Option Pricing under Multiscale Stochastic Volatility

Cristina Tessari* and Caio Almeida†

†EPGE-FGV

Ongoing Master’s Thesis
First Draft: November 30, 2015

Abstract

The stochastic volatility model proposed by Fouque, Papanicolaou, and Sircar (2000) explores a fast and a slow time-scale fluctuation of the volatility process to end up with a parsimonious way of capturing the volatility smile implied by close to the money options. In this paper, we test three different models of these authors using options on the S&P 500. First, we use model independent statistical tools to demonstrate the presence of a short time-scale, on the order of days, and a long time-scale, on the order of months, in the S&P 500 volatility. Our analysis of market data shows that both time-scales are statistically significant. We also provide a calibration method using observed option prices as represented by the so-called term structure of implied volatility. The resulting approximation is still independent of the particular details of the volatility model and gives more flexibility in the parametrization of the implied volatility surface. In addition, to test the model’s ability to price options, we simulate options prices using four different specifications for the Data generating Process. As an illustration, we price an exotic option.

Keywords: Option pricing; Stochastic volatility; Mean-reversion.

*tinatessari@gmail.com; Corresponding author.
†calmeida@fgv.br
1 Introduction

Asset return volatility is a central concept in finance, whether in asset pricing, portfolio selection, or risk management. Although many well-known theoretical models assumed constant volatility (see Merton, 1969; Black and Scholes, 1973), since Engle (1982) seminal paper on ARCH models, the financial literature has widely acknowledge that volatility is time-varying, in a persistent fashion, and predictable (see Andersen and Bollerslev, 1997).

From the point of view of derivative pricing and hedging, continuous time stochastic volatility models, which can be seen as continuous time versions of ARCH-type models, have become popular in the last twenty years. The idea of using stochastic volatility models to describe the dynamics of an asset price comes from empirical evidence indicating that asset price dynamics is driven by processes with time-varying volatility. In fact, two phenomena are observed: (i) a non-flat implied volatility surface when the Black and Scholes (1973) model is used to interpret financial data, and (ii) skewness and kurtosis are present in the asset price probability density function deduced from empirical data. These phenomena stands in empirical contradiction to the consistent use of a classical Black-Scholes approach to pricing options and similar securities, as the Black–Scholes model fails to correctly describe the market behavior.

Stochastic volatility models relax the constant volatility assumption of the Black-Scholes model by allowing volatility to follow a random process. In this context, the market is incomplete because the volatility is not traded and the volatility risk cannot be fully hedged using the basic instruments (stocks and bonds). To preclude arbitrage, the market selects a unique risk neutral derivative pricing measure, from a family of possible measures. As a result, in contrast to the Black-Scholes model, the stochastic volatility models are able to capture some of the well-known features of the implied volatility surface, such as the volatility smile and skew.

The presence of volatility factors is well documented in the literature using underlying returns data (see Alizadeh et al., 2002; Andersen and Bollerslev, 1997; Chernov et al., 2003; Engle and Patton, 2001; Fouque et al., 2003b; Hillebrand, 2005; LeBaron, 2001; Gatheral, 2006, for instance). While some single-factor diffusion stochastic volatility models such as Heston (1993) enjoy wide success, numerous empirical studies of real data have shown that the two-factor stochastic volatility models can produce the observed kurtosis, fat-tailed return distributions and long memory effect. For example, Alizadeh, Brandt, and Diebold (2002) used ranged-based estimation to indicate the existence of two volatility factors including one highly persistent factor and one quickly mean-reverting factor. Chernov, Gallant, Ghysels, and Tauchen (2003) used the efficient method of moments (EMM) to calibrate multiple stochastic volatility factors and jump components. They showed
that two factors are necessary for log-linear models. These evidences, among others, have motivated the development of multiscale stochastic volatility models as an efficient way to capture the principle effects on derivative pricing and portfolio optimization of randomly varying volatility.

In this paper, motivated by both the popularity and appeal of stochastic volatility models and by the difficulty associated with their estimation, we compare the performance of three different specifications of the Fouque, Papanicolaou, and Sircar (2000) stochastic volatility model for pricing European call options with the Black-Scholes model. We assume that the underlying asset evolves according to a geometric Brownian motion, as in the Black and Scholes (1973) model, but with stochastic volatility. We use S&P 500 data to demonstrate that the volatility of this index is driven by two diffusions: one fast mean-reverting and one slow-varying. In order to identify these scales, we analyze low- and high-frequency data using the empirical structure function, or variogram, of the log absolute returns. To retrieve the scale on the order of months, we use daily closing prices of the S&P 500 over several years. To extract the scale on the order of days, we follow Fouque et al. (2003a) and use high-frequency, intraday data.

In their work, Fouque, Papanicolaou, and Sircar (2000) show that multiscale stochastic volatility models lead to a first-order approximation of derivatives prices and of the implied volatility surface. By using a combination of singular and regular perturbations to approximate prices when volatility is driven by short and long time scale factors, they derive a first-order approximation for European options prices and their induced implied volatilities. This first-order approximation is composed by the Black-Scholes price and the first-order correction only involves Greeks. In terms of implied volatility, this perturbation analysis translates into an affine approximation in the log-moneyness to maturity ratio (LMMR). This perturbation method allows a direct calibration using real data of the group market parameters, which are exactly those needed to price exotic contracts at this level of approximation. The advantage of this approach is that the resulting formula for the option price does not depend on the unobserved current value of the volatility. However, calibration of this model requires information on near-the-money implied volatilities.

We have three different goals in this work. First, to demonstrate the existence of both a well-identified short time-scale (on the order of a few days) and a long time-scale (on the order of months) in the S&P 500 volatility. In order to do that, we show that the fast time-scale and the slow time-scale can be simultaneously estimated using two different datasets, making use of the empirical structure function, or variogram. To test the statistical significance of the mean-reverting speeds, we use the circular block bootstrap of Politis and Romano (1992). We bootstrap pairs of residuals using the automatic
block-length selection of Politis and White (2004). The second goal, conditional on the market supporting evidence of a multiscale stochastic volatility process, is related to the calibration of the model using option data on the underlying. At this stage, we compare the observed prices with the corrected Black-Scholes prices by using three different first-order corrections: (i) a fast time scale correction; (ii) a slow time scale correction; and (iii) a multiscale (fast and slow) correction. Finally, the third goal of this paper is to simulate both the trajectories of these three different corrected prices based on the model of Fouque, Papanicolaou, and Sircar (2000) and the Black-Scholes prices, and compare the prices of exotic derivatives based on these simulated prices.

The rest of this paper proceeds as follows. In Section 2, we describe the class of multiscale stochastic volatility models that we will work with. In Section 2.3, we present the empirical structure function, or variogram, which is used as an estimator of the speed of mean reversion. In Section 2.4, we present the first-order approximation derived in Fouque, Papanicolaou, and Sircar (2000), which is valid for any European-style option. Additionally, we present an explicit formula for the implied volatility surface induced by the option pricing approximation. In Section 3, we report the numerical results. In Section 4, we show how the corrected Black-Scholes prices can be used to price exotic options. Finally, Section 5 concludes.

2 A Model for the Stock Price Dynamics

2.1 The Model

Let \((\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \geq 0})\) be a complete probability space with a filtration satisfying the usual conditions. We consider a family of stochastic volatility models \((X_t, Y_t, Z_t)\), where \(X_t\) is the underlying price, and \(Y_t\) and \(Z_t\) are two mean-reverting Ornstein–Uhlenbeck (OU) processes. Under the physical probability measure \(\mathbb{P}\), our model can be written as the following system of stochastic differential equations (SDE):

\[
\begin{align*}
\frac{dX_t}{dt} &= \mu X_t dt + f(Y_t, Z_t) X_t dW^0_t, \\
\frac{dY_t}{dt} &= \frac{1}{\epsilon} \alpha(Y_t) dt + \frac{1}{\sqrt{\epsilon}} \beta(Y_t) dW^1_t, \\
\frac{dZ_t}{dt} &= \delta c(Z_t) dt + \sqrt{\delta g(Z_t)} dW^2_t,
\end{align*}
\]

(1)

where \((W^0, W^1, W^2)\) are correlated \(\mathbb{P}\)-Brownian motions with

\[
dW^0_t dW^i_t = \rho_{i} dt, \quad i = 1, 2, \quad dW^1_t dW^2_t = \rho_{12} dt,
\]

(2)
and $|\rho_1| < 1$, $|\rho_2| < 1$, $|\rho_{12}| < 1$, and $1 + 2\rho_1 \rho_2 \rho_{12} - \rho_1^2 - \rho_2^2 - \rho_{12}^2 > 0$ in order to ensure that the correlation matrix of the Brownian motions is positive-semidefinite. The underlying price $X_t$ evolves as a diffusion with geometric growth rate $\mu$ and stochastic volatility $\sigma_t = f(Y_t, Z_t).$\footnote{If $\sigma_t$ is constant, then $X_t$ is a geometric Brownian motion and corresponds to the classical model used in the Black and Scholes (1973) theory.} The fast volatility factor, $Y_t$, evolves with a mean reversion time $\epsilon$, while the slow volatility factor, $Z_t$, evolves with a time scale $1/\delta$. In order to have a separation of scales, we must assume that $\delta \ll 1/\epsilon$.

Here, it is worth mentioning that this model characterizes an incomplete market, in contrast to the Black and Scholes (1973) model, since the volatility presents its own independent sources of uncertainty, $Y_t$ and $Z_t$. The introduction of these two new sources of randomness give rise to a family of equivalent martingale measures that will be parameterized by the market price of risk, $\lambda(y, z) = (\mu - r)/f(y, z)$, and two market prices of volatility risk, which we denote by $\xi(y, z)$ and $\zeta(y, z)$, associated respectively with $Y_t$ and $Z_t$. All these market prices are not determined within the model, but are fixed exogenously by the market. To preclude arbitrage, we assume that the market chooses one measure $\mathbb{P}^*$ through the combined market price of volatility risk ($\xi, \zeta$).

Under the risk neutral measure $\mathbb{P}^*$, an application of the Girsanov’s theorem provides that our model can be written as

$$
\begin{align*}
\begin{cases}
    dX_t &= rX_tdt + f(Y_t, Z_t)X_t dW_t^{0*}, \\
    dY_t &= \left(\frac{1}{\epsilon} \alpha(Y_t) - \frac{1}{\sqrt{\epsilon}} \beta(Y_t) \Lambda_1(Y_t, Z_t)\right) dt + \frac{1}{\sqrt{\epsilon}} \beta(Y_t) dW_t^{1*}, \\
    dZ_t &= \left(\delta c(Z_t) - \sqrt{\delta} g(Z_t) \Lambda_2(Y_t, Z_t)\right) dt + \sqrt{\delta} g(Z_t) dW_t^{2*},
\end{cases}
\end{align*}
$$

where $f$ is a positive function, bounded above and away from zero, $r \geq 0$ is the risk-free rate of interest, and the combined market prices of volatility risk associated with $Y_t$ and $Z_t$ are

$$
\begin{align*}
\Lambda_1(y, z) &= \rho_1 \lambda(y, z) + \xi(y, z) \sqrt{1 - \rho_1^2}, \\
\Lambda_2(y, z) &= \rho_2 \lambda(y, z) + \xi(y, z) \rho_{1,2} + \zeta(y, z) \sqrt{1 - \rho_2^2 - \rho_{12}^2}.
\end{align*}
$$

In our model, the coefficients $\alpha(y) = m_Y - y$ and $\beta(y) = \nu_Y \sqrt{2}$ describe the dynamics of $Y$ under the physical measure $\mathbb{P}$ and the coefficients $c(z) = m_Z - z$ and $g(z) = \nu_Z \sqrt{2}$ describe the dynamics of $Z$ under $\mathbb{P}$. Their particular form does not play a role in the perturbation analysis provided that they are defined so that the processes $Y$ and $Z$ are mean-reverting and have a unique invariant distribution denoted by $\Phi_Y$ and $\Phi_Z$, respectively. Since $Y$ and $Z$ are OU processes, it can be shown that $\Phi_Y$ is the density of the normal distribution $N(m_Y, \nu_Y^2)$, while $\Phi_Z$ is the density of the normal
distribution $N(m_Z, \nu_Z)$.\(^2\) Finally, the $\mathbb{P}^*$-standard Brownian motions $(W_t^{0*}, W_t^{1*}, W_t^{2*})$ present the same correlation structure as between their $\mathbb{P}$-counterparts in Equation (2).

### 2.2 Pricing Equation

Consider a European option with smooth and bounded payoff function $h(x)$ and expiration date $T$. The fact that the discounted price $\hat{P}_t = e^{-rt}P_t$ is a $\mathbb{P}^*$-martingale guarantees that the no-arbitrage pricing function of this option at time $t < T$ is given by

$$P^{\epsilon, \delta}(t, X_t, Y_t, Z_t) = \mathbb{E}^* \{ e^{-r(T-t)}h(X_T)|X_t, Y_t, Z_t \},$$

(6)

where the expectation $\mathbb{E}^* \{ \cdot \}$ is taken under the risk-neutral pricing measure $\mathbb{P}^*$, and we have used the Markov property of $(X_t, Y_t, Z_t)$. By an application of the Feynman-Kac formula, we obtain a characterization of $P^{\epsilon, \delta}$ as the solution of the parabolic partial differential equation (PDE)

$$\frac{\partial P^{\epsilon, \delta}}{\partial t} + \mathcal{L}_{(X,Y,Z)}P^{\epsilon, \delta} - rP^{\epsilon, \delta} = 0$$

(7)

with terminal condition $P^{\epsilon, \delta}(T, x, y, z) = h(x)$, where $\mathcal{L}_{(X,Y,Z)}$ is the infinitesimal generator of the Markov processes $(X_t, Y_t, Z_t)$. Define the infinitesimal generator $\mathcal{L}^{\epsilon, \delta}$ as

$$\mathcal{L}^{\epsilon, \delta} = \frac{\partial}{\partial t} + \mathcal{L}_{(X,Y,Z)} - r,$$

(8)

so that (7) can be written as

$$\mathcal{L}^{\epsilon, \delta}P^{\epsilon, \delta} = 0,$$

(9)

with terminal condition $P^{\epsilon, \delta}(T, x, y, z) = h(x)$. It is convenient to re-write the operator $\mathcal{L}^{\epsilon, \delta}$ as a sum of components that are scaled by the different powers of the small parameters $(\epsilon, \delta)$ that appear in the infinitesimal generator of $(X, Y, Z)$ as

$$\mathcal{L}^{\epsilon, \delta} = \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 + \sqrt{\delta} \mathcal{M}_1 + \delta \mathcal{M}_2 + \sqrt{\delta \epsilon} \mathcal{M}_3,$$

(10)

where the operators are defined as

$$\mathcal{L}_0 = \frac{1}{2} \beta^2(y) \frac{\partial^2}{\partial y^2} + \alpha(y) \frac{\partial}{\partial y},$$

(11)

$$\mathcal{L}_1 = \beta(y) \left( \rho_1 f(y, z)x \frac{\partial^2}{\partial x \partial y} - \Lambda_1(y, z) \frac{\partial}{\partial y} \right),$$

(12)

\(^2\)For additional details about the derivation of the invariant distributions, see Fouque, Papanicolaou, Sircar, and Solna (2011).
\[
\mathcal{L}_2 = \frac{\partial}{\partial t} + \frac{1}{2} f^2(y, z) x^2 \frac{\partial^2}{\partial x^2} + r \left( x \frac{\partial}{\partial x} - \cdot \right),
\]
\[
\mathcal{M}_1 = g(z) \left( \rho_2 f(y, z) x \frac{\partial^2}{\partial x \partial z} - \Lambda_2(y, z) \frac{\partial}{\partial z} \right),
\]
\[
\mathcal{M}_2 = \frac{1}{2} g^2(z) \frac{\partial^2}{\partial z^2} + c(z) \frac{\partial}{\partial z},
\]
\[
\mathcal{M}_3 = \beta(y) \rho_{1,2} g(z) \frac{\partial^2}{\partial y \partial z}.
\]

Note that \(\mathcal{L}_2\) is the Black-Scholes operator, corresponding to a constant volatility level \(f(y, z)\). The Black-Scholes price \(C_{BS}(t, x; \sigma)\), the price of a European claim with payoff \(h\) at the volatility \(\sigma\), is given as the solution of the following PDE

\[
\mathcal{L}_{BS} C_{BS} = 0,
\]

with terminal condition \(C_{BS}(T, x; \sigma) = h(x)\). If we are able to calculate the solution of the PDE given by Equation (9), then we know how to price derivatives under the proposed model. However, for general coefficients \((f, \alpha, \beta, c, g, \Lambda_1, \Lambda_2)\), we do not have an explicit solution to this equation. In this context, Fouque, Papanicolaou, and Sircar (2000) developed an asymptotic approximation for the option price in the neighborhood of the Black-Scholes price that made the calibration problem computationally tractable.

In order to obtain the first-order approximation, they expand \(P_{\epsilon, \delta}\) in powers of \(\sqrt{\epsilon}\) and \(\sqrt{\delta}\) as follows:

\[
P_{\epsilon, \delta}(t, x, y, z) = \sum_{i \geq 0} \sum_{j \geq 0} \sqrt{\epsilon} \sqrt{\delta} P_{i,j}(t, x, y, z),
\]

where \(P_{0,0} = P_{BS}\) is the Black-Scholes price, \(P_{1,0} = \sqrt{\epsilon} P_{1,0}\) is the first-order fast scale correction, and \(P_{0,1} = \sqrt{\delta} P_{0,1}\) is the first-order slow scale correction. In the Appendix, we present the derivation of this first-order correction to the Black-Scholes solution.

If we define \(D_k\) as

\[
D_k = x^k \frac{\partial^k}{\partial x^k},
\]

then, for nice payoff functions \(h\), Fouque, Papanicolaou, and Sircar (2000) show that the corrected call price \(P^*\) for European options using a combination of singular and regular perturbation is given by

\[
P^*(t, x, y, z) = P_{BS}^* + (T - t) \left( V_0^\delta(z) \frac{\partial P_{BS}^*}{\partial \sigma} + V_1^\delta(z) D_1 \frac{\partial P_{BS}^*}{\partial \sigma} + V_3^\delta(z) D_1 D_2 P_{BS}^* \right),
\]

where \(P_{\epsilon, \delta} = P^* + O(\epsilon + \delta)\). Note that this formula involves only Greeks of the Black-Scholes price \(P_{BS}^*\) evaluated at \(\sigma^*(z)\). In addition, observe that only the group market parameters \((\sigma^*, V_0^\delta, V_1^\delta, V_3^\delta)\) are needed to compute the approximate price \(P^*\), where \(V_0^\delta\)
and \( V_1^\delta \) are of order \( \sqrt{\delta} \) and \( V_3^\epsilon \) is of order \( \sqrt{\epsilon} \). Thus, the main advantages of this price approximation are the ease of implementation and the parsimony in the number of parameters. In addition, if \( \delta = 0 \), then only the group market parameters \( \sigma^*(z) \) and \( V_3^\epsilon(z) \) are needed to compute the contribution due to the fast time scale rather than the full specification of the stochastic volatility model. Similarly, if \( \epsilon = 0 \), then the slow time scale contribution to \( \tilde{P}^* \) is contained in only two parameters, \( V_0^\delta(z) \) and \( V_1^\delta(z) \). Accordingly, the corrected price (20) obtained when we consider a two-factor stochastic volatility models includes the correction due to the fast time scale and the correction due to the slow time scale as its particular cases.

### 2.3 The Variogram and Time Scales in Market Data

In order to identify the time scales in the volatility, this section introduces the variogram, or empirical structure function, as a tool for estimating the mean-reversion times.

Consider a discrete version of time, with \( \Delta t = t_n - t_{n-1}, \) \( 0 \leq n \leq N \), and let \( X_n \) denote the S&P 500 data recorded at time \( t_n \). Adopting a discrete version of Equation (1), we define the normalized fluctuation of the data, \( \bar{D}_n \), as

\[
\bar{D}_n = \frac{1}{\sqrt{\Delta t}} \left( \frac{\Delta X_t}{X_t} - \mu \Delta t \right) = \sigma_t \frac{\Delta W_t}{\Delta t},
\]

where \( X_n, Y_n, Z_n \) and \( W_n \) represent the processes \( X_t, Y_t, Z_t \) and \( W_t \) sampled at time \( n\Delta t \), \( \sigma_t = f(Y_t, Z_t), \ t \geq 0 \), is the (positive) volatility process, \( \mu \) is a constant, and \( \Delta W_t \) is the increment of a Brownian motion. From the basic properties of a Brownian motion, we are able to model \( \bar{D}_n \) by

\[
\bar{D}_n = f(Y_n, Z_n)\epsilon_n,
\]

where \( \{\epsilon_n\} \) is a sequence of i.i.d. \( N(0,1) \) random variables. In order to obtain an expression similar to that found in Fouque, Papanicolaou, and Sircar (2000), we assume that the stochastic volatility can be written as \( f(y, z) = f_1(y)f_2(z) \). This functional form was proposed by Fouque, Papanicolaou, Sircar, and Solna (2003c) for the case in which the correlation between the instantaneous volatility components is zero.\(^3\) In this case, in order to obtain an additive noise process, we define the log absolute value of the normalized fluctuations as

\[
F_n = \log |\bar{D}_n| = \log |f_1(Y_n)| + \log |f_2(Z_n)| + \log |\epsilon_n|.
\]

\(^3\) Machuca (2010) demonstrated that this equation remains valid even when the correlation between \( Y \) and \( Z \) is different from zero.
Thus, the empirical structure function or variogram of $F_n$, which is a measure of the correlation structure of the sampled version of the process $\log (f(t, Z_t))$, is given by

$$V^N_j = \frac{1}{N_j} \sum_n (F_{n+j} - F_n)^2 ,$$

where $j$ is the lag for which we are measuring the correlation and $N_j$ is the total number of points for each lag.

In this paper, we consider a model with two components driving the volatility, one fast and one slow. In this case, the stochastic volatility model is given by

$$f(t, Z_t) = e^{Y_t + Z_t},$$

where $\{Y_t\}$ and $\{Z_t\}$ are independent OU processes. Fouque, Papanicolaou, and Sircar (2000) show that, for each $j = 1, 2, \ldots, J$, the empirical variogram $V^N_j$ is an unbiased estimator of

$$V_j = 2\gamma^2 + 2\nu_Y^2 \left(1 - e^{-\alpha j \Delta t}\right) + 2\nu_Z^2 \left(1 - e^{-\delta j \Delta t}\right),$$

where $\alpha \equiv 1/\epsilon$, $\gamma^2 = \text{Var}(\log(|\epsilon|))$, $\nu_Y^2$ is the variance of the stationary distribution of the process $\log f_1(Y_t)$ and $\nu_Z^2$ is the variance of the stationary distribution of the process $\log f_2(Z_t)$.

Note that, for the range of lags we are looking at, that is $j \Delta t$ up to a week, $\delta j \Delta t$ is small and the last term is negligible. Hence, we can use the following approximation in order to obtain the component of fast mean-reversion:

$$V^N_j \approx 2\gamma^2 + 2\nu_Y^2 \left(1 - e^{-\alpha j \Delta t}\right).$$

(27)

At the opposite extreme, if we had looked at a longer data sample and included larger lags, then the term related to $\alpha$ in (26) is sufficiently large. In this case, we can consider the following approximation to obtain the component of slow mean-reversion:

$$V^N_j \approx 2(\gamma^2 + \nu_Y^2) + 2\nu_Z^2 \left(1 - e^{-\delta j \Delta t}\right).$$

(28)

Using this result, we are able to estimate by non linear regression methods the rates of mean reversion of the volatility processes, $1/\epsilon$ and $\delta$. Note that Equations (27) and (28) share a parameter, $\gamma$. In order to take this fact into account, we fit two different datasets to different equations with a shared parameter. For the high frequency dataset, we use the empirical variogram from Equation (22) as the dependent variable, and adjust it to the functional form given in Equation (27). At the same time, for the slow frequency dataset, we adjust the empirical variogram to the functional form given by Equation (28).
The resulting estimates are reported in Section 3.

2.4 Implied Volatility Asymptotic and Calibration

It is common practice to quote option prices in units of implied volatility, by inverting the Black-Scholes formula for the European option with respect to the volatility parameter. This quantity is a convenient change of unit through which to view the departure of market data from the Black-Scholes theory.

The multiscale asymptotic theory developed in Fouque, Papanicolaou, and Sircar (2000) is designed to capture some of the important effects that fluctuations in the volatility have on derivative prices. In their work, they propose a first order approximation for the implied volatility function obtaining a very nice linear relation between implied volatility and the ratio log moneyness to time to maturity, from which we can estimate the parameters from Equation (20). This procedure is robust and no specific model of stochastic volatility is actually needed.

In order to obtain the implied volatilities, we consider a European call option with strike \( K \) and maturity \( T \). In this case, the corresponding payoff function is given by \( h(x) = (x - K)^+ \), and the Black-Scholes price at the corrected effective volatility \( \sigma^* \) is

\[
P^*_\text{BS} = xN(d_1^*) - Ke^{-r\tau}N(d_2^*),
\]

where \( \tau = T - t \) is the time to maturity, \( N \) is the cumulative standard normal distribution and

\[
d_{1,2}^* = \frac{\log(x/K) + (r \pm \frac{1}{2} \sigma^{*2}) \tau}{\sigma^* \sqrt{\tau}}.
\]

Given an observed European call option price \( C^{\text{obs}} \), the implied volatility \( I \) is defined to be the value of the volatility parameter that must go into the Black-Scholes formula (29) in order to have \( C_{BS}(t, x; K, T; I) = C^{\text{obs}} \). Thus, as \( \frac{\partial P_{BS}}{\partial \sigma} = \tau \sigma^* D_2 P_{BS}^* \) for European vanilla options, we can rewrite the corrected price (20) as

\[
P^* = P_{BS}^* + \left( \tau V_0^\delta + \tau V_1^\delta D_1 + \frac{V_2^\delta}{\sigma^*} D_1 \right) \frac{\partial P_{BS}^*}{\partial \sigma},
\]

where \( \frac{\partial P_{BS}}{\partial \sigma} = \frac{\tau^2 e^{-d_1^*/2}}{\sqrt{2\pi}} \).

Remember that \( \sigma^* \) solves \( P_{BS}^* = C_{BS}(\sigma^*) \). Thus, by expanding the difference \( I - \sigma^* \) between the implied volatility \( I \) and the volatility used to compute \( P_{BS}^* \) in powers of \( \sqrt{\epsilon} \) and \( \sqrt{\delta} \), Fouque, Papanicolaou, and Sircar (2000) show that the first-order approximation
for the implied volatility, \( I \approx \sigma^* + \sqrt{\epsilon} I_{1,0} + \sqrt{\delta} I_{0,1} \), takes the simple form

\[
I \approx b^* + \tau b^\delta + \left( a^\epsilon + \tau a^\delta \right) \text{LMR},
\]

where LMMR, the log-moneyness to maturity ratio, is defined by

\[
\text{LMMR} = \frac{\log (K/x)}{T - t} = \frac{\log (K/x)}{\tau}.
\]

The parameters \((b^*, b^\delta, a^\epsilon, a^\delta)\) depend on \(z\) and are related to the group parameters \((\sigma^*, V_0^\delta, V_1^\delta, V_3^\delta)\) by

\[
b^* = \sigma^* + \frac{V_3^\epsilon}{2\sigma^*} \left( 1 - \frac{2r}{\sigma^*2} \right),
\]

\[
b^\delta = V_0^\delta + \frac{V_1^\delta}{2} \left( 1 - \frac{2r}{\sigma^*2} \right),
\]

\[
a^\epsilon = \frac{V_3^\epsilon}{\sigma^*3},
\]

\[
a^\delta = \frac{V_1^\delta}{\sigma^*2}.
\]

The coefficients \(b^*\) and \(a^\epsilon\) are due to the fast volatility factor, while the coefficients \(b^\delta\) and \(a^\delta\) are due to the slow volatility factor, which becomes more important for large maturities. If we ignore the slow scale, then the implied volatility can be approximated by

\[
I \approx b^* + a^\epsilon \text{LMR},
\]

which corresponds to assuming \(\delta = 0\). On the other hand, if we assume \(\epsilon = 0\), then

\[
I \approx \sigma^* + b^\delta \tau + a^\delta \text{LM},
\]

where we denote LM = \(\tau\text{LMR}\), the log-moneyness. The estimated parameters for Equations (32), (38) and (39) will be presented in Section 3. In Section 4, we show that these same parameters are exactly those needed to price exotic derivatives. Thus, in the regime where \(\epsilon\) and \(\delta\) are small, corresponding to stochastic volatility models with fast and slow volatility factors, we can handle a wide class of pricing problems which would be very complicated without the perturbation theory.
3 Numerical Results

In this section, we present the dataset that will be used in this paper and the numerical results of our analysis. We divide our empirical study in two parts: the results related to the dynamics of the volatility and price processes of the S&P 500 index, and the results regarding the calibration of the model using option prices on this index.

3.1 Data

We are interested in estimating the fast and slow mean-reverting times of the S&P 500. In order to find the scale on the order of days, we use high frequency, 1-min data, from October 23, 2014, to October 23, 2015. To extract the fast scale, we average the data over 5-min intervals so that we have 72 data points per day. We collapse the time by eliminating overnights, weekends and holidays, so that we have 252 trading days with 18,072 data points per year. To estimate the time scale on the order of months, we use daily closing prices from January 02, 1980, to October 23, 2015. In this case, we average the data over 5-day intervals so that we have 50 data points per year.

To perform the fitting procedure described in Section 2.4, we use S&P 500 options data on November 03, 2015, obtained from the Bloomberg database. Our dataset contains S&P 500 index option quoted bid-ask prices, implied volatilities, and contract details such as strikes and expiration dates. We shall present results based on European call options, taking the average of bid and ask quotes to be the current price. Throughout the empirical analysis, we will restrict ourselves to options between 0.95 and 1.05 moneyness so as to be on the safe side of liquidity issues.

In order to remove inconsistencies and unreliable entries, we roughly follow the data-cleaning procedure described in Fouque, Papanicolaou, Sircar, and Solna (2011). We filter bid quotes less than $0.50 and options with no implied volatility value. We also remove the implied volatilities with the shortest maturity from our dataset. To smooth the jump between the implied volatility curves coming from calls or puts with the same maturity, we follow the general procedure described in Figlewski (2009), blending the implied volatility values for options with strikes between certain cutoffs $L$ and $H$, with $L < H$. We choose $L = 0.95$ and $H = 1.05$. For strikes less than $L$, we use only puts; for strikes greater than $H$, we use only calls; and for strikes between $L$ and $H$, we blend the two.
3.2 Mean-Reversion Rates and Block Bootstrap

In a previous study, Fouque, Papanicolaou, Sircar, and Solna (2003a) identify a fast mean-reversion time for the volatility of the returns of the S&P 500 index. Empirical evidences also show that the S&P 500 exhibits long memory, for example, the reader can refer to Chronopoulou and Viens (2012). In this context, our goal in this subsection is to verify if the volatility of the S&P 500 index is driven by two factors simultaneously: one factor mean-reverting on a short scale and the other on a long scale.

At this point, based on the procedure described in Section 2.3, we use the variogram to estimate the volatility mean-reversion times: $1/\epsilon$ and $\delta$. To do that, we fit the variogram given by Equation (24) by weighted least squares solving the following problem:

$$\min_\theta \sum_j \omega_j \left( V_j^N - V_j(\theta) \right)^2,$$

where $\theta$ is the vector of parameters to be estimated. We choose $\omega_j$ as the Cressie (1985) weights, given by an approximation for the variance of $V_j^N$:

$$\omega_j = \frac{1}{\text{Var}(V_j^N)} \approx \frac{N_j}{[V_j(\theta)]^2}.$$

Note that the Cressie’s weights put the highest emphasis on variogram estimates that are based on a large number of points (where $N_j$ is large) and on values near $j = 0$ (where $V_j(\theta)$ is small). In Figure 1, we present the empirical and the fitted variograms for the S&P 500, using the discrete model described in Section 2. The dashed line is a fitted exponential obtained by nonlinear weighted least squares regression. It is worth mentioning that, if the variograms in Figure 1 were flat, we would conclude there is no fast and/or slow time scale on the S&P 500 index.

To obtain the sample distribution of the estimated parameters, we bootstrap the pairs of residuals of the non linear regressions solved to obtain the variograms given by Equations (27) and (28). We use the circular block bootstrap of Politis and Romano (1992), with 10,000 bootstrap samples. The block size was chosen according to the procedure developed by Politis and White (2004). Table 1 presents the estimated values and standard deviations for the parameters, as well as the 95% bootstrapped confidence interval. The

---

4We refer to Fouque, Papanicolaou, Sircar, and Solna (2011) for further details about the properties of this estimator.
Figure 1: The empirical variogram for the S&P 500. The top graph shows the estimate of the variogram model for 5-min log normalized fluctuations after applying a 9 point median filter to compensate for the singular noise $\log|\epsilon_n|$, while the bottom graph shows the empirical variogram for 5-day log normalized fluctuations after applying a 3 point median filter. The dashed lines are exponential fits from which the rates of mean-reversion can be obtained. Since both variograms share a parameter, they were fitted simultaneously by using weighted least squares regression. The number of lags in the variogram model was chosen to be equal to $j = 1,275$. 
sample distribution of the estimated parameters is shown in Figure 2.

Table 1: Summary statistics of the estimates of the model variogram (36).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Mean</th>
<th>Std Dev</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>0.4643</td>
<td>0.4614</td>
<td>0.0233</td>
<td>0.4128 - 0.4998</td>
</tr>
<tr>
<td>$\nu_Y$</td>
<td>0.4752</td>
<td>0.4783</td>
<td>0.0199</td>
<td>0.4417 - 0.5174</td>
</tr>
<tr>
<td>$1/\epsilon$</td>
<td>0.2860</td>
<td>0.2933</td>
<td>0.0600</td>
<td>0.1908 - 0.4240</td>
</tr>
<tr>
<td>$\nu_Z$</td>
<td>0.3158</td>
<td>0.3091</td>
<td>0.0186</td>
<td>0.2690 - 0.3421</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.7114</td>
<td>0.7321</td>
<td>0.3104</td>
<td>0.2775 - 1.2590</td>
</tr>
</tbody>
</table>

Note: This table presents the summary statistics of the estimated parameters of the S&P 500 variogram given by Equation (26). The parameters were estimated using weighted least squares regression. We fit Equation (27) using 5-min data and Equation (28) using 5-day data simultaneously by performing the procedure described in Section 2.3. The mean, standard deviation and 95% confidence interval of each parameter were obtained using a circular block bootstrap with 10,000 bootstrap samples. LB indicates the lower bound of the confidence interval, while UB indicates the upper bound.

From Table 1, observe that the estimated fast time scale, $1/\epsilon$, is 0.286, with standard deviation equal to 0.023, meaning an average decoupling time of 3.497 days, with a 95% confidence interval [2.358, 5.241]. The estimated slow time scale, $\delta$, is 0.711, which corresponds to a mean-reversion time of 16.868 months, with a confidence interval [9.531, 43.243]. Notice that these mean-reversion times indicate that the volatility process of the S&P 500 is driven by a fast mean-reverting diffusion and a slowly varying mean-reverting process. The volatility of the invariant distribution of the OU process $Y$ was estimated to be 0.475, while the volatility of the OU process $Z$ was estimated to be 0.316.

Note the oscillatory nature of the variograms in Figure 1. This characteristic is related to intra-day variation of volatility, and this was already noticed in Fouque et al. (2000). For a discussion of this cycles from an implicit-volatility viewpoint, see Fouque et al. (2004).

3.3 Calibration Results

In this section, we report the results of the calibration of the group market parameters to the market implied volatility data on a specific day, following the procedure outlined in Section 2.4. Our goal is to demonstrate the improvement in fit of the models’s predicted
Figure 2: Bootstrap distributions of the sample parameters obtained by fitting Equations (27) and (28) to two different datasets. The red line corresponds to the estimated parameter, while the red lines correspond to the lower and upper bounds of the 95% bootstrapped confidence interval. In order to obtain the sample distributions, we use the circular block bootstrap of Politis and Romano (1992) with 10,000 bootstrap samples. The block length was chosen using the procedure suggested by Politis and White (2004).
implied volatility, given by Equation (32), to market data.

In order to do that, we consider three different models:

- **Model 1**: in this model, we assume that the volatility is driven by a single mean-reverting OU process, \( Z_t \), fluctuating on a slow time scale;

- **Model 2**: at the opposite extreme, in this model we assume that the volatility process is only driven by a fast mean-reverting OU process, \( Y_t \);

- **Model 3**: in this model, we assume that the volatility is driven by two different factors, one fluctuating on a fast time scale, \( Y_t \), and another fluctuating on a longer time scale, \( Z_t \).

We first look at the performance of Model 1, which considers only a slow volatility factor. In Figure 3a, we show the results of the calibration using only the slow-factor approximation given by Equation (39), which corresponds to assuming \( \epsilon = 0 \). Corroborating with the findings of Fouque, Papanicolaou, Sircar, and Solna (2011), we observe that Model 1 fails to capture the range of maturities.

In Figure 3b, we look at the performance of Model 2, which uses only the fast-factor approximation given by Equation (38). Again, note that Model 2 fails to capture the range of maturities.

Finally, in fitting the Model 3, which uses the two-factor volatility approximation (32), we follow the two-step procedure of Fouque, Papanicolaou, Sircar, and Solna (2011). First, we fit the skew to obtain \( \hat{a}_i \) and \( \hat{b}_i \) for different maturities \( \tau_i \). In the second stage, these estimates are then fitted to an affine function of \( \tau \) to give estimates of \( b^*, b^δ, a^ε, \) and \( a^δ \). A plot of this second term-structure fit on November 03, 2015, is shown in Figure 4a along with the calibrated multiscale approximation (32) to all the data on November 03, 2015 in Figure 4b.

From Figure 4b, note that the ability of Model 3 to capture the range of maturities is much improved when compared to Models 1 and 2. Thus, the two-scale volatility model with its additional parameters performs better than either of the one-scale models.

Table 2 reports the estimated parameters of the correction to the Black-Scholes price for each one of the three models presented in this section. Observe that in the regime where our approximation is valid, the parameters \( a^ε, a^δ, \) and \( b^δ \) are expected to be small, while \( b^* \) is the leading-order magnitude of volatility. The first parameter, \( \sigma^* \), can be interpreted as an indicator of the “average” S&P 500 volatility level. Here, \( r \) is the risk-free rate, which we assume to be known and constant. Throughout our analysis, we use \( r = 0.34% \), which is the value of the interest rate released by the Federal Reserve (FED)
(a) \( \tau \)-adjusted implied volatility \( I - b^f \tau \).

(b) S&P 500 implied volatilities.

**Figure 3:** Data and calibrated fit on November 03, 2015. In Figure (a), we plot the maturity adjusted implied volatility as a function of the log-moneyness, LM. The plus signs are from S&P 500 data and the line \( \sigma^* + a^f(LM) \) shows the fit using the estimated parameters from only a slow factor fit using ordinary least squares regression. In figure (b), we present the S&P 500 implied volatilities as a function of the LMMR. The plus signs are from S&P 500 data, and the line \( b^* + a^e(LMMR) \) shows the result using the estimated parameters from only a fast factor fit using ordinary least squares regression. In both figures, the maturities increase going counterclockwise from the top-leftmost strand.
Figure 4: Data and calibrated fit on November 03, 2015. In Figure (a), the plus signs in the top plot are the slope coefficients $a_i$ of LMMR fitted in the first step of the regression. The solid line is the straight line $(a^t + a^b \tau)$ fitted in the second step of the regression. The bottom plot shows the corresponding picture for the skew intercepts $b_i$ fitted to the straight line $b^s + b^b \tau$. In Figure (b), the plus signs are from S&P 500 data, and the lines are the formula (32) estimated from the full fast and slow factor fit.
Figure 5: Implied volatility fit to S&P 500 index options on November 03, 2015. In this Figure, we present the fit of Equation (38) to implied volatility data. Note that this is the result of a single calibration to all maturities and not a maturity-by-maturity calibration. Each panel shows the fit for a particular time to maturity.
on November 03, 2015.

Table 2: Estimated values of the group market parameters

<table>
<thead>
<tr>
<th></th>
<th>$\sigma^*$</th>
<th>$a^*$</th>
<th>$a^\delta$</th>
<th>$b^*$</th>
<th>$b^\delta$</th>
<th>$V^\delta_0$</th>
<th>$V^\delta_1$</th>
<th>$V^\delta_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.1257</td>
<td>-0.1223</td>
<td>0.0308</td>
<td>0.0313</td>
<td>-0.0019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>0.1434</td>
<td>-0.1192</td>
<td>0.1426</td>
<td></td>
<td>-0.0004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>0.1318</td>
<td>-0.1383</td>
<td>-0.2049</td>
<td>0.1311</td>
<td>0.0274</td>
<td>0.0285</td>
<td>-0.0035</td>
<td>-0.0003</td>
</tr>
</tbody>
</table>

Note: In this table, we report the estimated parameters from the fit of implied volatilities to log-moneyness to maturity ratio for every model on November 03, 2015. For Models 1 and 2, we fit Equations (38) and (39) to data using ordinary least squares regression. The estimated parameters for Model 3 result from the fit of Equation (32) to data following the 2-step procedure suggested by Fouque, Papanicolaou, Sircar, and Solna (2011). The last three columns were obtained through the inversion of Equations (34) to (37).

In Table 3, we compare the adjustment of the three models presented in this section to the market prices using the absolute percentage error $|P^\text{observed} - P^\text{model}|/P^\text{observed}$, where $P^\text{model}$ is the price given by the model and $P^\text{observed}$ is the observed market price of the option. Since the volatility used to obtain the Black-Scholes prices differs in every model, we are not able to compare them unless we use the same volatility in the Black-Scholes formula across all the models.

In order to compare the corrected prices for all the three models, we calculate the historical volatility of the returns of the S&P 500. First, we obtain a time-series of historical volatilities on this index by using a rolling-window of 63 days (three-months) during one year. Then, we use the average of this time-series as our estimator of the historical volatility, $\bar{\sigma}$. For the time period from October 23, 2014, to October 23, 2015, the estimated historical volatility in the S&P 500 was 14%. In Panel A, we report the percentage of times in which the model in question presented the best adjustment when compared to the observed market prices. Since our model provides only corrected call prices, we use the put-call parity in order to obtain the put prices with the same strike and time to maturity. Note that, for call prices, Model 1 presents the best adjustment to data, while for put prices, Model 3 provides the best fit, followed by the Black-Scholes model.

In Panel B, we compare the adjustment of Model 1 to data. To obtain the Black-Scholes prices, we use the effective volatility $\sigma^*$ obtained from Equation (34) when we fit Equation (38) to data. Note that Model 1 presents the best adjustment in most of the time for put prices. When we look at call prices, we find that the Black-Scholes model performs better most of the time. The same result is obtained in Panel D when
we compare the adjustment of Model 3 to data, although the magnitudes are different. Finally, in Panel C we compare the adjustment of Model 2 to data. In this case, the Black-Scholes model performs better when we look at put prices. For call prices, Model 2 presents the best adjustment.

Notice that, when we compare the prices given by every model to the Black-Scholes prices, we find opposite results for call and put prices. This may be an evidence that the put-call parity does not hold, even in perfectly liquid markets as the American market.

| Table 3: Comparison of prices generated by the models and observed market prices |

<table>
<thead>
<tr>
<th>Model</th>
<th>Call Prices</th>
<th>Put Prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Model Comparison using Historical Volatility</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>36.3636</td>
<td>13.5802</td>
</tr>
<tr>
<td>Model 2</td>
<td>16.2338</td>
<td>17.9012</td>
</tr>
<tr>
<td>Model 3</td>
<td>29.8701</td>
<td>37.6543</td>
</tr>
<tr>
<td>Black-Scholes</td>
<td>17.5325</td>
<td>30.8642</td>
</tr>
<tr>
<td>Panel B: Model 1 versus Black-Scholes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>4.5455</td>
<td>96.2963</td>
</tr>
<tr>
<td>Black-Scholes</td>
<td>95.4545</td>
<td>3.7037</td>
</tr>
<tr>
<td>Panel C: Model 2 versus Black-Scholes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>56.4935</td>
<td>47.5309</td>
</tr>
<tr>
<td>Black-Scholes</td>
<td>43.5065</td>
<td>52.4691</td>
</tr>
<tr>
<td>Panel D: Model 3 versus Black-Scholes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 3</td>
<td>31.1688</td>
<td>71.6049</td>
</tr>
<tr>
<td>Black-Scholes</td>
<td>68.8312</td>
<td>28.3951</td>
</tr>
</tbody>
</table>

Note: In this table, we compare the adjustment of each model to the observed market prices of the options. In Panel A, we compute the Black-Scholes prices using the historical volatility $\tilde{\sigma} = 0.14$. From Panel B to Panel D, we compute the Black-Scholes prices using the effective volatility $\sigma^*$ estimated from every model. All the values are expressed in percentage terms.

3.4 Simulating Derivative Prices

The goal of this section is to present a numerical experiment using synthetic data (i.e., synthetic option prices) in order to show the validity of the asymptotic theory of Fouque, Papanicolaou, and Sircar (2000) under different scenarios. We test the behavior of the algorithm suggested by these authors when the estimated model is the one generating the
data and also when we try to estimate using the wrong model (wrong number of scales, for example). We will also test the performance of the algorithm under different sample sizes and parameter settings.

We will denote as $N^*$ the true number of volatility scales that generate the data, while $N$ will be the value with which we do the estimation. We will generate synthetic data under four possible parameter settings:

- **Setting A**: The first data set will be obtained by generating the option prices with the Black–Scholes formula (29), assuming a constant volatility.

- **Setting B**: The second data set will be obtained by generating the option prices assuming that Model 1 is the Data Generating Process (DGP) and $N^* = 1$.

- **Setting C**: The third data set will be obtained by generating the option prices using Model 2 as the GDP, which assumes that $N^* = 1$.

- **Setting D**: Finally, the fourth data set will be obtained by generating the options prices with Model 3 as the DGP and $N^* = 2$.

In the estimation, we will use $N = 1$ or $N = 2$, with two different sample sizes: $T = 500$ or $T = 1000$. Note that any combination of the above forms an experiment.

4 Application: Pricing an Exotic Derivative

In this section, we present an application of the model for pricing a binary or digital option with no exercise option, using the calibration obtained in Section 3.3 for options in and close to the money.

Consider a cash-or-nothing call that pays a fixed amount $Q$ on the date $T$ if $X_T > K$, and zero otherwise. Its payoff function is

$$h(x) = Q\mathbf{1}_{\{x>K\}},$$

(42)

where $\mathbf{1}_A$ is the indicator function of a set $A$.

According to the perturbation theory developed in Fouque, Papanicolaou, Sircar, and Solna (2011), the stochastic volatility-corrected price is given by

$$\tilde{P}(t, x, z) = P_0(t, x, z) + P_1(t, x, z) = Qe^{-r(T-t)}N(d_2) + P_1(t, x, z),$$

(43)
\[ P_1(t,x,z) = (T - t) \left( V_0^\delta \frac{\partial}{\partial \sigma} + V_1^\delta D_1 \left( \frac{\partial}{\partial \sigma} \right) + V_3^\epsilon D_1 D_2 \right) P_{BS}^*, \] (44)

and \( P_{BS}^* = Qe^{-r(T-t)} N(d_2). \)

Using the simulated prices from Section 3.4, the goal of this section is to obtain the prices of this exotic derivative under the four different settings considered in the last section, and then compare the models across different scenarios.

5 Concluding Remarks

The main idea of the asymptotic models of Fouque, Papanicolaou, Sircar, and Solna (2011) is to use perturbation techniques to correct constant volatility models in order to capture the effects of stochastic volatility. In this paper, we compare the performance of three different specifications of the Fouque, Papanicolaou, and Sircar (2000) stochastic volatility model for pricing European call options with the Black-Scholes model. So far, we have found no evidence that these asymptotic models perform better than the Black-Scholes model. Moreover, we find opposite results when we apply this theory to pricing call and put options, what may indicate that the put-call parity does not hold.

Following Fouque, Papanicolaou, and Sircar (2000), we derive a first-order asymptotic approximation for European options under multiscale stochastic volatility models with fast and slow factors. The price approximation was translated to an implied volatility surface approximation which is linear in log-moneyness. We show that the extracted parameters (group market parameters) are small, in accordance with the asymptotic analysis developed by these authors.

We use S&P 500 data to demonstrate that the volatility of this index is driven by two diffusions: one fast mean-reverting and one slow-varying. We find a fast time scale of 3.5 days, and a slow time scale of 16.9 months. These time-scales are statistically significant, and therefore cannot be ignored in option pricing and hedging.

In what follows, we will perform the simulation study described in Section 3.4, and compare the prices of exotic derivatives based on these simulated prices. Our goal is to demonstrate that the first-order approximation fits the data well across strikes and maturities.
References


Appendix

A  Properties of the Variogram Estimator

In this section, we will study some properties of the variogram estimator given by Equation (26). In order to do that, consider the following lemmas:

**Lemma 1.** Let $s < t$, then

$$
E[(Y_t - E[Y_t]) (Y_s - E[Y_s])] = \nu_Y^2 e^{-\alpha(t-s)} \left(1 - e^{-2\alpha s}\right), \quad (A.1)
$$

$$
E[(Z_t - E[Z_t]) (Z_s - E[Z_s])] = \nu_Z^2 e^{-\delta(t-s)} \left(1 - e^{-2\delta s}\right). \quad (A.2)
$$

**Proof.**

$$
E[(Z_t - E[Z_t]) (Z_s - E[Z_s])] = E\left\{ \left( \nu_Y \sqrt{2\alpha} \int_0^t e^{-\alpha(t-u)}dW_u^{(1)} \right) \left( \nu_Y \sqrt{2\alpha} \int_0^s e^{-\alpha(s-u)}dW_u^{(1)} \right) \right\}
$$

$$
= 2\nu_Y^2 e^{-\alpha(t+s)} E\left\{ \left( \int_0^t e^{\alpha u}dW_u^{(1)} \right) \left( \int_0^s e^{\alpha u}dW_u^{(1)} \right) \right\}
$$

$$
= 2\nu_Y^2 e^{-\alpha(t+s)} E\left\{ \left( \int_0^t e^{\alpha u}dW_u^{(1)} \right) \left( \int_0^s e^{\alpha u}dW_u^{(1)} \right) \right\} +
$$

$$
+ 2\nu_Y^2 e^{-\alpha(t+s)} E\left\{ \left( \int_t^s e^{\alpha u}dW_u^{(1)} \right) \left( \int_0^s e^{\alpha u}dW_u^{(1)} \right) \right\}
$$

$$
= 2\nu_Y^2 e^{-\alpha(t+s)} E\left\{ \left( \int_0^t e^{\alpha u}dW_u^{(1)} \right)^2 \right\}
$$

$$
= 2\nu_Y^2 e^{-\alpha(t+s)} E\left\{ \left( \int_0^t e^{2\alpha u} dt \right)^2 \right\}
$$

$$
= \nu_Y^2 e^{-\alpha(t+s)(e^{2\alpha s} - 1)}
$$

$$
= \nu_Y^2 e^{-\delta(t-s)} \left(1 - e^{-2\delta s}\right).
$$

Similarly, we can get (A.1). □

By using the previous Lemma, we can conclude that

$$
E\left[ \log f_1(\bar{Y}_j) \log f_1(\bar{Y}_0) \right] \approx \nu_Y^2 e^{-\alpha j \Delta t},
$$

$$
E\left[ \log f_1(\bar{Y})^2 \right] \approx \nu_Y^2. \quad (A.3)
$$
Similarly, we have
\[
\mathbb{E}\left[ \log f_1(\tilde{Z}) \log f_1(\tilde{Z}_0) \right] \approx \nu_2^2 e^{-\alpha j \Delta t},
\]
\[
\mathbb{E}\left[ \log f_1(\bar{Z})^2 \right] \approx \nu_2^2,
\] (A.4)
where \( \bar{Y} \) and \( \bar{Z} \) denote the asymptotic distributions of the OU processes given by (1).

**Lemma 2.** For the model presented in this paper, if \( s < t \), then
\[
\mathbb{E}\left[ (Y_t - \mathbb{E}[Y_t]) (Z_s - \mathbb{E}[Z_s]) \right] = \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \tilde{\rho}_{1,2} \nu_Y \nu_Z \left( e^{-\alpha (t-s)} - e^{-\alpha s} e^{-\delta t} \right). \tag{A.5}
\]

Otherwise, if \( s > t \),
\[
\mathbb{E}\left[ (Y_t - \mathbb{E}[Y_t]) (Z_s - \mathbb{E}[Z_s]) \right] = \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \tilde{\rho}_{1,2} \nu_Y \nu_Z \left( e^{-\delta (s-t)} - e^{-\alpha s} e^{-\delta t} \right). \tag{A.6}
\]

**Proof.** Suppose that \( s < t \), then
\[
\mathbb{E}\left[ (Y_t - \mathbb{E}[Y_t]) (Z_s - \mathbb{E}[Z_s]) \right] = \mathbb{E}\left\{ \nu_Y \sqrt{2\alpha} \int_0^t e^{-\alpha (t-u)} d\tilde{W}_u^{(1)} \right\} \left( \nu_Z \sqrt{2\delta} \int_0^s e^{-\delta (s-u)} d\tilde{W}_u^{(1)} \right) + \mathbb{E}\left\{ \nu_Y \sqrt{2\alpha} \int_0^t e^{-\alpha (t-u)} d\tilde{W}_u^{(1)} \right\} \left( \nu_Z \sqrt{2\delta} \int_s^t e^{-\delta (s-u)} d\tilde{W}_u^{(2)} \right)
\]
\[
= \mathbb{E}\left\{ \nu_Y \sqrt{2\alpha} \int_0^s e^{-\alpha (t-u)} d\tilde{W}_u^{(1)} \right\} \left( \nu_Z \sqrt{2\delta} \int_0^s e^{-\delta (s-u)} d\tilde{W}_u^{(1)} \right) + \mathbb{E}\left\{ \nu_Y \sqrt{2\alpha} \int_0^s e^{-\alpha (t-u)} d\tilde{W}_u^{(1)} \right\} \left( \nu_Z \sqrt{2\delta} \int_s^t e^{-\delta (s-u)} d\tilde{W}_u^{(2)} \right)
\]
\[
= \sqrt{2\alpha} \nu_Y \sqrt{2\delta} \nu_Z \tilde{\rho}_{1,2} e^{-\alpha t} e^{-\delta s} \int_0^s e^{(\alpha + \delta) u} du
\]
\[
= \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \nu_Y \nu_Z \tilde{\rho}_{1,2} e^{-\alpha t} e^{-\delta s} \left( e^{(\alpha + \delta) s} - 1 \right)
\]
\[
= \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \nu_Y \nu_Z \tilde{\rho}_{1,2} \left( e^{-\delta (s-t)} - e^{-\alpha s} e^{-\delta t} \right).
\]

By following Fouque, Papanicolaou, Sircar, and Solna (2011) and using the previous Lemma, we get
\[
\mathbb{E}\left[ \log f_1(\bar{Y}) \log f_2(\bar{Z}) \right] \approx \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \tilde{\rho}_{1,2} \nu_Y \nu_Z, \tag{A.7}
\]
\[
\mathbb{E}\left[ \log f_1(\bar{Y}_j) \log f_2(\bar{Z}_0) \right] \approx \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \tilde{\rho}_{1,2} \nu_Y \nu_Z e^{-\alpha j \Delta t}, \tag{A.8}
\]
\[
\mathbb{E}\left[ \log f_1(\bar{Y}_0) \log f_2(\bar{Z}_j) \right] \approx \frac{2 \sqrt{\alpha} \sqrt{\delta}}{\alpha + \delta} \tilde{\rho}_{1,2} \nu_Y \nu_Z e^{-\delta j \Delta t}. \tag{A.9}
\]
Proposition 1.

\[\begin{align*}
\mathbb{E}\{ (F_{n+j} - F_n)^2 \} & \approx 2c_Y^2 + 2\nu_Y^2 (1 - e^{-\alpha j \Delta t}) + 2\nu_Z^2 (1 - e^{-\delta j \Delta t}) + \\
& \quad + \frac{4\sqrt{\alpha \sqrt{\delta}}}{\alpha + \delta} \hat{\rho}_{1,2} \nu_Y \nu_Z (2 - e^{-\alpha j \Delta t} - e^{-\delta j \Delta t}).
\end{align*}\]  
(A.10)

Proof.

\[\begin{align*}
\mathbb{E}\{ (F_{n+j} - F_n)^2 \} &= \mathbb{E}\{ (F_j - F_0)^2 \} \\
&= \mathbb{E}\left\{ \left( \log f_1(\bar{Y}_j) - \log f_1(\bar{Y}_0) \right) + \left( \log f_2(\bar{Z}_j) - \log f_2(\bar{Z}_0) \right) \right\} + \\
&\quad + \mathbb{E}\left\{ \left( \log |\epsilon_j| - \log |\epsilon_0| \right) \right\} \\
&= \mathbb{E}\left\{ \left( \log f_1(\bar{Y}_j) - \log f_1(\bar{Y}_0) \right) \right\} + \mathbb{E}\left\{ \left( \log f_2(\bar{Z}_j) - \log f_2(\bar{Z}_0) \right) \right\} + \\
&\quad + \mathbb{E}\left\{ \left( \log |\epsilon_j| - \log |\epsilon_0| \right) \right\} + \\
&\quad + 2\mathbb{E}\left\{ \left( \log f_1(\bar{Y}_j) - \log f_1(\bar{Y}_0) \right) \left( \log f_2(\bar{Z}_j) - \log f_2(\bar{Z}_0) \right) \right\} \\
&= 2\mathbb{E}\left\{ \log f_1(\bar{Y}_j)^2 \right\} - 2\mathbb{E}\left\{ \log f_1(\bar{Y}_j) \log f_1(\bar{Y}_0) \right\} + 2\mathbb{E}\left\{ \log f_2(\bar{Z})^2 \right\} \\
&\quad - 2\mathbb{E}\left\{ \log f_2(\bar{Z}_j) \log f_2(\bar{Z}_0) \right\} + 2\text{Var}(\epsilon) + 4\mathbb{E}\left\{ \log f_1(\bar{Y}_0) \log f_2(\bar{Z}) \right\} \\
&\quad - 2\mathbb{E}\left\{ \log f_1(\bar{Y}_j) \log f_2(\bar{Z}_0) \right\} - 2\mathbb{E}\left\{ \log f_1(\bar{Y}_j) \log f_2(\bar{Z}) \right\} \\
&\approx 2c_Y^2 + 2\nu_Y^2 (1 - e^{-\alpha j \Delta t}) + 2\nu_Z^2 (1 - e^{-\delta j \Delta t}) + 8\frac{\sqrt{\alpha \sqrt{\delta}}}{\alpha + \delta} \hat{\rho}_{1,2} \nu_Y \nu_Z \\
&\quad - \frac{4\sqrt{\alpha \sqrt{\delta}}}{\alpha + \delta} \hat{\rho}_{1,2} \nu_Y \nu_Z e^{-\alpha j \Delta t} - \frac{4\sqrt{\alpha \sqrt{\delta}}}{\alpha + \delta} \hat{\rho}_{1,2} \nu_Y \nu_Z e^{-\delta j \Delta t} \\
&\approx 2c_Y^2 + 2\nu_Y^2 (1 - e^{-\alpha j \Delta t}) + 2\nu_Z^2 (1 - e^{-\delta j \Delta t}) + \\
&\quad + \frac{4\sqrt{\alpha \sqrt{\delta}}}{\alpha + \delta} \hat{\rho}_{1,2} \nu_Y \nu_Z (2 - e^{-\alpha j \Delta t} - e^{-\delta j \Delta t}). \quad \Box
\end{align*}\]

In the previous proof, we used Equations (A.3), (A.4), and Equations (A.7) to (A.9). Since we assume that \( \delta \ll 1/\epsilon \), by Equation (A.10) we have that

\[
\frac{\sqrt{\alpha \sqrt{\delta}}}{\alpha + \delta} \approx \frac{\sqrt{\alpha \sqrt{\delta}}}{\alpha} \approx \frac{\sqrt{\delta}}{\sqrt{\alpha}} \approx 0. \quad \text{(A.11)}
\]

Therefore, the previous equation allows us to estimate the variogram model (24) by using two reduced models, one for each time scale.
B  First-Order Perturbation Theory

In this section, we derive the asymptotic approximation for options prices of Fouque, Papanicolaou, and Sircar (2000). In order to derive the corrected prices, as the terms associated with $\delta$ are small when $\delta \to 0$ and give rise to a regular perturbation problem, we first expand $P^{\epsilon,\delta}$ in powers of $\sqrt{\delta}$:

$$P^{\epsilon,\delta} = P_0^\epsilon + \sqrt{\delta} P_1^\epsilon + \delta P_2^\epsilon + \ldots.$$  (B.1)

We insert the above expansion in equation (9) (both the PDE and the terminal condition), grouping generators in term of powers of $\delta$ in the form

$$
\left(\frac{1}{\epsilon} L_0 + \frac{1}{\sqrt{\epsilon}} L_1 + \mathcal{L}_2\right) P_0^\epsilon + \sqrt{\delta} \left\{ \left(\frac{1}{\epsilon} L_0 + \frac{1}{\sqrt{\epsilon}} L_1 + \mathcal{L}_2\right) P_1^\epsilon + \left(\mathcal{M}_1 + \frac{1}{\sqrt{\epsilon}} \mathcal{M}_3\right) P_0^\epsilon \right\} + \ldots = 0.
$$  (B.2)

Write $\mathcal{L}^{\epsilon,\delta}$ and $P^{\epsilon,\delta}$ as

$$\mathcal{L}^{\epsilon,\delta} = \mathcal{L}^\epsilon + \sqrt{\delta} \mathcal{M}^\epsilon + \delta \mathcal{M}_2,$$  (B.3)

$$P^{\epsilon,\delta} = \sum_{j \geq 0} (\sqrt{\delta})^j P_j^\epsilon,$$  (B.4)

where

$$\mathcal{L}^\epsilon = \frac{1}{\epsilon} L_0 + \frac{1}{\sqrt{\epsilon}} L_1 + \mathcal{L}_2,$$  (B.5)

$$\mathcal{M}^\epsilon = \frac{1}{\sqrt{\epsilon}} \mathcal{M}_3 + \mathcal{M}_1,$$  (B.6)

$$P_j^\epsilon = \sum_{i \geq 0} (\sqrt{\epsilon})^i P_{i,j}^\epsilon.$$  (B.7)

Inserting (B.1) in (9) and collecting term of like-powers of $\sqrt{\delta}$, we find that the lowest order equations of the regular perturbation expansion are

$$\mathcal{O}(1) : 0 = \mathcal{L}^\epsilon P_0^\epsilon,$$  (B.8)

$$\mathcal{O}(\sqrt{\delta}) : 0 = \mathcal{L}^\epsilon P_1^\epsilon + \mathcal{M}^\epsilon P_0^\epsilon,$$  (B.9)

$$\mathcal{O}(\delta) : 0 = \mathcal{L}^\epsilon P_2^\epsilon + \mathcal{M}^\epsilon P_1^\epsilon + \mathcal{M}_2 P_0^\epsilon.$$  (B.10)

Therefore, we can choose $P_0^\epsilon$ and $P_1^\epsilon$ to satisfy the following PDEs with terminal con-
ditions

\[
\begin{align*}
\left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) P_0^\epsilon &= 0, \\
P_0^\epsilon(T, x, y, z) &= h(x). \\
\left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) P_1^\epsilon &= -\left( \mathcal{M}_1 + \frac{1}{\sqrt{\epsilon}} \mathcal{M}_3 \right) P_0^\epsilon, \\
P_1^\epsilon(T, x, y, z) &= 0,
\end{align*}
\]

(B.11) (B.12)

where the terminal payoff is assigned to the independent term \( P_0^\epsilon \). Next, we expand \( P_0^\epsilon \) and \( P_1^\epsilon \) in powers of \( \sqrt{\epsilon} \).

**B.0.1 Zeroth-Order Approximation \( P_0 \)**

From a fast factor expansion of equation (9), we will find the zeroth order term \( P_{0,0} \), in our approximation (18).

Consider the expansion of \( P_0^\epsilon \) in powers of \( \sqrt{\epsilon} \):

\[
P_0^\epsilon = P_0 + \sqrt{\epsilon} P_{1,0} + \epsilon P_{2,0} + \epsilon^{3/2} P_{3,0} + \ldots
\]

(B.13)

By inserting (B.13) in (9) and rearranging terms, we get

\[
\begin{align*}
\frac{1}{\epsilon} \mathcal{L}_0 P_0 + \frac{1}{\sqrt{\epsilon}} (\mathcal{L}_0 P_{1,0} + \mathcal{L}_1 P_0) + (\mathcal{L}_0 P_{2,0} + \mathcal{L}_1 P_{1,0} + \mathcal{L}_2 P_0) + \\
+ \sqrt{\epsilon} (\mathcal{L}_0 P_{3,0} + \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0}) + \ldots &= 0.
\end{align*}
\]

(B.14)

By collecting term of like-powers of \( \sqrt{\epsilon} \), we get

\[
\begin{align*}
\mathcal{O}\left( \frac{1}{\epsilon} \right) & : \quad \mathcal{L}_0 P_0 = 0, \\
\mathcal{O}\left( \frac{1}{\sqrt{\epsilon}} \right) & : \quad \mathcal{L}_0 P_{1,0} + \mathcal{L}_1 P_0 = 0, \\
\mathcal{O}(1) & : \quad \mathcal{L}_0 P_{2,0} + \mathcal{L}_1 P_{1,0} + \mathcal{L}_2 P_0 = 0, \\
\mathcal{O}\left( \sqrt{\epsilon} \right) & : \quad \mathcal{L}_0 P_{3,0} + \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0} = 0.
\end{align*}
\]

(B.15) (B.16) (B.17) (B.18)

We see from (11) and (12) that all terms in \( \mathcal{L}_0 \) and \( \mathcal{L}_1 \) take derivatives with respect to \( y \). Thus, if we choose \( P_{0,0} \) and \( P_{1,0} \) to be independent of \( y \), the above equations will
automatically be satisfied. Hence, we seek solutions of the form:

\[ P_{0,0} = P_{0,0}(t, x, z), \quad (B.19) \]
\[ P_{1,0} = P_{1,0}(t, x, z). \quad (B.20) \]

Note that Equation (B.17) is a Poisson equation for \( P_{2,0} \) with respect to \( y \). For there to be a solution, \( \mathcal{L}_2 P_0 \) must satisfy the solvability condition:

\[
\langle \mathcal{L}_2 P_0 \rangle = \frac{\partial P_0}{\partial t} + \frac{1}{2} \sigma^2 x^2 \frac{\partial^2 P_0}{\partial x^2} + r \left( x \frac{\partial P_0}{\partial x} - P_0 \right) = 0, \quad (B.21)
\]

which is the Black-Scholes pricing operator \( \mathcal{L}(\bar{\sigma}(z)) \) with effective averaged volatility \( \bar{\sigma}(z) \):

\[
\bar{\sigma}(z) = \langle f^2(\cdot, z) \rangle = \int f^2(y, z) \Phi(dy), \quad (B.22)
\]

where the level \( z \) of the slow factor appears as a parameter. Therefore, the leading-order term \( P_0(t, x, z) = P_{BS}(t, x; \bar{\sigma}(z)) \) is the Black-Scholes price at volatility \( \bar{\sigma}(z) \), which satisfies the following PDE and terminal condition:

\[
\begin{aligned}
\mathcal{L}_{BS}(\bar{\sigma}(z)) P_{BS} &= 0, \\
P_{BS}(T, x; \bar{\sigma}(z)) &= h(x).
\end{aligned} \quad (B.23)
\]

**B.0.2 Fast Time Scale Correction \( P_{1,0}^\epsilon \)**

The order \( \sqrt{\epsilon} \) in (B.18) give the following Poisson equation for \( P_{3,0} \):

\[
\mathcal{L}_0 P_{3,0} + \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0} = 0. \quad (B.24)
\]

The solvability condition for this equation is given by

\[
\langle \mathcal{L}_2 P_{1,0} + \mathcal{L}_1 P_{2,0} \rangle = 0. \quad (B.25)
\]

\(^5\)Note that Equations (B.17) and (B.18) are Poisson equations of the form

\[ 0 = \mathcal{L}_0 + \chi, \]

By the Fredholm alternative, the equation above, which is a linear ODE in \( y \), admits a solution \( P \) in \( L^2(\Phi) \) only if the following solvability, or centering, condition holds:

\[
\langle \chi \rangle \equiv \int \chi(y) \Phi(dy) = 0,
\]

where \( \Phi \) is the invariant distribution of \( Y \).
Let $D_k$ be defined as
\[ D_k = x^k \frac{\partial^k}{\partial x^k}, \] 
then, introducing a solution $\phi(y, z)$ to the Poisson equation $L_0 \phi = f^2 - \langle f^2 \rangle$, we deduce the following expression for $P_{2,0}$:
\[ P_{2,0}(t, x, y, z) = -\frac{1}{2} \phi(y, z) D_2 P_{0,0}(t, x, y, z) + c(t, x, z), \] 
up to an additive function $c(t, x, z)$ which does not depend on $y$. Then, by using Equations (11) and (12) and the fact that $P_{1,0}$ does not depend on $y$, we find that $P_1^\epsilon$ satisfied the following PDE:
\[ \begin{cases} 
L_{BS}(\sigma(z)) P_1^\epsilon = A^\epsilon P_{BS}; \\
P_1^\epsilon(T, x; z) = 0; 
\end{cases} \] 
where the $z$-dependent operator $A^\epsilon = \sqrt{\epsilon} (L_1 L_0^{-1} (L_2 - \langle L_2 \rangle))$. Using Equation (B.17), we can rewrite $A^\epsilon$ as
\[ A^\epsilon = -V_3^\epsilon(z) D_1 D_2 - V_2^\epsilon(z) D_2, \] 
where the two group parameters $V_2^\epsilon(z)$ and $V_3^\epsilon(z)$ are given by
\[ \begin{align*}
V_2^\epsilon(z) &= -\frac{\rho_1 \sqrt{\epsilon}}{2} \langle \beta f(\cdot, z) \frac{\partial \phi}{\partial y}(\cdot, z) \rangle, \\
V_3^\epsilon(z) &= -\frac{\sqrt{\epsilon}}{2} \langle \beta \Lambda_1(\cdot, z) \frac{\partial \phi}{\partial y}(\cdot, z) \rangle.
\end{align*} \] 

With the previous notation, we have
\[ L_{BS}(\sigma(z)) = \frac{\partial}{\partial t} + \frac{1}{2} \sigma(z)^2 D_2 + r(D_1 - \cdot). \] 
Therefore, the first-order fast scale correction term $P_{1,0}(t, x, z)$ is given in terms of $P_{BS}$ by
\[ P_{1,0}^\epsilon(t, x, z) = -(T - t) A^\epsilon P_{BS}(t, x; \sigma(z)). \]

**B.0.3 Slow Time Scale Correction $P_{0,1}^\delta$**

Proceeding as in the last section, we insert expansions (B.7) into (B.9) and collect terms of like-powers of $\sqrt{\epsilon}$. We find that the lowest order equations of the singular perturbation equation are given by
\[ \mathcal{O} \left( \sqrt{\frac{\delta}{\epsilon}} \right) : \quad \mathcal{L}_0 P_{0,1} = 0, \]  
\[ \mathcal{O} \left( \sqrt{\frac{\delta}{\sqrt{\epsilon}}} \right) : \quad \mathcal{L}_0 P_{1,1} + \mathcal{L}_1 P_{0,1} + \mathcal{M}_3 P_{0,0} = 0, \]  
\[ \mathcal{O} \left( \sqrt{\delta} \right) : \quad \mathcal{L}_0 P_{2,1} + \mathcal{L}_1 P_{1,1} + \mathcal{L}_2 P_{0,1} + \mathcal{M}_3 P_{1,0} + \mathcal{M}_1 P_{0,0} = 0, \]  
\[ \mathcal{O} \left( \sqrt{\delta} \sqrt{\epsilon} \right) : \quad \mathcal{L}_0 P_{3,1} + \mathcal{L}_1 P_{2,1} + \mathcal{L}_2 P_{1,1} + \mathcal{M}_3 P_{2,0} + \mathcal{M}_1 P_{1,0} = 0. \]  

Note that \( \mathcal{M}_3 P_{0,0} = 0 \) since \( \mathcal{M}_3 \) take derivatives in \( y \) and \( P_{0,0} \) is independent of \( y \). As all terms in \( \mathcal{L}_0 \) and \( \mathcal{L}_1 \) take derivatives in \( y \), we seek solutions \( P_{0,1} \) and \( P_{1,1} \) independent of \( y \).

Equations (B.36) and (B.37) are Poisson equations in \( y \) with respect to \( P_{2,1} \) and \( P_{3,1} \), respectively. Using the centering condition yields

\[ \mathcal{O} \left( \sqrt{\delta} \right) : \quad \langle \mathcal{L}_2 \rangle P_{0,1} + \langle \mathcal{M}_1 \rangle P_{0,0} = 0, \]  
\[ \mathcal{O} \left( \sqrt{\delta} \sqrt{\epsilon} \right) : \quad \langle \mathcal{L}_1 P_{2,1} \rangle + \langle \mathcal{L}_2 \rangle P_{1,1} + \langle \mathcal{M}_3 P_{2,0} \rangle + \langle \mathcal{M}_1 \rangle P_{1,0} = 0. \]

Again, since \( P_{0,0} \) and \( P_{0,1} \) do not depend on \( y \), condition (B.38), after multiplying by \( \sqrt{\delta} \) is given by

\[ \langle \mathcal{L}_2 \rangle P_{0,1}^\delta = -\sqrt{\delta} \langle \mathcal{M}_1 \rangle P_{0,0}. \]

This is an inhomogeneous Black-Scholes partial differential equation for \( P_{0,1}^\delta \). We can rewrite the expression above as

\[ \sqrt{\delta} \langle \mathcal{M}_1 \rangle P_{0,0} \equiv 2 \mathcal{A}^\delta P_{BS}, \]  
where we define the operator

\[ \mathcal{A}^\delta = V_0^\delta(z) \frac{\partial}{\partial \sigma} + V_1^\delta(z) D_1 \frac{\partial}{\partial \sigma}, \]

and the group parameters \( (V_0^\delta(z), V_1^\delta(z)) \) by

\[ V_0^\delta(z) = -\frac{g(z) \sqrt{\delta}}{2} \langle \Lambda_2(\cdot, z) \rangle \bar{\sigma}'(z), \]  
\[ V_1^\delta(z) = -\frac{\rho \sigma g(z) \sqrt{\delta}}{2} \langle f(\cdot, z) \rangle \bar{\sigma}'(z). \]
Therefore, the first order scale correction $P_{0,1}^\delta$ is the unique classical solution to the problem

\[
\begin{align*}
\mathcal{L}_{BS}(\bar{\sigma}(z))P_{0,1}^\delta &= -2A^\delta P_{BS}, \\
P_{0,1}(T, x, z) &= 0.
\end{align*}
\tag{B.44}
\]

The first-order slow scale correction $P_{0,1}^\delta(t, x, z)$ is given in terms of $P_{BS}(t, x; \bar{\sigma}(z))$ by

\[
P_{0,1}^\delta = (T-t)A P_{BS}. \tag{B.45}
\]

By replacing Equations (B.45) and (B.33) into (18), we get the following first-order approximation for a European-style option price:

\[
\hat{P}_{r,\delta} = P_{BS} + (T-t) \left( V_{0}^\delta(z) \frac{\partial}{\partial \sigma} + V_{1}^\delta D_1 \left( \frac{\partial}{\partial \sigma} \right) \right) P_{BS} + \\
+ (T-t) \left( V_{2}^\delta(z) D_2 + V_{3}^\delta(z) D_1 D_2 \right) P_{BS}.
\tag{B.46}
\]