by Demeterfi et al. (1999) and by Britten-Jones and Neuberger (2000) are identical.

In 2003, the CBOE began to calculate the VIX using the model-free method. However, there is still room for improvement, both theoretically and practically. For example, Jiang and Tian (2005) and Carr and Wu (2009) extend the previous work in model-free variance to a jump-diffusion stochastic volatility model and a semi-martingale process, respectively. Jiang and Tian (2007) discuss in depth the truncation and discretization errors of the CBOE VIX.

However, we find that the prior studies in this field are all based on the assumption of nonrandom interest rates. We know that when interest rates are not random, there is no difference between forward prices and futures prices, and there is also no difference among the volatilities of spot prices, forward prices and futures prices. When interest rates are not random, the risk-neutral measure and the T-forward measure[1] are the same. That is the reason why the previous research in this field does not distinguish between these two measures. For the same reason, when studying model-free volatilities, the objects of some studies are spot prices (e.g., Britten-Jones and Neuberger, 2000; Demeterfi et al., 1999), some are forward prices (e.g., Jiang and Tian, 2005), and others are futures prices (e.g., Carr and Wu, 2009). What if the interest rates are stochastic? Should the calculation of the VIX be adjusted? In the earlier literature, such as in Carr and Madan (1998), Demeterfi et al. (1999), Britten-Jones and Neuberger (2000), and Jiang and Tian (2005, 2007, 2010), the term "variance" is used. In the more recent literature, the term "variance" has been gradually substituted by the term "quadratic variation," such as in Carr and Wu (2006, 2009), Bollerslev and Todorov (2011, 2014), Bollerslev, Todorov, and Xu (2015), and Andersen, Fusari, and Todorov (2015). Which term is more accurate? All of these questions are worth exploring.

Although the VIX is described as "model-free" implied volatility, it still assumes that there is no market friction. In reality, especially in China's financial market, there are severe short-selling restrictions and high transaction costs. These restrictions can lead to substantial biases in the value of the VIX. In the China 50 ETF option market, the small number and wide increments of strike prices make the situation worse. For example, on June 26, 2015 and August 24, 2015, the B-S implied volatilities of most 30-day China 50 ETF options were between 60% and 70%, respectively, and the 30-day at-the-money (ATM) B-S implied volatilities were 60.00% and 68.77%. However, the iVIX values on these 2 days were only 48.08% and 35.61%! In addition, China 50 ETF options are dividend protected. That is, when dividends are paid before the expiration of options, the option contracts will be adjusted. In such markets, will the calculation of the model-free implied volatility be adjusted? And how will such adjustments occur?

In this article, an improved model-free implied variation index, the AVIX, is developed. The AVIX inherits the model-free feature of the VIX because its derivation is still independent of any particular option pricing models, such as the B-S model, the Heston model (Heston, 1993), and the Bates model (Bates, 1996). Compared with the VIX, the AVIX extends the

[1]That is the measure under which the volatility of the zero-coupon bond with maturity T is the market price of risk.

existing literature in the following three aspects. First, interest rates are allowed to be stochastic. We derive the theoretical formula of the AVIX under a more general semi-martingale process with stochastic interest rates, stochastic volatility and jumps. Only the assumption of no-arbitrage is needed. Second, we explore the differences between the AVIX formulas for cases with and without dividend protection. Third, to address market friction in reality, an adaptive screening mechanism that can automatically filter the options with the best information quality to compute the AVIX is proposed. The following interpolations and extrapolations are conducted on the basis of these filtered option prices. This mechanism helps to make the best use of the information that is contained in the option market. In addition, in each step of the computations, alternative methods are proposed to help the AVIX to better cope with different market situations.

We find that when interest rates are allowed to be stochastic, the AVIX can exclude the effect of interest rate variation and reflect the variation information of the underlying asset more purely. When interest rates are stochastic, what we obtain is the model-free expected quadratic variation, instead of model-free expected variance.[2] In the dividend protection case, the calculation of the model-free variation does not change much. With the introduction of the adaptive screening mechanism, the AVIX is designed to accommodate different market conditions, and it is a more adaptable measure of model-free implied variation than the traditional VIX.

Following Christensen and Prabhala (1998), Jiang and Tian (2005), and Huang and Zheng (2009), we estimate the AVIX for the China 50 ETF option market from February 9, 2015 (the first trading day of the China 50 ETF option market) to January 21, 2016, and we investigate the information content and the forecasting ability of the AVIX. It is not surprising that the AVIX performs much better than the iVIX, the computation procedure of which is similar to the CBOE VIX. Compared to the iVIX, the AVIX can better reflect market aggregate implied variation and investor sentiment and more sensitively and rapidly react to market changes. Furthermore, the AVIX forecasts more efficiently, and it contains more information regarding future realized volatility. These findings are robust to different regression models, alternative estimation methods of the AVIX, alternative measures for realized volatility, and samples over different horizons.

The remainder of the article proceeds as follows. Section 2 sets forth the model and derives the theoretical formulas of the AVIX. The effects of stochastic interest rates and dividend protection are discussed in depth in this section. Section 3 introduces the calculation methodology of the AVIX and discusses methods with which it can be made more adaptable. The adaptive screening mechanism is proposed in this section. Section 4 introduces the data. The information content and forecasting ability of the AVIX are investigated in the empirical studies presented in Section 5. Section 6 presents robustness tests that were conducted to ensure the generality of our results. Our conclusions are presented in the final section.

2. THE MODEL

2.1. The Setup

Although interest rates are usually less volatile than stock prices, the randomness of interest rates is not negligible if we want to calculate a more accurate implied variation. Inspired by Bakshi, Cao, and Chen (1997), who extend the option pricing models of Heston (1993) and

[2]Therefore, when we are discussing the model-free implied variations in this article, the terms of “variation” or “quadratic variation” are used. But for the B-S implied volatilities, historical volatilities and realized volatilities, the terms of “volatility” and “variance” are still used.

Bates (1996) by allowing more stochastic parameters and jumps, we generalize the model of model-free implied variation to processes with stochastic interest rates.

We suppose that there is no arbitrage. According to the fundamental theorems of asset pricing, it means the existence of an equivalent martingale measure. Delbaen and Schachermayer (1994) show that the absence of arbitrage also implies that a financial asset price is a semi-martingale. To obtain a more general result, we use an Itô semi-martingale with stochastic volatility,[3] stochastic interest rates and jumps to describe the random evolution of the underlying price.

Let S_t and r_t denote the spot price of a tradable underlying asset and an instantaneous risk-free interest rate, respectively. σ_t denotes the diffusion component of instantaneous volatility. Under the risk-neutral measure $\mathbb{Q}$, their dynamics are expressed as

$$dS_t = \left(r_t - q_t\right)S_{t-}dt + \sigma_t S_{t-}\sqrt{1-\rho_t^2}dW_{1t}^{\mathbb{Q}} + \sigma_t S_{t-}\rho_t dW_{2t}^{\mathbb{Q}} + \int_{R^0} S_{t-}(e^x - 1)\left[\mu(dx, dt) - \upsilon_t^{\mathbb{Q}}(dx)dt\right], \quad (1)$$

$$dr_t = \omega(\varphi - r_t)dt + v_{r,t}dW_{2t}^{\mathbb{Q}}, \quad (2)$$

$$d\sigma_t = \kappa(\theta - \sigma_t)dt + v_{\sigma,t}dW_{3t}^{\mathbb{Q}} + \int_{R^0} k(x)\mu_\sigma(dx, dt). \quad (3)$$

where $W_{1t}^{\mathbb{Q}}$, $W_{2t}^{\mathbb{Q}}$, and $W_{3t}^{\mathbb{Q}}$ are $\mathbb{Q}$-standard Brownian motions; $W_{1t}^{\mathbb{Q}}$ and $W_{2t}^{\mathbb{Q}}$ are independent of each other, whereas $W_{3t}^{\mathbb{Q}}$ can be correlated with $W_{1t}^{\mathbb{Q}}$ or $W_{2t}^{\mathbb{Q}}$.

As illustrated by Equation (1), the process of the underlying asset can be decomposed into a drift term and three martingales. q_t is the instantaneous dividend rate.[4] $W_{1t}^{\mathbb{Q}}$ and $W_{2t}^{\mathbb{Q}}$ are both purely continuous martingales, and ρ_t is the instantaneous correlation coefficient between S_t and r_t.[5] The last term of Equation (1) is a purely discontinuous (jump) martingale. S_{t-} denotes the price just prior to any jump at time t. $(e^x - 1)$ is the percentage jump size, where x is the log-jump size. R^0 denotes the real line excluding zero, which means that the jump size will not be zero. The random counting measure $\mu(dx, dt)$ realizes a nonzero value for a given x if and only if the log-jump size falls in the range of $(x, x + dx)$ over the time interval $(t, t + dt)$. The process $\upsilon_t^{\mathbb{Q}}(dx)dt$ denotes the compensator (stochastic intensity) of the jumps for $\upsilon_t^{\mathbb{Q}}(dx)$ to be predictable, so that the last term of Equation (1) is the increment of a $\mathbb{Q}$ -pure jump martingale. In addition, we assume $\int_{R^0}(|x| \wedge 1)\upsilon_t^{\mathbb{Q}}(dx) < \infty$ to ensure that the jump process has finite variation.[6]

Equation (1) is a very general process and is used in many articles, such as Carr and Wu (2006, 2009), Bollerslev and Todorov (2011, 2014), and Bollerslev et al. (2015). Obviously, the Poisson jump diffusion process is a special case of Equation (1) because it only allows a finite number of jumps in each finite time interval, that is, it satisfies the finite variation condition.

Equation (2) is the one-factor mean-reverting process that is followed by the instantaneous risk-free interest rate. The long-run mean, ω, and the speed of mean reversion, φ, can be constant or time-dependent. The volatility parameter, $v_{r,t}$, however, is not

[3]Jumps in the volatility are also allowed.

[4]Because there is no dividend protection in most options markets, we first derive the model without considering dividend protection and then extend it to the case of dividend protection. In this article, q_t is nonrandom but can be time-varying. The relationship between q_t and the present value of cash dividends, I_t, is $I_t = S_t - S_t e^{-\int_t^T q_u du}$. If there is no dividend before the expiration of the option, $q_t = 0$.

[5]In Bakshi et al. (1997), although stochastic interest rates are allowed, ρ_t is specified to be zero. Our article allows ρ_t to be different from zero. It can be constant, time-dependent or even stochastic.

[6]For jumps of finite variation and locally integrability conditions, see Proposition II.2.9 on P.77 of Jacod and Shiryaev (2003).

specified. It can be constant, time-dependent, or stochastic. Therefore, Equation (2) is a quite general model that encompasses many models, such as Vasicek (1977); Cox, Ingersoll, and Ross (1985); Chan, Karolyi, Longstaff, and Sanders (1992); and time-inhomogeneous models such as Hull and White (1994).[7]

Equation (3) is also a quite general model for instantaneous diffusion volatility, σ_t. This is also a mean-reverting process, where θ, κ, and $v_{\sigma,t}$ are the long-run mean, the speed of mean reversion and instantaneous volatility of volatility, respectively. All of these three parameters can be constant or time-dependent. The last term, $\int_{R^0} k(x)\mu_\sigma(dx, dt)$, captures the jumps. If $k(x)$ and $v_{\sigma,t}$ are both zero, σ_t will be constant or time-dependent.[8] If only $k(x)$ is zero, then Equation (3) will be a diffusion process, such as the Heston model (Heston, 1993). The dynamics, Equations (1–3), will encompass Poisson jump diffusion process such as Bates (1996) and Bakshi et al. (1997). If $k(x) \neq 0$, the counting measure, $\mu_\sigma(dx, dt)$, can be equal to $\mu(dx, dt)$ in Equation (1), as in Todorov (2010), Bollerslev et al. (2015) and Andersen et al. (2015), or be different from $\mu(dx, dt)$, as in Duffie, Pan, and Singleton (2000). As illustrated in Jiang and Tian (2005) and Carr and Wu (2009), the randomness of the instantaneous diffusion volatility, σ_t, will be reflected in the model-free implied variation. The calculation of model-free variation, however, is not related to the specification of the instantaneous diffusion volatility, and no restrictions are needed.

2.2. The Dynamics of Zero-Coupon Bond Prices and Forward Prices

Let $B_t = B(t, T)$ denote the price of a zero-coupon bond with maturity T $(T > t)$[9] and \$1 face value. By the Itô-Doeblin lemma, the stochastic process of B_t under the risk-neutral measure $\mathbb{Q}$ is

$$\frac{dB_t}{B_t} = r_t dt + \beta_t v_{r,t} dW_{2t}^{\mathbb{Q}}, \tag{4}$$

where $\beta_t = \frac{\partial B}{\partial r}\frac{1}{B_t}$.

Consider a forward contract on the same underlying asset with the same maturity T. Let $F_t = F(t, T)$ denote the corresponding forward price. According to the no-arbitrage principle, we have

$$F_t = \frac{S_t e^{-\int_t^T q_u du}}{B_t} \tag{5}$$

By the Itô-Doeblin lemma for semi-martingales,[10] we can obtain the dynamics of F_t under measure $\mathbb{Q}$,

$$\begin{aligned}\frac{dF_t}{F_{t-}} = &\left(\beta_t^2 v_{r,t}^2 - \sigma_t \rho_t \beta_t v_{r,t}\right)dt + \sigma_t \sqrt{1-\rho_t^2}\, dW_{1t}^{\mathbb{Q}} + \left(\sigma_t \rho_t - \beta_t v_{r,t}\right) dW_{2t}^{\mathbb{Q}} \\ &+ \int_{R^0} (e^x - 1)\left[\mu(dx, dt) - \upsilon_t^{\mathbb{Q}}(dx)dt\right].\end{aligned} \tag{6}$$

Note that the dividend rate, q_t, does not appear in the process of the forward price.

[7]Because interest rates are less volatile than stock prices, it is widely accepted that there is no need to include jump terms in the process.

[8]This depends on whether k^* and μ^* are constant or time-dependent.

[9]Note that T is also the maturity of the options concerned.

[10]See Shreve (2004) and Carr and Wu (2009).

Next, we would like to change from the risk-neutral measure $\mathbb{Q}$ to the T-forward measure $\mathbb{T}$. As we know, under the risk-neutral measure $\mathbb{Q}$, the market price of risk is zero, and the relative asset price with respect to a money market account is a martingale. Under the T-forward measure $\mathbb{T}$; however, the market price of risk is the volatility of B_t, and the relative asset price with respect to the zero-coupon bond is a martingale.[11] Observing Equation (4), we can find that the volatilities of B_t with respect to $W_{1t}^{\mathbb{Q}}$ and the jump risk are both zero, and the volatility of B_t with respect to $W_{2t}^{\mathbb{Q}}$ is $\beta_t v_{r,t}$. According to the Girsanov theorem, we have

$$dW_{1t}^{\mathbb{Q}} = dW_{1t}^{\mathbb{T}}, dW_{2t}^{\mathbb{Q}} = dW_{2t}^{\mathbb{T}} + \beta_t v_{r,t} dt$$
$$\int_{R^0} (e^x - 1)\upsilon_t^{\mathbb{T}}(dx) - \int_{R^0} (e^x - 1)\upsilon_t^{\mathbb{Q}}(dx) = 0, \tag{7}$$

where $W_{2t}^{\mathbb{T}}$ is a standard Brownian motion under the T-forward measure.[12]

Combining Equations (6) and (7), we have

$$\frac{dF_t}{F_{t-}} = \sigma_t\sqrt{1-\rho_t^2}dW_{1t}^{\mathbb{T}} + (\sigma_t\rho_t - \beta_t v_{r,t})dW_{2t}^{\mathbb{T}} + \int_{R^0}(e^x - 1)\left[\mu(\mathrm{d}x, dt) - \upsilon_t^{\mathbb{T}}(\mathrm{d}x)dt\right]. \tag{8}$$

That is, the T-forward price F_t is a martingale under the T-forward measure, although the spot price follows a complicated process. In other words, the relationship

$$F_t = \frac{S_t e^{-\int_t^T q_u du}}{B_t} = \mathbb{E}_t^{\mathbb{T}}(F_T) = \mathbb{E}_t^{\mathbb{T}}(S_T) \tag{9}$$

holds,[13] where $\mathbb{E}_t^{\mathbb{T}}(\cdot)$ denotes the conditional expectation under the T-forward measure.

2.3. The Quadratic Variation of Asset Prices

As mentioned above, different terms are used to describe the second-order moment of asset prices—sometimes "variance," sometimes "quadratic variation." When the stochastic process is relatively simple, the expectation of the quadratic variation is only the variance. However, this is not necessarily the case if the dynamics are complicated.[14] In the next two subsections, we will demonstrate that the term of quadratic variation is more precise when we are discussing the model-free implied second-order moment.

Following Bollerslev and Todorov (2011), we define the return quadratic variation of the underlying asset as

$$QV_{t,T}^S = [\ln S, \ln S](t,T) = \int_t^T (d\ln S_u)^2,$$

the return quadratic variation of a forward as

$$QV_{t,T}^F = [\ln F, \ln F](t,T) = \int_t^T (d\ln F_u)^2,$$

[11]See Shreve (2004).
[12]See Bollerslev and Todorov (2011).
[13]At maturity, the forward price equals the spot price.
[14]For more information, see Du and Kapadia (2012).

the return quadratic variation of a zero-coupon bond as

$$QV_{t,T}^{B} = [\ln B, \ln B](t,T) = \int_{t}^{T} (d \ln B_u)^2,$$

and the return cross variation of the underlying asset with the zero-coupon bond as

$$CV_{t,T}^{B,S} = [\ln S, \ln B](t,T) = \int_{t}^{T} d \ln S_u \, d \ln B_u.$$

Note that the quadratic variation for a random process usually depends on the path along which it is computed and is a random variable.

Among all of these variations, what we really care about is the return quadratic variation of the underlying asset, $QV_{t,T}^{S}$. *Proposition 1* illustrates the relationship among these quadratic variations.

Proposition 1: The return quadratic variation of a forward over horizon $[t, T]$ can be decomposed into three parts, the return quadratic variation of the underlying asset, the return quadratic variation of a zero-coupon bond, and the return cross variation of the underlying asset with the zero-coupon bond. That is,

$$QV_{t,T}^{F} = QV_{t,T}^{S} + QV_{t,T}^{B} - 2CV_{t,T}^{B,S} \tag{10}$$

Proof of Proposition 1: Applying the Itô-Doeblin lemma for semi-martingales to Equation (8), we have

$$\begin{aligned} d \ln F_t = &-\frac{1}{2}\left(\sigma_t^2 + \beta_t^2 v_{r,t}^2 - 2\sigma_t \rho_t \beta_t v_{r,t}\right)dt + \sigma_t \sqrt{1-\rho_t^2}\, dW_{1t}^{\mathbb{T}} + (\sigma_t \rho_t - \beta_t v_{r,t})dW_{2t}^{\mathbb{T}} \\ &+ \int_{R^0} x \mu(dx, dt) - \int_{R^0} (e^x - 1)\upsilon_t^{\mathbb{T}}(dx)dt \end{aligned} \tag{11}$$

This means that the return quadratic variation of a forward is

$$QV_{t,T}^{F} = \int_{t}^{T} \left(\sigma_u^2 + \beta_u^2 v_{r,u}^2 - 2\sigma_u \rho_u \beta_u v_{r,u}\right)du + \int_{t}^{T} \int_{R^0} x^2 \mu(dx, du)$$

Because the quadratic variation of one variable is the same under different probability measures, we can use the processes that are followed by the log underlying price and the log zero-coupon bond price under the risk-neutral measure[15] to obtain their quadratic variations,[16]

$$QV_{t,T}^{S} = \int_{t}^{T} \sigma_u^2 du + \int_{t}^{T} \int_{R^0} x^2 \mu(dx, du),$$

$$QV_{t,T}^{B} = \int_{t}^{T} \beta_u^2 v_{r,u}^2 du.$$

[15]Apply the Itô-Doeblin lemma to Equations (1) and (4), and we can obtain the processes of the log underlying price and the log zero-coupon bond price.

[16]Note that the dividend rate, q_t, does not appear in $QV_{t,T}^{S}$.

And the return cross variation of the underlying asset with the zero-coupon bond is

$$CV_{t,T}^{B,S} = \int_t^T \rho_u \sigma_u \beta_u v_{r,u} du.$$

Proposition 1 is proved.

Proposition 1 demonstrates that the introduction of stochastic interest leads to the existence of the return quadratic variation of a zero-coupon bond and the return cross variation of the underlying asset with the zero-coupon bond, and consequently, leads to the existence of a difference between the quadratic variations of the underlying asset and the forward. If interest rates are nonrandom, the quadratic variation of the underlying asset and the forward are identical. This is why the prior studies use the forward or futures prices to calculate the implied quadratic variation of the underlying asset. *Proposition 1* also shows that when interest rates are stochastic, we must exclude the effect of random interest rates to obtain clean variation information regarding the underlying asset.[17]

2.4. The Model-Free Implied Variation

Proposition 1 illustrates the relationship among quadratic variations of asset prices. In fact, what we care about is its ex-ante conditional expectation. *Proposition* 2 provides the formula for the conditional expectation of $QV_{t,T}^S$.

Proposition 2: Under the no-arbitrage condition, the time-t expected value of the return quadratic variation of the underlying asset over horizon $[t, T]$ under the T-forward measure, $\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right]$, can be approximated by the continuum of European call prices across all strikes $K > 0$ and at the same maturity date $T \geq t$,

$$\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right] = \mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^F\right] - \mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^B\right] + 2\mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right], \tag{12}$$

$$\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^F\right] = \frac{2}{B_t}\int_0^\infty \frac{c_t(T,K) - B_t \max(F_t - K, 0)}{K^2} dK - 2\varepsilon_t, \tag{13}$$

$$\varepsilon_t = \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \int_{R^0}\left(e^x - 1 - x - \frac{x^2}{2}\right)\upsilon_u^{\mathbb{T}}(dx)du\right], \tag{14}$$

where $c_t(T,K)$ denotes the time-t value of a call with strike price $K > 0$ and maturity $T \geq t$ and F_t is the forward price that is defined in Equation (5). The proof of *Proposition 2* is in the Appendix.

Using put-call parity (PCP), the first term on the RHS of Equation (13) can be rewritten as (Aït-Sahalia & Lo, 1998)

$$\begin{aligned} &\frac{2}{B_t}\int_0^\infty \frac{c_t(T,K) - B_t \max(F_t - K, 0)}{K^2} dK \\ &= \frac{2}{B_t}\left[\int_0^{F_t} \frac{c_t(T,K) - B_t(F_t - K)}{K^2} dK + \int_{F_t}^\infty \frac{c_t(T,K)}{K^2} dK\right], \\ &= \frac{2}{B_t}\left[\int_0^{F_t} \frac{p_t(T,K)}{K^2} dK + \int_{F_t}^\infty \frac{c_t(T,K)}{K^2} dK\right] \end{aligned} \tag{15}$$

[17] In addition, *Proposition 1* provides an explanation of a phenomenon in reality, that the realized variance of spot prices is different from that of futures prices. The main part of the difference originates from the randomness of interest rates.

which is the foundation of the CBOE VIX and the SSE iVIX. If the effect of interest rates is considered, a more precise ATM strike price for European options should be the current forward price, F_t, instead of the current spot price. Therefore, Equation (15) shows that the VIX is calculated by the set of European OTM option prices across all strikes $K > 0$ and at the same maturity date T.

Combining Equations (12–15), the economic meaning of *Proposition* 2 is clear. First, the VIX computation is essentially the expected quadratic variation. Next, with random interest rates and jumps, the main part of $\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right]$ is still the VIX. Three terms, however, are added. $-\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^B\right]$ and $2\mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right]$ reflect the influence of stochastic interest rates, whereas ε_t is the result of jumps. Finally, note that the calculation of $\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right]$ is not based any specific option pricing model. Instead, it is derived entirely from no-arbitrage conditions. It is thus still model-free.

2.5. What if There Is Dividend Protection?

Thus far, we have discussed the case of no dividend protection. In the China 50 ETF option market, the options are dividend protected. If the China 50 ETF pays dividends during the life of the option, the strike price and the contract size will be adjusted. For example, if a cash dividend of d dollars is paid on ex-dividend date τ ($t \leq \tau \leq T$), and the ex-dividend price at time τ is S^τ, then the option contract size will be $\lambda = 1 + d/S^\tau$ times the original one, and the strike price K will change to $\frac{K}{\lambda}$. Therefore, if there are several ex-dividend dates, $\tau_1, \tau_2, \cdots, \tau_n$, over horizon $[t,\ T]$, the payoff of a European call with dividend protection would be expressed as

$$\prod_i^n \lambda_i \cdot \max\left(S_T - K\Big/\left(\prod_i^n \lambda_i\right), 0\right) = \max\left(\prod_i^n \lambda_i S_T - K, 0\right) \tag{16}$$

Suppose that now is time t. Define the dividend-adjusted price at time u ($t \leq u \leq T$), S_u^* as

$$S_u^* = \begin{cases} S_u & u \in [t, \tau_1) \\ \lambda_1 S_u & u \in [\tau_1, \tau_2) \\ \lambda_2 \lambda_1 S_u & u \in [\tau_2, \tau_3) \\ \cdots & \\ \prod_i^n \lambda_i S_u & u \in [\tau_n, T] \end{cases}.$$

Note that at the initial time t, the dividend-adjusted price and the real price are the same. That is, $S_t^* = S_t$. The payoff of the European call in Equation (16) can be expressed by S_T^* as

$$\max\left(S_T^* - K, 0\right) \tag{17}$$

Equations (16) and (17) show that an option contract with dividend protection can be considered a contract that will be adjusted when the dividends are paid, or it can be regarded as an option contract on an underlying asset without dividends whose price is the dividend-adjusted price S_t^*. This is consistent with the analysis of Merton (1973a). Similar to Benninga, Björk, and Wiener (2002), the dynamics of the dividend-adjusted price under the risk-neutral measure $\mathbb{Q}$ in our article is[18]

$$\begin{aligned} dS_t^* &= r_t S_{t-}^* dt + \sigma_t S_{t-}^* \sqrt{1-\rho_t^2} dW_{1t}^{\mathbb{Q}} + \sigma_t S_{t-}^* \rho_t dW_{2t}^{\mathbb{Q}} \\ &+ \int_{R^0} S_{t-}^* (e^x - 1) \left[\mu(dx, dt) - \upsilon_t^{\mathbb{Q}}(dx) dt \right], \end{aligned} \tag{18}$$

With Equation (18), we can discuss the quadratic variation that is implied in the option prices with dividend protection. *Proposition* 3 provides the related conclusion.

Proposition 3: If one option is dividend protected, and the underlying asset will pay dividends during the life of the option, under the no-arbitrage condition, the time-*t* expected value of the return quadratic variation of the underlying asset over horizon $[t, T]$ under the T-forward measure, $\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right]$, can be approximated by the continuum of European call prices across all strikes $K > 0$ and at the same maturity date $T \geq t$ as

$$\begin{aligned} \mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right] &= \frac{2}{B_t} \int_0^\infty \frac{c_t(T,K) - B_t \max\left(F_t^* - K, 0\right)}{K^2} dK \\ &\quad - \mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^B\right] + 2\mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right] - 2\varepsilon_t, \end{aligned} \tag{19}$$

where

$$F_t^* = \frac{S_t^*}{B_t} = \frac{S_t}{B_t}. \tag{20}$$

The proof of *Proposition* 3 can be found in the Appendix.

Comparing *Propositions* 2 and 3, the formula for $\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^S\right]$ hardly changes. The only change is that the forward price, F_t, defined by Equation (5),

$$F_t = \frac{S_t e^{-\int_t^T q_u du}}{B_t},$$

is substituted by F_t^*, defined by Equation (20). The only difference between F_t and F_t^* is the dividend. If there is no dividend during the life of the option, *Propositions* 2 and 3 coincide.

In this section, we derive the formula of the expected value of the return quadratic variation of the underlying asset under a generalized semi-martingale process with stochastic interest rates, stochastic volatility and jumps. The effect of dividend protection is also discussed. It is shown that what we obtain is a model-free expected quadratic variation under a T-forward measure. When the interest rates are stochastic, the quadratic variation of spot price, forward price and futures price are different. With the option prices, we can only obtain the approximate expectation of the quadratic variation of the forward price. To obtain clean quadratic variation information about the spot price, we need to exclude the influence of random interest rates and jumps. The formulas that have been proposed by the previous

[18]Strictly speaking, the volatility of the real price and the dividend-adjusted price should not be the same. However, if the dividend is quite small, the difference can be ignored. For more information, see Benninga et al. (2002) and Geske, Roll, and Shastri (1983).

studies and the formula that have been used in the calculation of the CBOE VIX and the SSE iVIX are all special cases in our article.

3. THE CALCULATION OF THE AVIX

In Section 2, we derived a general theoretical formula for model-free implied quadratic variation. However, to convert this formula to a good barometer of aggregate implied variation and investor sentiment, much work remains. As mentioned in the introduction, the existence of market friction can lead to substantial biases. In this section, we will discuss the measures that can be taken to alleviate the biases as much as possible. The corresponding implied variation index is named the AVIX, where A means adaptive. We hope that the AVIX is more adaptable to different market conditions than the traditional VIX. In this article, The China 50 ETF option market is taken as an example. However, the AVIX could also be used in other option markets because the adjustments that we discuss are all optional.

3.1. Choosing Options with the Best Information Quality

In the original formulas of model-free implied quadratic variation, Equations (13) and (19), only the call prices are concerned. In the calculation of the CBOE VIX, PCP is used to convert Equations (13–15), and OTM calls and puts are employed. For the CBOE, this practice is reasonable. Just as Jiang and Tian (2010) argue, the liquidity of in-the-money (ITM) options is usually worse than that of the OTM and ATM options in the SPX option market. That means that OTMs and ATMs have better information quality. Thus the VIX calculated from OTM and ATM prices could reflect the investor sentiment quite well.

In the calculation of the SSE iVIX, a similar sampling method is employed. In the case of severe market friction in the China 50 ETF option market; however, this practice may suffer from serious biases. Take the case of August 24, 2015 as an example. Until that day, the underlying asset, the China 50 ETF, had decreased in value on three consecutive trading days and had declined by 11% in the past week. According to the contract specifications, only two OTM puts could be traded at the opening bell on August 24, 2015. When the China 50 ETF nosedived by 9.86% on that day, there were no OTM puts at all! In such a case, the sampling rule means that only OTM call prices are used to calculate the iVIX. During that period, the index futures prices were significantly discounted due to the restrictions of short-selling and pessimistic sentiment. The PCP arbitrage of

$$-S_t = p_t - c_t - KB_t$$

naturally resulted in the phenomenon that calls were relatively cheaper than puts. For these reasons, it is not surprising that the iVIX, which only used OTM call prices, underestimated the real market implied variation severely on August 24, 2015. It was only 35.61%, whereas most of the B-S implied volatilities were between 60% and 70%.

Moreover, even in the non-extreme market situation, the sampling method that only takes OTM prices is not suitable in the China 50 ETF option market. Strict trading restrictions and high transaction costs lead to inactive trading and irregularity in liquidity. It is uncertain which options are more liquid than others. OTM options are not necessarily more liquid. We can often observe that ITM options have greater trading volumes than OTM options with the same strike prices.

A possible solution to this problem is to change the middle point of Equation (15). Theoretically, this is acceptable. Taking the case without dividend protection as an example,[19] we can further convert Equation (15) to

$$\begin{aligned}&\frac{2}{B_t}\left[\int_0^{F_t}\frac{p_t(T,K)}{K^2}dK+\int_{F_t}^{\infty}\frac{c_t(T,K)}{K^2}dK\right]\\&=\frac{2}{B_t}\left[\int_0^{M_t}\frac{p_t(T,K)}{K^2}dK+\int_{M_t}^{\infty}\frac{c_t(T,K)}{K^2}dK+\int_{M_t}^{F_t}\frac{KB_t-F_tB_t}{K^2}dK\right]\\&=\frac{2}{B_t}\left[\int_0^{M_t}\frac{p_t(T,K)}{K^2}dK+\int_{M_t}^{\infty}\frac{c_t(T,K)}{K^2}dK+B_t\left(\ln\frac{F_t}{M_t}+1-\frac{F_t}{M_t}\right)\right]\end{aligned}\tag{21}$$

Similar to Equation (15), PCP is employed, and no essential changes are made in Equation (21). The result is that some of the ITMs with strike prices falling into the interval of $[F_t, M_t]$ or $[M_t, F_t]$ are included, whereas some of the OTMs are excluded.

There are two downsides to this solution. First, it does not demonstrate how to determine a reasonable M_t, which leads to poor maneuverability and objectivity. Second, once M_t is determined, we must use options with strike prices in a given interval. Remember that the main problem is irregularity in liquidity. There is no guarantee that these added options are more liquid and contain more useful information. This seemingly good solution does not work.

To address the irregularity in liquidity, we propose a mechanism to automatically screen the options with the best information quality in real time; this method is called an "adaptive screening mechanism." Take the case with dividend protection as an example. When calculating the first right term of Equation (19),

$$\frac{2}{B_t}\int_0^{\infty}\frac{c_t(T,K)-B_t\max\left(F_t^*-K,0\right)}{K^2}dK,$$

we replace $c_t(T,K)$ with $\bar{c}_t(T,K)$, which is defined by

$$\bar{c}_t(T,K)=\begin{cases}c_t(T,K)\ , & V_{\text{call}}>150\%\cdot V_{\text{put}}\\ \dfrac{c_t(T,K)+p_t(T,K)+F_t{}^*B_t-KB_t}{2}, & \dfrac{1}{150\%}\cdot V_{\text{put}}\leq \text{V}_{\text{call}}\leq 150\%\cdot V_{\text{put}}\\ p_t(T,K)+F_t{}^*B_t-KB_t, & V_{\text{call}}\leq\dfrac{1}{150\%}\cdot V_{\text{put}}\end{cases}\tag{22}$$

where V_{call} and V_{put} are the trading volumes of calls and puts with strike price K and maturity T, respectively.

As illustrated in Equation (22), if the liquidity of calls is much better than that of puts, call prices will be used (the first choice). If the liquidity of puts is much better, put prices will be used and be converted to call prices using PCP[20] (the third choice). If there is not much difference in the liquidity of calls and puts, the average of both prices will be employed (the second choice). Here, the proxy for the liquidity is the trading volume of the options, and the threshold value is 150%.[21]

[19]Similar results can be obtained for the case with dividend protection.

[20]The appropriate PCP when there are market friction and dividend protection will be discussed in subsection 3.4.

[21]In the robustness test, other proxies of liquidity and different threshold values are tested.

It can be observed that the criterion to select options is no longer the moneyness but rather the trading volume, which is a more direct measure of liquidity. That is the main difference between our adaptive screening mechanism and the CBOE VIX method. With this mechanism, the options with the best liquidity and the best information quality will always be chosen automatically in real time. That is the reason for the inclusion of the term "adaptive" in AVIX.

3.2. Dealing with Truncation Errors and Discretization Errors

As shown in Equations (13) and (19), in the calculation of model-free implied quadratic variation, in theory, we need to integrate the call prices over a continuum of the strike prices from zero to infinity. Market option prices, however, are available only over a subset of this range and at a finite set of values. For example, the integral in Equation (19) will be approximated with numerical integration techniques as follows:[22]

$$\int_0^{\infty} \frac{\bar{c}_t(T,K) - B_t \max\left(F_t^* - K, 0\right)}{K^2} dK \approx \sum_{i=1}^{m} \frac{\bar{c}_t(T,K_i) - B_t \max\left(F_t^* - K_i, 0\right)}{K_i^2} \Delta K_i, \tag{23}$$

where m is the number of the strike prices. Obviously, such approximation induces truncation errors and discretization errors.

Due to the large number of strike prices in the SPX option market, truncation errors and discretization errors are not addressed in the implementation of the CBOE VIX. Jiang and Tian (2007), however, show that these errors are significant, even for the CBOE VIX. The strike prices in the China 50 ETF option market are much less than those in the SPX option market. We assume that if we do not address the truncation and discretization problem, the approximation errors could be significant.

In principle, the more strike prices, the better. Jiang and Tian (2005), however, demonstrate that truncation errors are negligible if the truncation points are more than two standard deviations (SDs) from the forward price, and the discretization errors are negligible when the strike price increment (ΔK) is no more than 0.35 SDs. Interpolations and extrapolations could be used to obtain more option prices based on the above criteria.[23]

There are three key points in the implementation process. First, interpolations and extrapolations should be implemented on those option prices filtered by the adaptive screening mechanism. This ensures that those fitted values also contain useful market information. Second, following prior research (Aït-Sahalia & Lo, 1998; Carr & Wu, 2009; Jiang & Tian 2005; Shimko, 1993), interpolations and extrapolations are applied to the corresponding B-S implied volatilities instead of option prices. Then, the fitted B-S implied volatilities are translated into option prices using a B-S model.[24] Third, also following the prior research, we implement cubic spline interpolation between the available strike prices. Additionally, because the original number of strike prices is too small and the smoothing feature is not stable, we use flat extrapolation instead of smoothing extrapolation. That is, the extrapolation values are equal to the endpoint values.

[22]The numerical integration of Equation (23) uses the rectangle rule. Jiang and Tian (2005) use the trapezoidal rule. The CBOE VIX uses an adjusted rectangle rule. Our robustness test in Section 6 shows that this small difference does not affect the conclusions.
[23]The standard deviation that is used is the B-S implied volatility of the option with the highest trading volume.
[24]This does not mean that we think the B-S model is the true pricing model. Here, the B-S model is merely used as a tool to provide a one-to-one mapping between European option prices and implied volatilities.

3.3. Temporal Interpolation

As in the CBOE VIX and the SSE iVIX, the AVIX measures 30-day expected quadratic variation. Most of the time, there are no options with an exact 30-day maturity in the market. Similar to the CBOE procedure, we apply linear interpolation between two selected maturities using the formula

$$\frac{T_2 - 30}{T_2 - T_1} IV_{T_1} + \frac{30 - T_1}{T_2 - T_1} IV_{T_2},$$

where IV_{T_1} and IV_{T_2} are computed using the first term of Equations (13) or (19). Taking Equation (19) as an example, we have

$$IV_{T_1} \equiv \frac{2}{B(t, T_1)} \sum_{i=1}^{m} \frac{\bar{c}_t(T_1, K_i) - B(t, T_1)\max(F^*(t, T_1) - K_i, 0)}{K_i^2} \Delta K_i$$

$$IV_{T_2} \equiv \frac{2}{B(t, T_2)} \sum_{i=1}^{m} \frac{\bar{c}_t(T_2, K_i) - B(t, T_2)\max(F^*(t, T_2) - K_i, 0)}{K_i^2} \Delta K_i .$$

Usually, T_1 (T_2) is the maturity that is closest to but shorter than (longer than) 30 days. However, if the options with the shortest maturity will expire in 3 days, T_1 and T_2 will be the 2 maturities that are closest to and longer than 30 days.

3.4. Determining F_t and F_t^*

When calculating Equations (13) or (19), an issue that must be considered is how to determine F_t or F_t^*.

We first consider the case without dividend protection, Equation (13). There are two methods to fix F_t. One is to use the market forward price (or the approximate market futures prices), and the other is to obtain the option implied forward price F_t using PCP,

$$c_t + KB_t = p_t + S_t e^{-\int_t^T q_u du} = p_t + F_t B_t. \tag{24}$$

If there is not much market friction, the results for F_t that are obtained in these two manners should be similar. This is the case in markets with high efficiency, for example, the U.S. market. The second method is used more frequently because it is more convenient and more consistent with the objective of obtaining option-implied information. The CBOE VIX employs the PCP-implied forward price. If there is much market friction; however, there will be a great difference between the results for F_t that are obtained in the two manners. For the same reason, the PCP-implied forward price is again a better choice.

In theory, whatever strike price we use, the PCP-implied forward price should be equal. This is not the case in reality; however, due to the existence of market friction. The market practice is to use ATM options because the liquidity of ATM options is usually quite high. According to PCP, the ATM call price should be equal to the ATM put price.[25] Therefore, the strike price, X, which is used to calculate the PCP-implied forward price, is the one at which the absolute difference between the call and the put prices is smallest; that is,

[25]The ATM strike price is the forward price. As discussed in Section 2.4, the forward price is a better candidate for the ATM strike price.

$$X = \arg\min_{K} |c_t(T, K) - p_t(T, K)|.$$

Next we consider the case with dividend protection, Equation (19). There are three methods to obtain F_t^*. One is to use the spot price and compute the value of F_t^* using Equation (20),

$$F_t^* = \frac{S_t^*}{B_t} = \frac{S_t}{B_t}.$$

The second is to utilize the relationship between F_t and F_t^*,

$$F_t{}^* = F_t \cdot e^{\int_t^T q_u du},$$

to find F_t^* from the market forward price (or the approximate futures price). The last is to obtain the option implied F_t^* using PCP in the case of dividend protection, as was proposed by Merton (1973b):

$$c_t + KB_t = p_t + S_t = p_t + F_t^* B_t. \tag{25}$$

Theoretically, the three F_t^* s should be equal. However, for the same reason of market friction, when the arbitrage mechanism does not function well, the option-implied F_t^* is also a better choice. Note that in the case of dividend protection, the ATM strike price is $K = F_t^*$.

3.5. The Jump Error Term ε_t

In both Equations (13) and (19), there is an error term, ε_t. As shown in Equation (14), ε_t comes from the jump term in Equation (1) and reflects the effect of the jumps. Its calculation is complicated. However, we will fortunately show that it is small enough to be ignored. Equation (14) suggests that this term is a conditional expectation. According to the specifications of Equation (1), we have $\int_{R^0} 1_{|x|>1} \upsilon_t^{\mathbb{T}}(\mathrm{d}x) < \infty$ and $\int_{R^0} x 1_{|x|<1} \upsilon_t^{\mathbb{T}}(\mathrm{d}x) < \infty$. To be more specific, when the jump size is large, the probability of a jump is very small. When the jump size is small, although the probability of a jump is quite large, $e^x - 1 - x - \frac{x^2}{2} = o(x^3)$ converges to zero when $|x| \to 0$. Therefore, the average error that is caused by jumps is sufficiently small to be ignored.[26] In fact, Jiang and Tian (2005) find that the third-order approximation error that is produced by jumps could be ignored in their Monte Carlo simulations,[27] which supports our analysis.

3.6. The Quadratic Variations Related to Interest Rates

In both Equations (13) and (19), $\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^B\right]$ and $\mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right]$ reflect the effect of stochastic interest rates. To be more specific, these two terms are the expected quadratic variation of the zero-coupon bond and the expected cross variation of the underlying asset with the zero-coupon bond under the T-forward measure.

[26]As long as there are jumps, the error term, ε_t, cannot be strictly zero. It contains information about higher-order variations, such as $\int_t^T (d \ln S_u)^3$. Hence, if the error term, ε_t is ignored, the AVIX will contain information about not only quadratic variations but also higher-order variations.
[27]See Table I and Figure 1 in Jiang and Tian (2005).

It is not easy to estimate these two terms. Following Bollerslev, Tauchen, and Zhou (2009), we suppose that both quadratic variations are martingales under the T-forward measure. That is,

$$\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^{B}\right] = QV_{t-w,t}^{B}, \mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right] = \mathrm{C}V_{t-w,t}^{B,S}, w = T - t.$$

With this simplification, the estimation of the expectation becomes the estimation of historical realized values. An argument for such a simplification is that either $QV_{t,T}^{B}$ or $CV_{t,T}^{B,S}$ is very small. The changes of these two terms will be smaller. Hence, assuming that these variations are martingales is an acceptable hypothesis.[28]

There are two methods to determine $QV_{t,T}^{B}$ and $CV_{t,T}^{B,S}$. One is to apply the realized variation convergence theorem that was proposed by Andersen and Bollerslev (1998) and Andersen, Bollerslev, Diebold, and Labys (2003) and to calculate $QV_{t,T}^{B}$ and $CV_{t,T}^{B,S}$ as follows:

$$\sum_{u=t-w}^{t} d\ln B_u \cdot d\ln B_u \to \int_{t-w}^{t} (d\ln B_u)^2 = QV_{t-w,t}^{B},$$

$$\sum_{u=t-w}^{t} d\ln S_u \cdot d\ln B_u \to \int_{t-w}^{t} d\ln S_u \, d\ln B_u = CV_{t-w,t}^{B,S}.$$

The second method is to use the conclusion of *Proposition 1*,

$$QV_{t-w,t}^{B} - 2CV_{t-w,t}^{B,S} = QV_{t-w,t}^{F} - QV_{t-w,t}^{S},$$

to directly obtain $QV_{t-w,t}^{B} - 2CV_{t-w,t}^{B,S}$ according to the difference between the historical realized variation of the forward, $QV_{t-w,t}^{F}$, and that of the underlying, $QV_{t-w,t}^{S}$. Because forward prices with a particular maturity may not be available, the first method is more practical and will be used in our empirical studies.[29]

Note that the zero-coupon bond that is concerned in $QV_{t,T}^{B}$ and $CV_{t,T}^{B,S}$ will expire at time T, and the price of a bond must converge to its face value at maturity. Therefore, only bonds with maturity t can be chosen to calculate the historical realized variation, $QV_{t-w,t}^{B}$ and $CV_{t-w,t}^{B,S}$ to ensure the same convergence feature.

3.7. The AVIX

Now we can construct a model-free implied quadratic variation index that considers random interest rates, stochastic volatility, jumps, dividend protection and market friction. We call it the "AVIX," where "A" means adaptive. In the case with dividend protection,[30] the formula is

$$\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^{S}\right] \approx \frac{2}{B_t}\sum_{i=1}^{m}\frac{\bar{c}_t(T,K_i) - B_t \max\left(F_t^{*} - K_i, 0\right)}{K_i^2}\Delta K_i - \sum_{u=t-w}^{t} d\ln B_u \cdot d\ln B_u + 2\sum_{u=t-w}^{t} d\ln S_u \cdot d\ln B_u \tag{26}$$

[28]It is not reasonable to suppose that variations of stock prices follow martingales.

[29]In this article, daily data are employed to calculate the realized variations that are related to interest rates. In the robustness tests, other frequencies are tested.

[30]Similar results can be obtained for the case without dividend protection.

As discussed in the sections above, when estimating the first term on the RHS of Equation (26), we first employ Equation (22) to screen the options with the best information quality. We then address the discretization errors and truncation errors with cubic spline interpolation and flat extrapolation, and we obtain the 30-day result via linear temporal interpolation. The second and third terms on the RHS of Equation (26) are estimated by the historical realized variations. Finally, we take the square root and multiply by 100 to obtain the AVIX:

$$\mathrm{AVIX}_{t,T} = 100 \times \sqrt{\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^{S}\right] \times \frac{365}{T}}.$$

Compared with the CBOE VIX and the SSE iVIX, the AVIX has the following four features. First, the effect of random interest rates is considered. Second, an adaptive screening mechanism is introduced. With such a mechanism, the difference between ITM and OTM, in addition to the difference between calls and puts, are no longer important. What matters is the information quality. As long as live data are provided, a portfolio of options with the best information quality will be chosen automatically to calculate the AVIX. Third, the cases with and without dividend protection are both considered. Fourth, market friction—such as possible restrictions for short-selling, high transaction costs, in addition to the limited range and a sparse set of discrete strike prices—is considered, and corresponding adjustments are proposed. Simply put, the AVIX is a more general and adaptable model-free implied variation index. In the sections below, we aim to conduct empirical studies in the China 50 ETF option market and make comparisons between the AVIX and the iVIX. As an emerging option market, the China 50 ETF option market is obviously a good place to investigate whether the AVIX has the declared advantages.

4. DATA

The option data that are used in this study are daily closing prices, trading volumes, and open interests of China 50 ETF options. The sample period is from the first trading day of the China 50 ETF option market, February 9, 2015 to January 21, 2016. The corresponding underlying data are daily closing prices and high-frequency intraday data (at 5-, 15-, and 30-minute intervals) of the China 50 ETF (510050.SH). The risk-free interest rate data are from the ChinaBond Government Security Spot Rate Yield Curve.[31] To make comparisons, the iVIX values released by the SSE in the sample period are used. To calculate the historical realized quadratic variations, the sample period of spot prices and interest rates is from January 5, 2015 to January 21, 2016. In the empirical study of investor sentiment, the daily turnover rate and trading volume of China 50 ETF and the daily 1-month Shanghai Interbank Offered Rates (SHIBOR) are used. All of the data are from the Wind.

The option data filters that are applied in our study are as follows. Options with less than 3 days remaining to maturity are excluded because of liquidity concerns. Options with negative implied volatilities are also excluded because these prices are beyond the no-arbitrage upper and lower limit and may not reflect true option value. In the occasional case that closing prices are absent, settlement prices are used as substitutes. However, for calls and puts with the same strike prices and the same expiration dates, if the trading volumes in the closing auction are both zero, the data of that series will be excluded.

[31]Its original name is the inter-bank fixed-rate treasury bonds spot rate yield curve.

5. EMPIRICAL RESULTS

5.1. The AVIX in the China 50 ETF Option Market

Using Equation (26), we compute the AVIX for the China 50 ETF option market. We also calculate the 30-day ATM B-S implied volatility (multiplied by 100), which is referred to as the BSIV. Further, we plot the time series of the BSIV, the iVIX, and the AVIX in our sample period in Figure 1 to make a comparison.

It is evident that these three series track each other fairly closely, which verifies that they contain similar information and that the AVIX is a measure of option-implied variation. In the sample period, the average implied volatilities (variations) are approximately 40%, with the highest at approximately 80%, and the lowest at approximately 20%. It is much higher than the U.S. case and is consistent with the characteristics of most emerging markets. In addition, all three indexes exhibit mean-reverting properties, which is consistent with the results of previous research.

More importantly, there are differences among these three series. Next we will consider the difference between the iVIX and the AVIX, exploring which one is a better measure of model-free option-implied variation.

5.2. Which Index Is the Better Measure of Aggregate Implied Variation?

Calculating along a continuum of European option prices across all strikes K, model-free implied variation indexes are widely accepted as measures of aggregate option-implied variations. However, which one is better, the iVIX or the AVIX? To answer this question, we use the BSIV, the 30-day ATM B-S implied volatility as a benchmark.

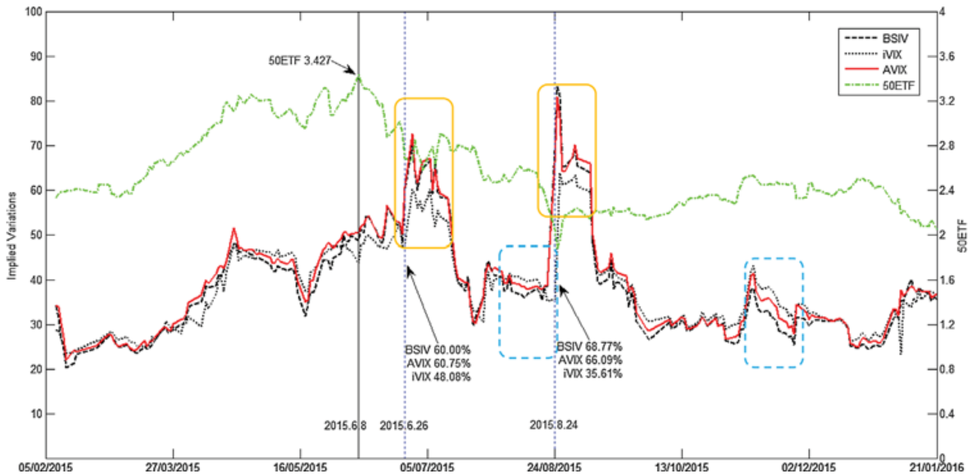

FIGURE 1

Implied Variations of China 50 ETF Options and China 50 ETF Closing Prices

Note: The figure plots the time series of the 30-day implied variations of China 50 ETF options and the China 50 ETF closing prices over the sample period from February 9, 2015 to January 21, 2016.The dashed line is the BSIV, the 30-day at-the-money Black-Scholes implied volatility (multiplied by 100). The dotted line is the iVIX, the model-free implied variation index of the Shanghai Stock Exchange (SSE). The solid line is the AVIX, calculated from Equation (26). The dash-dot line is the China 50 ETF closing prices. In two solid lined boxes, when the BSIV is particularly high, the iVIX is much lower. In two dashed lined boxes, when the BSIV is particularly low, the iVIX is higher. [Color figure can be viewed at wileyonlinelibrary.com]

From Figure 1, we can see that although it is quite close to the other two series, the iVIX tends to underestimate the variation when the benchmark (BSIV) is particularly high (see the periods in the two solid lined boxes). Meanwhile, the iVIX tends to overestimate the variation when the BSIV is particularly low (see the periods in the two dashed lined boxes). In other words, the more extreme the market condition, the more serious the deviation of the iVIX. However, it is in these periods that we most need an accurate variation index. In contrast, the AVIX is much closer to the BSIV, even in the extreme cases. The most typical cases are on August 24, 2015 and June 26, 2015. On these 2 days, most of the 30-day B-S implied volatilities were between 60% and 70%, and the 30-day ATM BSIVs were 68.77% and 60.00%, respectively. The iVIXs, however, were only 35.61% and 48.08%! The AVIX values exhibited a significant advantage and were 66.09% and 60.75%, respectively.

To further explore this issue, we compute the arithmetic average of B-S implied volatilities for all of the options with a 30-day maturity (multiplied by 100), referred to as the "ABSIV," and we use it as another benchmark. Table I summarizes the descriptive statistics of all four indexes, the BSIV, the ABSIV, the iVIX, and the AVIX.

As shown in Table I, as far as the mean values are concerned, the iVIX does not differ much from two benchmarks, and the AVIX has no obvious advantage. However, the standard deviation, the skewness, the excess kurtosis, and the maximum value of the iVIX are far less than those of two benchmarks, whereas the minimum is much higher. This verifies that the iVIX tends to deviate in extreme cases. In comparison, the AVIX succeeds. All of the descriptive statistics of the AVIX are closer to those of the two benchmarks. It is clearly demonstrated that the AVIX is a more accurate representative of market aggregate option-implied variation.

5.3. Which Index Is the Better Barometer of Investor Sentiment?

The CBOE VIX is referred to as the "fear gauge," and it is widely regarded as a good measure of investor sentiment. We aim to determine whether the corresponding indexes in China have the same function. Which one, the iVIX or the AVIX, better reflects investor sentiment?

To study this issue, a proxy for investor sentiment first needs to be constructed. A large number of studies have demonstrated that some economic and market variables can capture information regarding investor sentiment. The most famous is the investor sentiment index that was developed by Baker and Wurgler (2006). It is the first principal component (PC) of the closed-end fund discount, the NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium. It is, however, not suitable for a short sample such as the China 50 ETF option market because some of these variables are on weekly or monthly frequencies.[32] Following Chen, Chong, and She (2014), we choose the first PC of the following five daily variables as the proxy for investor sentiment.

The first two are VPCR, which is the ratio of trading volume of near-term puts to that of near-term calls, and OIPCR, which is the ratio of the open interest of near-term puts to that of near-term calls. The higher these two variables, the more pessimistic the market, and vice versa. The next two variables are the trading volume and turnover rate of the China 50 ETF, named EVOL and ETURN, respectively. Because investors in China's stock market tend to trade much more actively in bull markets than in bear markets, the higher these two variables, the more optimistic the market is. The last is the one-month SHIBOR, which is the short-term

[32]As introduced above, the China 50 ETF option market started on February 9, 2015 and our sample period is quite short.

TABLE I
Descriptive Statistics of Implied Variation Series

	Sample Size	*Mean*	*Standard Deviation*	*Skewness*	*Kurtosis*	*Minimum*	*Maximum*
BSIV	233	38.02	11.775	1.204	4.355	20.263	83.179
ABSIV	233	40.91	11.817	0.963	3.863	22.787	83.484
iVIX	233	37.99	9.241	0.668	2.867	23.330	63.788
AVIX	233	39.19	11.569	1.040	3.802	22.020	80.816

Note. BSIV is 30-day at-the-money Black-Scholes implied volatility (multiplied by 100). ABSIV is the arithmetic average of 30-day Black-Scholes implied volatility (multiplied by 100) of all options. iVIX is the model-free implied variation index of the Shanghai Stock Exchange (SSE). AVIX is calculated from Equation (26). The sample period is from February 9, 2015 to January 21, 2016.

financing cost. In China, the 1-month SHIBOR is inclined to increase in bullish markets due to a higher demand for money. That is, a higher 1-month SHIBOR is accompanied by more optimistic sentiment.

Applying the principal component analysis, we obtain the first PC, which contributes 44.97% of the total variance and is denoted by SENT_t. Based on the factor loadings, SENT_t can be expressed as

$$\mathrm{SENT}_t = -0.352 \times \mathrm{VPCR}_t - 0.303 \times \mathrm{OIPCR}_t + 0.633 \times \mathrm{EVOL}_t + 0.614 \times \mathrm{ETURN}_t + 0.078 \times \mathrm{SHIBOR}_t \tag{27}$$

As suggested by Equation (27), the proxy for investor sentiment, SENT_t, is negatively related to VPCRt_t and OIPCR_t, whereas it is positively related to EVOL_t, ETURN_t and SHIBOR_t. Therefore, the higher the value of SENT_t, the more optimistic the market, and vice versa.

Next, we employ encompassing regressions to analyze which is a better measure of investor sentiment, the iVIX or the AVIX. In encompassing regressions, both univariate and multivariate regressions are conducted to address the relative importance of competing explanatory variables. Those variables whose regression coefficients remain significant in both univariate and multivariate regressions and with higher goodness-of-fit are considered to contain more information about the dependent variable. Fair and Shiller (1990) were the first to use encompassing regressions.

Because a univariate regression is a restricted version of the corresponding multivariate regression, we only state the multivariate regression. It is

$$\mathrm{SENT}_t = \alpha + \beta^{\mathrm{iVIX}} \cdot \mathrm{iVIX}_t + \beta^{\mathrm{AVIX}} \cdot \mathrm{AVIX}_t + \varepsilon_t \tag{28}$$

Before the time-series analysis, we conduct the ADF test and find that for the three time series, SENT_t, iVIX_t, and AVIX_t, we reject the null hypothesis of the unit root at the 5%, 10%, and 10% significance levels, respectively. The regression results of Model (28) are reported in Table II.

We have obtained some significant findings. One is that both regression coefficients of the two univariate regressions (Model 28-1 and 28-2) are highly significantly different from zero, and the adjusted R^2 are both approximately 40%. This means that both iVIX and AVIX are good gauges of investor sentiment. More important, however, is that the AVIX is a better barometer of investor sentiment. The first evidence is the difference between the adjusted R^2 of the univariate regressions (Model 28-1 and 28-2). The goodness-of-fit of the AVIX is 42.8%, which is higher than that of the iVIX, which is 37.4%. This means that the correlation

TABLE II
Model-Free Implied Variations and Investor Sentiment

Model	α	*iVIX*	*AVIX*	*Adjusted R^2*
28-1	**−3.792***** **(0.000)**	**0.099*******(0.000)**		0.374
28-2	**−3.340***** **(0.000)**		**0.085***** **(0.000)**	0.428
28-3	**−3.292***** **(0.000)**	−0.006 (0.917)	**0.089*** **(0.082)**	0.426

Note. This table reports the results for univariate and multivariate regression models that study the relationship between investor sentiment and two model-free implied variation indexes, the iVIX and the AVIX in China's stock market. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $SENT_t$, which is obtained by Equation (27). The explanatory variables in Model 28-1 and 28-2 are iVIX and AVIX, respectively. The explanatory variables in Model 28-3 are iVIX and AVIX. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the *P*-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

between the AVIX and investor sentiment is higher. Second, in the univariate regressions, the regression coefficient of the AVIX is 0.085, which is lower than that of the iVIX (0.099). That is, $1/\beta^{\text{AVIX}} > 1/\beta^{\text{iVIX}}$. This result implies that for a given change in investor sentiment, the change in the AVIX is greater than that in the iVIX. In other words, the AVIX is more sensitive to changes in investor sentiment. The most important evidence comes from the last bivariate regression. Although the iVIX is significantly different from zero in the univariate regression (Model 28-1), it is no longer statistically significant in Model 28-3, the bivariate regression. In contrast, the AVIX is significantly different from zero in both the univariate regression (Model 28-2) and the bivariate regression (Model 28-3). Compared with Model 28-1, the addition of the AVIX (Model 28-3) improves the regression goodness-of-fit from 37.4% to 42.6%. However, the goodness-of-fit of Model 28-3 is not greater than that of Model 28-2. All of this evidence suggests that the iVIX is redundant and its information about investor sentiment has been impounded in the AVIX. There is no doubt that the AVIX is a better barometer of investor sentiment than the iVIX.

5.4. The Fear Index or the Greed Index?

In addition to the above results, Table II reveals another interesting phenomenon. Despite the difference in the sensitivity and correlation, the slope coefficients of the iVIX and the AVIX in the univariate regressions are both positive. In subsections 5.1 and 5.2, we have shown that both indexes could represent the market aggregate implied variation to some extent. This means that, in China, the implied variation is positively related to the market. Figure 1 also shows the positive relationship between three implied variation(volatility) indexes and the China 50 ETF (for example, the linear correlation coefficient between the AVIX and the China 50 ETF is 0.3135). This is opposite the case of the U.S. In the U.S. market, negative returns are associated with higher volatility. That is why the CBOE VIX is referred to as the "fear gauge."

To verify this phenomenon, we calculate realized volatility with 15-minute high-frequency data of 50 ETF, and we investigate its relationship with the closing price of the China 50 ETF from March 29, 2013 to March 25, 2016. A positive correlation is also found, and the linear correlation coefficient is 0.49. When the market rises, the volatility goes up, and vice versa. In other words, the iVIX and the AVIX in China should be referred to as the "greed index" rather than the "fear index"!

The study of this phenomenon is beyond the scope of this article. We would like to study this issue in another paper. However, we believe that the main reason may lie in investor structure. In

China's stock market, retail investors dominate. According to the 2015 yearbook of the SSE, retail investors accounted for 85.19% of the SSE turnover in 2014. This is the opposite of the case of the U.S. market. Different from institutional investors, retail investors are not obligated to obey strict investment disciplines, and they are not subject to risk management requirements. Once the market starts to fall, retail investors are not forced to reduce their position simultaneously and massively. Instead, they tend to exhibit myopic loss aversion, and will not sell their shares. Thus, when the market is falling, the volatility decreases. In a rising market, however, some retail investors want to realize their profits, and others want to buy on the upswing. Different investor opinions lead to high volatility in a rising market. Due to the above reasons, in a market in which retail investors dominate, the model-free implied variation index would become a greed index, which is precisely the case in China's stock market.

5.5. The iVIX and the AVIX: Which Plays a Leading Role?

Up to this point, the AVIX has been demonstrated to be a better index of aggregate implied variation and investor sentiment. Moreover, Figure 1 shows that the AVIX tends to change a bit earlier than the iVIX, that is, the AVIX may lead changes in the iVIX.

To verify this phenomenon, we perform pairwise Granger causality tests on the iVIX and the AVIX. We first take the logarithm of these two variables to make their values range from negative infinity to positive infinity.[33] The ADF tests show that both log series reject the null hypothesis of nonstationary at the 10% significance level. According to the AIC, SC, and HQ criteria, a lag length of two is chosen. Table III presents the results of the Granger causality tests.

As shown in Table III, the null hypothesis that $\ln(\text{iVIX})_t$ is not the Granger cause of $\ln(\text{AVIX})_t$ is rejected at the 1.E-21 significance level, whereas the P value of the null hypothesis that $\ln(\text{AVIX})_t$ is not the Granger cause of $\ln(\text{iVIX})_t$ reaches 8.92%. This asymmetric lead–lag relationship suggests that in the relationship between the iVIX and the AVIX, the latter plays a leading role. The AVIX is more sensitive to new information and tends to change more quickly, whereas the iVIX lags.

When we carefully observe Figure 1, we find that the AVIX leads the iVIX particularly significantly when there are extreme changes. For example, on August 24, 2015, the China 50 ETF went down 9.86%. On that day, the AVIX and the iVIX were 66.09% and 35.61%, respectively. In the next two trading days, the iVIX rose gradually. On August 26, the iVIX reached 63.79%, which is much closer to the AVIX (76.09%) and the BSIV (81.05%). A possible explanation is as follows. The crash of the market led to the absence of OTM puts, and the iVIX could only reflect the information of the OTM calls. Over the next two trading days; however, the market stabilized gradually, and new option contracts were added. The OTM puts appeared again. The iVIX could at last reflect the information

TABLE III
Granger Causality Test Between the iVIX and the AVIX

The Null Hypothesis	*F Statistics*	*P Value*
ln(iVIX) is not the Granger Cause of ln(AVIX)	2.44238	0.0892
ln(AVIX) is not the Granger Cause of ln(iVIX)	59.9508	1.E-21

Note. This table reports the lead-lag relationship between the 30-day iVIX and AVIX in the China 50 ETF option market from February 9, 2015 to Jan. 21, 2016. The lag length of 2 is determined by the AIC, SC and HQ criterion.

[33]The pairwise Granger causality tests on the original time series have similar results.

of the OTM puts and the true level of implied variation. In comparison, the AVIX utilizes the information inherent in the options with the best information quality. It can capture the market change without delay. For this reason, it is not surprising that the AVIX leads the iVIX.

5.6. Which Is the Better Predictor of Future Volatility?

Christensen and Prabhala (1998), Jiang and Tian (2005), and Huang and Zheng (2009) employ encompassing regressions to compare the forecasting ability of different volatility measures. In this subsection, we use the same method to investigate which measure is better at predicting future volatility, the iVIX or the AVIX.

The regression model is

$$RV_{t,T} = a + b^{\text{iVIX}} \cdot \text{iVIX}_t + b^{\text{AVIX}} \cdot \text{AVIX}_t + \varphi_t, \tag{29}$$

where $RV_{t,T}$ is the real volatility of the China 50 ETF from time t to T (30 days altogether), iVIX_t and AVIX_t are calculated from the option prices at time t.[34]

Similar to Model (28) in subsection 5.3, we compare the statistical significance of regression coefficients and the goodness-of-fit in univariate and multivariate regressions. In addition, χ^2 tests are introduced to compare the relative information content of the two explanation variables in Model (29). For example, if the null joint hypothesis that the slope coefficient is one for the AVIX and zero for the iVIX, it provides evidence that the AVIX fully subsumes the information that is contained in the iVIX.

Due to its good properties, such as direct observability, simplicity and ease of implementation, the realized volatility that is calculated from high-frequency intraday data has been recognized as one of the best measures of real volatility. Following French, Schwert, and Stambaugh (1987), Jiang and Tian (2005), Hansen and Lunde (2006), we use the adjusted realized volatility,

$$RV_{t,T} = \sqrt{\frac{1}{T-t}\sum_{i=1}^{n} R_i^2 + \frac{2}{T-t}\sum_{j=1}^{l}\left(\frac{n}{n-j}\right)\sum_{i=1}^{n-j} R_i R_{i+j}}, \tag{30}$$

as the proxy of true volatility, where R_i is the China 50 ETF return during the i-th interval, n is the total number of intervals in the period from t to T, and l is the order of autocorrelation. Hence, the second term on the right of Equation (30) is a correction term for autocorrelation in intraday returns. In this section, 5-minute returns are used. Because the 5-minute returns in our sample have a first-order autocorrelation of -0.059 and much smaller higher-order autocorrelations, the autocorrelation order (l) is set to zero.[35] The ADF test suggests that it is stationary at the 5% significance level.

Due to the dividend protection mechanism in the China 50 ETF option market, the underlying prices must be adjusted if dividends are paid. Specifically, if a cash dividend of d dollars is paid on ex-dividend date τ ($t \le \tau \le T$) and the ex-dividend price at time τ is S^τ, the prices during the period of $[\tau, T]$ should be multiplied by $1 + d/S^\tau$.

[34]Because the history of the China 50 ETF option market is quite short, we have to use overlapping samples. Newey-West heteroskedastic autocorrelated consistent standard errors are used in our regressions. We also examine the nonoverlapping samples. Because the nonoverlapping sample size is too small, we do not report the results here. However, the main conclusion of the nonoverlapping samples is consistent with that of the overlapping samples.

[35]In the robustness tests, we examine the effect of including one to three correction terms for 5-, 15-, and 30-minute returns. We also cleaned the data using an MA(1) filter and investigate the influence of this cleaning.

Table IV summarizes all of the univariate and multivariate regression results. We obtain two important findings.

First, in both of the univariate regressions (Model 29-1 and 2), the slope coefficient is positive and significantly different from zero at the 1% significance level. The adjusted R^2 is greater than 30%. This result implies that both the iVIX and the AVIX contain important information for future volatility.

The second finding is more important. The AVIX is more informationally efficient in the prediction of future volatility than the iVIX. There is sufficient evidence to support this conclusion. First, the adjusted R^2 of Model 29-2 (0.359) is greater than that of Model 29-1 (0.322). That means the AVIX explains more variations in future volatility. In addition, although the iVIX is significantly different from zero in the univariate regression (Model 29-1), it is no longer significant in the multivariate regressions (Model 29-3) . This is strong evidence that the iVIX is redundant when it is used to forecast future volatility with the AVIX. Further, the adjusted R^2 value of Model 29-3 is higher than that of Model 29-1, which means the addition of the AVIX is helpful in the improvement of the goodness-of-fit of the model. In contrast, the adjusted R^2 of Model 29-3 is lower than that of Model 29-2. In other words, the addition of the iVIX is of no help in the prediction of future volatility. Another important piece of evidence comes from the results of the χ^2 tests. The null hypothesis of the χ^2 test in Model 29-3 is that the slope coefficient is one for the AVIX and zero for the iVIX. The test statistics from χ^2 test fail to reject the null hypothesis. That is to say, the AVIX is not only a more efficient forecast for future realized volatility but also fully subsumes the information that is contained in the iVIX.

In summary, among these two implied variation measures, the AVIX contains more information and provides a more efficient forecast for future volatility. This is not unexpected, considering that the AIX aggregates information across options with the best information quality in the market.

6. ROBUSTNESS TESTS

The results from the previous section provide strong support for the advantages of the AVIX over the iVIX. Whether it is in the reflection of the aggregate implied variation and investor sentiment, in the reaction speed to the market changes, or in the forecast for future volatility, the AVIX beats the iVIX. The AVIX exhibits superior forecasting ability

TABLE IV

Model-Free Implied Variations and Future Volatility (the 5-Minute Frequency, $l = 0$)

Model	α	*iVIX*	*AVIX*	*Adjusted* R^2	χ^2 *Test*
29-1	−0.389 (0.960)	**0.990*** (0.000)**		0.322	0.096 (0.953)
29-2	4.579 (0.486)		**0.842*** (0.000)**	0.359	0.834 (0.661)
29-3	4.076 (0.570)	0.059 (0.868)	**0.798** (0.013)**	0.356	0.693 (0.707)

Note. This table reports the results for univariate and multivariate regression models that study the forecasting ability of 30-day model-free implied variations for future volatility in the China 50 ETF option market. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $RV_{t,T}$, the realized volatility calculated from 5-minute China 50 ETF returns from t to T (30 days) with an autocorrelation order (l) of 0. In the first two univariate regression models, the explanatory variables are the iVIX and the AVIX, respectively, and the χ^2 test is for the joint hypothesis $H_0: a = 0$ and $b^j = 1$ ($j =$ iVIX, AVIX). In Model 29-3, the explanatory variables are the iVIX and the AVIX, and the χ^2 test is for the joint hypothesis $H_0 : b^{iVIX} = 0$ and $b^{AVIX} = 1$. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the P-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

and is informationally more efficient. We now conduct robustness tests to ensure the generality of these findings.

6.1. Alternative Regression Models

In this subsection, we use squared and log variation measures to run regressions. Thus Models (28) and (29) are converted to

$$\mathrm{SENT}_t = \alpha + \beta^{\mathrm{iVIX}^2} \cdot \mathrm{iVIX}_t^2 + \beta^{\mathrm{AVIX}^2} \mathrm{AVIX}_t^2 + \varepsilon_t \tag{31}$$

$$\mathrm{SENT}_t = \alpha + \beta^{\ln(\mathrm{iVIX})} \cdot \ln(\mathrm{iVIX})_t + \beta^{\ln(\mathrm{AVIX})} \cdot \ln\left(\mathrm{AVIX}\right)_t + \varepsilon_t \tag{32}$$

$$RV_{t,T}^2 = a + b^{\mathrm{iVIX}^2} \cdot \mathrm{iVIX}_t^2 + b^{\mathrm{AVIX}^2} \cdot \mathrm{AVIX}_t^2 + \varphi_t \tag{33}$$

$$\ln(RV)_{t,T} = a + b^{\ln(\mathrm{iVIX})} \cdot \ln(\mathrm{iVIX})_t + b^{\ln(\mathrm{AVIX})} \cdot \ln(\mathrm{AVIX})_t + \varphi_t \tag{34}$$

Because the square of the implied variation can be understood as the implied quadratic variation, Models (31) and (33) are designed to examine which model-free implied quadratic variation can better reflect investor sentiment and forecast future volatility. Similarly, Models (32) and (34) are designed to investigate which model-free implied log variation measure is better.

An unreported descriptive statistics table reveals that the log variation is the most conformable with a normal distribution. Regressions that are based on the log variation are thus statistically better specified than those that are based on variation or quadratic variation. However, the regression results for the model-free implied quadratic variation and the model-free implied log variation are similar. For brevity, we only report the results for the model-free implied log variation in Tables V and VI.[36]

The results in Tables V and VI are consistent with those that are shown in Tables II and IV. Just as in the case of implied variation, the significance coefficient, the goodness-of-fit, and the result of the χ^2 test all indicate that the log AVIX performs better than the log iVIX. The change in the regression models does not affect our conclusion. The AVIX is the better barometer of investor sentiment as well as the better predictor of future volatility.

TABLE V
Model-Free Implied Log Variations and Investor Sentiment

Model	α	*ln(iVIX)*	*ln(AVIX)*	*Adjusted* R^2
32-1	−**13.14***** **(0.000)**	**3.641***** **(0.000)**		0.325
32-2	−**12.22***** **(0.000)**		**3.366***** **(0.000)**	0.381
32-3	−**11.56***** **(0.000)**	−0.981 (0.606)	**4.163**** **(0.027)**	0.381

Note. This table reports the results for univariate and multivariate regression models that study the relationship between investor sentiment and model-free implied log variations in China's stock market. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is SENT_t, which is obtained by Equation (27). The explanatory variables in Model 32-1 and 32-2 are ln(iVIX) and ln(AVIX), respectively. The explanatory variables in Model 32-3 are ln(iVIX) and ln(AVIX). Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the *P*-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

[36]Interested readers can contact the authors for the full results.

TABLE VI
Model-Free Implied Log Variations and Future Volatility (the 5-Minute Frequency, $l = 0$)

Model	α	*ln(iVIX)*	*ln(AVIX)*	*Adjusted* R^2	χ^2 *Test*
34-1	0.337 (0.577)	**0.891***** **(0.000)**		0.332	1.783 (0.409)
34-2	0.597 (0.248)		**0.815***** **(0.000)**	0.380	3.603 (0.165)
34-3	0.687 (0.228)	−0.135 (0.648)	**0.924***** **(0.000)**	0.378	1.673 (0.433)

Note. This table reports the results for univariate and multivariate regression models that study the forecasting ability of 30-day model-free implied log variations for future volatility in the China 50 ETF option market. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $\ln(RV_{t,T})$, the log realized volatility calculated from 5-minute China 50 ETF returns from t to T (30 days) with an autocorrelation order (l) of 0. In the first two univariate regression models, the explanatory variables are ln(iVIX) and ln(AVIX), respectively, and the χ^2 test is for the joint hypothesis $H_0 : a = 0$ and $b^j = 1$ ($j = \ln(\text{iVIX})_t, \ln(\text{AVIX})_t$). In Model 34-3, the explanatory variables are ln(iVIX) and ln(AVIX), and the χ^2 test is for the joint hypothesis $H_0 : b^{\ln(\text{iVIX})} = 0$ and $b^{\ln(\text{AVIX})} = 1$. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the P-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

6.2. Adjusting the Calculation Details of the AVIX

6.2.1. The trapezoidal rule and weekly interest rates

We have mentioned that the rectangle rule is employed in the numerical integration of Equation (23). Now we use the trapezoidal rule and examine whether this change affects the conclusions. In addition, when we calculate the historical quadratic variations that are related to the interest rates, $QV^{B}_{t-w,t}$ and $CV^{B,S}_{t-w,t}$, daily data are used. In theory, the higher the data frequency, the more accurate the quadratic variation. However, due to the poor liquidity of treasury bills in China's bond market, intraday high-frequency data are not available, and even the daily yield curves are obtained via interpolations. To reduce the noise that is induced by interpolations, we use weekly data to conduct the robustness test in this subsection.

Table VII provides summary statistics for the original AVIX and the adjusted AVIXs, where the AVIX1 is calculated by the trapezoidal rule, the AVIX2 uses weekly interest rates, and the AVIX3 reflects the influence of the trapezoidal rule and weekly interest rates simultaneously.

The results indicate that the main statistical characteristics of the AVIX do not change significantly after these adjustments. The rectangle rule has a greater impact: it results a slightly lower mean and higher standard deviation, skewness and kurtosis, in addition to more extreme extrema. These results are related to the shape of the integrand. After the introduction of the adaptive screening mechanism, the first right term of Equation (19) is replaced by

TABLE VII
Descriptive Statistics of Adjusted AVIX Series

	Sample size	*Mean*	*Standard Deviation*	*Skewness*	*Kurtosis*	*Minimum*	*Maximum*
AVIX	233	39.19	11.569	1.040	3.802	22.020	80.816
AVIX1	233	39.06	11.761	1.142	4.169	21.870	82.747
AVIX2	233	39.20	11.553	1.037	3.790	22.019	80.728
AVIX3	233	39.09	11.743	1.139	4.159	21.870	82.662

Note. AVIX is calculated from Equation (26). The difference between AVIX and AVIX1 is that AVIX1 uses the trapezoidal rule in the numerical integration. AVIX2 uses weekly data when computing historical quadratic variations related to interest rates. AVIX3 uses the trapezoidal rule and weekly data at the same time. The sample period is from February 9, 2015 to January 21, 2016.

$$\frac{2}{B_t}\int_0^{\infty}\frac{\bar{c}_t(T,K)-B_t\max\left(F_t^*-K,0\right)}{K^2}dK.$$

Unlike the CBOE VIX formula (Equation (15)), the first-order partial derivative of the numerator of the integrand with respect to K is no longer monotonically increasing. Nevertheless, it is difficult to judge between the rectangle rule and the trapezoidal rule.

Using these three adjusted AVIXs, we conduct the robustness tests. Table VIII reports the results of Model (28) for all of the adjusted AVIXs. Table IX provides the results of Model (29) for the AVIX3.[37] Not surprisingly, the significance of the coefficient, the goodness-of-fit, and the result of the χ^2 test in Table VIII and Table IX still indicate that the AVIX is better than the iVIX as a measure of investor sentiment and a predictor of future volatility.

6.2.2. *Alternative information quality criteria*

In this subsection, we aim to change the information quality criteria of the option screening mechanism and test the robustness of our findings.

In Equation (22), the proxy for the liquidity is the options trading volume, and the threshold level is specified at 150%. Now we consider five new criteria: 130% and 200% of the option trading volumes and 130%, 150%, and 200% of the option open interests. The corresponding new AVIXs are numbered from 4 to 8. Table X summarizes the descriptive statistics for these five AVIXs.

Compared with Table I, the AVIX index does not change much after new information quality criteria are employed. That means our adaptive screening mechanism is robust. Because the statistical characteristics remain almost the same, we can imagine that the subsequent regressions will have similar conclusions. This is the case. The robustness tests of Models (28) and (29) show that the AVIX is the better barometer of investor sentiment and more informationally efficient than the iVIX. Owing to limitations of space, the results are not reported here.[38]

TABLE VIII
Adjusted AVIXs and Investor Sentiment

Panel	α	*iVIX*	*AVIX*	*Adjusted* R^2
AVIX1	**−3.247***** **(0.000)**		**0.083***** **(0.000)**	0.420
	−3.316*** **(0.000)**	0.007 (0.879)	**0.077*** **(0.095)**	0.418
AVIX2	**−3.348***** **(0.000)**		**0.085***** **(0.000)**	0.428
	−3.296*** **(0.000)**	−0.006 (0.910)	**0.090*** **(0.080)**	0.426
AVIX3	**−3.325***** **(0.000)**		**0.083***** **(0.000)**	0.421
	−3.318*** **(0.000)**	0.006 (0.890)	**0.078*** **(0.093)**	0.419

Note. This table reports the results for univariate and multivariate regression models that study the relationship between investor sentiment and adjusted AVIXs in China's stock market. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $SENT_t$, which is obtained by Equation (27). In each panel, the explanatory variable of the first regression model is the AVIXi (i = 1,2,3). the explanatory variables of the second regression model are the iVIX and the AVIXi (i = 1,2,3). Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the *P*-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

[37]Unreported tables show similar results for the AVIX1 and the AVIX2.
[38]Interested readers can contact the authors for the full results.

6.3. Alternative Measures for Realized Volatility

Thus far, all of the reported results on forecasting ability are based on realized volatility calculated from 5-minute returns with an autocorrelation order of 0. We further investigate the robustness of our findings over alternative realized volatility calculated from 15- to 30-minute data with different autocorrelation orders (from 0 to 3). All of the parameter estimates and test statistics continue to support our key finding. There are also some noticeable differences as more correction terms are added. The regression R^2 is reduced. This is an indication of an increasing level of noise in the realized volatility series as more correction terms are included. Table XI summarizes the regression results of Model (29) with 15-minute returns and one correction term ($l = 1$).

Comparing Tables XI and IV, the results are not materially affected. The significance of the coefficients, the goodness-of-fit, and the χ^2 test still support our findings in Subsection 5.6. The AVIX contains more information regarding future volatility.[39]

In addition, Andersen, Bollerslev, Diebold, and Ebens (2001) suggest that returns should be cleaned up using an MA (1) filter before they are used to calculate realized volatility. Applying the MA (1) filter to 5-, 15-, and 30-minute returns, we recalculate realized volatility using the filtered series and rerun our encompassing regressions. The unreported results show that all of the parameter estimates and test statistics are similar, and there is no material change.[40]

6.4. The Subsamples

Due to the short history of China's option market, samples over different horizons are not available. What we can do is split the whole sample period into sub-periods and investigate the robustness of our findings over different sample periods. It is well known that China's stock market experienced an unprecedented crash in the summer of 2015. As Figure 1 shows, the price of the China 50 ETF rose in the first 4 months during our sample period and reached a peak of 3.427 RMB on June 8, 2015. After that, the China 50 ETF went down to 2.043 RMB on January 21, 2016, which was the last day of our sample. Based on these changes in

TABLE IX
AVIX3 and Future Volatility (the 5-Minute Frequency, $l = 0$)

Model	α	*iVIX*	*AVIX3*	*Adjusted* R^2	χ^2 *Test*
AVIX3-1	−0.389 (0.960)	**0.990***** **(0.000)**		0.322	0.096 (0.953)
AVIX3-2	5.219 (0.429)		**0.828***** **(0.000)**	0.357	0.845 (0.655)
AVIX3-3	4.077 (0.572)	0.123 (0.706)	**0.737**** **(0.013)**	0.355	0.979 (0.612)

Note. This table reports the results for univariate and multivariate regression models that study the forecasting ability of AVIX3 for future volatility in the China 50 ETF option market. AVIX3 uses the trapezoidal rule and weekly interest rate data. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $RV_{t,T}$, the realized volatility calculated from 5-minute China 50 ETF returns from t to T (30 days) with an autocorrelation order (l) of 0. In the first two univariate regression models, the explanatory variables are the iVIX and the AVIX3, respectively, and the χ^2 test is for the joint hypothesis $H_0 : a = 0$ and $b^j = 1$ ($j =$ iVIX, AVIX3). In Model AVIX3-3, the explanatory variables are the iVIX and the AVIX3, and the χ^2 test is for the joint hypothesis $H_0 : b^{\text{iVIX}} = 0$ and $b^{\text{AVIX3}} = 1$. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the P-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

[39]We also investigate the forecasting ability of squared and the log variation measures using alternative realized volatilities. The conclusions still hold.

[40]In fact the results are better with MA(1) adjustments.

TABLE X
Descriptive Statistics of Adjusted AVIXs with Different Information Quality Criteria

	Sample Size	*Mean*	*Standard Deviation*	*Skewness*	*Kurtosis*	*Minimum*	*Maximum*
AVIX4	233	39.18	11.573	1.045	3.817	22.054	80.817
AVIX5	233	39.19	11.587	1.037	3.788	22.122	80.576
AVIX6	233	39.17	11.577	1.053	3.831	22.147	80.961
AVIX7	233	39.17	11.569	1.051	3.826	21.988	80.855
AVIX8	233	39.16	11.556	1.048	3.819	22.023	80.758

Note. The information quality criteria of AVIX4 to AIX8 is 130% of option trading volume, 200% of option trading volume, 130% of option open interest, 150% of option open interest, 200% of option open interest. The sample period is from February 9, 2015 to January 21, 2016.

the market, we split the whole sample period into two sub-samples, the first from February 9, 2015 to June 8, 2015 and the second from June 9, 2015 to January 21, 2016. The descriptive statistics of the four indexes, the BSIV, the ABSIV, the iVIX, and the AVIX for the two subsamples are presented in separate panels in Table XII.

In both of the subsamples, the standard deviation, the excess kurtosis, and the maximum value of the iVIX are far less than those of the other three indexes. This result indicates that the iVIX tends to deviate in extreme cases. This is consistent with the findings for the full sample. In contrast, in both of the subsamples, the statistical characteristics of the AVIX are more similar to those of the BSIV and the ABSIV, which are the two benchmarks. This indicates that it is a robust conclusion that the AVIX is a better measure of aggregate implied variation.

Comparing the two subsamples, we find that the average level and the standard deviation of implied variation (volatility) indexes in subsample 2 are clearly higher. Most of the extrema appear in subsample 2. Further, the implied variation (volatility) indexes in subsample 2 are all positively skewed, which coincides with the prior research. Simply put, subsample 2 is statistically better. This is not surprising because the sample period of subsample 2 is much longer. Another important reason is that subsample 2 covers the whole crash period and contains more information.

In fact, when we conduct regressions for subsample 1, almost all of the coefficients are not statistically significant. This may be due to its small number of samples. Therefore, in Tables XIII and XIV, we only report the results of the robustness tests for subsample 2.

TABLE XI
Model-Free Implied Variations and Future Volatility (the 15-Minute Frequency, $l = 1$)

Model	α	*iVIX*	*AVIX*	*Adjusted* R^2	χ^2 *Test*
15m-1	4.130 (0.558)	**0.881***** **(0.000)**		0.285	0.348 (0.840)
15m-2	8.712 (0.134)		**0.738***** **(0.000)**	0.313	2.419 (0.298)
15m-3	7.813 (0.236)	0.105 (0.738)	**0.658**** **(0.012)**	0.310	2.889 (0.236)

Note. This table reports the results for univariate and multivariate regression models that study the forecasting ability of 30-day model free implied variations for future volatility in the China 50 ETF option market. The sample period is from February 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $RV_{t,T}$, the realized volatility calculated from 15-minute China 50 ETF returns from t to T (30 days) with an autocorrelation order (l) of 1. In the first two univariate regression models, the explanatory variables are the iVIX and the AVIX, respectively, and the χ^2 test is for the joint hypothesis $H_0 : a = 0$ and $b^j = 1$ ($j =$ iVIX, AVIX). In Model 15m-3, the explanatory variables are the iVIX and the AVIX, and the χ^2 test is for the joint hypothesis $H_0 : b^{\mathrm{iVIX}} = 0$ and $b^{\mathrm{AVIX}} = 1$. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the P-values are reported in parentheses below the coefficients. ***, ** and * indicate the 1%, 5%, and 10% significant levels, respectively.

TABLE XII
Descriptive Statistics of Implied Variation Series in Subsamples

	Sample size	*Mean*	*Standard Deviation*	*Skewness*	*Kurtosis*	*Minimum*	*Maximum*
Panel A: Subsample 1							
BSIV	79	35.66	8.850	−0.059	1.537	20.263	49.642
ABSIV	79	36.99	8.923	−0.073	1.590	22.787	55.553
iVIX	79	36.23	8.819	0.013	1.278	23.408	47.895
AVIX	79	37.06	9.252	−0.085	1.479	22.020	51.491
Panel B: Subsample 2							
BSIV	154	39.24	12.882	1.257	3.861	24.402	83.179
ABSIV	154	42.92	12.616	0.984	3.364	24.887	83.484
iVIX	154	38.90	9.351	0.942	3.131	23.330	63.788
AVIX	154	40.27	12.482	1.158	3.526	24.868	80.816

Note. BSIV is 30-day at-the-money Black-Scholes implied volatility (multiplied by 100). ABSIV is the arithmetic average of 30-day Black-Scholes implied volatility (multiplied by 100) of all options. iVIX is the model-free implied variation index of the Shanghai Stock Exchange (SSE). AVIX is calculated from Equation (26). The sample periods of two subsamples are from February 9, 2015 to June 8, 2015 and from June 9, 2015 to January 21, 2016, respectively.

The findings of Table XIII/XIV are similar to those of Table II/IV, which support the conclusion that the AVIX behaves better than the iVIX in reflecting investor sentiment and forecasting future volatility. Note that the average adjusted R^2 in Table XIII is approximately 0.54, whereas that in Table II is only approximately 0.41. This suggests that after the crash, options prices contained more information about investor sentiment.

6.5. Actual Spot Prices

As discussed in subsection 3.1, the spot price and the forward price that are used in the calculation of the AVIX are the PCP-implied prices. This is standard practice in both academics and professional practice. Doing so, the price biases that are caused by market frictions, measurement errors, or market inefficiency can be taken into account. However, what if the actual market prices are used?

TABLE XIII
Model-Free Implied Variations and Investor Sentiment (Subsample 2)

Model	α	*iVIX*	*AVIX*	*Adjusted* R^2
28-S2-1	**−5.392***** **(0.000)**	**0.133***** **(0.000)**		0.511
28-S2-2	**−4.375***** **(0.000)**		**0.103***** **(0.000)**	0.550
28-S2-3	**−4.683***** **(0.000)**	0.029 (0.633)	**0.083*** **(0.100)**	0.550

Note. This table reports the results for univariate and multivariate regression models that study the relationship between investor sentiment and two model-free implied variation indexes, the iVIX and the AVIX, in China's stock market. The sample period is from June 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $SENT_t$, which is obtained by Equation (27). The explanatory variables in Model 28-S2-1 and 28-S2-2 are iVIX and AVIX, respectively. The explanatory variables in Model 28-S2-3 are iVIX and AVIX. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the *P*-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

TABLE XIV
Model-Free Implied Variations and Future Volatility (the 5-Minute Frequency, $l=0$, Subsample 2)

Model	α	*iVIX*	*AVIX*	*Adjusted R^2*	*χ^2 Test*
29-S2-1	−5.428 (0.635)	**1.137***** **(0.000)**		0.330	0.348 (0.840)
29-S2-2	3.124 (0.705)		**0.886***** **(0.000)**	0.358	0.362 (0.834)
29-S2-3	0.763 (0.944)	0.223 (0.645)	**0.728**** **(0.025)**	0.355	0.767 (0.681)

Note. This table reports the results for univariate and multivariate regression models that study the forecasting ability of 30-day model free implied variations for future volatility in the China 50 ETF option market. The sample period is from June 9, 2015 to January 21, 2016. In all regression models, the dependent variable is $RV_{t,T}$, the realized volatility calculated from 5-minute China 50 ETF returns from t to T (30 days) with an autocorrelation order (l) of 0. In the first two univariate regression models, the explanatory variables are the iVIX and the AVIX, respectively, and the χ^2 test is for the joint hypothesis H_0 : $a=0$ and $b^j=1$ ($j=$ iVIX, AVIX). In Model 29-S2-3, the explanatory variables are the iVIX and the AVIX, and the χ^2 test is for the joint hypothesis $H_0 : b^{\text{iVIX}}=0$ and $b^{\text{AVIX}}=1$. Standard errors are adjusted for heteroskedasticity and serial correlation according to Newey and West (1987), and the P-values are reported in parentheses below the coefficients. ***, **, and * indicate the 1%, 5%, and 10% significant levels, respectively.

Some of the previous research (e.g., Jiang and Tian, 2005) finds that the choice of actual or implied index values has little impact on the calculation of the model-free implied variation in the SPX option market. However, we find opposite evidence in the China 50 ETF option market. When we try to use actual spot prices in this market, even reasonable B-S implied volatility could not be obtained, not to mention model-free implied variation indexes. B-S implied volatility calculated directly from actual market price is often negative or greater than 1. We believe that this is due to severe market frictions. In the China 50 ETF option market, transaction costs are very high, short-selling is limited or prohibited,[41] and the liquidity of the option market is poor. In view of these conditions, we believe that the iVIX and the AVIX should not be calculated from actual spot or forward prices in markets with severe market friction.

7. CONCLUSIONS

Based on the great works of Demeterfi et al. (1999), Britten-Jones and Neuberger (2000), Jiang and Tian (2005, 2007), and Carr and Wu (2009), we propose a more general and adaptable model-free implied variation index, the AVIX. The key properties of the AVIX are as follows. First, the introduction of stochastic interest rates excludes the effect of random interest rates and makes the AVIX a purer model-free implied variation index of the underlying asset. Second, the effect of dividend protection is considered. Formulas for the case with and without dividend protections are developed. Third, an adaptive screening mechanism is proposed to automatically select the options with the best information quality in real time. The following interpolations and extrapolations are conducted on the basis of these filtered option prices. In addition, in each step of the AVIX calculation, alternative methods are discussed to account for possible market friction.

We find that when interest rates are stochastic, the so-called model-free implied volatility is actually the square root of the expected quadratic variation of the underlying asset under the T-forward measure, where T is the maturity of the options used in the calculations. With fewer assumptions, the AVIX is a more general measure of implied variation. With more considerations about different market conditions, the AVIX provides an improvement on the traditional VIX to better address market frictions.

[41]Short-selling in China's stock market was virtually prohibited after the crash of 2015. What is more, strict restrictions have been imposed on the trading of index futures in China since the autumn of 2015.

We empirically test the information content and the forecasting ability of the AVIX using China 50 ETF option prices from February 9, 2015 to January 21, 2016. Unsurprisingly, we find that the AVIX is superior to the iVIX in all of the respects that we investigate. Relative to the iVIX, the AVIX can better reflect the market aggregate implied variation and investor sentiment. The AVIX reacts to market changes more sensitively and more quickly. In fact, the AVIX plays a leading role in the lead-lag relationship between the AVIX and the iVIX. In the prediction of future volatility, the AVIX behaves much better than the iVIX. The results of encompassing regressions suggest that the AVIX contains more information for future volatility and exhibits superior forecasting ability. These conclusions are robust for the squared and the log implied variation. Additional tests indicate that regardless of adjustments to the numerical integration rule, changes in the interest rates' data frequency, alterations to the information quality criteria, the use of alternative realized volatility, or the splitting of the sample, the relative advantages of the AVIX over the iVIX do not change. The AVIX is a more general and adaptable model-free implied variation index. Note that although our empirical study concerns the China 50 ETF option market, the AVIX can be used in any market.

APPENDIX

A. PROOF OF PROPOSITION 2

Applying integration by parts to the first term on the RHS of Equation (13), we have

$$\begin{aligned}
&\frac{2}{B_t}\int_0^\infty \frac{c_t(T,K)-B_t\max(F_t-K,0)}{K^2}dK \\
&=2\int_0^\infty \frac{\frac{c_t(T,K)}{B_t}-\max(F_t-K,0)}{K^2}dK \\
&=-\frac{2\left[\frac{c_t(T,K)}{B_t}-\max(F_t-K,0)\right]}{K}\Bigg|_{K=0}^{K=\infty}+2\int_0^\infty \frac{\frac{\partial\left(\frac{c_t(T,K)}{B_t}\right)}{\partial K}+1_{F_t>K}}{K}dK
\end{aligned} \tag{A1}$$

Note that as a relative asset price with respect to zero-coupon bond with maturity T, $\frac{c_t(T,K)}{B_t}$ is a martingale under the T-forward measure. When K is equal to zero, we have

$$\max(F_t-K,0)=F_t$$

and

$$\frac{c_t(T,K)}{B_t}=\mathbb{E}_t^{\mathbb{T}}[\max(F_T-K,0)]=\mathbb{E}_t^{\mathbb{T}}[F_T]=F_t.$$

This means that the first term on the RHS of Equation (A1) is zero when K is zero. When K goes to the positive infinity, the European call is a deep OTM option, and its price, $c_t(T,K)$, goes to zero. The term $\max(F_t-K,0)$ is also zero. Simply put, the first term on the RHS of Equation (A1) is always equivalent to zero.

Because

$$\frac{\partial\left(\frac{c_t(T,K)}{B_t}\right)}{\partial K} = \mathbb{E}_t^{\mathbb{T}}\left[\frac{\partial\ \max(F_T - K, 0)}{\partial K}\right] = -\mathbb{E}_t^{\mathbb{T}}(1_{F_T>K}),$$

the second term on the RHS of Equation (A1) can be simplified to

$$2\int_0^\infty \frac{\frac{\partial\left(\frac{c_t(T,K)}{B_t}\right)}{\partial K} + 1_{F_t>K}}{K} dK = 2\int_0^\infty \frac{1_{F_t>K} - \mathbb{E}_t^{\mathbb{T}}(1_{F_T>K})}{K} dK = 2\left[\ln F_t - \mathbb{E}_t^{\mathbb{T}}(\ln F_T)\right].$$

Hence, we have

$$\ln F_t - \mathbb{E}_t^{\mathbb{T}}(\ln F_T) = \frac{1}{B_t}\int_0^\infty \frac{c_t(T,K) - B_t \max(F_t - K, 0)}{K^2} dK.$$

Taking the integral of both sides of Equation (11) from t to T and computing its conditional expectation under the T-forward measure, we obtain

$$\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T d\ln F_u\right] = -\frac{1}{2}\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^F\right] + \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \int_{R^0}\left(1 + x + \frac{x^2}{2} - e^x\right)\upsilon_u^{\mathbb{T}}(dx)du\right].$$

Because

$$-\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T d\ln F_u\right] = \ln F_t - \mathbb{E}_t^{\mathbb{T}}(\ln F_T),$$

we further obtain

$$\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^F\right] = \frac{2}{B_t}\int_0^\infty \frac{c_t(T,K) - B_t \max(F_t - K, 0)}{K^2} dK - 2\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \int_{R^0}\left(e^x - 1 - x - \frac{x^2}{2}\right)\upsilon_u^{\mathbb{T}}(dx)du\right]$$

Thus, Equations (13) and (14) are proved.

Rearranging Equation (10) and computing the conditional expectations under the T-forward measure, we can obtain Equation (12). Thus, *Proposition 2* is proved.

B. PROOF OF PROPOSITION 3

Applying the Itô-Doeblin lemma to Equation (18), the dynamics of the dividend-adjusted price S_t^*, and Equation (4), the dynamics of the zero-coupon bond price B_t, we obtain

$$d\ln S_t^* = \left(r_t - \frac{1}{2}{\sigma_t}^2\right)dt + \sigma_t\sqrt{1-\rho_t^2}dW_{1t}^{\mathbb{Q}} + \sigma_t\rho_t dW_{2t}^{\mathbb{Q}}$$
$$+\int_{R^0} x\mu(\mathrm{d}x, dt) - \int_{R^0}(e^x - 1)\upsilon_t^{\mathbb{Q}}(dx)dt,$$

$$d \ln B_t = \left(r_t - \frac{1}{2} \beta_t^2 v_{r,t}^2 \right) dt + \beta_t v_{r,t} dW_{2t}^{\mathbb{Q}}.$$

Note that with the assumption that the volatility and the jumps of S_t and S_t^* is the same, we have

$$QV_{t,T}^S = QV_{t,T}^{S^*} = \int_t^T \sigma_u^2 du + \int_t^T \int_{R^0} x^2 \mu(dx, du) \tag{A2}$$

and

$$CV_{t,T}^{B,S} = CV_{t,T}^{B,S^*} = \int_t^T \rho_u \sigma_u \beta_u v_{r,u} du. \tag{A3}$$

Let $f(S_t^*)$ be a twice-differentiable function of S_t^*. By the Itô-Doeblin lemma for semi-martingales,

$$\begin{aligned} f(S_T^*) =& f(S_t^*) + \int_t^T f'(S_{u-}^*) dS_u^* + \frac{1}{2} \int_t^T f''(S_{u-}^*) \sigma_u^2 S_{u-}^{*2} du \\ &+ \int_t^T \int_{R^0} \left[f(S_{u-}^* e^x) - f(S_{u-}^*) - f'(S_{u-}^*) S_{u-}^* (e^x - 1) \right] \mu(dx, du) \end{aligned} \tag{A4}$$

Applying Equation (A4) to the function $f(S_t^*) = \ln S_t^*$ and noting that $S_t^* = S_t$, we obtain

$$\ln(S_T^*) = \ln(S_t) + \int_t^T \frac{1}{S_{u-}^*} dS_u^* - \frac{1}{2} \int_t^T \sigma_u^2 du + \int_t^T \int_{R^0} [x - e^x + 1] \mu(dx, du). \tag{A5}$$

Combining Equations (A2) and (A5), we obtain

$$QV_{t,T}^S = 2 \left[\ln(S_t) - \ln(S_T^*) + \int_t^T \frac{1}{S_{u-}^*} dS_u^* + \int_t^T \int_{R^0} \left[\frac{x^2}{2} + x + 1 - e^x \right] \mu(dx, du) \right]. \tag{A6}$$

Carr and Madan (2001) demonstrate that any payoff with bounded expectation can be spanned by a linear combination of the payoffs from a riskless asset, a single risky asset and options of all strikes. A Taylor expansion with remainder of $\ln(S_T^*)$ about any point M_t implies

$$\ln(S_T^*) = \ln(M_t) + \frac{1}{M_t} (S_T^* - M_t) - \int_0^{M_t} \frac{1}{K^2} (K - S_T^*)^+ dK - \int_{M_t}^\infty \frac{1}{K^2} (S_T^* - K)^+ dK \tag{A7}$$

Substituting Equation (A7) into Equation (A6) and taking the conditional expectation under the T-forward measure, we obtain

$$\begin{aligned} \mathbb{E}_t^{\mathbb{T}} \left[QV_{t,T}^S \right] =& 2 \left[\int_0^{M_t} \frac{1}{K^2} \mathbb{E}_t^{\mathbb{T}} \left[(K - S_T^*)^+ \right] dK + \int_{M_t}^\infty \frac{1}{K^2} \mathbb{E}_t^{\mathbb{T}} \left[(S_T^* - K)^+ \right] dK \right] \\ &+ 2 \left[\ln(S_t) - \ln(M_t) - \frac{1}{M_t} \mathbb{E}_t^{\mathbb{T}} [S_T^*] + 1 + \mathbb{E}_t^{\mathbb{T}} \left[\int_t^T \frac{1}{S_{u-}^*} dS_u^* \right] \right] \end{aligned}$$

$$-2\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \int_{R^0}\left(e^x - \frac{x^2}{2} - x - 1\right)\mu(dx, du)\right] \tag{A8}$$

Obviously, the last term on the RHS of Equation (A8) is identical to the last term on the RHS of Equation (19).

According to Equation (17), in the case of dividend protection, the payoff of a European call can be expressed by $\max\left(S_T^* - K, 0\right)$. Hence, we have

$$\frac{c_t(T,K)}{B_t} = \mathbb{E}_t^{\mathbb{T}}\left[\left(S_T^* - K, 0\right)^+\right]$$

$$\frac{p_t(T,K)}{B_t} = \mathbb{E}_t^{\mathbb{T}}\left[\left(K - S_T^*, 0\right)^+\right].$$

Thus the first term on the RHS of Equation (A8) can be expressed as

$$\frac{2}{B_t}\left[\int_0^{M_t}\frac{p_t(T,K)}{K^2}dK + \int_{M_t}^{\infty}\frac{c_t(T,K)}{K^2}dK\right]. \tag{A9}$$

Define

$$F_t^* \equiv \frac{S_t^*}{B_t} = \frac{S_t}{B_t}. \tag{A10}$$

According to PCP in the case of dividend protection, as proposed by Merton (1973b), we have

$$c_t(T,K) + KB_t = p_t(T,K) + S_t = p_t(T,K) + F_t^* B_t.$$

Let $M_t = F_t^*$, and Equation (A9) could be further written as

$$\frac{2}{B_t}\int_0^{\infty}\frac{c_t(T,K) - B_t\max\left(F_t^* - K, 0\right)}{K^2}dK.$$

The above expression is identical to the first term on the RHS of Equation (19).
Because F_t^* can also be regarded as a relative price with respect to B_t, it holds that

$$\mathbb{E}_t^{\mathbb{T}}\left[S_T^*\right] = \frac{S_t^*}{B_t} = \frac{S_t}{B_t} = F_t^*. \tag{A11}$$

Substituting Equations (A10) and (A11) into the second term on the RHS of Equation (A8) and noting that $M_t = F_t^*$, this term can be rewritten as

$$2\left[\ln(B_t) + \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \frac{1}{S_{u-}^*}dS_u^*\right]\right]. \tag{A12}$$

According to the Girsanov theorem, the dynamics of S_t^* and B_t under the T-forward measure are

$$\frac{dS_t^*}{S_{t-}^*} = \left(r_t + \rho_t\sigma_t\beta_t v_{r,t}\right)dt + \sigma_t\sqrt{1-\rho_t^2}dW_{1t}^{\mathbb{T}} + \sigma_t\rho_t dW_{2t}^{\mathbb{T}} + \int_{R^0} x\mu(dx,dt) - \int_{R^0}(e^x - 1)\upsilon_t^{\mathbb{T}}(dx)dt$$

and

$$d\ln B_t = \left(r_t + \frac{1}{2}\beta_t^2 v_{r,t}^2\right)dt + \beta_t v_{r,t} dW_{2t}^{\mathbb{T}},$$

respectively. Hence, we have

$$\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \frac{1}{S_{u-}^*} dS_u^*\right] = \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \left(r_u + \rho_u\sigma_u\beta_u v_{r,u}\right)du\right] = \mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right] + \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T r_u du\right] \tag{A13}$$

and

$$\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T d\ln B_u\right] = \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T \left(r_u + \frac{1}{2}\beta_u^2 v_{r,u}^2\right)du\right] = \frac{1}{2}\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^B\right] + \mathbb{E}_t^{\mathbb{T}}\left[\int_t^T r_u du\right] \tag{A14}$$

Because the price of the zero coupon bond at maturity is 1, we have

$$\mathbb{E}_t^{\mathbb{T}}\left[\int_t^T d\ln B_u\right] = \mathbb{E}_t^{\mathbb{T}}[\ln B_T] - \ln B_t = -\ln B_t. \tag{A15}$$

Combining Equations (A12–A15), we find that the second term on the RHS of Equation (A8) is identical to the sum of the second and third terms on the RHS of Equation (19)

$$-\mathbb{E}_t^{\mathbb{T}}\left[QV_{t,T}^B\right] + 2\mathbb{E}_t^{\mathbb{T}}\left[CV_{t,T}^{B,S}\right].$$

Thus, *Proposition* 3 is proved.

REFERENCES

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. Journal of Finance, 53, 499–547.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review, 39, 885–905.

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

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

Andersen, T. G., Fusari, N., & Todorov, V. (2015). The risk premia embedded in index options. Journal of Financial Economics, 117, 558–584.

Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. Journal of Finance, 61, 1645–1680.

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

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. Review of Financial Studies, 9, 69–107.

Benninga, S., Björk, T., & Wiener, Z. (2002). On the use of numeraires in option pricing. Journal of Derivatives, 10, 43–58.

Bollerslev, T., Tauchen, G., & Zhou, H. (2009). Expected stock returns and variance risk premia. Review of Financial Studies, 22, 4463–4492.

Bollerslev, T., & Todorov, V. (2011). Tails, fears, and risk premia. Journal of Finance, 66, 2165–2211.

Bollerslev, T., & Todorov, V. (2014). Time-varying jump tails. Journal of Econometrics, 183, 168–180.

Bollerslev, T., Todorov, V., & Xu, L. (2015). Tail risk premia and return predictability. Journal of Financial Economics, 118, 113–134.
Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. Journal of Business, 51, 621–651.
Britten-Jones, M., & Neuberger, A. (2000). Option prices, implied price processes, and stochastic volatility. Journal of Finance, 55, 839–866.
Carr, P., & Madan, D. (1998). Towards a theory of volatility trading. In R. Jarrow (Ed.), Risk book on volatility (pp. 417–427). New York: Risk Publications.
Carr, P., & Madan, D. (2001). Optimal positioning in derivative securities. Quantitative Finance, 1, 19–37.
Carr, P., & Wu, L. (2006). A tale of two indices. Journal of Derivatives, 13, 13–29.
Carr, P., & Wu, L. (2009). Variance risk premiums. Review of Financial Studies, 22, 1311–1341.
Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate. Journal of Finance, 47, 1209–1227.
Chen, H., Chong, T. T. L., & She, Y. (2014). A principal component approach to measuring investor sentiment in China. Quantitative Finance, 14, 573–579.
Christensen, B. J., & Prabhala, N. R. (1998). The relation between implied and realized volatility. Journal of Financial Economics, 50, 125–150.
Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. Econometrica, 53, 385–407.
Delbaen, F., & Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. Mathematische Annalen, 300, 463–520.
Demeterfi, K., Derman, E., Kamal, M., & Zou, J. (1999). A guide to volatility and variance swaps. Journal of Derivatives, 6, 9–32.
Du, J., & Kapadia, N. (2012). Tail and volatility indices from option prices. Working paper, University of Massachusetts, Amhurst.
Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. Econometrica, 68, 1343–1376.
Fair, R. C., & Shiller, R. J. (1990). Comparing information in forecasts from econometric models. The American Economic Review, 80, 375–389.
French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. Journal of Financial Economics, 19, 3–29.
Geske, R., Roll, R., & Shastri, K. (1983). Over-the-counter option market dividend protection and "biases" in the Black-Scholes model: a note. Journal of Finance, 38, 1271–1277.
Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. Journal of Business and Economic Statistics, 24, 127–161.
Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies, 6, 327–343.
Huang, Y., & Zheng, Z. (2009). Model-free implied volatility and its information content: Evidence from Hang Seng Index options. Systems Engineering-Theory & Practice, 11, 46–59.
Hull, J. C., & White, A. D. (1994). Numerical procedures for implementing term structure models I: Single-factor models. Journal of Derivatives, 2, 7–16.
Jacod, J., & Shiryaev, A. (2003). Limit theorems for stochastic processes (second ed.). Berlin: Springer Science and Business Media.
Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. Review of Financial Studies, 18, 1305–1342.
Jiang, G. J., & Tian, Y. S. (2007). Extracting model-free volatility from option prices: An examination of the VIX index. Journal of Derivatives, 14, 1–26.
Jiang, G. J., & Tian, Y. S. (2010). Misreaction or misspecification? A re-examination of volatility anomalies. Journal of Banking and Finance, 34, 2358–2369.
Merton, R. C. (1973a). Theory of rational option pricing. The Bell Journal of Economics and Management Science, 4, 141–183.
Merton, R. C. (1973b). The relationship between put and call option prices: Comment. Journal of Finance, 28, 183–184.
Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation-consistent covariance matrix. Econometrica, 55, 703–708.
Shimko, D. (1993). Bounds of probability. Risk, 6, 33–37.
Shreve, S. E. (2004). Stochastic calculus for finance II: Continuous-time models (Vol. 11). New York: Springer Science and Business Media.
Todorov, V. (2010). Variance risk-premium dynamics: The role of jumps. Review of Financial Studies, 23, 345–383.
Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics, 5, 177–188.