Option Profit and Loss Attribution and Pricing: A New Framework*

Peter Carr\textsuperscript{a}†, Liuren Wu\textsuperscript{b}‡

\textsuperscript{a}New York University Tandon School of Engineering, 12 MetroTech Center, New York, NY 11201, USA
\textsuperscript{b}Baruch College, Zicklin School of Business, One Bernard Baruch Way, New York, NY 10010, USA

Abstract

Marking to market dictates that investors worry not only about terminal payoffs, but also about daily price fluctuations. This paper develops a new valuation framework that links the pricing of an option investment to its daily profit and loss attribution. The new framework uses the explicit Black-Merton-Scholes option pricing formula to attribute the short term investment risk of the option to variations in the underlying security price and in the option’s implied volatility. This attribution highlights the key risk drivers affecting the short term investment return distribution and provides a basis for determining what risks to dynamically hedge and what risks to take exposure on. Taking risk-neutral expectation and applying dynamic no-arbitrage constraints results in a pricing relation that links the option’s fair implied volatility level to the underlying’s short term volatility level as well as corrections for the implied volatility’s own expected direction of movement, its variance, and its covariance with the underlying security return. Commonality assumptions on the implied volatility co-movements across strike price or maturity date allow one to generate cross-sectional pricing implications either for a selected number of option contracts or across the whole implied volatility surface.

\textit{JEL Classification:} C13; C51; G12; G13

\textit{Keywords:} Profit and loss attribution; Delta; Vega; Vanna; Volga; Implied volatility term structure; Implied volatility smile

*We thank Kris Jacobs, and seminar participants at Baruch College, Credit Suisse, the 2013 Derivatives Conference at New York University, and the 2017 6th IFSID Conference on Derivatives for comments. We welcome comments, including references we have inadvertently missed.

†E-mail address: pc73@nyu.edu
‡Tel: +1 646 312 3509; fax: +1 646 312 3451. E-mail address: liuren.wu@baruch.cuny.edu.
# Contents

1 Introduction ........................................... 1

2 P&L attribution on option investments ................. 7

3 Risk-neutral expectation and implied volatility valuation ........................................... 9
   3.1 P&L attribution based on deviations from pricing expectations .......................... 11
   3.2 Moment condition based implied volatility valuation of an option contract ............. 12
   3.3 The at-the-money implied variance term structure ........................................... 15
      3.3.1 Extract expected rate of implied volatility change from the at-the-money term structure 15
      3.3.2 Construct the at-the-money implied volatility term structure function ............. 17
   3.4 The implied volatility smile at a single maturity ........................................... 19
      3.4.1 A local quadratic smile matching locally parallel dynamics ......................... 20
      3.4.2 A 4th-order polynomial smile accommodating separate volatility and skew movements 22

4 Empirical analysis on S&P 500 index options ................. 24
   4.1 Construct floating series of fixed-contract implied volatility percentage changes ........... 25
   4.2 The at-the-money implied variance term structure ........................................... 27
   4.3 Extracting expected future movements from the at-the-money term structure ............. 28
   4.4 P&L attribution for at-the-money options ........................................... 31
   4.5 The implied volatility smile and its variations ........................................... 32
4.6 Extracting expected variance and covariance rates from the smile 34

4.7 Statistical arbitrage on the smile 35

4.8 P&L variance attribution across moneyness 37

5 Concluding remarks 38
1. Introduction

Different modeling frameworks serve different purposes. A major focus of the present option pricing literature is to derive option values that are internally consistent across all strikes and maturities. This is achieved by pricing all options based on a single yardstick linking risk to expected return. The standard practice starts with the specification of the full dynamics of the underlying security price, including the full dynamics of its instantaneous variance rate, and then performs valuation of all options on the same underlying by taking risk-neutral expectations of the option’s terminal payoff. The full underlying dynamical specification creates the single reference distribution of the relevant terminal random variable, which is then used for taking an expectation. Under this approach, even if the assumed dynamics are wrong, the valuations on the option contracts remain consistent with one another relative to this wrong reference.

It is good to have consistency, but it is not good to be wrong. Unfortunately, the assumed dynamics of the underlying security price and its instantaneous volatility often deviate strongly from reality. This is understandable, and sometimes unavoidable, because it is inherently difficult to accurately specify the full dynamics of everything. For example, to price long-dated options, this approach needs to make projections on the underlying security price and its instantaneous volatility far into the future. The accuracies of long-dated projections are understandably low. Furthermore, seemingly innocuous stationarity assumptions on the instantaneous volatility dynamics often generate much lower price variation in long-term contracts than actually observed in the data (Giglio and Kelly (2017)).

In practice, as long as one does not hold the contracts to maturity, one does not necessarily need to make long-run predictions to trade long-dated contracts. One can hold a very long-dated contract for a very short period of time. Over this very short holding period, the investor concern is much less about the terminal payoff far into the distant future, but much more about the factors that drive the profit and loss over the holding period, say, one day. To accurately assess the next day’s investment risk, one only needs to accurately estimate the current instantaneous variances and covariances of the driving factors. When the holding period is intended to be short, one need not be directly concerned with the factors’ long-run behavior.
and how these instantaneous variances and covariances vary after the holding period. In fact, the standard recommended practice of marking to market for financial securities makes it vitally imperative for investors to understand the magnitude and sources of the daily value fluctuation, regardless of the investor’s intended holding period. The process of attributing the profit and loss (P&L) of an investment at a given date to different risk exposures is commonly referred to as the P&L attribution process.

This paper develops a new valuation framework that links the pricing of a security at a given time to its P&L attribution, without directly referring to the terminal payoffs of the investment. The P&L attribution analysis highlights the risk sources that affect the investment return distribution over a short horizon (say the next day) and provides a basis for determining which risks to dynamically hedge (e.g., risks that one does not have a confident projection of its future direction) and which risks to take exposure on (e.g., the perceived alpha source of the investment). Taking subjective expectation on the return attribution shows how much one expects to earn from each source of risk exposure. Taking expectation on the short-term investment return under the risk-neutral measure and setting the expected excess return to zero leads to a dynamic no-arbitrage pricing relation that links the fair value of the investment to its risk exposures and risk magnitudes. The approach generates a valuation free of dynamic arbitrage without explicit reference to the terminal option payoff and accordingly without specifying the full underlying security dynamics to maturity. Instead, the valuation depends only on the estimates of the current risk exposures of the investment and the current magnitudes of the risks. The variation of risk exposures and risk magnitudes over time cause variations in value over time, but do not affect the level of the current valuation.

The P&L attribution necessitates the specification of a risk structure, with which one can compute the investment’s risk exposures and risk magnitudes. In this paper, we take the investment in a vanilla European-style option on an underlying stock (index) as an example and perform the P&L attribution on the option contract via the explicit Black-Merton-Scholes (BMS) option pricing formula. Black and Scholes (1973) and Merton (1973) derive their option pricing formula by assuming constant volatility geometric Brownian motion dynamics for the underlying security price. Their assumptions do not match reality as security return
volatilities tend to vary strongly over time; nevertheless, their pricing equation can be used, and has indeed
been used widely by practitioners, as a simple and intuitive representation of the option value in terms of
its major risk sources, i.e., variations in the underlying security price and its return volatility. In addition to
the underlying security price and contract terms (such as strike and maturity), the pricing equation takes a
volatility input that can be used to match the observed market price for the option. The volatility input that
matches the observed market price is commonly referred to as the BMS implied volatility.

A Taylor series expansion of the BMS option pricing formula attributes the option investment P&L to
partial derivatives in time, in the underlying stock price and in the option’s implied volatility. For a short
investment horizon, only the first order derivative in time (theta) is needed for the P&L attribution. When
the underlying stock price and the option’s implied volatility move continuously over time, expanding to the
second order in these variables is sufficient to bring the residual error to an order lower than the length of
the short investment horizon (heuristically $dt$).

In this expansion, the main risk source is variation in the price of the underlying security. Dynamic
delta hedging can hedge away this source of risk exposure and reduce the option return variance by over
90% (Figlewski (1989), Carr and Wu (2016)). The second major source of risk in holding an option for a
short time period comes from the variation in its implied volatility over this period. One can either perform
vega hedging to remove this part of the variation, or adjust the delta hedge to remove the part of the implied
volatility variation that is correlated with the security price movement.

Taking risk-neutral expectation on the P&L attribution via the BMS pricing formula and demanding no
dynamic arbitrage results in a simple pricing equation that relates the option’s implied volatility level to the
underlying instantaneous volatility as well as corrections due to the implied volatility’s expected direction,
its variance, and its covariance with the security return. The simplicity comes from the fact that for the
BMS pricing formula, all risk exposures including theta, cash vega, cash vanna, and cash volga, are all
proportional to its cash gamma. Factoring out the strictly positive cash gamma in the mean option P&L
expression and equating risk-neutral mean P&L’s leads to a simple algebraic equation linking the option’s
implied volatility level to the conditional risk-neutral first and second moments of the change in this level along with the underlying security return.

In contrast to the traditional option pricing approach which links the values of all option contracts to a single reference dynamical specification, the new approach links the current fair value of only one option contract’s implied volatility to current conditional moments of the near term change in this contract’s implied volatility and the security return. This subtle but vital shift in perspective is due to the use of the option’s implied volatility as a state variable, rather than the use of the underlying’s security’s instantaneous volatility as a state variable. Our new approach allows the moment conditions determining the single option’s fair implied volatility level to change over time. This time variation leads to corresponding changes in the fair value of the implied volatility level, but the valuation itself does not depend on how exactly these moment conditions vary over time or the full specification of their dynamics. The key strength of the traditional approach is to maintain cross-sectional consistency through the usage of one reference dynamics. However when the reference dynamics are mis-specified as they are most likely to be when forced to make projections far into the future and for rare scenarios, the valuations of all option contracts will be wrong consistently⁴. By contrast, our new approach values a single contract based on its current moment condition projections, which can be much easier to specify and more accurate by nature.

The new pricing relation guarantees the absence of dynamic arbitrage by demanding that the risk-neutral expected instantaneous return on an option matches that of a money market account. By pricing an option contract based only on its own risk-neutral moment conditions, the theory need not impose cross-sectional consistency across all option contracts. Thus, the approach is decidedly local not only in terms of the investment horizon it considers, but also in terms of the contracts it values. Under this approach, to compare the valuation of two or more distinct option contracts, one must first compare the risk projections on these contracts. In particular, one can impose common factor structures on the moment conditions of nearby contracts to link the valuations of these contracts together.

¹or as Ralph Waldo Emerson put it, “a foolish consistency is the hobgoblin of little minds”
Principal component analysis on the variation of an implied volatility surface often identifies three major sources of variation. In addition to the first common level component, the next two principal components capture the changes in the implied volatility term structure slope and changes in the implied volatility skew across strike, respectively. We perform separate analysis along these two dimensions. First, we define the at-the-money option at a given maturity as the particular option whose log ratio of strike to current forward is equal to half of the option’s total implied variance, which is the BMS risk-neutral mean of the log of the gross return to maturity. The fair valuation of this at-the-money option does not depend on the variance and covariance of the implied volatility change, but only depends on its risk-neutral drift and the instantaneous variance rate level. By imposing a common risk-neutral drift on two nearby at-the-money option contracts, we can extract the common risk-neutral drift from the term structure slope defined by these two nearby contracts. Further assuming a common one-factor mean-reverting structure on the at-the-money variance, we can define an at-the-money implied variance term structure function as an exponentially weighted average of the short- and the long- implied variance levels, analogous to implications from traditional stochastic volatility models (e.g., Heston (1993)).

At a fixed time to maturity, we propose to take the at-the-money implied variance as the reference point and perform a vega hedge of other contracts using the at-the-money option. We can then represent the implied volatility level of an option contract at this maturity as a function of the at-the-money implied volatility and the implied volatility’s variance and covariance with the security return. Assuming proportional movements have common variance and covariance within a certain strike range, we can build an implied volatility smile function relative to the at-the-money implied volatility level in terms of the two common second moments. The smile function can be approximately written as a simple quadratic function in relative strike, with the slope coefficient measuring the covariance with the security return and the curvature coefficient capturing the variance of the percentage implied volatility change of the contract. Thus, the parallel dynamics assumption provides a theoretical basis for the commonly used quadratic approximations of the implied volatility smile, e.g., Jarrow and Rudd (1982), Shimko (1993), and Backus, Foresi, and Wu (1997).
If we further decompose the percentage moves on the implied volatility smile at one maturity into the movement of the at-the-money implied volatility and the movement of the implied volatility skew, we obtain a more general cross-sectional pricing functional form that characterizes the smile iteratively as a 4th-order polynomial, with polynomial coefficients measuring the variances and covariances of the at-the-money implied volatility, the implied volatility skew, and the underlying security return. Allowing more flexible variations, such as different dynamics for out-of-the-money put and call implied volatility skews, leads to more flexible, piece-wise polynomial functional forms. These flexible functional forms provide theoretical guidance to the practice of polynomial smoothing of the implied volatility smile, e.g., Figlewski (2009).

In related literature, Israelov and Kelly (2017) recognize the limitations of standard option pricing models and propose to directly predict the distribution of the option investment return empirically. Our research shares the same shift of focus from terminal payoffs to the behavior of short-term investment returns and provides a theoretical foundation for how to analyze option investment returns and how to link the predicted investment return behavior to its pricing. Several studies strive to link option returns to various firm characteristics: An, Ang, Bali, and Cakici (2014) link future option implied volatility variation to past stock return performance; Boyer and Vorkink (2014) link ex post option returns to ex ante implied volatility skew; Byun and Kim (2016) link option returns to the underlying stock’s lottery-like characteristics; Hu and Jacobs (2017) identify linkages between option returns and the underlying stock’s volatility level. Our theory provides guidance on how to analyze the return on a derivative security investment and how to decompose this return along different risk drivers.

The remainder of this paper is organized as follows. Section 2 sets up our notation and establishes the P&L attribution for an option contract based on the BMS pricing formula. Section 3 takes risk-neutral expectations on the return attribution to generate the fair implied volatility level, both for an individual option contract and for the relative pricing of many option contracts based on a common factor dynamical specification. Section 4 performs empirical analysis on the S&P 500 index options. Section 5 provides concluding remarks and directions for future research.
2. P&L attribution on option investments

We consider a market with a riskfree bond, a risky asset, and in this section one vanilla European option written on the risky asset. For simplicity, we assume zero interest rates and zero carrying costs/benefits for the risky asset. In practical implementation, one can readily accommodate a deterministic term structure of financing rates by modeling the forward value of the underlying security and defining moneyness of the option against the forward. The risky asset can be any type of tradable security, but we will refer to it as the stock for concreteness. In the US, exchange-traded options on individual stocks are all American style. To apply our new theory to American options, a commonly used shortcut is to extract the BMS implied volatility from the price of an American option based on some tree/lattice method and use the implied volatility to compute a European option value for the same maturity date and strike. See Carr and Wu (2010) for a detailed discussion on data pre-processing of individual stock options.

We assume frictionless and continuous trading in the riskfree bond, the stock, and the option contract written on the stock. We assume no-arbitrage between the stock and the bond. As a result, there exists a risk-neutral probability measure $Q$, equivalent to the statistical probability measure $P$, such that the stock price $S$ is a martingale. We further assume that the option value we seek does not allow arbitrage against any portfolio of the stock and the riskless bond.

We start by considering a long position in a European call option. Holding the call to expiry generates a P&L dictated by the terminal payoff of the call. Classic option valuation often starts with the terminal payoff function and takes expectation of the terminal payoff based on assumptions governing the dynamics of the underlying asset price to maturity. Our new pricing approach focuses on the investment P&L over a short period of time, say, one day. The practice of marking to market on financial securities makes the daily P&L fluctuation vitally important regardless of the ultimate holding period. The short-term P&L fluctuation of the option investment is mainly driven by the option exposures to various risk sources and the variation of these risk sources. Accordingly, our analysis relies more on defining risk exposures and quantifying risk magnitudes than on describing terminal payoffs.
We propose to attribute the short-term investment P&L on the option contract by making use of the explicit Black-Merton-Scholes (BMS) pricing formula. In deriving their pricing formula, Black and Scholes (1973) and Merton (1973) assume that the underlying security price follows a geometric Brownian motion with constant volatility. Their assumptions have long been found to deviate from reality as financial security return volatilities tend to vary strongly over time. Nevertheless, the BMS pricing formula provides a simple and intuitive way of decomposing option returns and has been widely adopted in the industry as the starting point for option P&L attribution.

Let $B(t, S_t, I_t; K, T)$ denote the BMS pricing formula for a European call option at strike $K$ and expiry $T$,

$$B(t, S_t, I_t; K, T) \equiv S_t N\left(-\frac{k - \frac{1}{2} I_t^2 \tau}{I_t \sqrt{\tau}}\right) - KN\left(-\frac{k + \frac{1}{2} I_t^2 \tau}{I_t \sqrt{\tau}}\right),$$

where $N(\cdot)$ denotes the cumulative normal function, $\tau \equiv T - t$ denotes the time to maturity, and $k \equiv \ln(K/S_t)$ denotes the relative strike. The terms $z_{\pm} \equiv (k \pm \frac{1}{2} I_t^2 \tau)$ represent convexity-adjusted moneyness of the call under the risk-neutral measure and under the share measure, respectively, in the BMS model environment. Via the BMS pricing formula, the option value at time $t$ is represented as a function of the stock price $S_t$ and the option implied volatility $I_t$. As long as the option price does not allow arbitrage against the underlying risky stock and the riskless bond, one can always find a positive implied volatility input to the BMS pricing formula to match that price (Hodges (1996)). The BMS pricing formula builds a monotonic linkage between the option price and the option’s implied volatility, and captures all random shocks to the option, other than shocks to the underlying security price level, through the implied volatility.

Through the pricing equation, we can decompose the option variation over the next instant $(t+dt)$, or the instantaneous P&L of the option investment, through the variations in the stock price and the implied volatility,

$$dB(t, S_t, I_t; K, T) = \left[ B_t dt + B_S dS_t + B_I dI_t \right] + \left[ \frac{1}{2} B_{SS} (dS_t)^2 + \frac{1}{2} B_{III} (dI_t)^2 + B_{IS} (dS_t dI_t) \right] + J_t,$$
where $B_t, B_S, B_I, B_{SS}, B_{II}, B_{IS}$ denote the BMS sensitivity theta, delta, vega, gamma, volga, and vanna, respectively. The first bracket collects first-order effects, and the second bracket collects second-order effects. The last term $J_t$ captures the contribution of potential random jumps in the stock price and option implied volatility. When both the stock price and the option implied volatility show purely continuous movements, the first and second order effects capture all of the relevant movements for the option price over a short time interval. We henceforth assume continuous dynamics and link option pricing to its first and second-order exposures.

Equation (2) decomposes the short-term P&L of an option investment into different risk exposures. In particular, through the BMS model, each contract is exposed to variations in calendar time $t$, the underlying security price $S_t$, and the option’s implied volatility $I_t$. Since calendar time moves only deterministically, the risk arises only from random variation and covariation of $S_t$ and $I_t$. Standard option pricing models may specify multi-factor stochastic volatility dynamics, in which case the underlying security price dynamics and the stochastic volatility factors govern the variation of all options on the security. By contrast, equation (2) focuses on one option investment and represents its risk in terms of the variation and covariation of the underlying security price and its own BMS implied volatility.

3. Risk-neutral expectation and implied volatility valuation

To obtain the implications of the assumptions in the last section for the option price, we take expectation on the option P&L attribution in equation (2) under the risk-neutral measure $\mathbb{Q}$ and divide the expected P&L by the instantaneous investment horizon $dt$. We can link the time-$t$ annualized risk-neutral expected return on the option investment to the time-$t$ risk-neutral expectations of the variation and covariation of the underlying security price and the option’s BMS implied volatility.\(^3\)

\(^2\)We suppress the arguments, $(t, S_t, I_t; K, T)$, of the pricing function and its derivatives when no confusion shall occur.

\(^3\)Returns on derivative contracts are not as well-defined as on primary securities because the initiation cost of derivative contracts does not always reflect the risk level of the investment. Investors often define returns on derivative contracts not relative to its own value (which can be zero or even negative), but relative to the committed or required risk capital. We treat the profit or loss per unit time as the return on unit capital.
\[ \mathbb{E}_t \left[ \frac{dB}{dt} \right] = B_t + B_1 t \mu_t + \frac{1}{2} B_{SS} S_t^2 \sigma_t^2 + \frac{1}{2} B_{II} I_t^2 \omega_t^2 + B_{IS} I_t S_t \gamma_t, \]  

(3)

where \( \mathbb{E}_t [\cdot] \) denotes the expectation operation under the risk-neutral measure conditional on time-\( t \) filtration, \( \mu_t \) denotes the annualized risk-neutral expected rate of percentage change in the BMS implied volatility of the option contract,

\[ \mu_t \equiv \mathbb{E}_t \left[ \frac{dI_t}{I_t} \right] / dt, \]  

(4)

and \( \sigma_t^2, \omega_t^2, \gamma_t \) denote the time-\( t \) conditional variance and covariance rate of the stock return and the percentage implied volatility change:

\[ \sigma_t^2 \equiv \mathbb{E}_t \left[ \left( \frac{dS_t}{S_t} \right)^2 / dt \right], \quad \omega_t^2 \equiv \mathbb{E}_t \left[ \left( \frac{dI_t}{I_t} \right)^2 / dt \right], \quad \gamma_t \equiv \mathbb{E}_t \left[ \left( \frac{dS_t}{S_t}, \frac{dI_t}{I_t} \right) / dt \right]. \]  

(5)

Under the zero financing assumption, the risk-neutral expected stock return is zero and hence drops out of equation (3).

The zero financing assumption and no dynamic arbitrage also dictates that the risk-neutral expected return on the option investment is also zero:

\[ 0 = B_t + B_1 t \mu_t + \frac{1}{2} B_{SS} S_t^2 \sigma_t^2 + \frac{1}{2} B_{II} I_t^2 \omega_t^2 + B_{IS} I_t S_t \gamma_t. \]  

(6)

We can regard equation (6) as a pricing relation in the sense that as long as the option price satisfies no dynamic arbitrage and generates a risk-neutral expected return of zero, the option price must satisfy the constraints imposed by this equation. This pricing equation is not based on the full specification of the underlying stock price dynamics, but rather on the first and second conditional moments of the stock price and option implied volatility movements at time \( t \).
3.1. P&L attribution based on deviations from pricing expectations

Plugging the pricing equation in (6) into the original P&L attribution equation in (2) and ignoring the contribution of jumps, we can represent the P&L attribution purely in terms of deviations from pricing expectations,

\[
dB(t, S_t, I_t; K, T) = B_S S_t \frac{dS_t}{S_t} + B_I I_t \left( \frac{dI_t}{I_t} - \mu dt \right) + \frac{1}{2} B_{SS} S_t^2 \left( \left( \frac{dS_t}{S_t} \right)^2 - \sigma_t^2 dt \right) + \frac{1}{2} B_{II} I_t^2 \left( \left( \frac{dI_t}{I_t} \right)^2 - \omega_t^2 dt \right) + B_{IS} S_t I_t \left( \frac{dS_t}{S_t} \frac{dI_t}{I_t} - \gamma_t dt \right),
\]

For a naked option position, its main risk comes from the underlying security price movement \((dS_t)\).

For a delta hedged option, the remaining major variation comes from the implied volatility movement. One can further remove or limit this risk by vega hedging using, for example, a co-terminal variance swap or a co-terminal at-the-money option. Implied volatilities on same-maturity options tend to move together. Hence, vega hedging with co-terminal options can remove a large part of the implied volatility movement. At each maturity, at-the-money options tend to be the most actively traded and hence an ideal candidate as a hedging instrument. With delta and vega both hedged, the remaining variation in P&L comes from deviations of second-order variations from their respective expectations. When the security price and/or the implied volatility can jump by a randomly large amount, these jumps can induce significant P&L variations due to higher-order terms.

Several recent studies analyze the behavior of an option’s expected excess return. Taking expectations of equation (7) under the statistical measure \(P\) provides an attribution of the expected excess return on the option investment along the different risk sources. Due to a call’s positive delta exposure, the equity risk premium, \(\frac{1}{dt} \mathbb{E}_P[dS_t/S_t] > 0\) translates into a positive expected excess return on the option investment. Hence, returns on call options contain a positive equity risk premium component, while returns on puts have a negative equity risk premium component. There is a large literature devoted to the analysis of the time-series and cross-sectional variation in equity risk premiums. This literature is directly applicable to the delta exposure component of an option investment.
Professional option investors often delta hedge, thereby focusing their investment on the option’s volatility exposure. Carr and Wu (2004) document a strongly negative variance risk premium on several equity indices. This negative variance risk premium implies that shorting index options and delta hedging them results in positive excess returns on average. Under continuous dynamics, the variance risk premium shows up as the drift difference in the implied volatility dynamics, $\frac{1}{dt} \mathbb{E}_t^P [dI_t/I_t] - \mu$, which contributes to the expected option return via the option’s vega exposure. Cross-sectional analysis on the volatility risk premium behavior becomes the starting point for forming a volatility portfolio.

Under continuous dynamics of $S$ and $I$, the second-order moments in (5) remain the same under a measure change from $Q$ to $P$ and hence do not contribute to the $P$ expected return on an option investment. Nevertheless, they can contribute to the expected option return in the presence of random jumps in the security price and/or implied volatility movements.

### 3.2. Moment condition based implied volatility valuation of an option contract

A particularly nice feature of the BMS pricing equation is that the BMS theta, cash vega, cash vanna, and cash volga can all be represented in terms of the BMS cash gamma,

$$
B_t = -\frac{1}{2} I_t^2 B_{SS} S_t^2, \quad B_t I_t = I_t^2 \tau B_{SS} S_t^2, \quad B_{1SI} S_t = (k + \frac{1}{2} I_t^2 \tau) B_{SS} S_t^2, \quad B_{1II} I_t^2 = (k + \frac{1}{2} I_t^2 \tau) (k - \frac{1}{2} I_t^2 \tau) B_{SS} S_t^2.
$$

(8)

Since the option contract has strictly positive cash gamma, we can define the option investment return as the investment P&L per unit of cash gamma so that we can factor out the cash gamma component from the P&L attribution in (7),

$$
\frac{dB}{B_{SS} S_t} = l(\zeta) \frac{dS}{S} + I_t^2 \tau \left( \frac{dh}{h} - \mu dt \right) + \frac{1}{2} \left( \frac{dS}{S} \right)^2 - \sigma_t^2 dt + \frac{1}{2} \zeta^2 + \zeta \left( \frac{dh}{h} \right)^2 - \omega_t^2 dt + \zeta \left( \frac{dS}{S} \right) \left( \frac{dh}{h} \right) - \eta dt,
$$

(9)

---

4Hedge funds that mainly invest in options are often referred to as volatility strategies.
where \( I(\cdot) \equiv \frac{B_S}{\frac{d}{dS_t}} = N\left(-\frac{\cdot}{\sqrt{\tau}}\right) \frac{\sqrt{\tau}}{n}(-\frac{\cdot}{\sqrt{\tau}}) \) denotes the ratio of cash delta to cash gamma. With cash gamma scaling, the option investment return exposures can be written explicitly as a function of the implied variance level and the two moneyness \((z_+, z_-)\) measures.

Furthermore, substituting these partial derivatives into the pricing equation (6) and dividing both sides by the positive cash gamma, we have

\[
I_t^2 = \left[ 2\tau \mu_t I_t^2 + \sigma_t^2 \right] + \left[ 2\gamma_t z_+ + \omega_t^2 z_+ z_- \right].
\]

Equation (10) can be used to determine the time-\(t\) fair value of the implied volatility for an option contract, based on time-\(t\) projection of its risk-adjusted expected rate of return \(\mu_t\), its variance rate \(\omega_t^2\), the underlying security’s variance rate \((\sigma_t^2)\), and the covariance rate between the two \((\gamma_t)\).

**Theorem 1** Assuming continuous price and implied volatility movements and performing instantaneous P&L attribution on a European option investment based on the BMS pricing equation, the absence of dynamic arbitrage produces a link between the time-\(t\) fair value of the option’s BMS implied volatility level and the time-\(t\) risk-neutral conditional mean and variance of the implied volatility percentage change rate \((\mu_t\) and \(\omega_t^2\)), the conditional variance rate of the underlying security return \((\sigma_t^2)\), and the conditional covariance between the two \((\gamma_t)\).

When compared to traditional option pricing practice, equation (10) represents an extremely simple formulation of the option contract’s fair value. Traditional option pricing modeling emphasizes cross-sectional consistency as its chief objective. To achieve this objective, one specifies the full risk-neutral dynamics on the underlying security, and prices all options by taking expectations of their terminal payoff functions based on the same dynamics assumption. Thus, the assumed dynamics on the underlying security serves as a single reference, or yardstick, for all option contracts. The same yardstick guarantees that the valuations of all options are consistent with this yardstick and hence with one another. Even if the specification of the dynamics (the yardstick) is wrong, the option valuations are still internally consistent. They just become
By contrast, by starting with the P&L attribution of one option contract, the pricing equation (10) only guarantees that the time-\( t \) valuation of this contract is consistent with the first and second risk-neutral conditional moments on the variations of the underlying security price and the option’s implied volatility. It guarantees no dynamic arbitrage between this option contract and the underlying security and cash under the assumed moment conditions, but nothing more. Thus, instead of providing a yardstick for analyzing cross-sectional consistency among different option contracts, the approach provides a more direct linkage between the fair implied volatility valuation of an option contract and its conditional moment conditions of the next movements. The assumption of purely continuous dynamics allows us to build the linkage with only the first and second conditional moments. Thus, the new pricing relation allows us to assess the value of an option contract directly based on our view and pricing on its implied volatility variation.

The pricing relation in (10) applies to one particular option contract, with no reference to any other option contracts. The relation allows one to compare the valuation of an option contract to one’s projection of the first and second moment conditions of the underlying security price and the option’s implied volatility. To make relative value comparisons across different option contracts under this framework, one must first make commonality assumptions between the implied volatilities of these option contracts. For example, Carr and Wu (2016) derive the implications of no dynamic arbitrage for the whole implied volatility surface by assuming that the instantaneous variation of the whole surface is driven by a single Brownian factor.

Principal component analysis of option implied volatility surfaces often identifies separate variations in the maturity and the strike (moneyness) dimensions. Intuitively, the variation along the term structure reflects the expected direction and market pricing of the implied volatility movements, whereas the implied volatility smile across moneyness reflects the risk-neutral return distribution deviation from normality, induced by implied volatility variation and covariance with the security return. The two brackets in equation (10) provide an intuitive separation of these two effects, with the first bracket capturing the term structure effect and the second bracket capturing the moneyness effect. In what follows, we formalize this intuition and
separate the term structure behavior from the moneyness behavior at a certain maturity. We link each to a
separate set of moment conditions.

3.3. The at-the-money implied variance term structure

To separate out the term structure effect from the moneyness effect, we define the at-the-money option as
the option with \( z \) \(+\) \( = k + \frac{1}{2} I^2 \tau = 0 \), which corresponds to the strike price that equates the relative strike \( k \) to
the risk-neutral expected value of \( \ln(S_T/S_t) \) under the BMS model environment.

At \( z = 0 \), the P&L attribution in equation (9) shows that the option has both zero volga and zero vanna.
Investing in such an option only exposes the investor to delta, vega, and gamma risk. The pricing equation
in equation (10) also simplifies as the terms in the second bracket are reduced to zero. If we use \( A_t \) to denote
the at-the-money implied volatility, then we can write the pricing equation as,

\[
A^2_t = 2\mu A^2_t + \sigma^2_t.
\]  

(11)

The at-the-money option with \( z = 0 \) is the only strike point where the option has both zero volga and
zero vanna, and the corresponding at-the-money implied volatility is also the only strike point where the
volatility level only depends on the risk-neutral expected direction (drift) of the implied volatility, but not
on its variance and covariance with the security return. The separation presents a clean channel for us to
analyze expected volatility changes and the term structure without interference from higher-order effects.

3.3.1. Extract expected rate of implied volatility change from the at-the-money term structure

From (11), we can infer the risk-neutral expected rate of percentage changes in the at-the-money implied
volatility directly from the percentage difference between the at-the-money implied variance \( A^2_t \) and the
return variance \( \sigma^2_t \).
Proposition 1 The risk-neutral expected rate of percentage change $\mu_t$ of the implied volatility of the at-the-money option can be directly inferred from the distance between the at-the-money implied variance level $A_t^2$ and the return variance rate $\sigma_t^2$,

$$
\mu_t = \frac{A_t^2 - \sigma_t^2}{2A_t^2 \tau}.
$$

Equation (12) comes from a simple rearrangement of equation (11).

It is well known that security return volatility exhibits mean-reverting (clustering) behavior, which forms the basis for GARCH-type statistical modeling of volatility forecasts (Engle (2004)). Floating constant-maturity at-the-money implied volatility series also exhibit similar mean-reverting behavior. Nevertheless, it is important to realize that predictability of option investment return is not directly related to the predictability of the floating implied volatility series, but instead depends on the predictability of the changes in the implied volatility of the option contract with a fixed maturity date. The expected percentage rate of change $\mu_t$ is defined on a fixed option contract and is directly related to the option investment P&L.

Traditional statistical analysis often starts by interpolating to generate floating implied volatility series and then analyzes the floating implied volatility series behavior. To investigate the investment return behavior, it is also important to directly analyze the behavior of implied volatility changes on option contracts of fixed maturity dates. For ease of analysis, one can also perform interpolation to generate the implied volatility changes of fixed contracts at floating maturity and moneyness points. In the empirical section, we analyze the implied volatility changes of fixed contracts at floating moneyness and maturity points and link their behavior to the implied volatility surface shape. In particular, equation (11) provide a linkage between the expected rate of implied volatility change on a fixed at-the-money contract to the term structure slope.

Equation (12) infers the expected rate of change $\mu_t$ of a fixed-contract implied volatility from its slope against the instantaneous return variance rate. Yet, the instantaneous variance rate itself is not observable. Thus, the inference would depend crucially on the accurate estimation of an unobserved quantity. Alternatively, by making further local commonality assumptions on the term structure movements, we can extract
the expected rate of change from the observed term structure. In particular, if we assume that the expected rate of change is the same within the maturity range \([\tau_1, \tau_2]\) (or at these two maturity points), we can infer the common expected rate of change from the implied variance slope within this range.

**Proposition 2** When the implied volatilities of at-the-money option contracts within a maturity range \([\tau_1, \tau_2]\) share the same risk-neutral expected rate of change \(\mu_t\) at time \(t\), this rate can be inferred from the at-the-money implied variance slope within this maturity range,

\[
\mu_t = \frac{A_t^2(\tau_2) - A_t^2(\tau_1)}{2(A_t^2(\tau_2)\tau_2 - A_t^2(\tau_1)\tau_1)},
\]

Equation (13) can be readily derived by applying equation (11) twice at the two maturities \(\tau_1\) and \(\tau_2\), respectively, with the same expected rate of percentage change \(\mu_t\).

### 3.3.2. Construct the at-the-money implied volatility term structure function

The separation between term and moneyness effects also allows us to construct a smooth term structure function based on commonality assumptions on the risk-neutral expected rate of percentage change \(\mu_t\) across different maturities.

The expected percentage change rate \(\mu_t\) is for a fixed-expiry implied volatility. To enhance comparability with the traditional literature, we start by modeling the floating constant-maturity at-the-money implied variance rate via a simple mean-reversion structure with proportional volatility

\[
dA_t^2(\tau) = \kappa_t(\bar{\theta}_t - A_t^2(\tau))dt + 2\omega A_t^2(\tau)dZ_t,
\]

where we assume that the at-the-money implied variance rates across all maturities converge to the same long-run level \(\bar{\theta}_t\) with the same mean-reverting speed \(\kappa_t\). Note that although we assume strong commonality for the time-\(t\) conditional dynamics of the at-the-money implied variance across different maturities, we do
not require that the same conditional dynamics will hold at any other time. The mean-reversion speed \( \bar{\kappa}_t \), the long-run target \( \bar{\theta}_t \), and the volatility coefficient \( \omega_t \) can all change over time. How they will change in the future is irrelevant for the option investment P&L over the next instant and is hence irrelevant for our valuation at time \( t \).

From (14), we can derive the expected percentage rate of change on the fixed-expiry implied volatility as,

\[
\mu_t = \frac{\kappa_t (\theta_t - A_t^2(\tau))}{2A_t^2(\tau)} - \frac{A'(\tau)}{A(\tau)},
\]

with \( \kappa_t = \bar{\kappa}_t + \omega_t^2 \) and \( \kappa_t \theta_t = \frac{\kappa_t}{\bar{\kappa}_t + \omega_t^2} \bar{\theta}_t \). The first term denotes the proportional drift of the floating constant-maturity at-the-money implied volatility. The second term denotes the slope of the implied volatility term structure around time to maturity \( \tau \) and captures the sliding of the fixed-expiry implied volatility along the term structure slope.

Substituting (15) into (11), and letting \( y_t(\tau) \equiv A_t^2(\tau) \tau \) denote the total variance, we arrive at a first-order ordinary differential equation,

\[
y_t'(\tau) + \kappa_t y_t(\tau) = \kappa_t \theta_t \tau + \sigma_t^2,
\]

starting at \( y_t(0) = 0 \). Integrating out (16) gives the solution for the at-the-money implied variance term structure function as

\[
A_t^2(\tau) = \frac{1 - e^{-\kappa_t \tau}}{\kappa_t \tau} (\sigma_t^2 - \theta_t) + \theta_t,
\]

which represents the at-the-money implied variance rate at each time to maturity \( \tau \) as a weighted average of the instantaneous variance rate \( \sigma_t^2 \) and a long-run target variance rate \( \theta_t \), with \( \kappa_t \) controlling the transition speed. The at-the-money implied variance rate converges to \( \sigma_t^2 \) as \( \tau \to 0 \) and converges to \( \theta_t \) as \( \tau \to \infty \).

**Proposition 3**  
*Imposing a common mean-reverting behavior for the floating at-the-money implied variance rates across different maturities as in (14) produces a smooth at-the-money implied variance rate term structure function as in (17).*
Principal component analysis often decomposes term structure variation into level, slope, and curvature components. Equation (17) represents a more structural way of capturing these three components. The variance level can be represented by its level at the short end $\sigma_t^2$ or its long-maturity limit $\theta_t$. The slope is captured by the short-long difference ($\sigma_t^2 - \theta_t$). The curvature is controlled by the mean-reversion speed $\kappa_t$.

Equation (17) represents one particularly simple specification that is similar in form to results from standard stochastic volatility models (e.g., Egloff, Leippold, and Wu (2010)). It does not need to be. First, one can analyze each contract on its own according to the pricing relation in (11), and ask whether the expected direction of the implied volatility for the contract is consistent with its level relative to the instantaneous variance rate. Second, the term structure functional form in (17) is not derived from no-arbitrage arguments across the term structure based on instantaneous variance rate dynamics assumption, but rather a result of the structural constraints imposed in the dynamics specification in (14). Deviations from the term structure form can be directly regarded as a deviation from the structural assumptions. Third, the structural constraints in (14) are temporal as neither the long-run level and the mean-reversion speed are considered as fixed parameters, but rather are regarded as dynamic state variables. Identification of their values based on observable implied variance rate time series allows us to build a closer link between dynamics and the term structure, and potentially a stronger way to identify the variance risk premium.

3.4. The implied volatility smile at a single maturity

To separate out the moneyness effect from the term structure effect, we can consider vega hedging options with the at-the-money contract of the same maturity (in addition to delta hedging), assuming (largely) common implied volatility movements at the same maturity. In so doing, we can take the at-the-money implied volatility level $A_t$ as given, and represent the implied volatility levels of other option contracts at the same maturity relative to the at-the-money implied volatility level.

In particular, if we assume that the expected implied volatility moves scale proportionally with the at-
the-money contract at the same maturity according to,

\[ \mu_t I_t^2 = \mu_t A_t^2, \]  

we can subtract (11) from (10) to highlight the smile effect,

\[ I_t^2 - A_t^2 = 2\gamma_t z_+ + \omega_t^2 z_+ z_- \]  

where the implied variance rate smile relative to the at-the-money contract is represented in terms of the two convexity-adjusted moneyness measures, with the slope and curvature of the smile determined by its covariance rate with the security return \( \gamma_t \) and its variance rate \( \omega_t^2 \), respectively.

If we assume that the conditional covariance and variance rates \((\gamma_t, \omega_t^2)\) are the same within a strike range \((k_1, k_2)\) at the same maturity, we can regard equation (19) as a continuous smile function in \( k \) within this strike range, and infer the common variance and covariance rates from the smile function within this maturity range. Given implied variance observations within this strike range, we can estimate the variance and covariance rates by regressing \((I_t^2 - A_t^2)\) against the two convexity-adjusted moneyness measure \([2z_+, z_+ z_-]\).

### 3.4.1. A local quadratic smile matching locally parallel dynamics

Earlier literature has proposed quadratic approximations of the implied volatility smile via expansions of the return distribution, e.g., Jarrow and Rudd (1982) and Backus, Foresi, and Wu (1997). Our framework allows us to derive quadratic or even higher-order polynomial approximations of the implied volatility smile under different assumptions on the commonality of the implied volatility smile movements.

To start, we assume equal proportional shifts on the smile with the same covariance and variance rates \((\gamma_t, \omega_t^2)\) across strikes. Then we can treat equation (19) as an implied volatility smile function across strikes. To make the smile function more explicit, we can factor out the convexity adjustment terms, and construct
the implied volatility smile purely as a function of the relative strike $k$,

$$I_t^2 = \mathcal{A}_t^2 + 2\gamma t k + \omega_t^2 k^2,$$  \hspace{1cm}  (20)

where

$$\mathcal{A}_t^2 = A_t^2 + \gamma_t I_t^2 \tau - \frac{1}{4} \omega_t^2 I_t^4 \tau^2.$$  \hspace{1cm}  (21)

If we further assume that $\mathcal{A}_t^2$ is flat across strike and is equal to the at-the-money forward implied variance

$$\mathcal{A}_t^2 = A_t^2 + \gamma_t I_t^2 \tau - \frac{1}{4} \omega_t^2 I_t^4 \tau^2 = A_t^2 + \gamma_t \mathcal{A}_t^2 \tau - \frac{1}{4} \omega_t^2 \mathcal{A}_t^4 \tau^2,$$  \hspace{1cm}  (22)

we can regard equation (20) as a quadratic approximation of the implied variance smile in log strike $k$.

More generally, equation (20) can be the starting point for constructing smooth implied volatility smile functions based on locally smooth assumptions on the conditional variance and covariance rates across moneyness. Assuming equal proportional shifts across the whole strike range of the smile may deviate too much from reality, but limiting this assumption to a narrow strike range appears to be a good starting point. Such a local proportional shift assumption leads to a locally quadratic smile. Thus, when one needs to smooth and interpolate implied volatility smiles, a local quadratic regression of the implied variance against the log strike is a good starting place.

In a classic paper, Breeden and Litzenberger (1978) show that one can construct the risk-neutral conditional density of the underlying security by taking the second derivative of the option price against the strike price. In practice, since option quotes are available only at a discrete number of strikes, it becomes imperative to perform smoothing interpolation across the discrete number of observations to come up with a smooth (and twice differentiable) price function against strike so that one can compute the second derivative. Figlewski (2009) shows that directly taking differences on the discrete option observations results in unstable density estimates. Aït-Sahalia and Lo (1998) apply locally constant Gaussian kernel nonparametric smoothing to the pooled implied volatility quotes over a sample period to come up with a smoothed implied
volatility surface, from which they compute the risk-neutral density across different horizons. Equation (20) suggests that applying a locally quadratic function along the log strike dimension in the nonparametric smoothing would be particularly useful as it not only guarantees twice differentiability against log strike but also generates local quadratic coefficients with specific meanings in terms of variance and covariance rates for the implied volatility movements.

3.4.2. A 4th-order polynomial smile accommodating separate volatility and skew movements

In exploring different smoothing methods on the implied volatility smile across strikes, Figlewski (2009) finds that a 4th-order spline works well for interpolating the implied volatility smile. We show that a 4th-order polynomial function on the implied variance smile can be made consistent with two main sources of variation on the smile, that is, the level and the slope.

Assuming these two sources of variation, we can decompose the percentage moves on the implied volatility smile into the movement of the at-the-money forward implied volatility and the movement of the implied volatility skew:

\[
\ln I_t(k) = \ln A_t + L_t k, \tag{23}
\]

where the skew measure \( L \) is defined in log implied volatility change per log relative strike distance,

\[
L_t = \frac{\ln I(k)/A_t}{k}. \tag{24}
\]

With this decomposition, the log percentage change of each implied volatility series for a fixed-strike contract can be represented as

\[
d \ln I_t(k) = d \ln A_t + k d L_t - L_t d \ln S_t, \tag{25}
\]

where the last term captures sliding along moneyness as the underlying security price moves. Accordingly, the variance and covariance of percentage implied volatility changes at each fixed moneyness \( k \) can be
expanded as,
\[
\gamma_t = C_{AS} + kC_{LS} - L_t \sigma_t^2, \\
\omega_t^2 = C_{AA} + 2kC_{AL} + k^2C_{LL} + L_t^2 \sigma_t^2 - 2L_t C_{AS} - 2L_t kC_{LS}.
\]
with \( C_{AS}, C_{LS}, C_{AA}, C_{AL}, C_{LL} \) denoting the corresponding covariances of in log percentage changes in spot, at-the-money-forward implied volatility, and the skew. Plugging equation (26) into the pricing relation in (20) results in a 4th-order polynomial.

To enhance identification, we further simplify the polynomial coefficients by making the following observations and assumptions. First, the percentage changes in the implied volatility level and skew show little correlation in practice. Thus, we set \( C_{LS} = 0 \). Second, from (20), the implied volatility skew at \( k = 0 \) is related to the covariance estimates by \( L_t = \frac{f'(0)}{A} = \frac{\gamma_t}{A_t^2} \). Third, we further consolidate the coefficients by approximating \( \sigma_t^2 \) with at-the-money implied variance \( A_t^2 \). With these assumptions, we can consolidate the coefficients as,
\[
\gamma_t = \frac{1}{2} C_{AS}, \\
\omega_t^2 = C_{AA} \left(1 - \frac{3}{4} \rho^2\right) + 2kC_{AL} + k^2C_{LL},
\]
where the sliding along the skew reduces the covariance of fixed-strike implied volatility series by half of the covariance with the floating series, and also reduces the variance of the implied volatility series. We use \( \tilde{C}_{AA} \equiv C_{AA} \left(1 - \frac{3}{4} \rho^2\right) \) to denote the adjusted variance estimate. With these consolidations, each coefficient of the polynomial captures the variance or covariance of the three common risk sources: the security return, the at-the-money implied volatility, and the skew.

**Proposition 4** When the movements of implied volatilities at one maturity can be summarized by two common factors: the at-the-money implied volatility and the implied volatility skew, the implied volatility smile at this maturity at any time \( t \) can be represented approximately as a fourth-order polynomial function of the relative log strike \( k \),
\[
I_t^2(k) = A_t^2 + C_{AS}k + \tilde{C}_{AA}k^2 + 2C_{AL}k^3 + C_{LL}k^4.
\]
the security return, the at-the-money implied volatility, and the implied volatility skew.

Sometimes for certain securities, investors observe distinct behaviors for out-of-the-money call and put options. The observations prompt them to model the call segment \((k > 0)\) and the put segment \((k < 0)\) of the implied volatility smile with different functional forms, subject to smooth pasting conditions between the two segments. We can readily accommodate such distinct behaviors by introducing separate variations for put and call skews, and accordingly separate variance and covariance coefficients, \((C_{AL}^c, C_{LL}^c)\) for calls and \((C_{AL}^p, C_{LL}^p)\) for puts. The implied volatility smile in this case is characterized by two segments of 4th-order polynomials:

\[
I^2_t(k) = \begin{cases} 
A^2_t + C_{AL}^c k + \tilde{C}_{AL}^c k^2 + 2C_{AL}^c k^3 + C_{LL}^c k^4, & \text{for calls } k > 0 \\
A^2_t + C_{AL}^p k + \tilde{C}_{AL}^p k^2 + 2C_{AL}^p k^3 + C_{LL}^p k^4, & \text{for puts } k < 0 
\end{cases} 
\]

(29)

where the implied volatility smile is continuous and smooth to the second order at \(k = 0\).

When empirically necessary, one can divide the skew into as many regions as desired and use separate covariance coefficients to capture the higher-order effects of each region while maintaining the global level, slope, and curvature around the at-the-money strikes. The net result would be close to a 4th order spline interpolation, but with specific interpretations on the coefficients.

4. Empirical analysis on S&P 500 index options

We perform an empirical analysis of our new approach using S&P 500 index (SPX) options. SPX options are actively traded on the Chicago Board of Options Exchange (CBOE). We obtain the history of closing option prices and implied volatilities on SPX options, as well as the underlying index level and interest rate series, from OptionMetrics. The sample period is from January 4, 1996 to April 29, 2016, spanning 5,111 business days. Over the sample period, the index level started around 617 and ended around 2,065, with an annualized daily return volatility of 19.5%.
4.1. Construct floating series of fixed-contract implied volatility percentage changes

The new theory relates the fair level of the implied volatility of an option contract to the first and second moment conditions of percentage changes in the implied volatility of this contract. To estimate the statistical moment conditions of such implied volatility changes, we need to construct time series of implied volatility changes at fixed time to maturity and moneyness points. The option contracts that trade on the exchange have fixed strike prices and expiry dates. Thus, their time to maturity declines as time moves forward and their moneyness changes as the underlying index level fluctuates. In this section, we use local averaging to generate the time series of implied volatility changes at fixed moneyness and time to maturity points.

It is important to point out that most empirical studies in the option pricing literature examine the behavior of floating implied volatility time series at fixed time to maturities and fixed moneyness. Changes in such floating implied volatility time series can be quite different from the fixed-contract implied volatility changes that we are about to construct. The difference can be particularly large when the implied volatility surface has a steep term structure and/or a strong implied volatility skew or smile. Even if the floating implied volatility surface remains the same, sliding along the term structure (due to time running forward) and along the skew (due to spot price movement) can lead to large changes in the implied volatility of a fixed option contract.

At each business date \( t \), we retrieve data on both date \( t \) and the next business day, and compute the log percentage implied volatility changes on each option contract. OptionMetrics provides a unique option identification number on each option contract that allows fast matching of option contracts. From the matched option contracts and computed implied volatility changes, we interpolate to generate the change at floating time to maturity and moneyness points. We consider five fixed time to maturities (\( \tau \)) at one month (30 days), two months (60 days), three months (91 days), six months (182 days), and one year (365 days). At each maturity, we compute the change for the at-the-money implied volatility (at \( z_+ = 0 \)). We also compute changes at three relative strike points at \( k/\sqrt{\tau} = 0, \pm 10\%, \pm 20\% \), where recall that 20% approximates the average annual volatility level. We also interpolate to obtain the date-\( t \) implied volatility levels at the same floating
points.

We obtain the floating implied volatility changes and levels via local averaging, with the following weighting schemes. First, at each strike, there can be two quotes, one from the call option and the other from the put option. We put more weight on the out-of-the-money option as they tend to be more actively traded and hence their quotes more reliable. For this purpose, we use one minus the absolute value of the option’s forward delta as the weight, and further truncate the weight to zero when the absolute delta is greater than 80%. The truncation essentially ignores deep in-the-money options when the absolute delta is over 80%, where the quotes tend to become unreliable.

Second, we weigh each observation based on its distance to the target relative strike and its distance to the target log time to maturity based on a Gaussian kernel with default bandwidth choices. Taking logs on time to maturity gives more resolution to the shorter time to maturity, where the term structure changes the most rapidly.

Taken together, to construct the implied volatility change and level at \((\bar{k}, \bar{\tau})\), we take a weighted average of the changes and levels of all option contracts, with the weighting on contract \(i\) by

\[
w_i = (1 - |\delta_i|)I_{|\delta_i| < 8}\exp\left(-\frac{(k_i - \bar{k})^2}{2h_k^2}\right)\exp\left(-\frac{(\ln \tau_i - \ln \bar{\tau})^2}{2h_\tau^2}\right),
\]

where \(\delta_i\) denotes the BMS forward delta of the option, and \((h_k, h_\tau)\) denote the two bandwidths. We use the same weighting to generate both the implied volatility level and the fixed-contract log percentage implied volatility changes.

To perform option P&L attribution, we also interpolate the option mid price change at the same floating time to maturity and moneyness. Since the call delta and put delta have opposite signs, we interpolate the call and put option changes separately. Since we assume zero financing rates in our derivation, we convert the spot option value into forward value in computing the changes.

Our interpolation scheme is based on the current best industry practice. Common variations in the
interpolation scheme include the relative weighting between call and put options at the same strike\(^5\) and the degree of smoothing applied to the quotes.\(^6\) Small variations in the interpolation methodology do not affect the general conclusion of our analysis, but can add noise to particular estimates.

### 4.2. The at-the-money implied variance term structure

Table 1 reports in Panel A the summary statistics of the interpolated floating at-the-money implied volatility levels at the five time to maturities. The sample average of the implied volatility levels increases as the time to maturity increases from one month to one year. The one-month at-the-money implied volatility averages at 19.6%, very close to the full-sample return standard deviation estimate of 19.5%. The sample average increases with maturity and reaches 21.5% at one-year maturity.

The standard deviation of the at-the-money implied volatility series declines with maturity from 7.4% at one-month maturity to 5.6% at one-year maturity. The range between the historical minimum and maximum also narrows with increasing maturity, consistent with the lowering standard deviation. The last row of the panel reports the autocorrelation of the floating series. All five series show mean reversion behavior, stronger at short maturities than at long maturities.

Panel B Table 1 reports the summary statistics for the daily percentage changes of the floating series. The statistics on the daily changes are annualized. As the floating series show no obvious trend during our sample period, the sample average of the daily change is very small, at merely about 3% per year across all maturities. By contrast, the annualized standard deviation estimates of the daily percentage change are very large, from 91.2% at one month maturity to 36.5% at one-year maturity. The minimum and maximum daily percentage changes also show a wide band. The autocorrelation estimates on the daily changes are all negative due to the mean reverting behavior of the floating series. The last row reports the contemporaneous correlation between the at-the-money volatility change and the security return. The estimates are strongly

\(^5\)Some may, for example, take only the out-of-the-money option quote at each strike and essentially set the weight on the in-the-money options to zero.

\(^6\)One can, for example, perform simple linear interpolation across log strike at each maturity and then perform linear interpolation in total variance across maturity at each fixed log strike or standardized moneyness level.
negative and in similar absolute magnitudes across all maturities.

Panel C Table 1 reports the summary statistics of the daily percentage implied volatility changes of the at-the-money contracts, which differs from the daily percentage changes of the floating series in Panel B due to maturity and moneyness changes as time moves forward and as the index level changes. Compared to Panel B, the daily percentage changes on the fixed contracts show a higher sample mean, from 13.3% at one-month maturity to 8.4% at one-year maturity. In absence of variance risk premium, these positive mean estimates imply an upward sloping at-the-money implied variance term structure, as shown equation (13).

The annualized standard deviation estimates in Panel C are smaller than that on the changes of the floating series in Panel B. The estimates range from 52.9% at one-month maturity to 20% at one-year maturity. The minimum and maximum also form a narrower band than in Panel B. Our derivations in equation (27) show that the variance of fixed-contract implied volatility changes are smaller than the variance of floating-series changes by approximately a factor of $(1 - \frac{3}{4} \rho^2)$. The ratios of the variance estimates between the two panels range from 0.30 to 0.34, implying a strong return-volatility correlation.

The autocorrelation estimates on the implied volatility changes of at-the-money contracts are negative at short maturities, but become positive at long maturities. Thus, mean reversion in the floating series does not always translate into mean reversion in the implied volatility changes of the fixed maturity date contracts.

The last row reports the contemporaneous correlation with the index return. The estimates are much smaller in absolute magnitude that the correlation estimates on the changes of the floating series in Panel B. Thus, sliding along the term structure and moneyness structure significantly alters the behavior of the fixed-contract implied volatility changes.

4.3. Extracting expected future movements from the at-the-money term structure

Assuming that the proportional movements in fixed maturity date implied volatilities are the same for a pair of nearby contracts, Proposition 2 allows us to extract the conditional risk-neutral forecast of the implied
volatility percentage changes ($\mu_t$) of the at-the-money contracts from the at-the-money implied variance term structure slope. We compute the forecast using implied variance estimates of adjacent time-to-maturities. Specifically, we estimate $\mu_t$ at 2-month maturity based on the term structure slope defined by 1-month and 3-month variances, estimate $\mu_t$ at 3-month maturity based on the term structure defined by 2- and 6-month variances, and so on. For maturities at the edge (one and 12 months), we use the nearest two maturities to estimate the expected direction.

Table 2 reports in Panel A the summary statistics of the $\mu_t$ estimates. The sample averages of the risk-neutral expectation are in the same range as the sample average of the actual changes in Panel C of Table 1, which can be regarded as an estimate for the physical expectation. The small difference suggests that the risk premium induced by the implied volatility variation is small. The large difference observed between variance swap rates and realized variance is, then, mostly driven by other risk factors such as concerns for price jumps.

The standard deviation of the estimates are about twice as large as the mean estimates, suggesting that the conditional expectation does vary significantly over time. To examine whether the risk-neutral conditional expectation predicts future percentage implied volatility change of the at-the-money contracts, Panel A also reports the forecasting correlation of the estimates on future changes, as well as the $t$-statistics of the correlation estimates. The correlation estimates are small, but are positive across all maturities. The $t$-statistics show that the estimates are statistically significant for all but the one-year maturity.

In Panel B of Table 2, we report the results from an out-of-sample forecasting exercise. On each date starting January 2006, we perform an autoregressive regression of the percentage changes against its past value using data over the past 10 years and we generate a forecast for the next daily percentage change based on the regression estimates. We compare this regression-based forecasting performance with the forecasting performance of the expected risk-neutral drift estimated from the term structure. Panel B reports the out-of-sample forecasting R-squared from the two approaches. To test the statistical significance of the forecasting performance difference between the two approaches, we compute the Diebold and Mariano (1995) (DM) $t$-
statistics on the squared forecasting error difference between the regression approach and the term structure estimates,

\[
DM = \frac{\bar{\delta}}{s_{\delta}} \left( \frac{T + 1 - 2(h + 1) + h(h + 1)/T}{T} \right)^{0.5},
\]

(31)

where \(\bar{\delta}\) denotes the sample mean of the difference, \(s_{\delta}\) denotes the Newey and West (1987) standard error estimate, computed with a lag of 5 days. The statistics adjust for the small-sample size bias according to Harvey, Leybourne, and Newbold (1997). Under the null hypothesis that each rolling-window estimated model and the random walk benchmark have equal finite-sample forecast accuracy, Clark and McCracken (2012) find that the thus-computed Diebold and Mariano (1995) test statistic can be compared to standard normal critical values. The last row of the panel reports the DM statistics.

For one-month contracts, the regression approach generates an R-squared estimate of 1.32% whereas the term structure approach generates an R-squared estimate of 0.33%. The DM test shows that the regression approach significantly outperforms the term structure approach. The regression approach is able to effectively take advantage of the mean-reversion in the daily changes to generate better forecasts. At the two-month maturity, the two approaches perform similarly with R-squared estimates at 0.48% and 0.488%, respectively. As maturity increases further, the mean reversion behavior becomes weaker and the performance of the regression approach starts to deteriorate, generating negative R-squared estimates at the six-month and the one-year maturity. By contrast, the term structure approach generates positive R-squared estimates across all maturities. The DM tests show that the term structure approach significantly outperforms at the 6-month and the one-year maturities, and marginally so at the three-month maturity.

---

4.4. P&L attribution for at-the-money options

For at-the-money options with \( z_+ = 0 \), s vanna and volga both vanish. As a result, the P&L attribution reduces to three terms, viz delta, vega, and gamma,

\[
\frac{dB}{B S S T} = l_A \frac{dS_i}{S_i} + A_i^2 \tau \left( \frac{dA_i}{A_i} - \mu dt \right) + \frac{1}{2} \left( \frac{dS_i}{S_i} \right)^2 - \sigma_i^2 dt, \tag{32}
\]

where the at-the-money ratio of cash delta to cash gamma is \( l_A = N(A_t \sqrt{\tau}) A_t \sqrt{\tau} / n(A_t \sqrt{\tau}) \).

We examine the variance contribution from each of the three risk sources. The forward call option value changes \((dB)\) and the implied volatility changes of the at-the-money contract \((dA_i/A_i)\) are interpolated from the option price and implied volatility quotes. The daily index return \((dS_i/S_i)\) are from the same data source. The drift \( \mu_t \) is estimated from the at-the-money implied variance term structure as described in the previous section, and we take the at-the-money implied variance at the short maturity (one month) as a proxy for \( \sigma_i^2 \).

To gauge the variance contribution of each risk component, we compare the variance of the unhedged option return per unit dollar gamma at each maturity with the variance of the “hedged” return series with progressively more risk sources removed,\(^8\)

\[
\begin{align*}
R_0 &= \frac{dB}{B S S T} \\
R_1 &= \frac{dB}{B S S T} - l_A \frac{dS_i}{S_i} \\
R_2 &= \frac{dB}{B S S T} - l_A \frac{dS_i}{S_i} - A_i^2 \tau \left( \frac{dA_i}{A_i} - \mu dt \right) \\
R_3 &= \frac{dB}{B S S T} - l_A \frac{dS_i}{S_i} - A_i^2 \tau \left( \frac{dA_i}{A_i} - \mu dt \right) - \frac{1}{2} \left( \frac{dS_i}{S_i} \right)^2 - \sigma_i^2 dt,
\end{align*}
\tag{33}
\]

with \( R_0 \) denoting the unhedged series and \( R_j \) for \( j = 1, 2, 3 \) denoting the return series with \( j \)-sources of risk removed. We measure the variance contribution of each risk component via an R-squared measure, \( R_j^2 \), defined as one minus the variance ratio of the corresponding hedged return series \( R + j \) to the unhedged return series \( R_0 \). The higher the R-squared estimates, the higher a proportion of variance can be attributed

\(^8\)We put a quotation market on the “hedged return” series as we are not performing a practical hedging exercise, but a simple variance attribution.
to the associated risk sources.

Table 3 reports the annualized standard deviation ($SD_0$) of the unhedged option return series ($R_0$) on the at-the-money contracts across different maturities, as well as the proportion of variance contribution from progressively more risk sources. The statistics on the left are for call options and on the right are for put options. As expected, the two sets of return series show similar behaviors.

The standard deviation estimates on the unhedged return series increase with option maturity. The variation is somewhat larger for the call options than for the put options. The row under $R^2_1$ measures the percentage variance contribution from the delta risk. The estimates are between 91.4% and 94.6%, suggesting that delta hedging removes a significant component of the option return variation. Indeed, under the BMS model environment, delta hedging can remove all unforeseen option return variation when the hedge updating frequency approaches the continuous limit. Our daily return measure induces some discretization error (Carr and Wu (2014)), and ignores the correlated component of movements induced by other risk sources. Still, despite the many documented deviations from the BMS assumptions, hedging the delta risk exposure based on the BMS suggestion can remove over 90% of the option investment risk. Figlewski (1989) documents similar magnitudes of variance reduction.

The row under $R^2_2$ measures the cumulative variance contribution from delta and vega risk. The estimates show that the implied volatility variation contributes to another one to two percentage to the option return variation. Further adjusting for gamma variation accounts for another half percentage of variation.

4.5. The implied volatility smile and its variations

Table 4 reports in Panel A the sample mean and standard deviation of the interpolated implied volatility smile series across different relative strikes. At each maturity, the average implied volatility level is higher at lower strikes than at higher strikes, forming the well-known implied volatility skew pattern. The standard deviation estimates are similar across different strike levels, but decline with increasing maturity.
Panel B reports the annualized mean and the annualized standard deviation statistics on the daily percentage implied volatility changes of each corresponding option contract. The mean estimates are strongly positive for out-of-the-money put options, especially at short maturities, but close to zero for long-term out-of-the-money call options. The implied volatility skew tends to be steeper at lower strikes, generating stronger sliding effects for out-of-the-money puts. The standard deviation estimates decline with increasing maturity. At each maturity, the standard deviation estimates on the daily changes for out-of-the-money calls are larger than that for out-of-the-money puts, even though the standard deviation estimates of the implied volatility levels in Panel A are smaller for out-of-the-money calls. Out-of-the-money calls tend to show more idiosyncratic movements, which are exacerbated in the daily changes.

In Panel C, we estimate the covariance rate ($\gamma_t$) between the daily index return and the daily percentage implied volatility changes with a 21-business day (one month) rolling window, and report the sample mean and standard deviation at each floating point. The mean covariance estimates are negative across all maturities and strikes, in line with the negatively skewed implied volatility smile. The mean estimates decline in absolute magnitude with increasing maturities. The standard deviation of the estimates also decline with maturity.

Panel D reports the sample mean and standard deviation of the 63-day rolling window variance estimates ($\omega_t$) for the daily percentage implied volatility changes. Similar to the covariance estimates, the variance estimates also decline as maturity increases.

At the same maturity, both the mean covariance and the variance estimates are larger in absolute magnitude at higher strikes than at lower strikes. While there is systematic institutional demand for out-of-the-money index put options, the demand for out-of-the-money call options tend to be more retail driven and less systematic, leading to more idiosyncratic movements and data noises for out-of-the-money calls. We suspect that the higher variance estimates are partially driven by such noise.
4.6. Extracting expected variance and covariance rates from the smile

From the implied volatility smile, by assuming common variance and covariance rates at each maturity, we can extract the risk-neutral variance and covariance rates. At each date and maturity, we regress the difference between the implied variance at the five relative strike points \( k = \sqrt{\tau}(0, \pm 10\%, \pm 20\%) \) and the at-the-money implied variance at \( z = 0 \) against \([2z_+, z_+ z_-]\). We can readily convert the relative strike \( k \) into the convexity-adjusted moneyness \( z_+ \) and \( z_- \) given the implied volatility estimate at each strike. The slope estimates from this regression are estimates for \( \gamma \) and \( \omega^2 \), respectively. We constrain the regression coefficient on \( \omega^2 \) to be positive.

Table 5 reports the mean and standard deviation of the regression-based risk-neutral conditional variance and covariance estimates in Panel A and compares in Panel B with the corresponding statistics on the rolling-window statistical estimates on the implied volatility changes of the at-the-money options. The average covariance estimates are negative in both panels across all maturities, but the risk-neutral estimates are more negative. The two sets of covariance estimates have similar standard deviation estimates, and show strongly positive co-movements as shown by the correlation estimates in Panel C.

The risk-neutral variance estimates extracted from the smile are larger than the statistical estimates at short maturities, but smaller at long maturities. The average magnitudes are in a similar range for the two sets of estimates. The risk-neutral estimates show much larger standard deviation, highlighting stability issues in the curvature estimates. Furthermore, different from the covariance estimates, the two sets of variance rate estimates show negative correlation. When inspecting the time series, we find that during volatile times when the statistical variance estimates become large, the smile tends to become highly negatively skewed and the curvature of the smile becomes harder to identify, leading to the negative correlation.
4.7. Statistical arbitrage on the smile

We explore the implications of the differences between the risk-neutral and historical estimates from several angles. First, we examine whether the risk-neutral estimates extracted from the smile can be used to enhance the forecast on future realizations of the variance and covariance rates. In Table 6, we regress the future one-month realized covariance rate $\gamma_t$ and variance rate $\omega_t$ of the at-the-money implied volatility series against the corresponding one-month historical estimator and the risk-neutral estimator extracted from the implied volatility smile. Entries report the full-sample regression coefficient estimates, the Newey and West (1987) standard errors (in parentheses, with 63-day lag), and the R-squared. Panel A reports the forecasting regression results for the covariance rate at each maturity. The R-squared estimates of the regressions range from 20.56% at one-month maturity to 26.09% at one-year maturity. Across all maturities, the coefficient estimates on both the historical estimator and the risk-neutral estimator are strongly positive. The coefficient estimates on both estimators are strongly significant statistically.

Panel B reports the forecasting regression results on the variance rate. The R-squared estimates are much lower for the variance rate prediction, increasing with maturity from merely 3.53% at one-month maturity to 13.67% at one-year maturity. The coefficient estimates on the historical estimator are strongly positive and statistically significant across all maturities. By contrast, the coefficient on the risk-neutral estimators are all negative. The negative coefficients are not statistically significant at short maturities, and only become strongly significant for the one-year implied volatility series. Thus, while the slope of the implied volatility smile contains strong information about future realized covariance rate between the implied volatility series and the index return, the curvature of the implied volatility smile is less informative about the variance rate.

Second, we want to understand the pricing implications of the average difference between the risk-neutral and historical estimates. Under continuous price and implied volatility dynamics, the instantaneous covariance rate $\gamma_t$ and variance rate $\omega_t$ should be the same under the physical and the risk-neutral measure. The average differences between the two sets of estimates suggest that either there are other risk sources that our pricing does not account for, such as random arriving price and/or implied volatility jumps, or there...
are potentially profitable statistical arbitrage opportunities.

We examine the latter via an out-of-sample investment exercise. Starting in January 2000, at each date and for each option maturity, we use a 4-year rolling window to estimate a forecasting relation on the one-month ahead covariance and variance rate on the at-the-money implied volatility by regressing against their corresponding historical estimators and risk-neutral estimators extracted from the observed smile. From the regression, we generate out-of-sample forecasts on the covariance rate $\gamma_t$ and the variance rate $\omega_t$. With the forecasts, we construct the fair implied volatility spread at the five relative strikes according to the pricing equation,

$$S_t = \hat{I}_t^2 - \hat{A}_t^2 = 2\gamma_t z_+ + \omega_t^2 z_+ z_-, \tag{34}$$

where we take the observed at-the-money implied volatility $A_t$ as given and generate fair valuations on the five relative strikes relative to this at-the-money implied volatility level. The two moneyness measures $z_+$ and $z_-$ also depend on the implied volatility level, making the equation iterative. We approximate the relation in one iteration by constructing the two moneyness measures using the observed implied volatility level.

At each of the five relative strike level, we form vega-neutral spread option portfolios against the at-the-money contract and normalize the weight on the at-the-money contract to one.\footnote{In principle, it does not matter whether we use call or put options to construct the vega-neutral spread as we maintain delta-neutral through the index futures. Nevertheless, since out-of-the-money options tend to be more actively traded than in the money options, we form put spreads at low strikes and call spreads at high strikes.} Then, we decide whether we enter a long or short position based on the difference between the spread valuation in (34) and the observed spread $S_t$ from the market, with the dynamic weight on the vega-neutral spread given by $w = (S_t - \hat{S}_t)/A_t^2$. That is, when the market observed spread is higher than the fair value, we long $w_t$ dollars of the at-the-money contract and short the contract at the corresponding relative strike to make it vega neutral. We then hold the position for 21 days (one month) while performing daily delta hedge with the underlying index futures.

We track the delta-hedged profit and loss of each spread portfolio and compute its annualized information ratio, i.e., the ratio of annualized profit to its annualized standard deviation. Table 7 reports the
annualized information ratio at each of the five relative strikes and across the five time to maturities. Each column represent one relative strike level, and each row represents one maturity. The last column reports the information ratio for investing in all the spreads across the five strikes at each maturity.

The information ratio is low for the investments in one-month options and even becomes slightly negative for the investment in the 10% out-of-the-money put option. The information ratio becomes increasingly large as the option maturity increases and reaches as high as 3.79 for investing in the five strikes at one-year maturity. Short-term options are more sensitive to price jumps, a component absent from our pricing relation. By contrast, for long-dated contracts, the variation and covariation of the implied volatility series dominate the profit and loss of the option investments. As a result, investing based on the difference between the variance/covariance estimators extracted from the observed smile and that from our statistical forecasts can generate very high information ratios. While the analysis itself is quite stylized and does not adjust for transaction costs, it highlights the importance of the implied volatility variation in pricing the implied volatility smiles, a cornerstone of our new pricing approach.

4.8. P&L variance attribution across moneyness

Table 8 examines the P&L variance attribution on options across different moneyness levels. Panel A reports the annualized standard deviation estimates on each unhedged option return series. At each moneyness level, the return variation increases with the option maturity. At each maturity, the variation of call option returns decreases with strike and the variation of put option returns increases with strike, highlighting the difference in the delta risk contribution.

Panel B shows the percentage variance contribution from the delta risk. As expected, the underlying index variation explains a higher percentage of variation for in-the-money options than for out-of-the-money options, around 80% for 20% out-of-the-money options, but up to 95% for 20% in-the-money options.

Panel C shows that the first-order underlying index movement and the implied volatility movement can
together explain over 90% of the variation for all put return series and for most call return series. Taking the difference between the estimates in Panels C and B show that the volatility variation contributes about one to two percentage points to the option return variation. Panel D further shows that the gamma variation contributes another half a percentage point to this variation. We have also analyzed the vanna and volga contributions, and found that their contributions are negligible.

5. Concluding remarks

We develop a new option pricing framework based on the profit and loss attribution analysis of option investments. The analysis starts with the Black-Merton-Scholes option pricing equation and attributes the instantaneous return of an option investment to calendar decay, to the underlying security price movement, to the option’s implied volatility movement, and to higher-order effects. The attribution highlights the key risk sources of option investment, and illustrates the role of dynamic hedging in neutralizing these risk sources. Taking risk-neutral expectation and applying dynamic no-arbitrage constraints results in a pricing relation that links the option’s fair implied volatility level to the implied volatility’s own expected direction of movement, its variance, and its covariance with the security return. The valuation does not need to specify where the first and second moment conditions come from and how they vary in the future, thus allowing us to make localized valuations based on what we know best about the particular contract, without interference from assumptions on other contracts or long term dynamics.

Nevertheless, the localized nature does not preclude us from either making local commonality assumptions on the implied volatility co-movements across nearby option contracts, or making global commonality assumptions based on principal component analysis or some other factor structures. Imposing these commonality assumptions on our pricing equation leads to cross-sectional pricing implications either for a selected number or range of option contracts, or across the whole implied surface.

By shifting the focus from terminal payoffs to short-term P&L fluctuation, our new theoretical frame-
work not only provides simple and flexible pricing solutions for derivative contracts of interest, but also tightly links the pricing to risk management practices on marking to market risks, and lays out a foundation for performing classic risk-return analysis on derivative contracts. In particular, all findings from traditional risk-return analysis on the underlying security, such as the various risk factors (e.g., BARRA) and the pricing anomalies identified in the stock market (e.g., Fama and French (1993), Jegadeesh and Titman (2001)), are directly applicable to the component of the option investment return due to its delta exposure. One can develop analogous risk structures on the option implied volatilities, and construct derivative investments in the same framework as the underlying securities. The framework is also useful in directing pricing anomaly research on options to different risk exposures.
References


Table 1
Summary statistics of at-the-money implied volatility levels and fixed-contract daily changes
Entries report the summary statistics of the interpolated at-the-money implied volatility level (Panel A), its
daily percentage change (Panel B), and the daily changes of the at-the-money contracts (Panel C) at time-to-
maturities of 1, 2, 3, 6, and 12 months. The statistics include the sample average (“Mean”), sample standard
deviceion (“Std”), minimum (“Min”), maximum (“Max”) and the daily autocorrelation (“Auto”) estimates.
In Panel B and C, we annualize the mean, standard deviation, minimum, and maximum statistics, and also
report the correlation with the index return in the last row (“Corr”).

<table>
<thead>
<tr>
<th>Maturity</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: At-the-money implied volatility level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.196</td>
<td>0.199</td>
<td>0.202</td>
<td>0.209</td>
<td>0.215</td>
</tr>
<tr>
<td>Std</td>
<td>0.074</td>
<td>0.070</td>
<td>0.068</td>
<td>0.061</td>
<td>0.056</td>
</tr>
<tr>
<td>Min</td>
<td>0.086</td>
<td>0.097</td>
<td>0.103</td>
<td>0.113</td>
<td>0.120</td>
</tr>
<tr>
<td>Max</td>
<td>0.736</td>
<td>0.702</td>
<td>0.670</td>
<td>0.587</td>
<td>0.526</td>
</tr>
<tr>
<td>Auto</td>
<td>0.983</td>
<td>0.987</td>
<td>0.989</td>
<td>0.992</td>
<td>0.994</td>
</tr>
<tr>
<td><strong>Panel B: Daily percentage change of the floating at-the-money implied volatility series</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.036</td>
<td>0.033</td>
<td>0.033</td>
<td>0.032</td>
<td>0.030</td>
</tr>
<tr>
<td>Std</td>
<td>0.912</td>
<td>0.734</td>
<td>0.634</td>
<td>0.470</td>
<td>0.365</td>
</tr>
<tr>
<td>Min</td>
<td>-82.543</td>
<td>-67.682</td>
<td>-65.370</td>
<td>-55.132</td>
<td>-58.003</td>
</tr>
<tr>
<td>Max</td>
<td>121.716</td>
<td>99.542</td>
<td>86.031</td>
<td>74.590</td>
<td>55.347</td>
</tr>
<tr>
<td>Auto</td>
<td>-0.106</td>
<td>-0.081</td>
<td>-0.063</td>
<td>-0.038</td>
<td>-0.032</td>
</tr>
<tr>
<td>Corr</td>
<td>-0.749</td>
<td>-0.776</td>
<td>-0.787</td>
<td>-0.795</td>
<td>-0.781</td>
</tr>
<tr>
<td><strong>Panel C: Daily percentage implied volatility changes of at-the-money contracts</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.133</td>
<td>0.096</td>
<td>0.101</td>
<td>0.102</td>
<td>0.084</td>
</tr>
<tr>
<td>Std</td>
<td>0.529</td>
<td>0.407</td>
<td>0.348</td>
<td>0.256</td>
<td>0.200</td>
</tr>
<tr>
<td>Min</td>
<td>-40.088</td>
<td>-37.851</td>
<td>-37.726</td>
<td>-29.608</td>
<td>-22.723</td>
</tr>
<tr>
<td>Max</td>
<td>78.000</td>
<td>61.714</td>
<td>52.921</td>
<td>37.376</td>
<td>35.131</td>
</tr>
<tr>
<td>Auto</td>
<td>-0.090</td>
<td>-0.052</td>
<td>-0.030</td>
<td>0.015</td>
<td>0.056</td>
</tr>
<tr>
<td>Corr</td>
<td>-0.437</td>
<td>-0.474</td>
<td>-0.488</td>
<td>-0.498</td>
<td>-0.473</td>
</tr>
</tbody>
</table>
Table 2
Forecasting future movement based on the term structure shape
Panel A reports the summary statistics of the risk-neutral drift ($\mu_t$) inferred from the at-the-money implied variance term structure, including the sample average (“mean”), standard deviation (“Std”), its forecasting correlation(“Corr”, in percentages) with future daily percentage implied volatility changes of the corresponding at-the-money contracts, and the $t$-statistics of the correlation estimates. Panel B reports the out-of-sample forecasting R-squared estimates (in percentages) from an autoregressive regression with a 10-year rolling window and the term structure approach. The last row reports the Diebold-Mariano (DM) test statistics between the two approaches. A negative statistic suggests outperformance from the regression approach.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.243</td>
<td>0.219</td>
<td>0.178</td>
<td>0.104</td>
<td>0.048</td>
</tr>
<tr>
<td>Std</td>
<td>0.497</td>
<td>0.367</td>
<td>0.241</td>
<td>0.147</td>
<td>0.081</td>
</tr>
<tr>
<td>Corr</td>
<td>5.542</td>
<td>5.833</td>
<td>5.808</td>
<td>4.040</td>
<td>0.620</td>
</tr>
<tr>
<td>$t$-stats</td>
<td>3.964</td>
<td>4.172</td>
<td>4.154</td>
<td>2.887</td>
<td>0.442</td>
</tr>
</tbody>
</table>

Panel B: Out-of-sample forecasting performance

<table>
<thead>
<tr>
<th></th>
<th>Regression $R^2$</th>
<th>Term structure $R^2$</th>
<th>DM-statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.323</td>
<td>0.327</td>
<td>-1.987</td>
</tr>
<tr>
<td></td>
<td>0.480</td>
<td>0.488</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>0.020</td>
<td>0.525</td>
<td>1.577</td>
</tr>
<tr>
<td></td>
<td>-0.700</td>
<td>0.387</td>
<td>2.832</td>
</tr>
<tr>
<td></td>
<td>-1.102</td>
<td>0.095</td>
<td>2.406</td>
</tr>
</tbody>
</table>

Table 3
P&L variance attribution for at-the-money options
Entries report the annualized standard deviation ($SD_0$) of the unhedged return series on at-the-money options as well as the proportion of variance contribution from progressively more risk sources: the delta risk ($R^2_1$), the vega risk ($R^2_2$), and the gamma risk ($R^2_3$). The statistics on the left are for call options and on the right are for put options.

<table>
<thead>
<tr>
<th></th>
<th>Call options</th>
<th>Put options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>$SD_0$</td>
<td>0.387</td>
<td>0.544</td>
</tr>
<tr>
<td>$R^2_1$</td>
<td>0.944</td>
<td>0.945</td>
</tr>
<tr>
<td>$R^2_2$</td>
<td>0.959</td>
<td>0.964</td>
</tr>
<tr>
<td>$R^2_3$</td>
<td>0.965</td>
<td>0.969</td>
</tr>
</tbody>
</table>
Table 4
Implied volatility smile and historical variance/covariance estimates
Entries report the sample average (left side) and standard deviation (right side) estimates for four sets of time series in four panels. Panel A reports the statistics for the floating implied volatility level at fixed time to maturities $\tau$ (1, 2, 3, 6, and 12 months) and fixed relative strikes ($-20\% \sqrt{\tau}, 0, 20\% \sqrt{\tau}$). Panel B reports the annualized statistics for the daily percentage implied volatility change of the corresponding option contracts. Panel C reports the statistics for the 63-day rolling historical covariance estimates between the index return and the percentage implied volatility change. Panel D reports the statistics for the 63-day variance estimates for the percentage implied volatility change.

<table>
<thead>
<tr>
<th>Stats</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{k}{\sqrt{\tau}} \backslash \tau$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Panel A: Implied volatility level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-10%</td>
<td>0.231</td>
<td>0.235</td>
</tr>
<tr>
<td>-20%</td>
<td>0.210</td>
<td>0.213</td>
</tr>
<tr>
<td>0</td>
<td>0.191</td>
<td>0.194</td>
</tr>
<tr>
<td>10%</td>
<td>0.177</td>
<td>0.178</td>
</tr>
<tr>
<td>20%</td>
<td>0.166</td>
<td>0.166</td>
</tr>
<tr>
<td>Panel B: Daily percentage implied volatility changes of option contracts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-10%</td>
<td>0.981</td>
<td>0.580</td>
</tr>
<tr>
<td>-20%</td>
<td>0.531</td>
<td>0.317</td>
</tr>
<tr>
<td>0</td>
<td>0.232</td>
<td>0.120</td>
</tr>
<tr>
<td>10%</td>
<td>0.164</td>
<td>0.023</td>
</tr>
<tr>
<td>20%</td>
<td>0.431</td>
<td>0.090</td>
</tr>
<tr>
<td>Panel C: Annualized covariance estimates $\gamma_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-10%</td>
<td>-0.034</td>
<td>-0.032</td>
</tr>
<tr>
<td>-20%</td>
<td>-0.042</td>
<td>-0.037</td>
</tr>
<tr>
<td>0</td>
<td>-0.053</td>
<td>-0.044</td>
</tr>
<tr>
<td>10%</td>
<td>-0.070</td>
<td>-0.055</td>
</tr>
<tr>
<td>20%</td>
<td>-0.093</td>
<td>-0.071</td>
</tr>
<tr>
<td>Panel D: Annualized variance estimates $\omega_t$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-10%</td>
<td>0.218</td>
<td>0.132</td>
</tr>
<tr>
<td>-20%</td>
<td>0.252</td>
<td>0.153</td>
</tr>
<tr>
<td>0</td>
<td>0.316</td>
<td>0.188</td>
</tr>
<tr>
<td>10%</td>
<td>0.430</td>
<td>0.249</td>
</tr>
<tr>
<td>20%</td>
<td>0.595</td>
<td>0.340</td>
</tr>
</tbody>
</table>
Table 5
Implied volatility smile and historical variance/covariance estimates
Entries report the sample average (mean) and standard deviation (Stdev) of the covariance rate $\gamma_t$ and variance rate $\omega_t^2$. The estimates in Panel A are extracted from the implied variance smile at each maturity. The estimates in Panel B are 63-day rolling window estimates on implied volatility changes of at-the-money options. Panel C reports the cross correlation between the two sets of estimates.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>$\gamma_t$</th>
<th>$\omega_t^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Mean</td>
<td>0.118</td>
<td>0.090</td>
</tr>
<tr>
<td>Stdev</td>
<td>0.067</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>0.047</td>
<td>0.039</td>
</tr>
<tr>
<td></td>
<td>0.080</td>
<td>0.068</td>
</tr>
<tr>
<td>Correlation</td>
<td>0.664</td>
<td>0.693</td>
</tr>
</tbody>
</table>
Table 6
Predicting realized variance/covariance rates with historical and risk-neutral estimators
We predict one-month ahead realized covariance rate $\gamma_t$ and variance rate $\omega^2_t$ for the at-the-money implied volatility series with their corresponding one-month historical estimator (HE) and the risk-neutral estimator (RN) extracted from the implied volatility smile. Entries report the full-sample forecasting regression coefficient estimates, Newey-West standard errors (in parentheses), and R-squared for each regression. Panel A reports the regression on the covariance rates. Panel B reports the regression on the variance rates. Each row is for each maturity.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Intercept</th>
<th>HE</th>
<th>RN</th>
<th>$R^2$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A. Predicting covariance rate $\gamma_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.011 (0.005)</td>
<td>0.356 (0.141)</td>
<td>0.152 (0.070)</td>
<td>20.56</td>
</tr>
<tr>
<td>2</td>
<td>-0.010 (0.005)</td>
<td>0.407 (0.141)</td>
<td>0.141 (0.090)</td>
<td>23.37</td>
</tr>
<tr>
<td>3</td>
<td>-0.010 (0.006)</td>
<td>0.438 (0.136)</td>
<td>0.118 (0.104)</td>
<td>24.37</td>
</tr>
<tr>
<td>6</td>
<td>-0.007 (0.005)</td>
<td>0.462 (0.128)</td>
<td>0.105 (0.110)</td>
<td>25.43</td>
</tr>
<tr>
<td>12</td>
<td>-0.003 (0.003)</td>
<td>0.446 (0.120)</td>
<td>0.167 (0.090)</td>
<td>26.09</td>
</tr>
<tr>
<td>B. Predicting variance rate $\omega_t$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.231 (0.030)</td>
<td>0.186 (0.054)</td>
<td>-0.004 (0.029)</td>
<td>3.53</td>
</tr>
<tr>
<td>2</td>
<td>0.134 (0.017)</td>
<td>0.221 (0.058)</td>
<td>-0.018 (0.036)</td>
<td>5.21</td>
</tr>
<tr>
<td>3</td>
<td>0.098 (0.013)</td>
<td>0.252 (0.062)</td>
<td>-0.039 (0.043)</td>
<td>7.04</td>
</tr>
<tr>
<td>6</td>
<td>0.053 (0.008)</td>
<td>0.286 (0.071)</td>
<td>-0.088 (0.051)</td>
<td>9.64</td>
</tr>
<tr>
<td>12</td>
<td>0.032 (0.005)</td>
<td>0.332 (0.090)</td>
<td>-0.180 (0.069)</td>
<td>13.67</td>
</tr>
</tbody>
</table>

Table 7
The annualized information ratio of out-of-sample option investments
Entries report the annualized information ratio of the out-of-sample option investment exercise on each contract. The last column denotes the information ratio for investing all the five strikes at each maturity.

<table>
<thead>
<tr>
<th>$\tau/\sqrt{k}$</th>
<th>-20%</th>
<th>-10%</th>
<th>0</th>
<th>10%</th>
<th>20%</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.32</td>
<td>-0.09</td>
<td>1.63</td>
<td>0.22</td>
<td>0.07</td>
<td>0.22</td>
</tr>
<tr>
<td>2</td>
<td>1.53</td>
<td>1.14</td>
<td>2.50</td>
<td>1.51</td>
<td>1.12</td>
<td>1.77</td>
</tr>
<tr>
<td>3</td>
<td>2.16</td>
<td>1.71</td>
<td>2.75</td>
<td>2.09</td>
<td>1.55</td>
<td>2.45</td>
</tr>
<tr>
<td>6</td>
<td>2.97</td>
<td>2.18</td>
<td>2.94</td>
<td>3.05</td>
<td>2.40</td>
<td>3.46</td>
</tr>
<tr>
<td>12</td>
<td>3.35</td>
<td>2.18</td>
<td>2.89</td>
<td>3.34</td>
<td>2.82</td>
<td>3.79</td>
</tr>
</tbody>
</table>
Table 8
P&L variance attribution across moneyness
Entries report the annualized standard deviation ($SD_0$) of the unhedged return series in Panel A, and the proportion of variance contribution from different risk sources in Panel B, C, and D. The statistics on the left side are for call options and on the right side are for put options.

<table>
<thead>
<tr>
<th>$\frac{k}{\sqrt{T}} \setminus \tau$</th>
<th>Call options</th>
<th>Put options</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>-10%</td>
<td>0.661</td>
<td>0.931</td>
</tr>
<tr>
<td>-20%</td>
<td>0.453</td>
<td>0.629</td>
</tr>
<tr>
<td>0</td>
<td>0.334</td>
<td>0.456</td>
</tr>
<tr>
<td>10%</td>
<td>0.264</td>
<td>0.355</td>
</tr>
<tr>
<td>20%</td>
<td>0.232</td>
<td>0.309</td>
</tr>
</tbody>
</table>

Panel B: Proportion of variance contribution from delta risk

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10%</td>
<td>0.944</td>
<td>0.941</td>
<td>0.938</td>
<td>0.932</td>
<td>0.919</td>
</tr>
<tr>
<td>-20%</td>
<td>0.932</td>
<td>0.928</td>
<td>0.925</td>
<td>0.918</td>
<td>0.902</td>
</tr>
<tr>
<td>0</td>
<td>0.920</td>
<td>0.916</td>
<td>0.912</td>
<td>0.904</td>
<td>0.887</td>
</tr>
<tr>
<td>10%</td>
<td>0.897</td>
<td>0.898</td>
<td>0.894</td>
<td>0.884</td>
<td>0.869</td>
</tr>
<tr>
<td>20%</td>
<td>0.802</td>
<td>0.806</td>
<td>0.802</td>
<td>0.790</td>
<td>0.777</td>
</tr>
</tbody>
</table>

Panel C: Proportion of variance contribution from delta and vega risk

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10%</td>
<td>0.957</td>
<td>0.958</td>
<td>0.958</td>
<td>0.953</td>
<td>0.945</td>
</tr>
<tr>
<td>-20%</td>
<td>0.954</td>
<td>0.956</td>
<td>0.957</td>
<td>0.952</td>
<td>0.945</td>
</tr>
<tr>
<td>0</td>
<td>0.949</td>
<td>0.955</td>
<td>0.955</td>
<td>0.951</td>
<td>0.946</td>
</tr>
<tr>
<td>10%</td>
<td>0.928</td>
<td>0.938</td>
<td>0.940</td>
<td>0.934</td>
<td>0.930</td>
</tr>
<tr>
<td>20%</td>
<td>0.814</td>
<td>0.821</td>
<td>0.819</td>
<td>0.803</td>
<td>0.794</td>
</tr>
</tbody>
</table>

Panel D: Proportion of variance contribution from delta, vega, and gamma risk

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10%</td>
<td>0.961</td>
<td>0.961</td>
<td>0.960</td>
<td>0.955</td>
<td>0.946</td>
</tr>
<tr>
<td>-20%</td>
<td>0.961</td>
<td>0.962</td>
<td>0.961</td>
<td>0.955</td>
<td>0.946</td>
</tr>
<tr>
<td>0</td>
<td>0.957</td>
<td>0.961</td>
<td>0.961</td>
<td>0.954</td>
<td>0.948</td>
</tr>
<tr>
<td>10%</td>
<td>0.933</td>
<td>0.944</td>
<td>0.944</td>
<td>0.935</td>
<td>0.931</td>
</tr>
<tr>
<td>20%</td>
<td>0.812</td>
<td>0.822</td>
<td>0.819</td>
<td>0.800</td>
<td>0.790</td>
</tr>
</tbody>
</table>