Implied Remaining Variance with Application to Bachelier Model

JIAN SUN, QIANKUN NIU, SHINAN CAO, AND PETER CARR

JIAN SUN
is a managing director at Jurun Capital in Marlboro, NJ.
jiansun@juruncapital.com

QIANKUN NIU
is an analyst at a large financial institution in New York, NY.
qniu@princeton.edu

SHINAN CAO
is an associate professor in the School of Finance at the University of International Business and Economics in Beijing, China.
shinancao@uibe.edu.com

PETER CARR
is the chair of the Finance and Risk Engineering Department at New York University in New York, NY.
pcarr@nyc.rr.com

Implied volatility is widely used to quote the European options. After the Black–Scholes–Merton model was introduced (Black and Scholes [1973]; Merton [1973]), it became critically important to model the implied volatilities for theoretical and practical applications. As a consequence, many models were invented to capture the characteristics of the implied volatility shape (e.g., Heston [1993]; Schweizer and Wissel [2008]; Carmona and Nadtochiy [2009]; Carr and Wu [2016]). Different models may focus on different characters of implied volatility surface.

Recently, Carr and Sun [2013] introduced the concept of implied remaining variance and used it to price options. For convenience, we call it IRV model. But if we focus on the interest rate derivatives, we can see that the rates are often close to zero, and even negative in some countries. So practitioners give up the log normal (or so called geometric Brownian Motion) assumption and go back to the normal assumption for the underlying interest rate; see the work of Gorovoi and Linetsky [2004] for one-dimensional Vesicek model and Ichiue and Ueno [2007] for two-factor Gaussian model. Bachelier was the first one to introduce the normal model for the financial assets (Bachelier [1900]; Sullivan and Weithers [1991]). See the work of Schachermayer and Teichmann [2008] for the difference between the Bachelier model and Black–Scholes–Merton model.

When building option pricing models, it is important to ensure that no-arbitrage condition holds. For instance, Carr and Madan [2005] introduced a sufficient condition to exclude all static arbitrage: absence of call spread, butterfly spread, and calendar spread arbitrages; Lee [2004] introduced a model-free condition. It is broadly known that no arbitrage is a more delicate issue for the market models approach. Arbitrage opportunities can be classified into two types: the dynamic arbitrage, which occurs when there is a risk-free portfolio with returns exceeding risk-free rate (see Carr and Wu [2010] for more details), and the static arbitrage, which may occur within relative price among different options, such as negative butterfly (see the discussion of Carr and Madan [2005] and the examples of models with static arbitrage problem in the work of Roper [2010]). A recent article by Huitema and Peeters [2014] contains a good summary of related literature. The model offered here under Bachelier setting belongs to the market model approach, which is derived from no-dynamic-arbitrage condition (see Peeters [2013]; Huitema and Peeters [2014]; Carr and Wu [2016]), but fails at the no-static-arbitrage conditions. However, we will see later that under clear closed-form condition of the parameters, the static arbitrage can be completely avoided in our models. To our best knowledge, this is the first market model for the evolution of the implied volatility surface that is full arbitrage free. Further, the advantages of having joint dynamics for the implied remaining variance and underlying will allow us calculate better risk Greeks.

The remainder of this article is organized as follows. We will first generalize the approach used by Carr and Sun [2013] to Bachelier setting with its parameterization. To avoid confusion, we name the original parameterization IRV3 which will be proved in a later section to imply arbitrage in extremely low or high strikes. Then, we will propose two new parameterizations, named IRV4 and IRV5. Calibration results show that our models work very well. Finally, we will prove that, under certain conditions, only our IRV4 and IRV5 could be full arbitrage free. The proofs of the theorems and conclusions are left in the appendix.

Our starting point is the following SDE under the forward measure

$$dF_t = \sigma dW_t. \tag{1}$$

where volatility σ is a constant, F_t is the forward price at time t, and W_t is the standard Brownian motion. It is well known that the absence of arbitrage between options of maturity T and underlying implies the existence of the forward martingale measure, under which the forward stock price F_t is a martingale leading to the zero-drift process in (1). For brevity, we denote the well-known two normal distribution functions as follows:

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt. \tag{2}$$

According to Bachelier model, a call option at time t with strike K and maturity time T can be priced (ignoring the discounting) as

$$\begin{aligned} C(K, t; F) &= E_t[(F_T - K)^+] \\ &= \sigma\sqrt{T-t}\,\phi\left(\frac{F-K}{\sigma\sqrt{T-t}}\right) + (F-K)\Phi\left(\frac{F-K}{\sigma\sqrt{T-t}}\right). \end{aligned} \tag{3}$$

Now, if we use the concept of implied remaining variance from Carr and Sun [2013]

$$I_t = \sigma^2(T-t), \tag{4}$$

we can rewrite the Bachelier formula (3) as

$$C(K, t; F) = \sqrt{I}\,\phi\left(\frac{F-K}{\sqrt{I}}\right) + (F-K)\Phi\left(\frac{F-K}{\sqrt{I}}\right). \tag{5}$$

with the following Greeks calculation:

$$\begin{aligned} &\frac{\partial C}{\partial F} = \Phi\left(\frac{F-K}{\sqrt{I}}\right), \quad \frac{\partial^2 C}{\partial F^2} = \frac{1}{\sqrt{I}}\phi\left(\frac{F-K}{\sqrt{I}}\right), \\ &\frac{\partial C}{\partial I} = \frac{1}{2\sqrt{I}}\phi\left(\frac{F-K}{\sqrt{I}}\right), \\ &\frac{\partial^2 C}{\partial I^2} = \left(-\frac{1}{4I\sqrt{I}} + \frac{(F-K)^2}{4I^2\sqrt{I}}\right)\phi\left(\frac{F-K}{\sqrt{I}}\right), \\ &\frac{\partial^2 C}{\partial F\,\partial I} = -\frac{F-K}{2I\sqrt{I}}\phi\left(\frac{F-K}{\sqrt{I}}\right). \end{aligned} \tag{6}$$

In an arbitrage-free market consisting of a single option maturing at T, the parameter σ in (3) can always be chosen to achieve a perfect fit to that option's price. In a market with two or more options maturing at T, we cannot choose a single numerical value σ to fit the two options values. Hence we have to deviate from the classic Bachelier model for a moment and introduce the general stochastic volatility process,

$$dF_t = \sqrt{v_t}\, dW_t \tag{7}$$

where $\sqrt{v_t}$ is stochastic normal volatility, which satisfies a certain process to be specified. Under this assumption, the option price is given by

$$C^{true}(K, t; F) = E[(F_T - K)^+ \mid F_t]$$

which is no longer consistent with the Bachelier formula. However, we can still use the Bachelier formula to quote the true price by choosing the right implied remaining variance in

$$C^B(I, K, F) = \sqrt{I}\,\phi\left(\frac{F-K}{\sqrt{I}}\right) + (F-K)\Phi\left(\frac{F-K}{\sqrt{I}}\right), \tag{8}$$

where $I(t, K)$ could be uniquely determined by inverting

$$C^B(I(t, K), K, F) = C^{true}(K, t; F)$$

Following the similar approach employed by Carr and Sun [2007; 2013], we bypass the unobserved ν_t to directly specify the dynamics for $I(t, K)$ in (8). We assume that the implied remaining variance $I(t, K)$ is governed by the following dynamics

$$dI_t = a(I_t)\nu_t dt + b(I_t)\sqrt{\nu_t}\, dZ_t \tag{9}$$

where $dW_t dZ_t = \rho(I_t)\, dt$, and I_t stands for $I(t, K)$. Since

$$C^B(I(t, K), K, F) = E[(F_T - K)^+ \mid F_t]$$

is a martingale, we can use Itô's formula to obtain the drift term of C^B and equate it to zero (or we could use Feymann-Kac theorem directly). This implies that the call price with strike K has to satisfy

$$\frac{1}{2}\frac{\partial^2 C^B}{\partial F^2} + \frac{1}{2}b(I)^2\frac{\partial^2 C^B}{\partial I^2} + \rho(I)b(I)\frac{\partial^2 C^B}{\partial F\,\partial I} + a(I)\frac{\partial C^B}{\partial I} = 0, \tag{10}$$

where I stands for $I(t, K)$. We now plug in all the Greeks computation above in (6) and get the following basic equation:

$$\frac{1}{2}(1 + a(I)) + \frac{1}{2}b(I)^2\left(\frac{(F-K)^2}{4I^2} - \frac{1}{4I}\right) + \rho(I)b(I)\left(-\frac{F-K}{2I}\right) = 0. \tag{11}$$

At this stage, we can freely choose certain simple parameterizations of the functions $a(I)$ and $b(I)$, plug into (11) and produce an explicit formula for implied remaining variance I as a function of K.

However, we will be facing with two fundamental tasks. One is to produce the implied volatility matching the market implied volatility. Usually this can be done by optimizing the underlying parameters. Secondly, we need to make sure the model of the implied volatility curve does not introduce arbitrage, even if we can produce a relatively good fit to the market, in particular at extremely low or high strikes. In the next several sections, we will introduce different sets of parametrizations and analyze their implied volatility shape. Later in our article, we will show not only that our choice of parametrization produces a very good fit to the market but also that we can root out the arbitrage.

IRV3. Borrowed from the work of Carr and Sun [2013], the following parametric functional form for $a(I)$, $b(I)$ and a constant ρ will be introduced here:

$$a(I) = -a_1 I + a_0 - 1,\quad b(I) = bI. \tag{12}$$

Plug (12) into our basic equation (11), we will have

$$-\frac{1}{2}a_1 I + \frac{1}{2}a_0 + \frac{1}{8}b^2(F-K)^2 - \frac{1}{8}b^2 I - \frac{1}{2}\rho b(F-K) = 0. \tag{13}$$

Dividing b^2 on both sides yields

$$-\frac{1}{2}\frac{a_1}{b^2}I + \frac{1}{2}\frac{a_0}{b^2} + \frac{1}{8}(F-K)^2 - \frac{1}{8}I - \frac{1}{2}\frac{\rho}{b}(F-K) = 0. \tag{14}$$

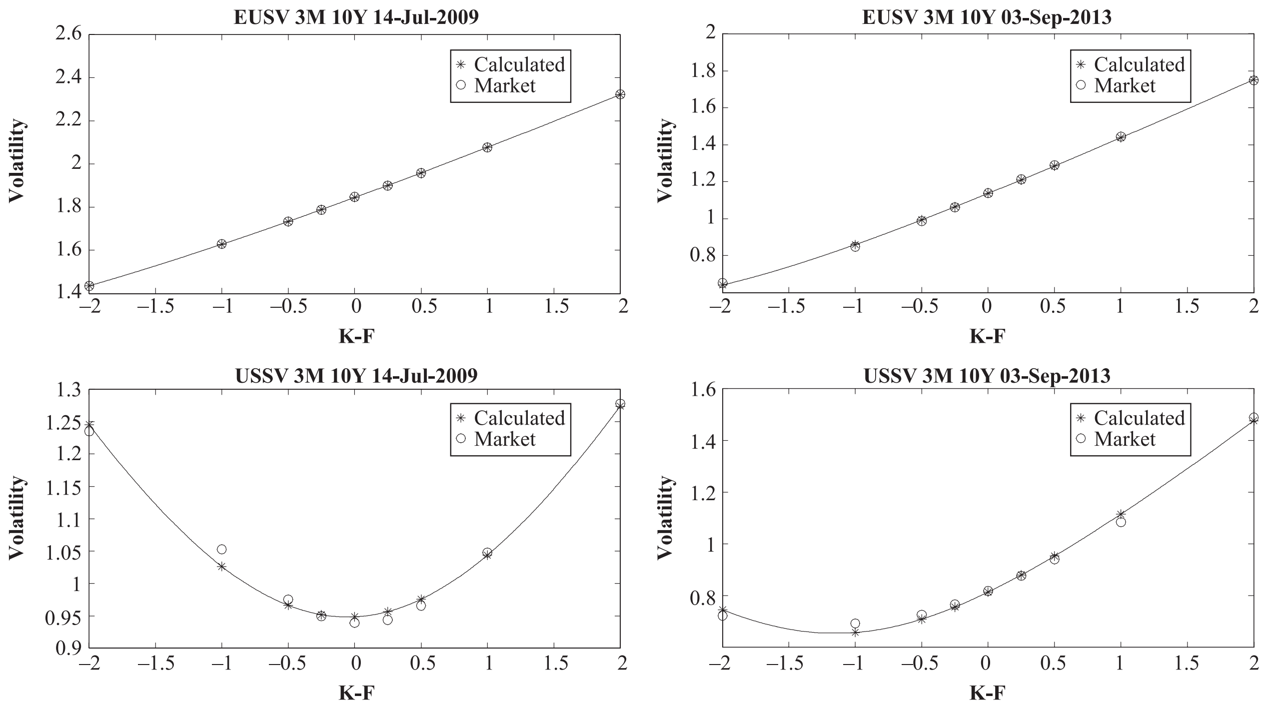
Where we can formalize the parameters by setting $b = 1$. Thus each implied volatility curve becomes the solution to a quadratic equation with three unknown parameters (ρ, a_0, a_1). This implies

$$I(K, F) = \frac{\frac{1}{4}(F-K)^2 - \rho(F-K) + a_0}{a_1 + \frac{1}{4}}. \tag{15}$$

Note that $\rho^2 < a_0$, $a_1 > -\frac{1}{4}$ will ensure that I is always positive. We call this model IRV3.

We will now show the calibration of this model to the market implied volatility curves. For this purpose, we are using data from interest rate derivatives market, in particular the swaption market. We perform the calibration based on two swaption markets: U.S. market and EU market. For each market, we calibrate market data over the sample period from July 14, 2009, to July 14, 2014. We also use a range of standard filters to remove illiquid or erroneous quotes and ensure that each day has nine reliable contracts with relative strikes ranging from −0.02 to +0.02 to the spot forward rate if that strike is available. As a result, we have 923 days and 4,615 contracts for the U.S. market and 792 days and 3,960 contracts for the EU market. The mean of all root mean squared error (RMSE) of the calibration of each day is 0.0134 for the U.S. market and 0.0323 for the EU market. We give four examples below in Exhibit 1. In each case, we can still see some errors near the money

Exhibit 1
3M by 10Y Implied Volatility (IRV3)



points, which are sacrificed to fit the deep in the money or out of money strikes.

IRV4. Now we generalize the parameterization of (12) into a four-parameter case, by allowing that ρ varies along with strikes, where we call it IRV4, with the following specification:

$$a(I) = -a_1 I + a_0 - 1,\ \ b(I) = bI,\ \ \rho(I) = cI + d, \tag{16}$$

The basic equation becomes

$$(a_0 - a_1 I) + \frac{1}{4}b^2((F-K)^2 - I) - b(cI + d)(F-K) = 0. \tag{17}$$

Similar to IRV3, we set $b = 1$ and get

$$I = \frac{a_0 + \frac{1}{4}(F-K)^2 - d(F-K)}{a_1 + \frac{1}{4} + c(F-K)} \tag{18}$$

where $c < 0$, $d^2 < a_0$, $a_1 + \frac{1}{4} + cF > 0$ will ensure that I is always positive.

In Exhibit 2, we see the calibration results of IRV4 from swaption market with 3-month by 10-year contracts in both the U.S. market and the EU market. We are using the same data from IRV3 so that we don't repeat the data description here. The mean of all RMSE of the calibration of each day is 0.0033 for the U.S. market and 0.0104 for the EU market. IRV4 reduces 75% in sample pricing error of IRV3 for the U.S. market and 68% for the EU market.

IRV5. To introduce the second order term into $a(I)$, we have the following IRV5 model:

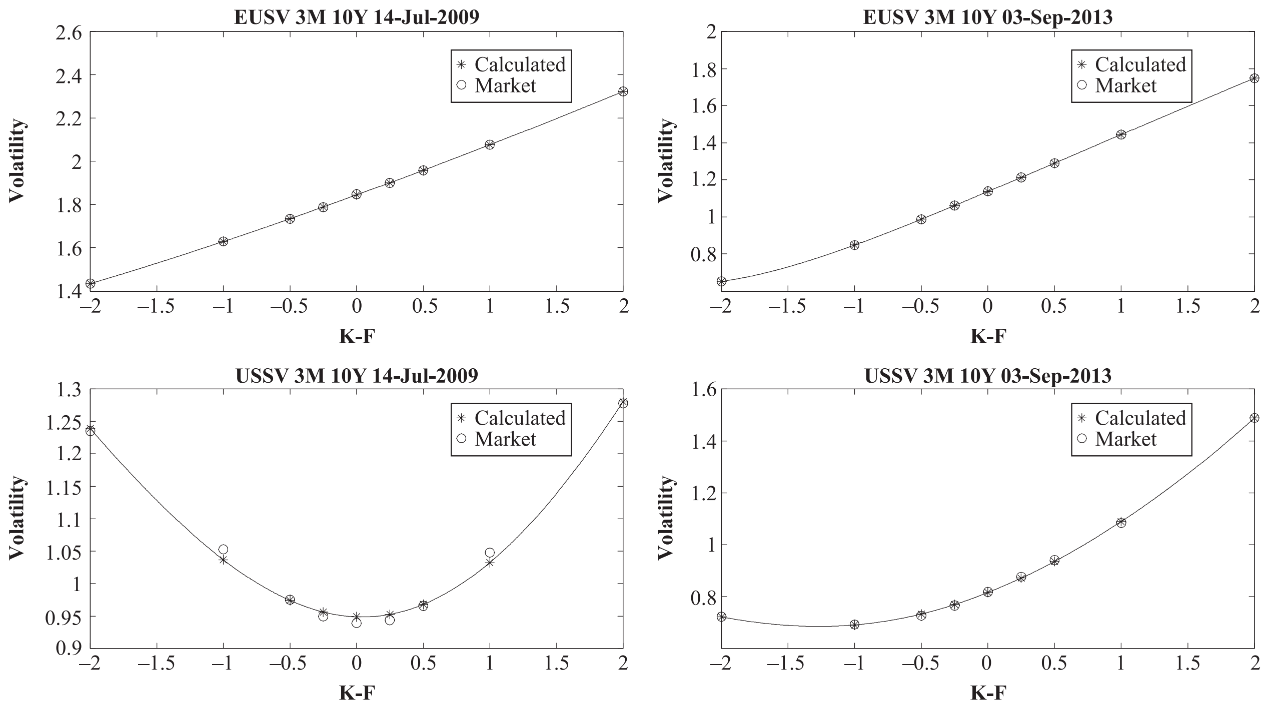
$$a(I) = a_2 I^2 - a_1 I + a_0 - 1,\ \ b(I) = bI,\ \ \rho(I) = cI + d, \tag{19}$$

with the basic equation:

$$a_2 I^2 + \left(-a_1 - \frac{1}{4}b^2 - bc(F-K)\right)I + a_0 + \frac{1}{4}b^2(F-K)^2 - bd(F-K) = 0. \tag{20}$$

EXHIBIT 2

3M by 10Y Implied Volatility (IRV4)



For the same reason, we still set $b = 1$. Then the IRV5 model comes to

$$I = \frac{a_1 + \frac{1}{4} + c(F-K) \pm \sqrt{\left(a_1 + \frac{1}{4} + c(F-K)\right)^2 + 4a_2\left(d(F-K) - \frac{1}{4}(F-K)^2 - a_0\right)}}{2a_2}. \tag{21}$$

According to calibration results, we prefer to choose − rather than +:

$$I = \frac{1}{2a_2}\left(\beta - \sqrt{\beta^2 + 4a_2\left(d(F-K) - \frac{1}{4}(F-K)^2 - a_0\right)}\right),$$
$$\beta = a_1 + \frac{1}{4} + c(F-K). \tag{22}$$

where $a_2 < 0$ and $d^2 < a_0$ will ensure that I is always positive.

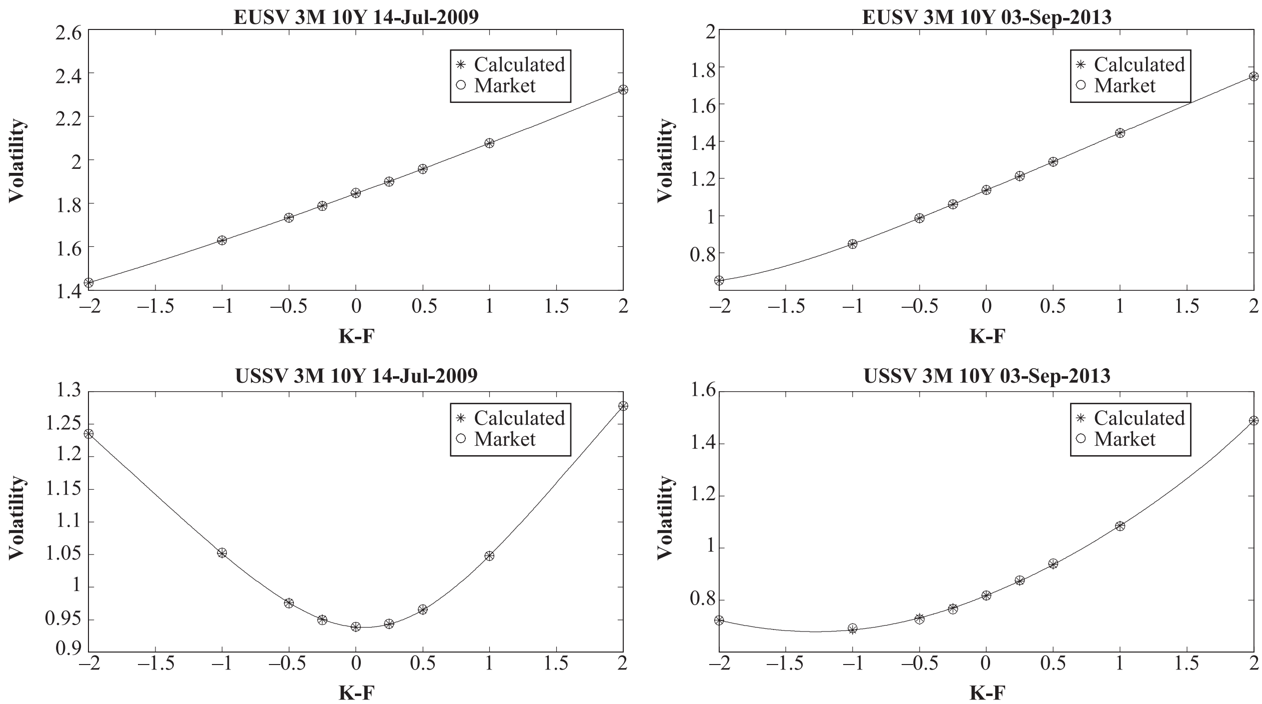
In Exhibit 3, we show the calibration results of IRV5 from swaption market with 3-month by 10-year contracts in both the U.S. and EU markets. We are using the same data from IRV3 so that we don't repeat the data description here. The mean of all RMSE of the calibration of each day is 0.0015 for the U.S. market and 0.0024 for the EU market. IRV5 reduces 55% in-sample pricing error of IRV4 for the U.S. market and 77% for the EU market, 89% of IRV3 for the U.S. market and 93% for the EU market. We can see from the samples plot here that IRV5 fit much better than IRV3.

It is obvious that, with most parameters, IRV5 offers the best fit to market. Thus practitioners can choose each of these three parametrizations by trading off between flexibility and over-parameters. However, IRV4 and IRV5 do have a unique advantage in the proof of being totally arbitrage free, which IRV3—which can be regarded as a special case of IRV5 with some restriction of parameters value—and IRV4 don't have.

IMPLIED REMAINING VOLATILITY

So far we have focused on the so-called implied remaining variances. However, we can also use implied remaining volatility (as used by Carr and Wu [2016], Peeters [2013], and Huitema and Peeters [2014]) in our

EXHIBIT 3
3M by 10Y Implied Volatility (IRV5)



framework, which can be defined by $\omega_t = \sqrt{I_t}$. But this approach will lose some advantages for no-arbitrage, which will be discussed later. Following an approach similar to that in the previous section and assuming

$$d\omega_t = a(\omega_t)v_t dt + b(\omega_t)\sqrt{v_t}\,dZ_t, \tag{23}$$

we have the following basic equation:

$$1 + b(\omega)^2 \frac{(F-K)^2}{\omega^2} - 2\rho(\omega)b(\omega)\frac{(F-K)}{\omega} + 2a(\omega)\omega = 0. \tag{24}$$

Now, we can use the same parameterization of $a(I)$ and $b(I)$ in IRV3 and IRV4 to obtain the following IRVV4 and IRVV5.[1]

- In the IRVV4 model, we first set

$$a(\omega) = -a_1\omega + a_0,\ \ b(\omega) = b\omega,\ \ \rho(\omega) = \rho \tag{25}$$

and then plug into (24),

$$\omega = \frac{a_0 + \sqrt{a_0^2 + 2a_1(b^2k^2 + 2\rho bk + 1)}}{2a_1}. \tag{26}$$

Here $a_1 > 0$ and $\rho^2 < 1$ will ensure that ω is always positive. We call this model IRVV4 since we have 4 parameters (a_0, a_1, b, ρ).

- In the IRVV5 model, we first set

$$a(\omega) = -a_1\omega + a_0,\ \ b(\omega) = b\omega,\ \ \rho(\omega) = c\omega + d \tag{27}$$

and then plug into (24)

$$\omega = \frac{(a_0 + bck) + \sqrt{(bck + a_0)^2 + 2a_1(b^2k^2 + 2bdk + 1)}}{2a_1}. \tag{28}$$

We can choose $a_1 < 0$ and $d^2 < 1$ to ensure that ω is always positive.

The linkage between implied remaining volatility approach and implied remaining variance approach is clear. If we assume the implied remaining volatility has the following dynamics

$$d\omega_t = a(\omega_t)v_t dt + b(\omega_t)\sqrt{v_t}dZ_t, \tag{29}$$

and use the Itô's formula, we will have

$$\begin{aligned} dI_t &= d\omega_t^2 = 2\omega_t d\omega_t + d\omega_t d\omega_t \\ &= 2\omega_t a(\omega_t)v_t dt + 2\omega_t b(\omega_t)\sqrt{v_t}dZ_t + b^2(\omega_t)v_t dt \\ &= (2\omega_t a(\omega_t) + b^2(\omega_t))v_t dt + 2\omega_t b(\omega_t)\sqrt{v_t}dZ_t. \end{aligned} \tag{30}$$

These two approaches are somehow equivalent but tend to suggest different parametric functional forms of implied volatility curves. However, we will show that the implied remaining variance approach has better performance in both curve fitting and no-arbitrage conditions.

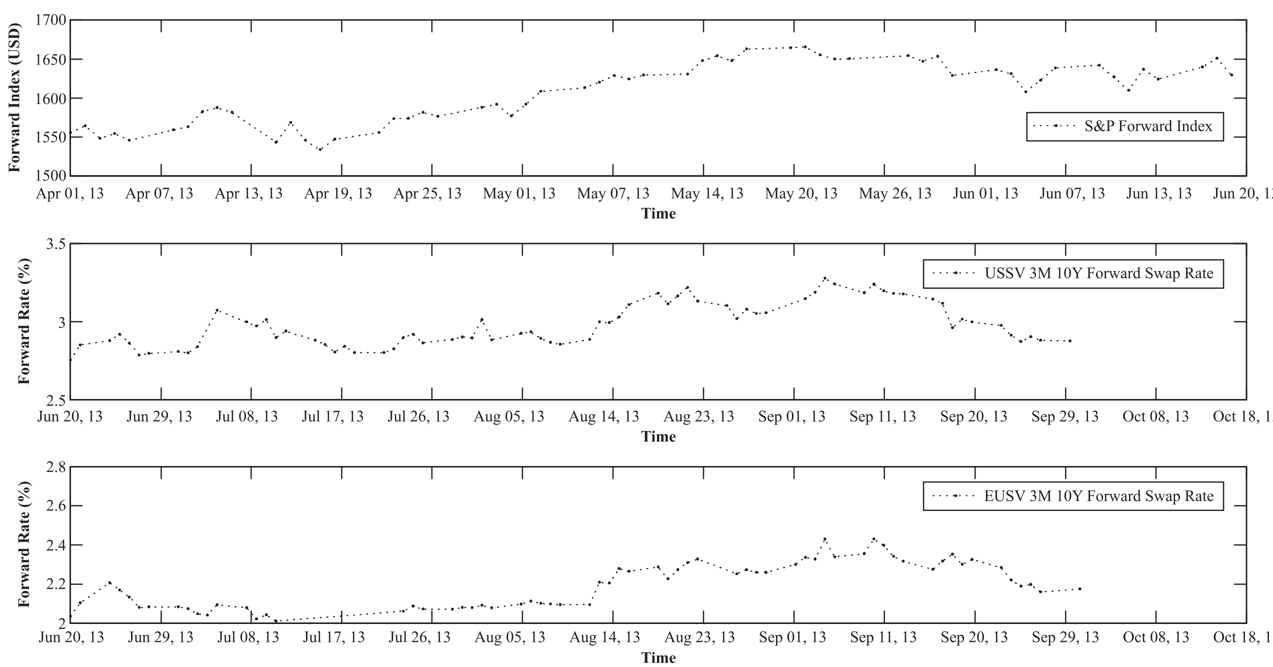
EMPIRICAL ANALYSIS

We use S&P 500 call option prices and U.S./EU swaption prices for empirical analysis for two reasons. First, the options written on the S&P 500 are the most actively traded European-style contracts. Second, the forward swaption rate is usually close to zero in the swaption market, which satisfies our motivation. For the equity option, we employ the end-of-day call option data on the S&P 500 index from April 1, 2013, to June 19, 2013, for the same contract maturating at June 20, 2013. For the swaption market, we employ the U.S. market with swaption contract of 3 months by 10 years from June 20, 2013, to October 1, 2013, and the EU market with swaption contract of 3 months by 10 years from June 20, 2013, to October 1, 2013. The reason to choose this time interval for the swaption market is that every weekday has nine strikes contract with reliable prices. Since the market is always changing, using a three-month period may lead to a cleaner comparison among the five option models.

In Exhibit 4, we provide a plot of the underlyings. The sample periods run from April 1, 2013, to June 19, 2013, for the S&P index and from June 20, 2013, to October 1, 2013, for swap rate in both the U.S. and EU markets. The underlyings are adjusted to forward value consistent with the numerare of options or swaptions, especially for the product of swaption market we use here—the three-month by 10-year swaption in which the numerare is a forward starting annuity.

EXHIBIT 4

S&P 500 Forward Index, from April 1, 2013, to Jun 19, 2013; Forward Swap Rate in U.S. and EU Markets, from June 20, 2013, to October 1, 2013



For the option market, the strikes range from 1,390 to 1,950. Since not all the strikes have liquid options, we eliminate the illiquid points. For swaption market, we always have nine strikes for each day, with ±2%, ±1%, ±0.5%; and ATM relative to the underlying. Note that the forward swap rate in EU market is close to 2%; so for some previous days or later days, the relative −2% strike contract becomes illiquid, which is also a reason for the data choice.

To calibrate the parameters under forward measure of our five models, we use the European Vanilla option data and the 3-month by 10-year swaption datasets described before. We calibrate the five models using daily data in the sample periods and hence obtain sets of parameters for every calendar day and every model. We calculated the in-sample pricing errors via RMSE.

For the daily calibrations, we need to adopt a calibration objective function to minimize squared implied volatility errors as the standard approach. But this approach may force the estimation to assign more weight to relatively high implied volatilities (e.g., ITM and OTM volatilities). We should keep this influence in mind.

1. Collect N swaptions/options (9 for swaptions, about 30~40 for options) implied volatility on the same underlying for every day. Let $\hat{\sigma}_{market}(K_n)$ be the observed implied volatility and $I(K_n)$ its model implied remaining variance as determined by formulas (15), (18), (22), (26), and (28), for $n = 1, 2, \ldots, n$. The difference between $\hat{\sigma}^2_{market}(K_n)T$ and $I(K_n) = \sigma^2_{imp}(K_n)T$ is a function of the values taken by parameters $\Lambda = (b, c, d, \rho, a_0, a_1, a_2)$ in which each model takes its own parameters. For each n, define

$$\varepsilon_n[\Lambda] = \hat{\sigma}^2_{market}(K_n)T - I(K_n).$$

2. Find parameters Λ to solve

$$\min_{\Lambda} \sum_{n=1}^{N} \left|\varepsilon_n[\Lambda]\right|^2.$$

Since there are large nonlinear properties in the objective function caused by square roots in the models, the main computing burden is to find the appropriate initial value. One can refine the objective function by some transformation to eliminate the square roots (but this may generate extra solutions), or use numerical methods to search for better points to try, or change this problem into a quasilinear form. Then, this step results in an estimation of the parameter values at date t. We have repeated these two steps for each day in the sample periods.

Here we are plotting some of the sample dates to show the fit visually and offer a table to report the average performance for each model. Note that all the implied volatility is in normal setting rather than log-normal setting, so the implied normal volatilities are sometimes big or small than log-normal case depending on the underlying scale where the log-normal implied volatilities are scaleless.

Since each implied volatility curve has only nine points even close to a straight line in the U.S. market in sample period, it is difficult to see the difference between the five models; this is why we have chosen to present the EU market. For the option market, we have 30−40 points for each day, and it is easier to present the difference visually. The results in Exhibits 5–7 show that in both options and swaptions markets, the IRV5 fits the market best where IRVV4 and IRVV5 have some unpleasant restriction on their shapes, especially for OTM points. Since these curves are only representative samples within whole sample sets, we show the mean RMSE for each model in each market in Exhibit 8.

Exhibit 8 reports the mean of RMSE from daily calibration of the S&P index option, U.S. 3-month by 10-year swaption, and EU 3-month by 10-year swaption. RMSE is the root mean squared error. We calibrate the market daily implied volatilities with different strikes and collect the errors through RMSE to calculate their means for each model in each market. The models are abbreviated as outlined in previous sections. Note that the implied volatilities are calculated from the Bachelier model and thus are consistent with the dimensionality of their underlying: dollar for S&P index, percentage for swap rate. Exhibit 8 shows that, in most cases, the IRVV4 has the largest pricing errors and the IRV5 has the least. Even though IRVV4 and IRVV5 have the same numbers of parameters as IRV4 and IRV5, respectively, errors of IRVV4 and IRVV5 are still much bigger than those of IRV4 and IRV5. We also find that it is much more difficult for IRVV4 and IRVV5 to find stable and reliable solutions since the optimization is too sensitive to the initial value.

EXHIBIT 5

Date: April 1, 2013. Data: S&P 500 Index Options' Implied Volatilities, Maturing at June 20, 2013

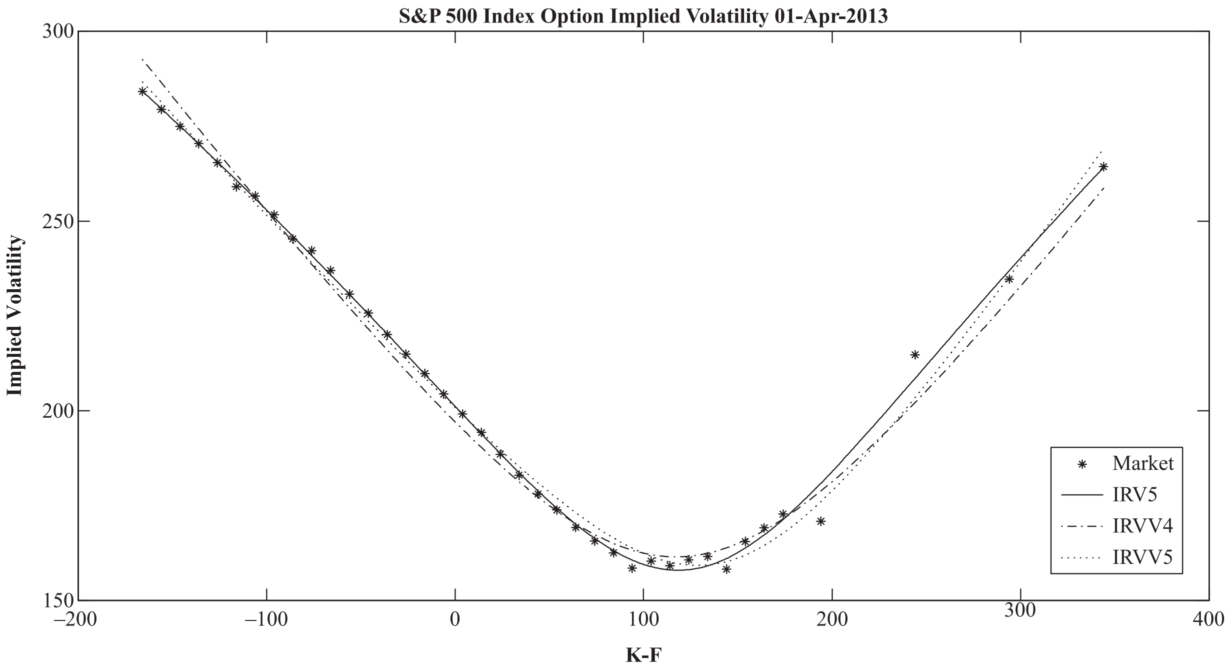
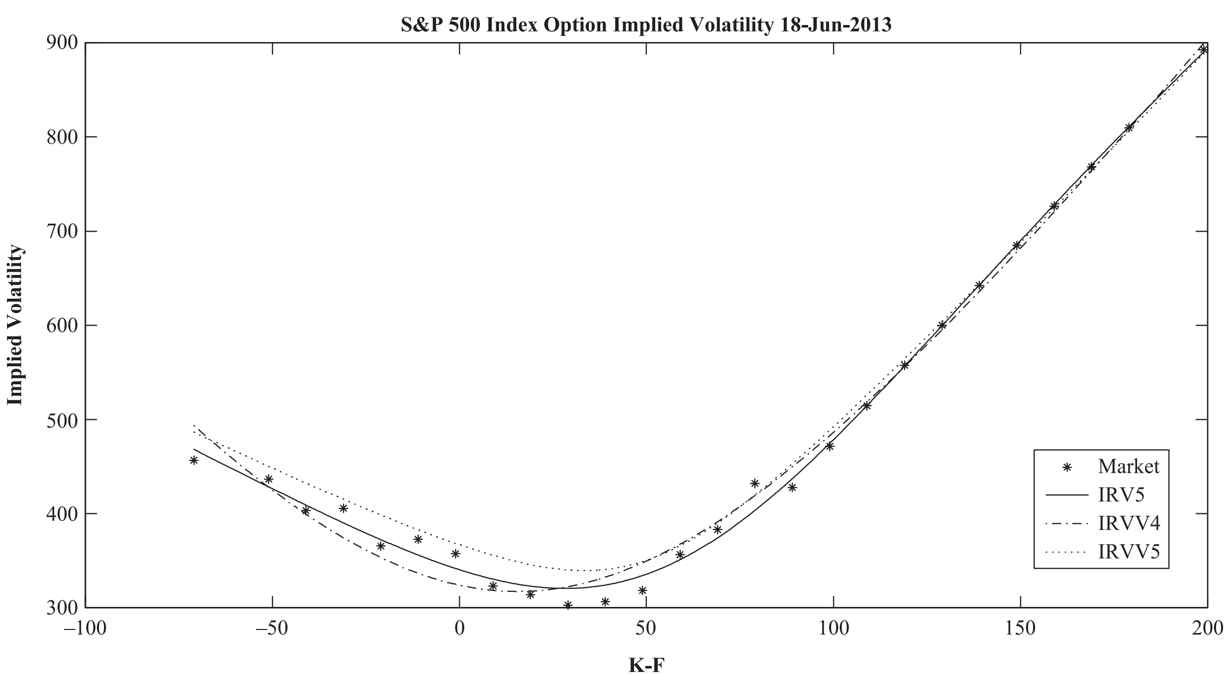


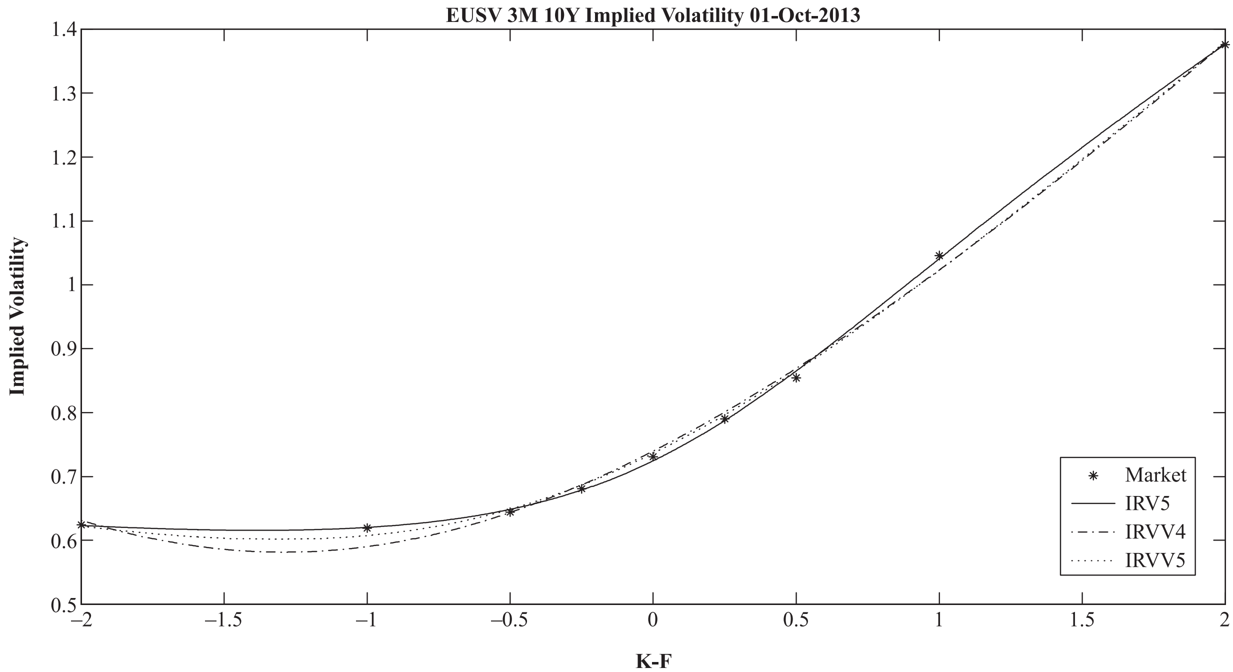
EXHIBIT 6

Date: June 18, 2013. Data: S&P 500 Index Options, Maturing at June 20, 2013



E X H I B I T 7

Date: October 1, 2013. Data: 3M by 10Y Call Swaption in EU Market



E X H I B I T 8

Mean of RMSE for Each Model in Each Market

Models Markets	S&P Option	U.S. Swaption	EU Swaption
IRV3	8.6148	4.3e–03	2.44e–02
IRV4	7.3730	6.6884e–04	1.42e–02
IRV5	4.2406	4.8151e–04	6.6e–03
IRVV4	20.1279	3.1e–03	1.49e–02
IRVV5	17.2203	2.0e–03	1.20e–02

We will see in the next section that IRVV4 and IRVV5 actually can't fulfil the no-arbitrage conditions. This implies that, even if we could use both volatility or variance approaches, we could not choose the parametric functional form arbitrarily. There are some "invisible" structures that control the shape of the implied volatility surface. It seems that the simple parametric functional form of variance approach is closer to the "true structure." The next section will address the no-arbitrage conditions.

No Arbitrage

Researchers are sparing no effort to tackle the problem of finding conditions that may be necessary and/or sufficient to ensure that prices/vols are free of arbitrage. For example, Lee [2004] investigated the large/low strikes linear behavior of time scaled implied variance with the boundary condition, and an article by Carr and Madan [2005] presented a simple algorithm based on the observation that the absence of call spread, butterfly spread, and calendar spread arbitrages is sufficient to exclude all static arbitrage from a set of option price quotes across strikes and maturities. Inspired by these simple conclusions, researchers have also presented several but similar model free conditions for no-arbitrage implied volatility surface parameterization; for example, Roper's [2010] work showed sufficient conditions and nearly necessary conditions to be free from static arbitrage, which is consistent with Roger's bound, and Carr and Wu's [2010] work showed a summary about no-arbitrage conditions. Here, our approach to find the no-arbitrage conditions for parameters selection is based on the research of Roper [2010] for no static arbitrage and Carr and Wu [2010] for no dynamic arbitrage.

No dynamic arbitrage. Our framework is consistent with the "PDE" in Carr and Wu [2010] where equation (10) is exactly from no dynamic arbitrage, which was rigorously defined and proven by Carr and Wu [2016].

This consistency puts constraints on the shape of the initial volatility curve if we specify the parametric functional form of dynamics of implied remaining variance. So the curve generated in this approach is originally free of dynamic arbitrage. Let's see the static arbitrage.

No static arbitrage. In this subsection, we follow Roper [2010] in excluding the static arbitrage. Although our setting is different than Roper's, our goal is the same—to transfer conditions for call price surface into implied volatility surface. First, let us state the definition here

Definition 1. There is no static arbitrage in a call price surface C if there exists a non-negative martingale X on some stochastic basis $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, P)$ with

$$C(K, \tau) = E\big((X_\tau - K)^+ \mid \mathcal{F}_0\big)$$

for each $(K, \tau) \in [0, \infty) \times [0, \infty)$, with $X_0 = S_0$, (i.e., the current stock price). If such a martingale and probability space exists, we say that the call price surface is free of static arbitrage.

Then, we will give a sufficient and nearly necessary (truly necessary for IRV models) condition for the call price surface to be free from static arbitrage.

Theorem 2. A call price surface $C(K, \tau)$ is free of static arbitrage (exists consistent martingale) if and only if

1. $C(K, \tau)$ is a convex function w.r.t. K.
2. $C(K, \tau)$ is non-decreasing w.r.t. τ.
3. $\lim_{K\to\infty} C(K, \tau) = 0$.
4. $C(K, 0) = (S_T - K)^+$
5. $(S_t - K)^+ \leq C(K, \tau) \leq S_t$.

The proof can be found in Roper's work [2010].

Actually, these conditions are quite fundamental. Although in the literature there are many other conditions for no arbitrage, it is easy to check some examples to show that they could be implied by the conditions in Theorem 2. For example, see Lee's condition (Lee [2004]) and Durrleman's conditions (Durrleman [2010]) under BSM setting. These conditions could be simply derived from the properties of call price function in Theorem 2. Since we are under Bachelier model, things will be a little different with BSM setting. Now, let us translate these conditions to our IRV surface under Bachelier Model. Denoting $k = K - F$, we have the following Theorem 3. Note that the proof can also be found in the Appendix.

Theorem 3. Assume that our call price surface $C(K, \tau)$ is smooth—twice continuous differentiable[2] w.r.t. strike K, then we could state that the implied remaining variance surface $I(k, \tau)$[3] is also smooth—twice continuous differentiable w.r.t. moneyness k. Then, the no-arbitrage conditions in Theorem 2 are equivalent to the following conditions one by one:

1. Condition 1 in Theorem 2 is equivalent to
$$g(k) = 1 - \frac{k}{2I} I_k - \frac{1}{4I} I_k^2 + \frac{k^2}{4I^2} I_k^2 + \frac{1}{2} I_{kk} \geq 0,$$
$$\text{for any } k < F, \ \tau > 0; \tag{31}$$
2. condition 2 in Theorem 2 is equivalent to that $I(k, \tau)$ is non-decreasing w.r.t. τ;
3. assume $\lim_{k\to\infty} \frac{1}{|k|^\alpha} =: \gamma$ where $\gamma < +\infty$ and $\alpha > 0$ exist, then, the condition 3 in Theorem 2 is equivalent to $\alpha < 2$;
4. condition 4 in Theorem 2 is equivalent to $\lim_{\tau\to 0} I(k, \tau) = 0$;
5. condition 5 in Theorem 2 is equivalent to $0 \leq I_t \leq \lambda_0^2 F_t^2$ where λ_0 is a constant.

Remark. Note that condition 5 in Theorem 3 may not be necessary for us under Bachelier model. This condition is actually from the restriction that the martingale process of underlying asset cannot be negative. Under Bachelier model, however, this is not the case.

Model check. We will see that some of the models we described before are able to be free of static arbitrage while others are not. In the case of fixed maturity time T at time t in Bachelier model, we need to check only conditions 1 and 3 in Theorem 3. The proofs of the following conclusions can be found in the Appendix.

Conclusion 4. IRV3, IRVV4, IRVV5 can't satisfy condition 3 in Theorem 3, while IRV4, IRV5 can.

According to Conclusion 4, IRV3, IRVV4, IRVV5 are not able to be free of arbitrage, while IRV4 and IRV5 survive. Investigation of condition 1 in Theorem 3 is still needed. Note that the implied remaining volatility approaches described above are similar to the proportional volatility approach proposed in Carr and Wu [2016]. Consistent with our Conclusion 4, it is worth noting that the proportional volatility approach also implies arbitrage in the same way. It seems that the implied remaining volatility approach performs much worse than the implied remaining variance approach in satisfying no-arbitrage. Now, let's see the following conclusions for IRV4 and IRV5 respectively.

Conclusion 5. Under IRV4 model with (a_0, a_1, c, d), for any fixed $\tau > 0$; if $c < 0$; $d^2 \leq a_0$, $a_1 + \frac{1}{4} + cF > 0$, $6a_1 + 12cd + 1 > 0$ are satisfied, then the implied remaining variance $I(k)$ is free of static arbitrage.

Conclusion 6. Under IRV5 model with (a_0, a_1, a_2, c, d), for any fixed $\tau > 0$; if $a_2 < 0$; $d^2 \leq a_0$, $a_1 \leq -\frac{1}{3a_2^2}(2c^2 - a_2)(c^2 - a_2) - 2cd - \frac{1}{4}$ are satisfied, then the implied remaining variance $I(k)$ is free of static arbitrage.

Remark. Though the conditions in Conclusions 5 and 6 are only sufficient, not necessary, it still shows that for both IRV4 and IRV5, there exists a class of parameters where the consistent implied volatility curves are free of static arbitrage. For application, practitioners could simply plot $g(k)$ in a finite interval as the indicator for static arbitrage. (The asymptotic behavior of $g(k)$ is restricted to be positive under IRV4 and IRV5 models; see the proofs in Appendix.) Note that $g(k) \geq 0$ is equivalent to convexity $\frac{\partial^2 C}{\partial K^2} \geq 0$, which is equivalent to a positive implied transition density.[4]

SUMMARY

The negative behavior of some underlying assets is important for practitioners. To deal with this problem, we follow the approach used by Carr and Sun [2013] and then by Carr and Wu [2016] and generalize it to the Bachelier model, which is under normal setting instead of log-normal setting, so that it admits negativity. We first propose three parameterizations of the dynamics which lead to three models—IRV3, IRV4, IRV5—with calibration results. We further show that, although we could also choose remaining volatility instead of remaining variance to obtain some other models (such as IRVV4 and IRVV5), the choice of implied remaining variance is intrinsic since IRV4 and IRV5 provide better fitting than other models. Finally, we show that the choice of the parameterization is not arbitrarily by the no-arbitrage analysis. The condition $\lim_{K\to\infty} C(K,T) = 0$ puts strict constraints on the large strike behavior of implied volatility and excludes the choice of IRV3, IRVV4, and IRVV5, which is stated in Theorem 3 and Conclusion 4. Further, Conclusions 5 and 6 show that under some constraints of parameters, our IRV4 and IRV5 are free of arbitrage. Further discussion about generalization and application of this framework can be found in the work of Peeters [2013], Huitema and Peeters [2014], and Carr and Wu [2016]. Future research can still be conducted on our no-arbitrage models. First, we can explore the hedging behavior of our new models, according to our joint dynamics of the implied remaining variance and underlying, where we believe better risk Greeks could be generated. Second, this approach could be used for interpolation and extrapolation of the implied volatility surface, or be used as indicator of the kinds of level of implied volatility surface. Third, most of the consequences we have derived here under Bachelier setting could also be straightly transferred to a BSM setting. Finally, since we have fixed time to maturity, our current framework considers only one slice of the implied volatility surface. Future research can investigate how to add term structure into this framework to generate the whole implied volatility surface free of arbitrage and how to introduce jump in this framework.

Proofs

Proof of Theorem 3. First, let us prove that the volatility surface is also smooth if the call price surface is smooth. Let us define the function

$$G(K, x, y) = x - \sqrt{y}\phi\left(\frac{F-K}{\sqrt{y}}\right) - (F-K)\Phi\left(\frac{F-K}{\sqrt{y}}\right). \tag{32}$$

Then we have

$$\frac{\partial G(K, x, y)}{\partial y} = -\frac{1}{2\sqrt{y}}\phi\left(\frac{F-K}{\sqrt{y}}\right) < 0,$$
$$\frac{\partial G(K, x, y)}{\partial K} = \Phi\left(\frac{F-K}{\sqrt{y}}\right), \quad \frac{\partial G(K, x, y)}{\partial x} = 1. \tag{33}$$

For some strike $K_0 > 0$; we have the consistent option price C_0 and implied remaining variance I_0. So, by definition

$$G(K_0, C_0, I_0) = 0, \tag{34}$$

and $G(K, x, y)$, $\dfrac{\partial G(K, x, y)}{\partial y}, \dfrac{\partial G(K, x, y)}{\partial K}, \dfrac{\partial G(K,x,y)}{\partial x}$ are all continuous for $\mathbb{R} \times \mathbb{R} \times (0, +\infty)$. Thus, the implicit function theorem implies that, there exists a continuous function $\gamma(K, x)$ such that

$$G(K,x,\gamma(K,x)) = 0, \;\; I_0 = \gamma(K_0,C_0). \tag{35}$$

where the derivatives

$$\frac{\partial \gamma(K, x)}{\partial K} := \gamma_1(K, x), \;\; \frac{\partial \gamma(K, x)}{\partial x} := \gamma_2(K, x) \tag{36}$$

are existent and continuous. For every $K > 0$ we have consistent call price $C(K)$ and implied remaining variance $I(K) = \gamma(K, C(K))$. Since $\dfrac{\partial C(K)}{\partial K}$ exists and is continuous, using the chain rules,

$$\frac{\partial I(K)}{\partial K} = \gamma_1(K, C(K)) + \gamma_2(K, C(K))\frac{\partial C(K)}{\partial K}$$

also exists and is continuous for any $K > 0$. Following the same procedure, we can also prove that $\dfrac{\partial^2 I(K)}{\partial K^2}$ exists and is continuous, hence

$$I_k := \frac{\partial I}{\partial k} = -\frac{\partial I}{\partial K}, \;\; I_{kk} := \frac{\partial^2 I}{\partial k^2} = \frac{\partial^2 I}{\partial K^2}$$

also exist and are continuous. With the smoothness of the implied remaining variance, let's consider those five conditions.

1. Since the call price surface is smooth, the condition 1 is equivalent to $\dfrac{\partial^2 C}{\partial K^2} \geq 0$. Note that

$$\frac{\partial^2 C}{\partial K^2} = \frac{\partial^2 C^B}{K^2} + \frac{\partial^2 C^B}{\partial I\, \partial K} I_K + \frac{\partial^2 C^B}{\partial I^2} I_K^2 + \frac{\partial C^B}{\partial I} I_{KK} \tag{37}$$

$$= \frac{1}{\sqrt{I}}\phi\left(\frac{F-K}{\sqrt{I}}\right) \left[1 + \frac{F-K}{2I} I_K - \frac{1}{4I} I_K^2 + \frac{(F-K)^2}{4I^2} I_K^2 + \frac{1}{2} I_{KK}\right]. \tag{38}$$

So $\dfrac{\partial^2 C}{\partial K^2} \geq 0$ is equivalent to

$$1 + \frac{F-K}{2I} I_K - \frac{1}{4I} I_K^2 + \frac{(F-K)^2}{4I^2} I_K^2 + \frac{1}{2} I_{KK} \geq 0 \tag{39}$$

for every $K > 0$; which is equivalent to

$$1 - \frac{k}{2I} I_k - \frac{1}{4I} I_k^2 + \frac{k^2}{4I^2} I_k^2 + \frac{1}{2} I_{kk} \geq 0 \quad \text{for every } k > -F, \tag{40}$$

where $k = K - F$.

2. Since $\dfrac{\partial C^B}{\partial I} = \dfrac{1}{2\sqrt{I}}\phi\left(\dfrac{F-K}{\sqrt{I}}\right) \geq 0$, so $C^B(K, I)$ is non-decreasing w.r.t. I. I is actually the only way for τ to enter into $C^B(K, I)$.
3. First, let us prove the following statement: assume that

$$\lim_{k\to\infty} \frac{I(k)}{k^2} =: \beta^2, \;\; \beta > 0 \text{ exists,} \tag{41}$$

then $\lim_{k\to\infty} C(I(k), k) = 0$ can't hold. In fact,

$$\begin{aligned}\lim_{k\to\infty} \frac{C(I(k),k)}{k} &= \beta\phi\left(-\frac{1}{\beta}\right) - \Phi\left(-\frac{1}{\beta}\right) \\ &= \beta\left(\phi\left(-\frac{1}{\beta}\right) - \frac{1}{\beta}\Phi\left(-\frac{1}{\beta}\right)\right) \\ &= \beta\int_{-\infty}^{-\frac{1}{\beta}} \Phi(u)du > 0.\end{aligned} \tag{42}$$

So that $\lim_{k\to\infty} C(I(k), k) = +\infty$. Now we could find that for any $\gamma > 0$; if

$$\lim_{k\to\infty} \frac{1}{|k|^2} = \gamma$$

we must have $\lim_{k\to\infty} C(I, k) = +\infty$, but what we need is $\lim_{k\to\infty} C(I, k) = 0 < +\infty$. Considering the $C(\cdot, k)$ is non-decreasing, we can conclude that for any $\alpha \geq 2$; if

$$\lim_{k\to\infty} \frac{1}{|k|^\alpha} = \gamma, \gamma < +\infty$$

exists, then $\lim_{k\to\infty} C(I, k) = +\infty$. Inversely, if $\alpha < 2$ and

$$\lim_{k\to\infty}\frac{I}{|k|^{\alpha}}=\gamma, \gamma<+\infty,$$

then

$$C(I,k)=k\left(\sqrt{\frac{I}{k^2}}\phi\left(-\sqrt{\frac{k^2}{I}}\right)-\Phi\left(-\sqrt{\frac{k^2}{I}}\right)\right)$$
$$\to k\left(\frac{1}{|k|^{1-\frac{\alpha}{2}}}\sqrt{\gamma}\phi\left(-|k|^{1-\frac{\alpha}{2}}\frac{1}{\sqrt{\gamma}}\right)-\Phi\left(-|k|^{1-\frac{\alpha}{2}}\frac{1}{\sqrt{\gamma}}\right)\right)$$
$$\to 0. \tag{43}$$

4. Since $\frac{\partial C^B}{\partial I}>0$ for any $K>0$; and $\lim_{I\to 0}C^B(K,I)=C(K,0)=(S_T-K)^+$.
5. Actually, the condition 5 in the Bachelier setting is a little bit difficult to transform to I. Denote $\omega=\sqrt{I}$. Note that in BSM setting, since the underlying process is non-negative, it is easy to prove that condition 5 is equivalent to a positive volatility which is natural. But under our Bachelier model, it needs more things. First, we want $(F_t-K)^+\le C(K,\tau)$, which means

$$C_t=\omega\phi\left(-\frac{k}{\omega}\right)-k\Phi\left(-\frac{k}{\omega}\right)$$
$$=\omega\int_{-\infty}^{-\frac{k}{\omega}}\Phi(u)du\ge(-k)^+. \tag{44}$$

This is equivalent to

$$\int_{-\infty}^{-\frac{k}{\omega}}\Phi(u)du\ge-\frac{k}{\omega},\ \ -F<k<0. \tag{45}$$

Denote λ_0 is the real value which satisfies $\int_{-\infty}^{\lambda}\Phi(u)\,du=\lambda$. It is easy to prove that the solution is unique and numerical result shows that $\lambda_0\approx 9.2258$. And we could prove that if $0\le\omega_t\le\lambda_0F_t$, the condition 5 must hold.

Proof of Conclusion 4. For IRV3 model, we have

$$\lim_{k\to\infty}\frac{I}{k^2}=\lim_{k\to\infty}\frac{\frac{1}{4}k^2+\rho k+a_0}{a_1+\frac{1}{4}}=\frac{1}{4a_1+1}. \tag{46}$$

So, in IRV3, $\alpha=2$.
For IRV4 model, we have

$$\lim_{k\to\infty}\frac{I}{k}=\lim_{k\to\infty}\frac{\frac{1}{4}k^2+dk+a_0}{\left(a_1+\frac{1}{4}-ck\right)k}=-\frac{1}{4c}. \tag{47}$$

So, in IRV4, $\alpha=1<2$.
For IRV5 model, we have

$$\lim_{k\to\infty}\frac{I}{k}=\lim_{k\to\infty}\frac{a_1+\frac{1}{4}-ck-\sqrt{\left(a_1+\frac{1}{4}-ck\right)^2-4a_2\left(\frac{1}{4}k^2+dk+a_0\right)}}{2a_2k}$$
$$=-\frac{c+\sqrt{c^2-a_2}}{2a_2}. \tag{48}$$

So, in IRV5, $\alpha=1<2$.
For IRVV4 model, we have

$$\lim_{k\to\infty}\frac{I}{k^2}=\left(\lim_{k\to\infty}\frac{\omega}{k}\right)^2$$
$$=\left(\lim_{k\to\infty}\frac{a_0+\sqrt{a_0^2+2a_1\left(b^2k^2+2\rho bk+1\right)}}{2a_1k}\right)^2$$
$$=\frac{b^2}{2a_1}. \tag{49}$$

So, in IRVV4, $\alpha=2$.
For IRVV5 model, we have

$$\lim_{k\to\infty}\frac{I}{k^2}=\lim_{k\to\infty}\frac{\omega^2}{k^2}=\left(\lim_{k\to\infty}\frac{\omega}{k}\right)^2=\left(\frac{bc+\sqrt{b^2c^2+2a_1b^2}}{2a_1}\right)^2. \tag{50}$$

So, in IRVV5, $\alpha=2$. According to condition 3 in Theorem 3, we could easily get our conclusion.

Proof of Conclusion 5. Now, we only need to prove that IRV4 satisfies the positive property and condition 1 in Theorem 3. Denote $k=K-F$, $\alpha=\alpha_1+\frac{1}{4}$, we have

$$I(k) = \frac{\frac{1}{4}k^2 + dk + a_0}{\alpha - ck}. \tag{51}$$

Note that $a_1 + \frac{1}{4} + cF > 0$ implies $\alpha - ck > 0$; $d^2 \le a_0$ implies $\frac{1}{4}k^2 + dk + a_0 = \left(\frac{k}{2} + d\right)^2 + a_0 - d^2 \ge 0$. So the positive property is straightforward. Then, let's see the convexity. Denote

$$\begin{aligned} A &= \frac{1}{4}k^2 + dk + a_0 = \left(\frac{k}{2} + d\right)^2 + a_0 - d^2 \ge 0, \\ B &= \alpha - ck, \ C = a_0 - d^2 \ge 0, \\ D &= \frac{1}{2}(\alpha + 2cd)^2 + 2c^2(a_0 - d^2) \ge 0. \end{aligned} \tag{52}$$

Then, we can find

$$\begin{aligned} I_k &= \frac{\partial I(k)}{\partial k} = \frac{-\frac{1}{4}ck^2 + \frac{1}{2}\alpha k + \alpha d + a_0 c}{(\alpha - ck)^2} \\ &\to = -\frac{1}{4c}\,(k \to \infty), \\ I_{kk} &= \frac{\partial^2 I(k)}{\partial k^2} = \frac{D}{B^3} = \frac{\frac{1}{2}(\alpha + 2cd)^2 + 2c^2(a_0 - d^2)}{(\alpha - ck)^3} \\ &\to 0\,(k \to \infty), \\ I &= \frac{A}{B}, \frac{I}{k} \to -\frac{1}{4c}(k \to \infty). \end{aligned} \tag{53}$$

We could first state that the asymptotic behavior of the function $g(k)$ under IRV4 is positive. Actually,

$$\begin{aligned} g(k) &= 1 - \frac{k}{2I}I_k - \frac{1}{4I}I_k^2 + \frac{k^2}{4I^2}I_k^2 \\ &\quad + \frac{1}{2}I_{kk} \to \frac{3}{4} > 0\,(k \to \infty). \end{aligned} \tag{54}$$

So when k goes to infinity, $g(k)$ is always positive. Further, we could see that

$$\begin{aligned} g(k) &= 1 - \frac{k}{2I}I_k - \frac{1}{4I}I_k^2 + \frac{k^2}{4I^2}I_k^2 + \frac{1}{2}I_{kk} \\ &= \left(\frac{k}{2I}I_k - \frac{1}{2}\right)^2 + \frac{3}{4} + \frac{1}{2}I_{kk} - \frac{1}{4I}I_k^2. \end{aligned} \tag{55}$$

If we define $h(k) = \frac{3}{4} + \frac{1}{2}I_{kk} - \frac{1}{4I}I_K^2$, it is straightforward that if $h(k)$ is always positive, so is $g(k)$. Now let us investigate $h(k)$.

$$\begin{aligned} h(k) &= \frac{3}{4} + \frac{1}{2}I_{kk} - \frac{1}{4I}I_k^2 \\ &= \frac{3AB^3 + 2AD - \left(\left(\frac{1}{2}k + d\right)B + cA\right)^2}{4AB^3}, \end{aligned} \tag{56}$$

$$\begin{aligned} 4AB^4 h(k) &= C(\alpha - ck)^3 + 3\left(\left(\frac{k}{2} + d\right)^2 + C\right)^2 c^2(\alpha - ck) \\ &\quad + \left(\left(\frac{k}{2} + d\right)^2 + C\right)(\alpha - ck)^2\,[ck + 2cd + 3(\alpha - ck)^2], \end{aligned} \tag{57}$$

The conditions $c < 0$; $a_1 + \frac{1}{4} + cF > 0$ lead to

$$B = \alpha - ck = a_1 + \frac{1}{4} + cF - cK > 0, \tag{58}$$

while the condition $d^2 \le a_0$ lead to

$$C = a_0 - d^2 \ge 0. \tag{59}$$

Further, with $6a_1 + 12cd + 1 > 0$ which is equivalent to $12\alpha + 24cd - 1 > 0$; we have

$$\begin{aligned} &\frac{1}{3}(ck + 2cd + 3(\alpha - ck)^2) \\ &= \left(ck + \frac{1 - 6\alpha}{6}\right)^2 + \frac{24cd + 12\alpha - 1}{36} > 0. \end{aligned} \tag{60}$$

Then, we could state that $4AB^4h(k) \ge 0$, $A = \left(\frac{k}{2} + d\right)^2 + C \ge 0$; $B > 0$; which leads to $h(k) \ge 0$; which lead to $g(k) \ge 0$. The proof is complete.

Proof of Conclusion 6. Now, we only need to prove that IRV5 satisfies the positive property and condition 1 in Theorem 3. Denote $k = K - F$, we have

$$\begin{aligned} I(k) &= \frac{1}{2a_2}\left(\beta - \sqrt{\beta^2 - 4a_2\gamma}\right), \\ \alpha &= a_1 + \frac{1}{4},\ \beta = \alpha - ck,\ \gamma = \frac{1}{4}k^2 + dk + a_0. \end{aligned} \tag{61}$$

Denote $k = K - F$, $\alpha = \alpha_1 + \frac{1}{4}$, we can rewrite our IRV5 model into

$$I_t = \frac{1}{2a_2}\left(a_1 + \frac{1}{4} - ck - \sqrt{\begin{array}{l}\left(a_1 + \frac{1}{4} - ck\right)^2 \\ -a_2(k^2 + 4dk + 4a_0)\end{array}} \right)$$
$$= \frac{-\alpha - 2cd}{2(c^2 - a_2)} - \frac{\sqrt{c^2 - a_2}}{2a_2}$$
$$\left\{ \begin{array}{c} \frac{c}{\sqrt{c^2 - a_2}}\left(k - \frac{\alpha c + 2a_2 d}{c^2 - a_2}\right) \\ + \sqrt{\begin{array}{l}\left(k - \frac{\alpha c + 2a_2 d}{c^2 - a_2}\right)^2 \\ + \frac{4a_0a_2^2 - a_2\alpha^2 - 4a_0a_2c^2 - 4a_2^2d^2 - 4a_2\alpha cd}{(c^2 - a_2)^2}\end{array}} \end{array} \right\}, \tag{62}$$

where

$$(\tilde{a}, \tilde{b}, \tilde{\rho}, \tilde{m}, \tilde{\sigma}^2) = \left(\begin{array}{c} \frac{\frac{\alpha}{2} + cd}{a_2 - c^2}, \frac{\sqrt{c^2 - a_2}}{-2a_2}, \frac{c}{\sqrt{c^2 - a_2}}, \frac{\alpha c + 2a_2 d}{a_2 - c^2}, \\ \frac{4a_0a_2^2 - a_2\alpha^2 - 4a_0a_2c^2 - 4a_2^2d^2 - 4a_2\alpha cd}{(c^2 - a_2)^2} \end{array} \right). \tag{63}$$

Then, we have to ensure that $\tilde{a} \in \mathbb{R}, \tilde{b} \geq 0, |\tilde{\rho}| < 1$, $\tilde{m} \in \mathbb{R}, \tilde{\sigma} > 0$ thus I is always positive. Actually our IRV5 has the following parameters constraints $a_2 < 0$, $d^2 < a_0$ which is enough for it, since $d^2 < a_0$ ensures that $\gamma > 0$ by

$$\gamma = \frac{1}{4}k^2 + dk + a_0 = \left(\frac{k}{2} + d\right)^2 + a_0 - d^2$$

then $a_2 < 0$ leads to $\sqrt{\beta^2 - 4a_2\gamma} > \beta$ which leads to $I(k) > 0$ for every $k \in \mathbb{R}$. These constraints also lead to

$$\tilde{b} = \frac{\sqrt{c^2 - a_2}}{-2a_2} \geq 0, |\tilde{\rho}| = \frac{|c|}{\sqrt{c^2 - a_2}} \leq 1,$$
$$\tilde{\sigma}^2 = \frac{4a_0a_2^2 - a_2\alpha^2 - 4a_0a_2c^2 - 4a_2^2d^2 - 4a_2\alpha cd}{(c^2 - a_2)^2}$$
$$= -\frac{1}{a_2}\frac{(\alpha + 2cd)^2 + 4(c^2 - a_2)(a_0 - d^2)}{(c^2 - a_2)^2} \geq 0 \tag{64}$$

which is well defined. For the convexity, we first check the asymptotic behavior of $g(k)$. Since

$$\frac{I}{k} \to \frac{c + \sqrt{c^2 - a_2}}{2a_2}\ (k \to \infty),$$
$$I_k = \frac{1}{2a_2}\left(-c + \frac{c(\alpha - ck) + a_2(k + 2d)}{\sqrt{(\alpha - ck)^2 - 4a_2\left(\frac{1}{4}k^2 + dk + a_0\right)}} \right)$$
$$\to -\frac{1}{2a_2}\left(c + \sqrt{c^2 - a_2}\right)\ (k \to \infty),$$
$$I_{kk} = \left(\frac{(c^2 - a_2)(4a_0a_2 - \alpha^2) + (c\alpha + 2da_2)^2}{2a_2\sqrt{(\beta^2 - 4a_2\gamma)^3}} \right)$$
$$\to 0\ (k \to \infty), \tag{65}$$

so that

$$g(k) = 1 - \frac{k}{2I}I_k - \frac{1}{4I}I_k^2 + \frac{k^2}{4I^2}I_k^2 + \frac{1}{2}I_{kk}$$
$$\to 1 - \frac{1}{2} - 0 + \frac{1}{4} + 0 = \frac{3}{4} > 0. \tag{66}$$

Further, we could see that

$$g(k) = \left(\frac{k}{2I}I_k - \frac{1}{2}\right)^2 + \frac{1}{2}I_{kk} + \frac{3}{4} - \frac{1}{4I}I_k^2. \tag{67}$$

Note that

$$I_{kk} = \frac{(\alpha + 2cd)^2 + 4(c^2 - a_2)(a_0 - d^2)}{2\sqrt{(\beta^2 - 4a_2\gamma)^3}}, \tag{68}$$

then $a_2 < 0$, $d^2 < a_0$ will ensure that I_{kk} is always positive which will make I_k non-decreasing. Now, the only part in $g(k)$ which could be negative is $\frac{3}{4} - \frac{1}{4I}I_k^2$. From the following result

$$\lim_{k \to \pm\infty} I_k = -\frac{1}{2a_2}\left(c \pm \sqrt{c^2 - a_2}\right), \tag{69}$$

we could state that

$$-\frac{1}{2a_2}\left(c - \sqrt{c^2 - a_2}\right) \leq I_k \leq -\frac{1}{2a_2}\left(c + \sqrt{c^2 - a_2}\right), \tag{70}$$

which leads to

$$I_k^2 \le \frac{1}{2a_2^2}(c^2 + c^2 - a_2) = \frac{1}{2a_2^2}(2c^2 - a_2). \tag{71}$$

On the other hand, according to previous proof, our IRV5 model could be rewritten in the following form

$$I = \tilde{a} + \tilde{b}\left\{\tilde{\rho}(k - \tilde{m}) + \sqrt{(k - \tilde{m})^2 + \tilde{\sigma}^2}\right\} \tag{72}$$

with $\tilde{a} \in \mathbb{R}$, $\tilde{b} \ge 0, |\tilde{\rho}| < 1, \tilde{m} \in \mathbb{R}$, $\tilde{\sigma} > 0$. And with

$$a_1 \le -\frac{1}{3a_2^2}(2c^2 - a_2)(c^2 - a_2) - 2cd - \frac{1}{4}, \tag{73}$$

it is straightforward that

$$\begin{aligned} I \ge \tilde{a} &= \frac{\frac{\alpha}{2} + cd}{a_2 - c^2} = -\frac{\frac{1}{2}\left(a_1 + \frac{1}{4}\right) + cd}{c^2 - a_2} \\ &\ge \frac{1}{6a_2^2}(2c^2 - a_2) \ge \frac{1}{3} I_k^2 \end{aligned} \tag{74}$$

which is equivalent to $\frac{3}{4} - \frac{1}{4I} I_k^2 \ge 0$, so that

$$g(k) = \left(\frac{k}{2I} I_k - \frac{1}{2}\right)^2 + \frac{1}{2} I_{kk} + \frac{3}{4} - \frac{1}{4I} I_k^2 \ge 0 \tag{75}$$

for every *k*. The proof is complete.

ENDNOTES

[1]In terms of IRVV models, we can't kill *b* as what we did in IRV models. Thus, IRVV4 and IRVV5 are consistent with IRV3 and IRV5 regarding to parametrization. However, we can't add ω^2 or even higher order terms into $a(\omega)$ in order to the simple quadratic form.

[2]Here, twice continuous differentiable means having continuous second-order derivatives.

[3]Note that now for convenience, we just simply regard *I* as function of *k* and τ instead of *t* and *K*. Actually, the functional forms of *I* w.r.t (k, τ) and (t, K) are different.

[4]Denote the price of a put option with strike *K* maturity *T* as $P(K, T)$, then we have $P(K, T) = E^T(K - F_T)^+$, which implies $\frac{\partial P}{\partial K} = E^T 1_{\{K \ge F_T\}}$, which implies $\frac{\partial^2 C}{\partial K^2} = \frac{\partial^2 P}{\partial K^2} = E^T \partial(K - F_T) = q(K)$, where $q(K)$ is the transition density which is defined by $P^T(F_T \in dK) = q(K)dK$. So that $\frac{\partial^2 C}{\partial K^2} \ge 0$ is equivalent to $q(K) \ge 0$.

REFERENCES

Bachelier, L. *Théorie de la Spéculation*. Gauthier-Villars, 1900.

Black, F., and M. Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, (1973), pp. 637-654.

Carr, P., and D.B. Madan. "A Note on Sufficient Conditions for No Arbitrage." *Finance Research Letters*, Vol. 2, No. 3 (2005), pp. 125-130.

Carr, P., and J. Sun. "A New Approach for Option Pricing Under Stochastic Volatility." *Review of Derivatives Research*, Vol. 10, No. 2 (2007), pp. 87-150.

——. "Implied Remaining Variance in Derivative Pricing." *The Journal of Fixed Income*, Vol. 23, No. 4 (2013), pp. 19-32.

Carr, P., and L. Wu. "A New Simple Approach for Constructing Implied Volatility Surfaces." Working Paper, New York University and Baruch College, 2010.

——. "Analyzing volatility risk and risk premium in option contracts: A new theory." *Journal of Financial Economics*, Vol. 120, No. 1 (2016), pp. 1-20.

Carmona, R., and S. Nadtochiy. "Local Volatility Dynamic Models." *Finance and Stochastics*, Vol. 13, No. 1 (2009), pp. 1-48.

Durrleman, V. "Implied Volatility: Market Models." *Encyclopedia of Quantitative Finance*, Wiley Online Library, 2010.

Gorovoi, V., and V. Linetsky. "Black's Model of Interest Rates as Options, Eigenfunction Expansions and Japanese Interest Rates." *Mathematical Finance*, Vol. 14, No. 1 (2004), pp. 49-78.

Heston, S.L. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options." *Review of Financial Studies*, Vol. 6, No. 2 (1993), pp. 327-343.

Huitema, R., and B. Peeters. "Risk Premiums in a Multi-Factor Jump-Diffusion Model for the Joint Dynamics of Equity Options and Their Underlying." Working Paper, University of Zurich and ING Investment Management, 2014.

Ichiue, H., and Y. Ueno. "Equilibrium Interest Rate and the Yield Curve in a Low Interest Rate Environment." No. 07-E-18. Bank of Japan, 2007.

Lee, R.W. "The Moment Formula for Implied Volatility at Extreme Strikes." *Mathematical Finance*, Vol. 14, No. 3 (2004), pp. 469-480.

Merton, R.C. "Theory of Rational Option Pricing." *The Bell Journal of Economics and Management Science* (1973), pp. 141-183.

Peeters, B. "Risk Premiums in a Simple Market Model for Implied Volatility." *Quantitative Finance*, Vol. 13, No. 5 (2013), pp. 739-748.

Roper, M. "Arbitrage Free Implied Volatility Surfaces." Working Paper, The University of Sydney, 2010.

Schachermayer, W., and J. Teichmann. "How Close are the Option Pricing Formulas of Bachelier and Black–Merton–Scholes?" *Mathematical Finance*, Vol. 18, No. 1 (2008), pp. 155-170.

Sullivan, E.J., and T.M. Weithers. "Louis Bachelier: The Father of Modern Option Pricing Theory." *Journal of Economic Education*, (1991), pp. 165-171.

Schweizer, M., and J. Wissel. "Term Structures of Implied Volatilities: Absence of Arbitrage and Existence Results." *Mathematical Finance*, Vol. 18, No. 1 (2008), pp. 77-114.