Rev Deriv Res (2007) 10:87–150
DOI 10.1007/s11147-007-9014-6

A new approach for option pricing under stochastic volatility

Peter Carr · **Jian Sun**

Published online: 20 February 2008

Abstract We develop a new approach for pricing European-style contingent claims written on the time T spot price of an underlying asset whose volatility is stochastic. Like most of the stochastic volatility literature, we assume continuous dynamics for the price of the underlying asset. In contrast to most of the stochastic volatility literature, we do not directly model the dynamics of the instantaneous volatility. Instead, taking advantage of the recent rise of the variance swap market, we directly assume continuous dynamics for the time T variance swap rate. The initial value of this variance swap rate can either be directly observed, or inferred from option prices. We make no assumption concerning the real world drift of this process. We assume that the ratio of the volatility of the variance swap rate to the instantaneous volatility of the underlying asset just depends on the variance swap rate and on the variance swap maturity. Since this ratio is assumed to be independent of calendar time, we term this key assumption the stationary volatility ratio hypothesis (SVRH). The instantaneous volatility of the futures follows an unspecified stochastic process, so both the underlying futures price and the variance swap rate have unspecified stochastic volatility. Despite this, we show that the payoff to a path-independent contingent claim can be perfectly replicated by dynamic trading in futures contracts and variance swaps of the same maturity. As a result, the contingent claim is uniquely valued relative to its underlying's futures price and the assumed observable variance swap rate. In contrast to standard models of stochastic volatility, our approach does not require specifying the market price of volatility risk or observing the initial level of instantaneous volatility.

P. Carr (✉)
Bloomberg LP, 731 Lexington Avenue, New York, NY 10022, USA
e-mail: pcarr4@bloomberg.com

J. Sun
XE Capital Management, 24 West 40th Street, New York, NY 10018, USA
e-mail: jsun@xecapital.com

As a consequence of our SVRH, the partial differential equation (PDE) governing the arbitrage-free value of the contingent claim just depends on two state variables rather than the usual three. We then focus on the consistency of our SVRH with the standard assumption that the risk-neutral process for the instantaneous variance is a diffusion whose coefficients are independent of the variance swap maturity. We show that the combination of this maturity independent diffusion hypothesis (MIDH) and our SVRH implies a very special form of the risk-neutral diffusion process for the instantaneous variance. Fortunately, this process is tractable, well-behaved, and enjoys empirical support. Finally, we show that our model can also be used to robustly price and hedge volatility derivatives.

Keywords Option pricing · Stochastic volatility

1 Introduction

In this article, we consider the standard problem of valuing and hedging a contingent claim written on the price at expiry of some underlying asset. In contrast to the standard model of Black and Scholes (1973) and Merton (1973), we assume that both the spot price and the instantaneous volatility of the claim's underlying asset are stochastic and imperfectly correlated. The standard approach to derivative security valuation under stochastic volatility specifies the statistical dynamics and derives the risk-neutral dynamics of both quantities. As is well known, this approach requires specifying the market price of volatility risk. This specification is fraught with difficulty since this market price is not directly observable. Even if one manages to achieve the correct parametrization of the market price of volatility risk, the identification of these parameters and the initial instantaneous volatility from option prices can be problematic in practice.

Fortunately, there is an alternative approach which bypasses the need to specify the dynamics of the market price of volatility risk. It also bypasses the need to observe or infer the instantaneous volatility. The approach is to model the statistical dynamics of some process which is a known function of option prices. As the instantaneous volatility of the underlying asset is intrinsic to option valuation, this function should have the property that this instantaneous volatility can be expressed in terms of this process. Since the risk-neutral relative drift of an option price is just the riskfree rate, the risk-neutral drift of the process can be calculated through Itô's formula. If the statistical process describing the function of options prices is assumed to be continuous over time, then all that remains is to model the statistical volatility of the process.

This approach was pioneered in Dupire (1992). Inspired by the pioneering contribution of Heath et al. (1992) to the analysis of interest rate derivatives, the function of the option prices which Dupire chose was the entire term structure of forward variance swap rates. Assuming only positivity and continuity of the underlying asset price, Dupire showed that a forward variance swap rate can be determined from the cost of forming a particular static position in options involving a continuum of strike prices. As a result, the risk-neutral drift of the forward variance swap rate is zero. Once one specifies the volatility of all forward variance swap rates, one also determines the risk-neutral dynamics of the instantaneous variance of the underlying.

Unfortunately, the determination of the initial curve of forward variance swap rates can be tricky in practice due to the discreteness of strikes and maturities in options markets. Now that variance swaps trade outright, one can overcome the discrete strikes issue by direct observation of variance swap rates. However, the discreteness of maturities in the relatively nascent variance swap market still makes observation of the initial continuum of variance swap rates tricky in practice.

To circumvent this problem, Duanmu (2004) proposes modelling the spot variance swap rate of a single maturity.[1] He assumes a particular diffusion process for the variance swap rate and shows that the payoff to volatility derivatives of the same maturity can be replicated by dynamic trading in variance swaps.

Like Duanmu, Potter (2004) also proposes completing markets by dynamic trading in variance swaps of a single maturity. Like us, Potter examines the implications of this assumption for the valuation of contingent claims on price as well as for volatility derivatives. To value these claims, he assumes that the instantaneous variance of the underlying asset is an affine function of the variance swap rate. He then shows that this assumption is a consequence of the dynamics assumed in several popular stochastic volatility models. He also analyzes Duanmu's model and shows that it is a special case of his framework in which the instantaneous variance of the underlying asset is just the variance swap rate.

Our analysis is similar to that of Duanmu and Potter in that we model the dynamics of a variance swap rate of a single maturity. Like them and Dupire, we do not have to specify the market price of volatility risk. The major difference between our work and all previous work is that we impose a special structure on the assumed dynamics of the variance swap rate. In particular, we assume that the ratio of the volatility of the variance swap rate to the instantaneous volatility of the underlying asset just depends on the variance swap rate and the variance swap maturity. Since this ratio is assumed to be independent of calendar time, we term this key assumption the stationary volatility ratio hypothesis (SVRH).

The instantaneous volatility of the futures follows an unspecified stochastic process, so both the underlying futures price and the variance swap rate have unspecified stochastic volatility. Despite this, we show that the payoff to a path-independent contingent claim can be perfectly replicated by dynamic trading in futures contracts and variance swaps of the same maturity. As a result, no arbitrage implies that the contingent claim is uniquely valued relative to its underlying's futures price and the assumed observable variance swap rate. Under the SVRH, parsimony is achieved in that our valuation PDE depends only on two independent variables rather than the usual three. This speeds up numerical solution by an order of magnitude. The PDE to be numerically solved is a second order linear elliptic PDE and hence standard solution methods are available.

A standard assumption in the stochastic volalatility literature is that the statistical process for instantaneous variance and the market price of variance risk are such that

[1] The spot variance swap rate corresponds to the area under Dupire's forward variance rate curve between the valuation time and the option maturity.

the risk-neutral process for instantaneous variance is a diffusion. Assuming that a money market account acts as numeraire, the coefficients of this risk-neutral diffusion process are independent of the variance swap maturity. In order to determine whether our approach can be rendered consistent with this now standard approach, we investigate the implications of this maturity independent diffusion hypothesis (MIDH) and our SVRH for the risk-neutral process followed by the instantaneous variance. We show that MIDH and SVRH together dictate that the risk-neutral drift of the instantaneous variance must be a quadratic function of the instantaneous variance (with zero intercept). Furthermore, the normal volatility of the instantaneous variance must be proportional to its level raised to the power 3/2. Fortunately, we document that this quadratic drift 3/2 process is tractable, well behaved, and enjoys a surprising amount of empirical support.

Although the MIDH and our SVRH determine the form of the risk-neutral drift of the instantaneous variance, they do not specify the statistical drift of this process. As a result, our pricing model places no restrictions on the market price of volatility risk. This is a big advantage of our approach over standard stochastic volatility models which require that the market price of volatility risk be specified in order to uniquely price contingent claims.

The quadratic drift 3/2 process for instantaneous variance has many desirable properties. For example, the instantaneous variance is always positive and never explodes. Also, the process is mean-reverting, where the speed of mean-reversion is proportional to the level of the process. The process yields closed form formulas for the joint Fourier Laplace transform of returns and their quadratic variation. As a result, many derivative securities on price and/or realized volatility can be valued. In particular, standard options on price can be valued via (fast) Fourier inversion. The quadratic drift 3/2 process also yields closed form formulas for the variance swap rate and its volatility. Since the general formulas for these quantities are complicated, we focus on the proportional drift subcase, which has very simple formulas for the variance swap rate and its volatility. Although this proportional drift risk-neutral process does not mean-revert to a positive level, we show that its statistical counterpart can have this property, where the speed of mean-reversion is proportional to the level.

Finally, we examine the pricing of volatility derivatives in our model. Like contingent claims on price, these derivatives can be priced without specifying the market price of volatility risk or the initial level of the instantaneous variance. In contrast to contingent claims on price, one need only dynamically trade variance swaps in order to replicate the payoff of these claims. As a result, the price dynamics for the underlying asset need not be specified.

An overview of this paper is as follows. The next section lays out our notations and assumptions including our critical SVRH. The following section shows that a European-style payoff for a path-independent claim can be replicated by dynamic trading in futures and variance swaps of the same maturity. It also derives a fundamental elliptic PDE which governs the values of all European-style claims in our model. The subsequent section deals with the issue that variance swaps may not trade by showing that both the level of the variance swap rate and the gain on a variance swap position can be accessed through options. The next section addresses the issue of calibrating the model to market options prices. The subsequent section shows how Monte

Carlo simulation can be used to efficiently determine both values and Greeks. The following section examines the implications of further assuming that the risk-neutral process for instantaneous variance is a diffusion whose coefficients are independent of the variance swap maturity. We show that this maturity independent diffusion hypothesis (MIDH) and our SVRH imply a condition on the risk-neutral drift and diffusion coefficients of the instantaneous variance. The next section shows that this condition implies that the risk-neutral process for the instantaneous variance is a quadratic drift 3/2 process. The following section focusses on a subcase that yields simple formulas for the variance swap rate and its volatility. The penultimate section extends our results to volatility derivatives. The final section summarizes the paper and makes suggestions for future research.

2 Assumptions and notation

Our objective is to price a path-independent claim of a fixed maturity T. To accomplish this objective, we assume that over this period, the underlying asset trades continuously in a frictionless market. For simplicity, we assume zero interest rates over this period. When one introduces positive interest rates, one needs to model the forward or futures price of the underlying asset to achieve our results and hence we will model one of these.

Let F_t be the time t futures price for maturity T, where we assume continuous marking-to-market for simplicity. Let $\mathbb{P}$ denote the real world probability measure, also known as the statistical or physical measure. Under this measure, we assume that the underlying asset's futures price process $\{F_t, t \in [0, T]\}$ is positive and continuous over time. The martingale representation theorem then implies that there exists stochastic processes μ and σ such that:

$$\frac{dF_t}{F_t} = \mu_t dt + \sigma_t dB_t, \quad t \in [0, T], \tag{1}$$

where B_t is standard Brownian motion under $\mathbb{P}$. We refer to σ as the instantaneous volatility. Since futures contracts are costless, μ is compensation for σ differing from zero. We leave the processes μ and σ unspecified for the time being.

Instead, we will partially specify the dynamics of a variance swap rate. A variance swap is an over-the-counter contract which now trades liquidly on several stock indices and stocks. This contract has a single payoff occuring at a fixed time, which we require to be T. The floating part of the payoff on a continuously monitored variance swap on one dollar of notional is:

$$\frac{1}{T}\int_0^T \left(\frac{dF_t}{F_t}\right)^2 dt = \frac{1}{T}\int_0^T \sigma_t^2 dt, \tag{2}$$

from (1). At initiation, a variance swap has zero cost to enter. Since the floating part of the payoff is positive, a positive fixed amount is paid at expiration. When expressed as an annualized volatility, this fixed payment is called the variance swap rate. Letting

s_0 be the initial variance swap rate, the final payoff of a variance swap on one dollar of notional is:

$$\frac{1}{T}\int_0^T \sigma_t^2 dt - s_0^2. \tag{3}$$

Neuberger (1990) and Dupire (1992) independently show that if the underlying price process is positive and continuous as in (1), then the payoff to a variance swap can be replicated without making any assumption on the dynamics of the instantaneous volatility σ. However, the replicating strategy requires a static position in European options of all strikes $K > 0$.

Following Duanmu (2004), we reverse the approach taken in Neuberger (1990) and Dupire (1992). We treat a variance swap of a fixed maturity as the fundamental asset whose price process is to be modelled. We treat a path-independent claim maturing with the variance swap as the asset whose payoff is to be replicated by dynamic trading in variance swaps and the option's underlying asset. For now, we assume that a variance swap of maturity T trades continuously over $[0, T]$ in a frictionless market. We do not require that European options of any strike or maturity be available for trading. We will relax the requirement that variance swaps trade in the section after next.

At any time t prior to the common maturity T of the option, the futures, and the variance swap, let $s_t(T)$ denote the fixed rate for a newly issued variance swap. Let $w_t(T) = s_t^2(T)(T-t)$ be the time t value of the claim which pays out a continuous cash flow of $\sigma_u^2 du$ for each $u \in [t, T]$. Given the close relationship between w and s, we will henceforth abuse terminology by referring to w as the variance swap rate.

Under probability measure $\mathbb{P}$, we assume that the variance swap rate process $\{w_t, t \in [0, T]\}$ is continuous over time and given by the solution to the following SDE:

$$\frac{dw_t}{w_t} = \left(\pi_t^w - \frac{\sigma_t^2}{w_t}\right)dt + \sigma_t^w dW_t, \quad t \in [0, T), \tag{4}$$

where W_t is standard Brownian motion under $\mathbb{P}$. Here, π^w is an unspecified stochastic process which represents compensation for the process σ^w differing from zero. The expected growth rate in w is the difference of π^w and the time t stochastic dividend yield $\frac{\sigma_t^2}{w_t}$.

A key assumption which enables practically all of our results is that the ratio of the volatility σ_t^w of the variance swap rate to the instantaneous volatility σ_t of the futures is independent of time:

$$\frac{\sigma_t^w}{\sigma_t} = \alpha(w_t; T), \quad t \in [0, T]. \tag{5}$$

As the notation indicates, the volatility ratio $\alpha(w; T) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a known function of w and T, but is independent of t and σ. We refer to this assumption repeatedly as the stationary volatility ratio hypothesis (SVRH). Since T is fixed in our setting, we henceforth suppress the notational dependence of $\alpha(w)$ on T.

We close the partial specification of our two stochastic processes F and w by requiring that:

$$dB_t dW_t = \rho dt, \quad t \in [0, T], \tag{6}$$

where the correlation parameter $\rho \in [-1, 1]$ is constant. Our final assumption is that there is no arbitrage.

Substituting (5) in (4) implies that the assumed dynamics of F and w are given by:

$$\begin{aligned} \frac{dF_t}{F_t} &= \mu_t dt + \sigma_t dB_t, \\ \frac{dw_t}{w_t} &= \left(\pi_t^w - \frac{\sigma_t^2}{w_t}\right) dt + \alpha(w_t)\sigma_t dW_t \quad t \in [0, T). \end{aligned} \tag{7}$$

Notice that the volatilities of F and w share a common component σ, whose dynamics are unspecified. A motivation for the sharing of this component is stochastic time change. If business time runs at a different rate than calendar time, then σ becomes a proxy for business time and hence affects both volatilities. However, in contrast to other work on option pricing with stochastic time change, we do not specify the dynamics of σ under $\mathbb{P}$. The next section shows that we can nonetheless hedge path-independent claims perfectly and hence price them uniquely.

3 Hedging and pricing path-independent claims

In this section, we show that the terminal payoff $f(S_T)$ of a European-style path-independent claim maturing at T can be replicated by dynamic trading in futures and variance swaps of maturity T. Consider some $C^{2,2}$ function $\Pi(F, w) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}$ and let Π_t denote the stochastic process induced by evaluating the function Π at F_t and w_t:

$$\Pi_t \equiv \Pi(F_t, w_t), \quad t \in [0, T]. \tag{8}$$

Note that the function Π does not depend on time or time to maturity. We can write the assumed statistical dynamics in (7) for F and w as:

$$\begin{aligned} dF_t &= \mu_t F_t dt + \sigma_t F_t dB_t, \\ dw_t &= (\pi_t^w w_t - \sigma_t^2) dt + g(w_t)\sigma_t dW_t \quad t \in [0, T], \end{aligned} \tag{9}$$

where $g(w) \equiv w\alpha(w)$ and $dB_t dW_t = \rho dt$. Itô's formula implies that:

$$\Pi(F_T, w_T) = \Pi(F_0, w_0) + \int_0^T \frac{\partial}{\partial F}\Pi(F_t, w_t)dF_t + \int_0^T \frac{\partial}{\partial w}\Pi(F_t, w_t)dw_t$$
$$+ \int_0^T \left[\frac{F_t^2}{2}\frac{\partial^2 \Pi}{\partial F^2}(F_t, w_t) + \rho F_t g(w_t)\frac{\partial^2 \Pi}{\partial F \partial w}(F_t, w_t) \right.$$
$$\left. + \frac{g^2(w_t)}{2}\frac{\partial^2 \Pi}{\partial w^2}(F_t, w_t)\right] \sigma_t^2 dt. \tag{10}$$

Note that the instantaneous gain on a long position in a futures contract is dF_t, while the instantaneous gain on a long position in a variance swap is $dw_t + \sigma_t^2 dt$. Recognizing, this, suppose that we add and subtract $\int_0^T \frac{\partial}{\partial w}\Pi(F_t, w_t)\sigma_t^2 dt$ to the right hand side of (10). Then $\Pi(F_T, w_T)$:

$$= \Pi(F_0, w_0) + \int_0^T \frac{\partial}{\partial F}\Pi(F_t, w_t)dF_t + \int_0^T \frac{\partial}{\partial w}\Pi(F_t, w_t)(dw_t + \sigma_t^2 dt)$$
$$+ \int_0^T \left[\frac{F_t^2}{2}\frac{\partial^2 \Pi}{\partial F^2}(F_t, w_t) + \rho F_t g(w_t)\frac{\partial^2 \Pi}{\partial F \partial w}(F_t, w_t) + \frac{g^2(w_t)}{2}\frac{\partial^2 \Pi}{\partial w^2}(F_t, w_t) \right.$$
$$\left. - \frac{\partial \Pi(F_t, w_t)}{\partial w}\right] \sigma_t^2 dt. \tag{11}$$

The last term in (11) represents the cash flow generated through time by the dynamic trading strategy consisting of holding $\frac{\partial}{\partial F}\Pi(F_t, w_t)$ futures and $\frac{\partial}{\partial w}\Pi(F_t, w_t)$ variance swaps at each $t \in [0, T)$. Since the path-independent claim which we wish to value has no intermediate payouts, suppose that the function $\Pi(F, w)$ is chosen to solve the following second order linear elliptic PDE:

$$\frac{F^2}{2}\frac{\partial^2 \Pi}{\partial F^2}(F, w) + \rho F g(w)\frac{\partial^2 \Pi}{\partial F \partial w}(F, w) + \frac{g^2(w)}{2}\frac{\partial^2 \Pi}{\partial w^2}(F, w)$$
$$- \frac{\partial \Pi}{\partial w}(F, w) = 0, \quad F > 0, w > 0. \tag{12}$$

Further suppose that we restrict Π by the boundary condition:

$$\Pi(F, 0) = f(F), \quad F > 0, \tag{13}$$

where the contingent claim payoff function f need not be C^2. Since zero is a natural boundary for the futures price, suppose we further require that:

$$\Pi(0, w) = f(0), \quad w > 0. \tag{14}$$

The solution of (12) subject to (13), (14), and growth conditions at $w = \infty$ and $F = \infty$ is unique. Numerical methods such as finite differences or finite elements can be used to efficiently determine $\Pi(F, w)$. In Sect. 6, we also explore Monte Carlo simulation and how this problem can be reduced to an ODE for the characteristic function of the log price.

Since $F_T = S_T$ and $w_T = 0$, substitution of (12) and (13) in (11) implies:

$$f(S_T) = \Pi(F_0, w_0) + \int_0^T \frac{\partial}{\partial F}\Pi(F_t, w_t)dF_t + \int_0^T \frac{\partial}{\partial w}\Pi(F_t, w_t)(dw_t + \sigma_t^2 dt). \tag{15}$$

Hence, the payoff $f(S_T)$ can be perfectly replicated by charging $\Pi(F_0, w_0)$ initially and being long $\frac{\partial}{\partial F}\Pi(F_t, w_t)$ futures and $\frac{\partial}{\partial w}\Pi(F_t, w_t)$ variance swaps at each $t \in [0, T)$. Since time 0 was arbitrary, we refer to $\Pi(F, w)$ as the valuation function for the contingent claim.

Notice that the boundary value problem to be solved for the claim value just involves two independent variables rather than the usual three. This speeds up numerical solution by an order of magnitude compared to the usual boundary value problem arising in SV models. Furthermore, the PDE (12) in this boundary value problem is just a standard linear second order elliptic PDE so standard solution methods are available.

As usual, the claim value and hedge ratios are independent of the processes μ and π^w appearing in the statistical drifts of F and w respectively. In contrast to standard models of stochastic volatility, the option value and hedge ratios are also independent of the market price of volatility risk and the stochastic process for σ, even though the latter affects the dynamics of both assets.

Notice that our arguments fail if the volatility of w depends on time for then the option price must also depend on time. The PDE gains a third independent variable and the departure from zero of $\frac{\partial \Pi}{\partial t}$ further induces dependence of Π on the unknown σ dynamics. Even if the statistical σ dynamics are assumed to be known, the fact that σ is not in general[2] a known function of the price of a long-lived asset will re-introduce dependence on the market price of volatility risk. Hence, our ability to hedge and price under unspecified stochastic volatility and risk premia hinges on our crucial assumption that the volatility of the variance swap rate w be independent of time. The existence of standard diffusion models of stochastic volatility with this property is addressed in Sect. 8.

It may appear that all of the advantages accruing to variance swap rate modelling vanish if variance swaps are not available for trading. Fortunately, the next section shows that the variance swap level can be determined from option prices. Furthermore, the gain on a variance swap position can be accessed by a position in a delta-hedged option. It follows that the advantages outlined in this section can be realized even when variance swaps do not trade.

4 Illiquid variance swaps

In this section, we drop the assumption that variance swaps trade continuously. We propose two different methods by which one can observe the variance swap rate. The first method assumes that one can observe the price of T maturity options of all

[2] However, it is well known that assuming mean-reversion for the risk-neutral process for σ^2 suffices to make σ a known function of the variance swap rate.

strikes. In practice, only discrete strikes are available, but market participants routinely determine a complete implied volatility smile. This smile can be used to determine the prices of options of all strikes and hence value variance swaps. For a given specification of the ratio of the volatility of the variance swap to the volatility of the futures, there is no guarantee that the model value of the option reproduces the market price. The next section shows how this ratio can be chosen so that the model reproduces market option prices.

Let $C_t(K,T)$ and $P_t(K,T)$ respectively denote the prices at time $t \in [0,T]$ of European calls and puts of strike $K > 0$ and fixed maturity date T. Assuming only continuity of the underlying asset price, the payoff to a variance swap can be replicated by holding a static position in options of all strikes and furthermore dynamically trading the underlying futures. It follows that at any time $t \in [0,T]$, the variance swap rate can be determined from the prices of all out-of-the-money options:

$$w_t(T) = \int_0^{F_t} \frac{2}{K^2} P_t(K,T)dK + \int_{F_t}^{\infty} \frac{2}{K^2} C_t(K,T)dK, \quad t \in [0,T]. \tag{16}$$

When w is calculated by (16), we refer to it as the synthetic variance swap rate.

Suppose that we define:

$$\theta_t(m,T) \equiv \frac{P_t(K,T)1(K < F_t) + C_t(K,T)1(K > F_t)}{K}, \tag{17}$$

where:

$$m \equiv \ln(F_t/K). \tag{18}$$

Financially, $\theta_t(k,T)$ is the value at time $t \in [0,T]$ of an out-of the money option per unit of strike expressed in terms of log moneyness m. Note that θ and m are both dimensionless, so this transformation just removes the (arbitrary) dimensions from the dependent and independent variables. Performing the change of variables given by (17) and (18) in the integrals in (16) implies:

$$w_t(T) = 2\int_{-\infty}^{\infty} \theta_t(m,T)dm, \quad t \in [0,T]. \tag{19}$$

Hence, the variance swap rate is just twice the simple average of nondimensionalized out-of-the-money option prices. By modelling the dynamics of this synthetic variance swap rate, one can in turn value options, as shown in the last section. The end result relates the value of a given T maturity option to its underlying asset price and to the simple average in (19). The dependence of each option price on the average is analogous to the situation in the CAPM where an individual stock is priced relative to the market portfolio. We note that as a zero strike call is just the underlying asset, which has zero vega, it plays the role of the zero beta asset in the Black CAPM.

The second method for determining the variance swap rate is to imply it from the market price of a claim with a convex payoff such as a single European option.

Let $\mathcal{C}_t^m$ be the market price of a claim at time $t \in [0, T]$, which has a convex payoff $\gamma(F)$ at T, such as $(F-K)^+$ or $(K-F)^+$. Then the variance swap rate w_t is defined implicitly by:

$$\mathcal{C}_t^m = \mathcal{C}(F_t, w_t), \quad t \in [0, T], \tag{20}$$

where the function $\mathcal{C}(F, w)$ solves the PDE (12) subject to a boundary condition $\mathcal{C}(F, 0) = \gamma(F)$. In Sect. 6, we prove that the convexity of the payoff in F implies that $\frac{\partial}{\partial w}\mathcal{C}(F, w)$ is positive. Hence, the implied variance swap rate is well-defined so long as the market price of the option is arbitrage-free.

As we have two methods for determining the variance swap rate, the question arises as to whether one should use the synthetic variance swap rate or the implied variance swap rate. When a market has many liquid options trading, the synthetic rate is preferred as it is relatively robust. When a market does not have many liquid options trading, one is forced to use the implied rate.

Similarly, we have two ways to observe and access the instantaneous gain on a long position in a variance swap $dw_t + \sigma_t^2 dt$. The first method is to simply replace the variance swap position by the static option component of its replicating portfolio. In this case, the replication strategy for a path-independent claim involves dynamic trading in all options of maturity T. Given the bid ask spread operating in practice, such a strategy would be ruinous if implemented. Fortunately, the martingale representation (15) implies that for any claim with a convex payoff:

$$d\mathcal{C}(F_t, w_t) = \frac{\partial}{\partial F}\mathcal{C}(F_t, w_t) dF_t + \frac{\partial}{\partial w}\mathcal{C}(F_t, w_t)(dw_t + \sigma_t^2 dt), \quad t \in [0, T). \tag{21}$$

Dividing by $\frac{\partial}{\partial w}\mathcal{C}(F_t, w_t) > 0$ implies that the gain on the variance swap position can be accessed by a position in a delta-hedged convex claim:

$$dw_t + \sigma_t^2 dt = \frac{1}{\frac{\partial}{\partial w}\mathcal{C}(F_t, w_t)}\left[d\mathcal{C}(F_t, w_t) - \frac{\partial}{\partial F}\mathcal{C}(F_t, w_t) dF_t\right], \quad t \in [0, T). \tag{22}$$

Thus, the payoff on a path-independent claim can be replicated by dynamic trading in futures and another claim of the same maturity which has a convex payoff, e.g. an option. The same conclusion holds in standard models of stochastic volatility, but there are three major differences in our analysis. First, the requirement that one can imply the instantaneous variance gets replaced by the requirement that one can observe the synthetic variance swap rate or the implied variance swap rate. Second, the market price of volatility risk never has to be modelled. Third, the assumption on the drift and diffusion coefficients of the instantaneous variance gets replaced by the modelling of how the volatility of the variance swap rate depends on the variance swap rate and its maturity.

The next section shows that this dependence can be determined from the market prices of standard options at a fixed time. Hence, the model can be calibrated to the market prices of standard options and then used to determine the dependence of these option prices or other path-independent claims on the futures price and the variance swap rate. It can also be used to price path-dependent claims such as volatility derivatives as we will show in the penultimate section.

5 Calibration

In the last two sections, we assumed that the actual, synthetic, or implied variance swap rate for a fixed maturity was observable and we used it to price path-independent contingent claims. The analysis assumed that the ratio of the variance swap volatility to the underlying futures volatility was a known function of the variance swap rate and its maturity. Knowledge of this function is critical for valuing contingent claims and determining their dependence on the futures price and the variance swap rate. In this section, we take market prices of standard options as given and use this information to determine this critical function. In particular, we assume that market option prices are observable for all strikes $K > 0$ and all maturities $T > 0$. As we continue to make all of the assumptions of prior sections, it follows that market variance swap rates are observable for all maturities $T > 0$.

We exploit the fact that options have payoffs that are linearly homogeneous in their underlying futures price F and their strike K. In fact, we define a contingent claim to be an option so long as its terminal payoff $h(F, K)$ is linearly homogeneous in F and K, i.e.:

$$h(\lambda F, \lambda K) = \lambda h(F, K), \tag{23}$$

for all $\lambda > 0$. Let $\mathcal{O}(F, w; K) : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}$ be the $C^{2,2,2}$ function which relates the price of an option to the contemporaneous futures price F, the variance swap rate w, and the option strike K. For each fixed $K > 0$, $\mathcal{O}(F, w; K)$ solves the elliptic PDE (12):

$$\begin{aligned}\frac{F^2}{2}\frac{\partial^2\mathcal{O}(F, w; K)}{\partial F^2} + \rho F g(w; T)\frac{\partial^2\mathcal{O}(F, w; K)}{\partial F\partial w} + \frac{g^2(w; T)}{2}\frac{\partial^2\mathcal{O}(F, w; K)}{\partial w^2}\\ = \frac{\partial\mathcal{O}(F, w; K)}{\partial w},\end{aligned} \tag{24}$$

for $F > 0$, $w > 0$ subject to the lower Dirichlet boundary condition:

$$\mathcal{O}(F, 0; K) = h(F, K). \tag{25}$$

For any such payoff, it is easy to see that $\mathcal{O}(F, w; K)$ is also linearly homogeneous in F and K, i.e.:

$$\mathcal{O}(\lambda F, w; \lambda K) = \lambda\mathcal{O}(F, w; K), \tag{26}$$

for all $\lambda > 0$. This is proved by showing that the PDE (24) is invariant to the change of variables $(F', K') = (\lambda F, \lambda K)$. Euler's Theorem then implies:

$$F \frac{\partial}{\partial F} \mathcal{O}(F, w; K) = \mathcal{O}(F, w; K) - K \frac{\partial}{\partial K} \mathcal{O}(F, w; K). \tag{27}$$

Differentiating w.r.t. w implies:

$$F \frac{\partial^2}{\partial F \partial w} \mathcal{O}(F, w; K) = \frac{\partial}{\partial w} \mathcal{O}(F, w; K) - K \frac{\partial^2}{\partial K \partial w} \mathcal{O}(F, w; K). \tag{28}$$

It is also easy to show that:

$$F^2 \frac{\partial^2}{\partial F^2} \mathcal{O}(F, w; K) = K^2 \frac{\partial^2}{\partial K^2} \mathcal{O}(F, w; K). \tag{29}$$

Substituting (27) to (29) in (24) implies that:

$$\frac{\partial^2 \mathcal{O}(F, w; K)}{\partial w^2} \frac{g^2(w; T)}{2} + \rho \left[\frac{\partial}{\partial w} \mathcal{O}(F, w; K) - K \frac{\partial^2 \mathcal{O}(F, w; K)}{\partial K \partial w} \right] g(w; T)$$
$$= \frac{\partial \mathcal{O}(F, w; K)}{\partial w} - \frac{K^2}{2} \frac{\partial^2 \mathcal{O}(F, w; K)}{\partial K^2}, \quad K > 0, w > 0. \tag{30}$$

Since the term structure of variance swap rates is assumed to be observable, the function $w(T)$ relating initial variance swap rates to their maturity T is known. This function is monotonically increasing in T and we further assume it is C^2. Let $T(w)$ be the inverse of w, which is also observable, increasing, and C^2. For F fixed at F_0, let:

$$H(K, T) \equiv \mathcal{O}(F_0, w(T); K), \quad K > 0, T > 0, \tag{31}$$

be the initial option price as a function of strike and maturity. Differentiating (31) w.r.t. w implies:

$$\frac{\partial}{\partial w} \mathcal{O}(F_0, w; K) = \frac{\partial}{\partial T} H(K, T) T'(w), \quad K > 0, T > 0. \tag{32}$$

Differentiating (32) w.r.t. K implies:

$$\frac{\partial^2}{\partial w \partial K} \mathcal{O}(F_0, w; K) = \frac{\partial^2}{\partial T \partial K} H(K, T) T'(w), \quad K > 0, T > 0. \tag{33}$$

Differentiating (32) w.r.t. w implies:

$$\frac{\partial^2}{\partial w^2} \mathcal{O}(F_0, w; K) = \frac{\partial^2}{\partial T^2} H(K, T)(T'(w))^2 + \frac{\partial}{\partial T} H(K, T) T''(w), \; K > 0, T > 0. \tag{34}$$

Substituting (32) to (34) in (30) implies that:

$$\left[\frac{\partial^2}{\partial T^2}H(K,T)(T'(w))^2 + \frac{\partial}{\partial T}H(K,T)T''(w)\right]\frac{g^2(w;T)}{2} \\ +\rho\left[\frac{\partial}{\partial T}H(K,T) - K\frac{\partial^2}{\partial T\partial K}H(K,T)\right]T'(w)g(w;T) \\ = \frac{\partial}{\partial T}H(K,T)T'(w) - \frac{K^2}{2}\frac{\partial^2 H(K,T)}{\partial K^2}, \quad K>0, w>0. \tag{35}$$

This is a quadratic equation for $g(w; T)$ which is easily solved. Hence, given that ρ is known, the function $g(w; T)$ can be determined for all $w > 0$ and $T > 0$ since the dependence of the initial option prices H on their strike $K > 0$ and their maturity $T > 0$ has been assumed to be observable. We note that our analysis generalizes to the case where the correlation ρ between F and w depends on w and T, provided that this dependence is known. However, the correlation cannot depend on F as this would cause $\mathcal{O}(F, w; K)$ to no longer be linearly homogeneous in F and K.

6 Monte Carlo simulation for values and greeks

In this section, we show how Monte Carlo simulation can be used to numerically solve the boundary value problem governing the value of the path-independent claim We also investigate how the value of the path-independent claim $\Pi(F, w)$ varies with the futures price F for fixed w. As in standard SV models, we find that Π inherits its behavior from its payoff $f(F)$. Specifically, the n-th partial derivative of $\Pi(F, w)$ w.r.t. F has the same sign as $f^{(n)}(F)$, for $n = 0, 1, 2 \ldots$. We are also interested in how the value of the path-independent claim $\Pi(F, w)$ varies with the variance swap rate w for fixed F. Not surprisingly, we find that path-independent claims with convex payoffs have values that are increasing in w. Hence, for a call, $\Pi(F, w)$ is increasing and convex in F and increasing in w.

6.1 Monte Carlo simulation

By the Feynman Kac theorem, there is an explicit probabilistic representation for the solution to the boundary value problem consisting of the second order linear elliptic PDE (12) and the boundary condition (14):

$$\Pi(F,w) = E^{\hat{Q}_0}\left[f(\hat{F}_\tau)|\hat{F}_0 = F, \hat{w}_0 = w\right]. \tag{36}$$

where under the measure $\hat{Q}_0$, the process $\hat{F}_t$ solves the SDE:

$$\frac{d\hat{F}_t}{\hat{F}_t} = \rho dZ_{1t} + \sqrt{1-\rho^2}dZ_{2t}, \quad t>0, \tag{37}$$

and the process $\hat{w}_t$ solves the SDE:

$$d\hat{w}_t = -dt + g(\hat{w}_t)dZ_{1t}, \quad t > 0. \tag{38}$$

Here, Z_1 and Z_2 are independent standard Brownian motions under the probability measure $\hat{Q}_0$ and τ is the first passage time of $\hat{w}$ to the origin.

Thus, a finite-lived path-independent claim under stochastic instantaneous variance has the same value as a perpetual claim under constant instantaneous variance of 1. The perpetual claim is a down-and-out claim which knocks out when the variance swap hits zero. At the random knockout time τ, the claim pays a rebate of $f(F_\tau)$. Monte Carlo simulation can be used to numerically find the value and futures price greeks of the claim. To speed up computations, one can take advantage of the fact that $\hat{F}_t$ follows geometric Brownian motion under the probability measure $\hat{Q}_n$. Adapting the mixing argument in Romano and Touzi (1997) to the present setting, let $B(F, T) \equiv E^{\hat{Q}_0}[f(\hat{F}_T)|\hat{F}_0 = F]$ be the Black model value for the forward price of a path-independent claim paying $f(\hat{F}_T)$ at the fixed time T, when $\hat{F}$ is the geometric Brownian $\hat{Q}_0$ martingale with unit volatility in (37). The solution to the SDE in (37) is:

$$\hat{F}_T = Fe^{\left(-\frac{1}{2}\right)T+\rho Z^{(n)}_{1,T}+\sqrt{1-\rho^2}dZ^{(n)}_{2,T}}. \tag{39}$$

Hence, $\hat{F}_T$ is lognormally distributed with mean F and variance of $\ln F_T$ given by T. Notice that these are the arguments of the Black model value function.

If we condition on the Z_1 path in (38), then we learn the $\hat{w}$ path and hence τ and $Z_{1,\tau}$. Evaluating (39) at τ rather than T implies:

$$\hat{F}_\tau = Fe^{\left(-\frac{1}{2}\right)\tau+\rho Z^{(n)}_{1,\tau}+\sqrt{1-\rho^2}dZ^{(n)}_{2,\tau}} \tag{40}$$

$$= Fe^{-\frac{\rho^2}{2}\tau+\rho Z^{(n)}_{1,\tau}}e^{-\frac{1-\rho^2}{2}\tau+\sqrt{1-\rho^2}dZ^{(n)}_{2,\tau}}. \tag{41}$$

Hence if we condition on $\tau = T$ and $Z_{1,\tau} = z$, then $\hat{F}_\tau$ is lognormally distributed with mean $Fe^{-\frac{\rho^2}{2}T+\rho z}$ and variance of $\ln F_\tau$ given by $(1-\rho^2)T$. Thus, the conditional mean of $\hat{F}_\tau$ is obtained from the mean of F_T by multiplying by the factor $e^{-\frac{\rho^2}{2}T+\rho z}$. Likewise, the conditional variance of $\ln F_\tau$ is obtained from the variance of $\ln F_T$ by multiplying by $1-\rho^2$. This motivates the following representation for $\Pi(F, w)$:

$$\Pi(F, w) = \int_0^\infty \int_{-\infty}^\infty B(Fe^{-\frac{\rho^2}{2}T+\rho z}, (1-\rho^2)T)\phi_{0,\tau}(T, w)dzdT, \tag{42}$$

for $n = 0, 1, 2\ldots$, where $\phi_{0,\tau}(T; w) \equiv \frac{\hat{Q}_0\{\tau \in dT|\hat{w}_0 = w\}}{dT}$ is the probability density function under $\hat{Q}_0$ for the first passage time τ of the process $\hat{w}$ to zero, given that $\hat{w}_0 = w$.

For many path-independent claims such as calls and puts, the Black model value in (42) is known in closed form. For such claims, one need only simulate the $\hat{Q}_0$ dynamics of $\hat{w}$ in (38). For each simulated $\hat{w}$ path terminating at $\hat{w}_\tau = 0$, one just needs τ and $Z_{1,\tau}$ to evaluate this closed form expression. The claim value is approximated by averaging over paths. Since the barrier at zero is monitored continuously, a naive discrete time Monte Carlo will tend to overvalue τ, producing upward bias in convex claims. To remedy this, one can use Brownian bridges as discussed in Beaglehole et al. (1997), El Babsiri and Noel (1998), and in Andersen and Brotherton-Ratcliffe (1996). Alternatively, one can use large deviations as discussed in Baldi et al. (1999).

6.2 Partial derivatives

Since the process $\hat{w}$ in (38) is a univariate diffusion, a coupling argument implies that a rise in its initial value w weakly increases the first passage time to the origin. If the payoff function f is convex, then since $\hat{F}$ is a $\hat{Q}_0$ martingale, the process $\{f(F_t), t > 0\}$ is a $\hat{Q}_0$ submartingale. It follows that a rise in w causes Π to increase. Hence, when f is convex, then the hedge ratio $\frac{\partial \Pi}{\partial w}(F, w)$ is positive.

Similarly, one can use a coupling argument to show that the hedge ratio $\frac{\partial \Pi}{\partial F}(F, w)$ is nonnegative when f is increasing in F. More generally, the following theorem is proved in Appendix 1:

Theorem 1 *Let $\Pi(F, w)$ defined by (36) be the value function for a path-independent claim. Then:*

$$\frac{\partial^n}{\partial F^n}\Pi(F, w) = E^{\hat{Q}_n}\left[f^{(n)}(\hat{F}_\tau)e^{\frac{n(n-1)}{2}\tau}|\hat{F}_0 = F, \hat{w}_0 = w\right], n = 0, 1, 2 \ldots, \quad (43)$$

where under the measure $\hat{Q}_n$, the process $\hat{F}_t$ solves the SDE:

$$\frac{d\hat{F}_t}{\hat{F}_t} = n dt + \rho dZ_{1t}^{(n)} + \sqrt{1-\rho^2} dZ_{2t}^{(n)}, \quad t > 0, \quad (44)$$

and the process $\hat{w}_t$ solves the SDE:

$$d\hat{w}_t = [n\rho g(\hat{w}_t) - 1]dt + g(\hat{w}_t)dZ_{1t}^{(n)}, \quad t > 0, \quad (45)$$

where $Z_1^{(n)}$ and $Z_2^{(n)}$ are independent standard Brownian motions under the probability measure $\hat{Q}_n$ and τ is still the first passage time of $\hat{w}$ to the origin.

Theorem 1 shows that futures price greeks also have a probabilistic representation. Since $e^{\frac{n(n-1)}{2}\tau} > 0$, Theorem 1 implies that the n-th partial derivative of $\Pi(F, w)$ w.r.t. F has the same sign as $f^{(n)}(F)$. Hence, the value of a European put is a decreasing convex function of F. One can also use a coupling argument on (43) to show that the cross partial $\frac{\partial^{n+1}\Pi}{\partial F^n \partial w}(F, w)$ is nonnegative when $f^{(n+1)}(F)$ and $f^{(n+2)}(F)$ are both increasing in F. In particular, for $n=1$, the vanna $\frac{\partial^2\Pi}{\partial F\partial w}(F,w)$ is nonnegative when $f^{(2)}(F)$ and $f^{(3)}(F)$ are both increasing in F.

We can adapt the above mixing argument to also value futures price Greeks using simulation. As a special case of Theorem 1 with $w=T$ and $g(w)=0$:

$$\frac{\partial^n}{\partial F^n}B(F,T)=E^{\hat{Q}_n}\left[f^{(n)}(\hat{F}_T)e^{\frac{n(n-1)}{2}T}|\hat{F}_0=F\right]$$

is the Black model value of the n-th partial derivative of $B(F,T)$ w.r.t. F. In the case of calls and puts, an explicit formula for this partial derivative is given in Carr (2001). It is straightforward to derive the following generalization of (42) to $\frac{\partial^n}{\partial F^n}\Pi(F,w)$:

$$\frac{\partial^n}{\partial F^n}\Pi(F,w)=\int_0^\infty\int_{-\infty}^\infty\frac{\partial^n}{\partial F^n}B(Fe^{-\frac{\rho^2}{2}T+\rho z},\\ \times(1-\rho^2)T)e^{\frac{n(n-1)}{2}\rho^2 T}\phi_\tau(T,w)dzdT, \quad (46)$$

for $n=0,1,2\ldots$, where now $\phi_{n,\tau}(T;w)\equiv\frac{\hat{Q}_n\{\tau\in dT|\hat{w}_0=w\}}{dT}$ is the probability density function under $\hat{Q}_n$ for the first passage time τ of the process $\hat{w}$ to zero, given that $\hat{w}_0=w$.

For many path-independent claims such as calls and puts, the function multiplying $\phi_\tau(T,w)$ in (46) is known in closed form. For such claims, one need only simulate the $\hat{Q}_n$ dynamics of $\hat{w}$ in (45). The futures price greek is approximated by averaging over paths.

To summarize the results of this section and the last, we can use the observed variation of market option prices across strike and maturity to determine the dependence of the volatility ratio on w and T. This function in turn determines the dependence of the model value on the underlying futures price and the variance swap rate. This dependence is determined without having to specify risk premia, the market price of volatility risk, or the starting value or dynamics of the instantaneous volatility σ. The model values and greeks can be efficiently computed by finite differences, finite elements, or by Monte Carlo simulation.

Despite these compelling advantages, it would be extremely disturbing if there was no way to specify a stochastic process for σ which is consistent with our key assumption (5) that the volatility ratio is independent of calendar time. We address the formulation of this consistency problem in the next section.

7 Consistency with risk-neutral diffusion

It is well known that no arbitrage implies the existence of a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which the prices of all nondividend paying assets are martingales. The standard terminology when describing a stochastic processes under the measure $\mathbb{Q}$ is to refer to it as a risk-neutral process. Taking the money market account as numeraire, the risk-neutral process for the underlying futures price is:

$$dF_t = \sqrt{v_t} F_t d\tilde{B}_t, \quad t \in [0, T], \tag{47}$$

where $\tilde{B}$ is standard Brownian motion under $\mathbb{Q}$. The risk-neutral drift of the futures price is zero, since futures contracts are costless. The process v in (47) is commonly referred to as the instantaneous variance. Under $\mathbb{Q}$, the variance swap rate is the risk-neutral expected value of the remaining integral of v:

$$w_t = E^{\mathbb{Q}}\left[\int_t^T v_u du \middle| v_t = v\right], \quad v \geq 0, t \in [0, T]. \tag{48}$$

Under our SVRH, the risk-neutral process for w is given by:

$$dw_t = -v_t dt + g(w_t)\sqrt{v_t} d\tilde{W}_t, \quad t \in [0, T], \tag{49}$$

where $\tilde{W}$ is standard Brownian motion under $\mathbb{Q}$. The risk-neutral drift of $-v_t dt$ in (55) reflects the fact that a long position in a variance swap results in a cash inflow of $v_t dt$ at each instant of time. By Girsanov's theorem, the two Brownian motions have the same correlation under $\mathbb{Q}$ as they have under $\mathbb{P}$:

$$d\tilde{B}_t d\tilde{W}_t = \rho dt. \tag{50}$$

Suppose that in addition to (47) and (49), the risk-neutral dynamics of the instantaneous variance rate are assumed to be given by:

$$dv_t = a(t, v_t)dt + b(t, v_t)\sqrt{v_t} d\tilde{W}_t, \quad t \in [0, T]. \tag{51}$$

Notice that the Brownian motion $\tilde{W}$ driving v is the same as the one driving w. Since the coefficients a and b in (51) are assumed to be independent of T, we refer to (51) as the maturity independent diffusion hypothesis (MIDH).

As we argue later, there is no economic justification for the MIDH. We believe that the only reason why the option pricing literature has adopted the MIDH is its inherent tractability and the lack of any plausible alternatives. Nonetheless, a great deal of empirical work has been done testing the consistency of this class of models with the data. Hence, it is of some interest to discern whether or not there is at least some subclass of our stationary models for the w dynamics which is consistent with the MIDH (51). In the remainder of this section, we find a condition which the risk-neutral drift and diffusion coefficients of v must satisfy in order to meet both the MIDH and our SVRH. In the next section, we find the only drift and diffusion functions for v which meet this condition.

Since we have Markovian dynamics for the instantaneous variance rate v, there exists a $C^{2,1}$ function $w(v, t) : \mathbb{R}^+ \times [0, T] \mapsto \mathbb{R}^+$ such that the variance swap rate w_t is given by:

$$w_t = w(t, v_t), \quad t \in [0, T]. \tag{52}$$

By standard arbitrage arguments, this function $w(t, v)$ solves the following second order linear parabolic PDE:

$$\frac{b^2(t, v)v}{2}\frac{\partial^2 w}{\partial v^2}(t, v) + a(t, v)\frac{\partial w}{\partial v}(t, v) + \frac{\partial w}{\partial t}(t, v) = -v, \tag{53}$$

on the domain $t \in [0, T], v \geq 0$. The function w is also subject to the terminal condition:

$$w(T, v) = 0, \quad v \geq 0. \tag{54}$$

The solution to the Cauchy problem consisting of (53) and (54) is unique.

Applying Itô's formula to (52) implies that the risk-neutral dynamics of the variance swap rate w are given by:

$$dw_t = -v_t dt + \frac{\partial w}{\partial v}(t, v_t)b(t, v_t)\sqrt{v_t}d\tilde{W}_t, \quad t \in [0, T]. \tag{55}$$

Since the form of the volatility function is invariant to the measure change, our objective is to restrict the risk-neutral v dynamics so that the risk-neutral dynamics of w obey:

$$dw_t = -v_t dt + g(w_t)\sqrt{v_t}d\tilde{W}_t, \quad t \in [0, T], \tag{56}$$

where the function $g(w)$ is independent of t and v. Comparing the diffusion coefficients in (55) and (56), the question is whether one can specify the functions $a(t, v)$ and $b(t, v)$ governing the risk-neutral dynamics of v in (51) so that:

$$\frac{\partial w}{\partial v}(t, v)b(t, v) = g(w(t, v)), \quad t \in [0, T], v \geq 0, \tag{57}$$

where $w(t, v)$ solves (53) and (54) and $g(w)$ is independent of t and v.

Even if there exist solutions to (57), then a further open question is whether the chosen g function gives sensible properties to w. For example, since v is a nonnegative process, (48) implies that $g(w)$ must be chosen so that w is also a nonnegative process:

$$w_t \geq 0, \quad t \in [0, T]. \tag{58}$$

Furthermore, right at maturity, the variance swap rate is related to the instantaneous variance rate by the following consistency condition:

$$\lim_{t \uparrow T} \frac{w_t}{T - t} = v_T. \tag{59}$$

Finally, since v_T is finite, (59) implies the following terminal condition:

$$w_T = 0. \tag{60}$$

Since the nonnegativity condition (58), the consistency condition (59), and the terminal condition (60) all refer to T, we will refer to them jointly as the T-conditions.

Since the coefficients of the w process in (56) are independent of t and the coefficients of the v process in (51) are independent of T, it is not at all obvious whether the three T-conditions can be achieved. Fortunately, the next section shows that the set of models satisfying both the SVRH and the MIDH is not empty. However, imposing the SVRH on the class of maturity independent diffusion processes for v reduces the risk-neutral process for v to a diffusion with quadratic drift of the form $p(t)v_t + qv_t^2$ and with normal volatility proportional to $v_t^{3/2}$. Conversely, imposing the MIDH on the class of time homogeneous continuous processes for w imposes a very specific structure on the function $g(w; T)$ governing its normal volatility. Hence, the next section shows that only very specific risk-neutral processes for v and w are consistent with the joint hypotheses of SVRH and MIDH.

8 The general solution to the consistency problem

This section shows that there exists a (maturity independent) risk-neutral diffusion process for v which is consistent with our SVRH. Fortunately, this process is both empirically supported and tractable. The tractability arises from the fact that the joint Fourier Laplace transform of returns and their quadratic variation can be derived in closed form. We present this transform along with valuation formulas for the variance swap rate and its volatility. The next section presents simpler formulas for these quantities that arise in a subcase of the process presented in this section.

To emphasize which quantities are maturity dependent, the notation in this section will indicate maturity dependence whenever it is present. The following theorem shows that the SVRH and the MIDH completely determine the form of the risk-neutral process for v:

Theorem 2 *The SVRH (5) and the MIDH (51) jointly imply that the risk-neutral process for the instantaneous variance is given by:*

$$dv_t = [p(t)v_t + qv_t^2]dt + \epsilon v_t^{3/2} d\tilde{W}_t, \quad t \in [0, T], \tag{61}$$

where p is an arbitrary function of time and $\epsilon > 0$ and $q < \frac{\epsilon^2}{2}$ are arbitrary constants. Furthermore, in the risk-neutral process for w:

$$dw_t(T) = -v_t dt + g(w_t(T); T)\sqrt{v_t} d\tilde{W}_t, \quad t \in [0, T], \tag{62}$$

we must have $g(0; T) = 0$ and $g_w(0; T) = \epsilon$.

The proof of the above theorem is in Appendix 2. The content of Theorem 2 is that only a small set of risk-neutral processes for v are consistent with both our SVRH and the MIDH. Theorem 4 of this section will also show that the function g governing the volatility of w is much more restricted than as indicated in Theorem 2.

However, we caution that the small set of allowed processes for v and w is just as much due to the MIDH as our SVRH. The usual mechanism by which a maturity independent risk-neutral diffusion process for v is derived is to assume a diffusion process for v under $\mathbb{P}$ and to assume that the market price of variance risk is a function of just time t and the instantaneous variance v. To our knowledge, there is no economic justification for either assumption. Rather, these assumptions are made solely for the purpose of gaining the tractability associated with the risk-neutral v process being a diffusion. While tractability is a worthy objective, our approach of directly modelling the variance swap rate dynamics already provides a tractable option pricing model. As a result, this justification for MIDH is absent in our setting. Consequently, any unwelcome restrictions which the MIDH introduces to our setting can be banished by simply rejecting it. We furthermore note that the maturity independence of the risk-neutral v process is only a consequence of the standard assumption that the money market account serves as the numeraire. If the numeraire were instead some asset with a strictly positive payoff at T, then the drift and diffusion coefficients of v can depend on T. Hence, there is no need to have either maturity independence or a diffusion specification for the risk-neutral process for the instantaneous variance.

Fortunately, the existing empirical literature examining the structure of the risk-neutral process for instantaneous variance is quite supportive of the one kind of process consistent with both hypotheses. In particular, there is strong evidence in favor of specifying the diffusion coefficient of v as proportional to $v_t^{3/2}$. There is also mildly supportive evidence that the risk-neutral drift of v has the form $p(t)v_t + qv_t^2$. The theoretical advantages of our SVRH models suggest that further empirical investigation along these lines is warranted. However, should further testing reject the consistency of the quadratic drift 3/2 process with the data, our view is that the theoretically and numerically inferior MIDH should be jettisoned. Of course, we would advocate disposing of SVRH despite its advantages if direct empirical evidence were mounted against it.

Recall that the risk-neutral process for the underlying futures price is:

$$dF_t = \sqrt{v_t} F_t d\tilde{B}_t, \quad t \in [0, T]. \tag{63}$$

Theorem 2 implies that the v process in (63) satisfies the MIDH and our SVRH if and only if:

$$dv_t = [p(t)v_t + qv_t^2]dt + \epsilon v_t^{3/2} d\tilde{W}_t, \quad t \in [0, T], \tag{64}$$

where $p(t)$ is an arbitrary function of time, $q < \frac{\varepsilon^2}{2}$ and $\varepsilon > 0$, and where:

$$d\tilde{B}_t d\tilde{W}_t = \rho dt. \tag{65}$$

The reason for the upper bound on q becomes clear if we examine the process followed by:

$$R_t \equiv \frac{1}{v_t}, \quad t \in [0, T]. \tag{66}$$

Employing Itô's formula:

$$dR_t = \left[\varepsilon^2 - q - p(t)R_t\right]dt - \epsilon\sqrt{R_t}d\tilde{W}_t, \quad t \in [0, T]. \tag{67}$$

Thus, the reciprocal of V is just a special case of the affine drift square root process. It is well known that this process avoids zero if $\varepsilon^2 - q > \frac{\varepsilon^2}{2}$ or equivalently $q < \frac{\varepsilon^2}{2}$. If q violates this upper bound, then the v process can explode. As a result, we henceforth assume that $q < \frac{\varepsilon^2}{2}$. In fact, if $q < 0$, then by defining $p(t) = \kappa\theta(t)$ and $q = -\kappa$ for $\kappa > 0$, then (64) can be re-written as:

$$dv_t = \kappa v_t[\theta(t) - v_t]dt + \epsilon v_t^{3/2}d\tilde{W}_t, \quad t \in [0, T]. \tag{68}$$

Hence the risk-neutral process for v becomes mean-reverting, where the speed of mean-reversion is proportional to v.

As (64) is actually more general than (68), we now further analyze the risk-neutral v process given in (64). With q respecting its upper bound, the results on the reciprocal of v imply that the F and v processes respectively given by (63) and (61) never explode. Furthermore, for $F_0 > 0$ and $v_0 > 0$, the two processes are always positive. If $F_0 = 0$ or $v_0 = 0$, then the corresponding process is trapped at the origin. As we assume that $F_0 > 0$ and $v_0 > 0$, this feature is of no economic consequence, but it does have the virtue of providing simple boundary conditions when employing finite differences. Although the v process is well-behaved under $\mathbb{Q}$, Lewis (2000) shows that option pricing also requires that the v process have zero explosion probability under the measure induced by taking the underlying asset as the numeraire. Feller's explosion test implies that there is zero explosion probability under this measure if and only if $\rho < \frac{1}{2\epsilon}$. We henceforth impose this constraint which is automatically met if $\rho \leq 0$ since $\epsilon > 0$.

The probability density function of R_T is given in closed form in Cox et al. (1985), when (67) is assumed to be time-homogeneous. It follows from (66) that the probability density function for v_T is also known by the change of variable formula, under time-homogeneity. Other properties of the time-homogeneous version of the quadratic drift 3/2 process are discussed in Ahn and Gao (1999), Andreasen (2001), Cox et al. (1980, 1985), Heston (1997), Lewis (2000), and Spencer (2003).

In particular, Heston (1997) points out that (64) is actually a generalization of the well-known CEV process, first proposed in Cox (1975). To see this, write the risk-neutral futures price process as:

$$\frac{dF_t}{F_t} = \eta F_t^{\frac{\gamma}{2}}d\tilde{B}_t, \quad t \in [0, T]. \tag{69}$$

To ensure that F is a martingale rather than a strict local martingale, we require that $\gamma \in [0, 2]$. Now let $V_t \equiv V(F_t)$ be the instantaneous variance rate, where from (69) the function V is defined by $V(F) \equiv \eta^2 F^\gamma$ for $F > 0$. Using Itô's formula:

$$dV_t = \frac{\gamma}{2}(\gamma - 1)V_t^2 dt + \gamma V_t^{\frac{3}{2}}d\tilde{B}_t, \quad t \in [0, T]. \tag{70}$$

Hence, the CEV process can be obtained from (63) and (64) by setting $p(t) = 0$, $q = \frac{\gamma}{2}(\gamma - 1)$, $\epsilon = \gamma$, and $\rho = 1$.

By Girsanov's theorem, the volatility of the instantaneous variance must have the same form under $\mathbb{P}$ as it has under $\mathbb{Q}$. Hence, Theorem 2 implies that the joint hypotheses of SVRH and MIDH force the normal volatility of v to be proportional to $v_t^{3/2}$ under both $\mathbb{P}$ and $\mathbb{Q}$. While this specification may seem unnatural, we note that if we let $\sigma_t \equiv \sqrt{v_t}$ be the volatility, then Itô's formula implies that the martingale part of $\frac{d\sigma_t}{\sigma_t}$ is $\frac{\epsilon}{2}\sigma_t dW_t$. In words, the volatility of volatility is proportional to volatility. Hence, the 3/2 specification is natural when posed in terms of the usual lognormal volatility.

Although the 3/2 specification was derived purely from theoretical considerations, it has received a surprisingly large amount of empirical support for both the statistical process and the risk-neutral process. We first summarize the literature on the statistical process. Using affine drift and a CEV specification for the volatility of instantaneous variance, Ishida and Engle (2002) estimate the CEV power to be 1.71 for S&P500 daily returns measured over a 30-year period. Javaheri (2004) also estimates this process on the time series of S&P500 daily returns, but with the CEV power constrained to either be 0.5, 1.0, or 1.5. He concludes that for most of his filters, a power of 1.5 outperforms the other two possible choices. Chacko and Viceira (1999) use Spectral GMM estimation on the same affine drift CEV process as in Ishida and Engle. Using the CRSP value weighted portfolio, they estimate the CEV power at 1.10 using weekly data over a 35-year period and at 1.65 using monthly data over a 71-year period.

Two studies examine both the statistical process and the risk-neutral process. Poteshman (1998) examines S&P500 index option prices over a 7-year period and finds that both the statistical and the risk-neutral drift of the instantaneous variance are not affine. He also finds that the volatility of the instantaneous variance is an increasing convex function of the instantaneous variance. Jones (2003) examines daily S&P100 returns and implied volatilities over a 14-year period. Using this data, he first estimates the statistical version of the affine drift CEV process for the instantaneous variance. Using his time series of S&P100 daily returns, he finds the CEV power to be 1.33. Comparing the fit of this CEV model with the square root model on the option price data, he finds much better option pricing under the CEV model for 3- and 6- month options. For shorter maturity options, all of his (purely continuous) models fail and so he concludes that jumps may be needed. Using the same data as in Jones (2003), Bakshi et al. (2004) look at a time series of S&P100 implied volatilities as captured by VIX^2. They estimate several models including an SDE of the form

$$dv_t = \left(\alpha_0 + \alpha_1 v_t + \alpha_2 v_t^2 + \frac{\alpha_3}{v_t}\right) dt + \beta_2 v_t^{\beta_3} dW_t, \quad t \in [0, T]. \tag{71}$$

They find that α_2 and α_3 are highly significant. They find that a linear drift model is rejected in favor of a nonlinear drift model. They find that the CEV power parameter is highly significant and estimate it at 1.27. They conduct a one sided t test on the null hypotheseis that $\beta_3 \leq 1$ and reject this null. They conclude that $\beta_3 > 1$ is needed to match the time series properties of the VIX index with CEV models.

We conclude that there is substantial evidence supporting a 3/2 power specification for the normal volatility of v. The evidence supporting the deterministically time varying quadratic drift in Theorem 2 is weaker, but not inconsistent. The theoretical and numerical advantages of the quadratic drift 3/2 process promulgated in Theorem 2 suggests that further empirical work on this process is warranted.

Besides enjoying empirical support, the quadratic drift 3/2 process is also extremely tractable. Let $X_t \equiv \ln(F_t/F_0)$ be the return process and let $\langle X\rangle_t \equiv \int_0^t v_s ds$ denote its quadratic variation. Let $E^{\mathbb{Q}}[e^{iuX_T - s\langle X\rangle_T}]$ be the joint Fourier Laplace transform of X_T and $\langle X\rangle_T$. Let $E^{\mathbb{Q}}[e^{iuX_T - s(\langle X\rangle_T - \langle X\rangle_t}]|X_t = x, v_t = v]$ be the joint conditional Fourier Laplace transform of X_T and $\langle X\rangle_T - \langle X\rangle_t$. We now show that there is a closed form formula for the latter expression.

Theorem 3 *Suppose that the risk-neutral process for the instantaneous variance is given by:*

$$dv_t = [p(t)v_t + qv_t^2]dt + \epsilon v_t^{3/2} d\tilde{W}_t, \quad t \in [0, T], \tag{72}$$

where $q < \frac{\epsilon^2}{2}$. Then the joint conditional Fourier Laplace transform of X_T and $\langle X\rangle_T - \langle X\rangle_t$ is given by:

$$E^{\mathbb{Q}}\left[e^{iuX_T - s(\langle X\rangle_T - \langle X\rangle_t}\right]|X_t = x, v_t = v] = e^{iux}\frac{\Gamma(\gamma-\alpha)}{\Gamma(\gamma)}\left(\frac{2}{\epsilon^2 y(t,v)}\right)^{\alpha} \times M\left(\alpha; \gamma; \frac{-2}{\epsilon^2 y(t,v)}\right), \tag{73}$$

where:

$$y(t, v) \equiv \int_t^T e^{\int_t^{t'} p(u)du} dt' \times v, \tag{74}$$

the confluent hypergeometric function M is defined as:

$$M(\alpha;\gamma;z) \equiv \sum_{n=0}^{\infty} \frac{(\alpha)_n}{(\gamma)_n}\frac{z^n}{n!},$$

$$\alpha \equiv -\left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right) + \sqrt{\left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right)^2 + 2\frac{\lambda}{\epsilon^2}},$$
$$\gamma \equiv 2\left[\alpha + 1 - \frac{\tilde{q}}{\epsilon^2}\right], \tag{75}$$

and where:

$$\tilde{q} \equiv q + \rho\epsilon iu \qquad \lambda \equiv s + \frac{iu}{2} + \frac{u^2}{2}.$$

The proof of Theorem 3 is in Appendix 3. The result in Theorem 3 makes it straightforward to numerically value many derivative securities written on X_T and/or $\langle X\rangle_T$. In particular, the prices of European options on F_T can be quickly obtained via (fast) Fourier inversion. Assuming time homogeneity, Heston (1997) and Lewis (2000) have independently derived the conditional characteristic function of X_T. Our result extends these previously derived results in two ways. First, we have the joint transform which allows us to value both options on price and options on realized variance, or even options on both. Second, we have obtained this joint Fourier Laplace transform despite the fact that the risk-neutral drift of v contains an arbitrary function of time. This feature allows one to calibrate the quadratic drift to an arbitrarily given term structure of say ATM implied volatilities or variance swap rates. It also allows one to generalize the model by randomizing the level towards which the v process reverts. If the risk-neutral process for this level evolves independently of F and v then the risk-neutral density for a functional of this process can be chosen so as to gain consistency with an implied volatility skew.

We now turn to the implication of our SVRH and the MIDH for the form of $w(v, t; T)$ and the form of the function $g(w; T)$ determining its normal volatility. The following theorem shows that both functions are determined in closed form.

Theorem 4 *Suppose that the risk-neutral process for the instantaneous variance is given by:*

$$dv_t = [p(t)v_t + qv_t^2]dt + \epsilon v_t^{3/2} d\tilde{W}_t, \quad t \in [0, T], \tag{76}$$

where $q < \frac{\epsilon^2}{2}$*. Then the variance swap rate valuation function* $w(t, v; T) \equiv E^{\mathbb{Q}}\left[\int_t^T v_u du \middle| v_t = v\right]$ *is given by:*

$$w(t, v; T) = h\left(v \times \int_t^T e^{\int_t^{t'} p(u)du} dt'\right), \quad t \in [0, T], v \geq 0, T > 0, \tag{77}$$

where:

$$h(y) = \int_0^y e^{-\frac{2}{\varepsilon^2 z}} z^{-\frac{2q}{\varepsilon^2}} \int_z^\infty \frac{2}{\varepsilon^2} e^{\frac{2}{\varepsilon^2 u}} u^{\frac{2q}{\varepsilon^2}-2} du dz. \tag{78}$$

The risk-neutral process for the variance swap rate is given by:

$$dw_t(T) = -v_t dt + g(w_t(T))\sqrt{v_t} d\tilde{W}_t, \quad t \in [0, T], \tag{79}$$

where:

$$g(w) = \epsilon \frac{\partial}{\partial y} h(h^{-1}(w))h^{-1}(w), \quad w > 0, \tag{80}$$

where $h^{-1}(w)$ *is the inverse in* w *of* $h(w)$.

The proof of Theorem 4 is in Appendix 4. This proof makes it clear that the 3 T-conditions (58) to (60) are all met. By the equivalence of $\mathbb{Q}$ and $\mathbb{P}$, these conditions also hold for the statistical process for w provided that it retains the same form.

Clearly, the general forms for w and g are complicated. To address this deficiency, we offer three solutions. First, Appendix 4 shows that h solves the simple linear ODE:

$$\frac{\epsilon^2 y^2}{2}\frac{\partial^2}{\partial y^2}h(y) + (qy-1)\frac{\partial}{\partial y}h(y) + 1 = 0, \quad y > 0, \tag{81}$$

subject to $h(0) = 0$ and $\lim_{y\downarrow 0}\frac{\partial}{\partial y}h(y) = 1$. Furthermore, Appendix 4 also shows that $g(w)$ solves the following nonlinear ODE:

$$\frac{g^2(w)}{2}g''(w) + \frac{\epsilon}{2}g(w)g'(w) + \left(q - 1 - \frac{\epsilon^2}{2}\right)g(w) + \epsilon = 0, \tag{82}$$

subject to $g(0) = 0$ and $g'(0) = \epsilon$. In applied work, it may be easier to numerically evaluate these boundary value problems for h and g rather than their solution given in Theorem 4.

A second approach for dealing with the complexity of w and its volatility is to switch attention from w to a related process. Suppose we define a stochastic process $\{y_t; t \in [0, T]\}$ by:

$$y_t \equiv E_t^{\mathbb{Q}}\int_t^T e^{-q\int_t^{t'} v_u du} v_{t'} dt', \quad t \in [0, T]. \tag{83}$$

Note that the parameter q in (83) is the same as the one used in Theorem 2 and hence we require $q < \varepsilon^2/2$. A financial interpretation of y is that it represents the value at time t of the floating part of a dynamic trading strategy in variance swaps. For each $t' \in [t, T]$, this strategy holds $e^{-q\int_t^{t'} v_u du}$ variance swaps. Since the v process is Markovian in itself and time, it follows that there exists a $C^{1,2}$ function $y(t, v) : [0, T] \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ such that:

$$y_t = y(t, v_t), \quad t \in [0, T]. \tag{84}$$

By the Feynman Kac Theorem, the function $y(t, v)$ solves the following PDE:

$$\frac{\epsilon^2 v^3}{2}\frac{\partial^2 y}{\partial v^2}(t, v) + \left[p(t)v + qv^2\right]\frac{\partial y}{\partial v}(t, v) + \frac{\partial y}{\partial t}(t, v) = qvy(t, v) - v, \tag{85}$$

on the domain $t \in [0, T], v \geq 0$. The function y is also subject to the terminal condition:

$$y(T, v) = 0, \quad v \geq 0. \tag{86}$$

The solution to the Cauchy problem consisting of (85) and (86) is unique.

Given that q and ϵ are both in the above Cauchy problem, it is remarkable that given $v_t = v$, the solution for the value of y_t is always independent of both q and ϵ. Intuitively, the independence of y from q arises because the dynamic trading strategy in variance swaps offsets the dependence of w on q. The independence of y from ε arises because y is just proportional to v:

$$y_t = \int_t^T e^{\int_t^{t'} p(u)du} dt' \times v_t, \quad t \in [0, T]. \tag{87}$$

It is straightforward to verify that (87) solves (85) and (86). Applying Itô's formula to (87), (76) implies that the risk-neutral dynamics of y are given by:

$$dy_t = (q y_t - 1) v_t dt + \varepsilon y_t \sqrt{v_t} d\tilde{W}_t, \quad t \in [0, T]. \tag{88}$$

Hence, the level of y in (87) is independent of q and ε, but this level does depend on the function p. Conversely, the dynamics of y in (88) depends on q and ε, but these dynamics are independent of the function p. Notice also that the y dynamics in (88) are time homogeneous in both calendar time and business time defined by $\langle X \rangle_t = \int_0^t v_s ds$. Comparing (79) and (88), we see that although y has a slightly more complicated risk-neutral drift than w, it has much simpler volatility.

The definition (83) of y makes it clear that it is positive before T and vanishes right at T. Since the dynamics of y are independent of t and the dynamics of v are independent of T, it is somewhat surprising that the process y is positive before T and vanishes right at T. The intuition for this result is best seen when $p(t) = p$. At first, this seems even more puzzling since v is then also a time homogeneous diffusion. However from (87):

$$y_t = v_t \frac{e^{p(T-t)} - 1}{p}. \tag{89}$$

Hence, the ratio of the two time-homogeneous diffusion processes is just a function of the time to maturity.

The process y has the same qualitative behavior as w in that they both proxy for both time to maturity and the risk-neutral expectation of remaining volatility. Recall that under SVRH, it is sufficient to use the pair (F, w) as state variables in place of the triple (F, v, t). Likewise, the time homogeneity of the y process makes it possible to use use the pair (F, y) as state variables in place of this triple. Furthermore, the much simpler form of the volatility of y suggests that the pair (F, y) should be used in place of (F, w).

The definition (83) of y makes it clear that it reduces to w when $q = 0$. This suggests that the complexity of w vanishes if we simply assume that the risk-neutral process for v has no quadratic component. Compared to modelling the y process, this third approach for dealing with the complexity of w has the disadvantage of imposing structure on the risk-neutral drift of v that may not be supported by the (options) data. However, it has the conceptual advantage of focussing our assumptions concerning price dynamics on a static position in variance swaps, rather than a dynamic one.

As the traditional approach in the asset pricing literature has been to place structure on the price dynamics of a static position, the next section elaborates on this final approach for dealing with the complexity of w and its volatility.

9 Simple variance swap rate dynamics

This section shows that setting $q = 0$ in the risk-neutral drift of v leads to much simpler valuation formulas and dynamics for w. In particular, the value function $w(v, t; T)$ becomes proportional to v, while the function $g(w; T)$ governing w's volatility becomes proportional to w. Setting $q = 0$ in the risk-neutral drift of v does imply that it no longer mean reverts to a positive level. However, it should be remembered that this property applies only to the risk-neutral process. We show that a standard specification for the market price of v risk results in a specification of the statistical v process which is just the quadratic drift 3/2 process discussed in the last section. Hence, the tractability of the risk-neutral process discussed in the last section extends to the statistical process to be discussed in this one. The quadratic drift in the statistical v process can be interpreted as providing mean-reversion to a positive level. The novelty compared to standard affine models is that the speed of mean-reversion is proportional to the level of v. As already mentioned, the martingale component of this statistical process has received much empirical support.

In this section, we assume that the dynamics of v under $\mathbb{Q}$ are given by:

$$dv_t = p(t)v_t dt + \epsilon v_t^{3/2} d\tilde{W}_t, \quad t \in [0, T]. \tag{90}$$

In other words, we have set $q = 0$ in the only v dynamics consistent with both our SVRH and the MIDH. Setting $q = 0$ in the ODE (81) implies that it simplifies to:

$$\frac{\epsilon^2 y^2}{2} \frac{\partial^2}{\partial y^2} h(y) - \frac{\partial}{\partial y} h(y) + 1 = 0, \quad y > 0, \tag{91}$$

A simple solution which meets the boundary conditions $h(0) = 0$ and $\lim_{y \downarrow 0} \frac{\partial}{\partial y} h(y) = 1$ is:

$$h(y) = y, \quad y > 0. \tag{92}$$

As a consequence, (77) implies that the variance swap rate is proportional to v:

$$w(t, v; T) = \int_t^T e^{\int_t^{t'} p(u)du} dt' v, \quad t \in [0, T], v \geq 0, T > 0. \tag{93}$$

Recall from the last section that the probability density function of v_T is known from the work of Cox et al. (1985), when the risk-neutral v process is time-homogeneous. It follows from (93) that the probability density function for w_T is also known by the change of variable formula, under time-homogeneity of the risk-neutral v process.

As (92) implies that $\frac{\partial}{\partial w}h(w) = 1$ and $h^{-1}(w) = w$, (80) implies that:

$$g(w) = \epsilon w, \quad w \geq 0, T > 0. \tag{94}$$

Hence, the risk-neutral w dynamics simplify to:

$$dw_t(T) = -v_t dt + \epsilon w_t(T)\sqrt{v_t} d\tilde{W}_t, \quad t \in [0, T]. \tag{95}$$

Since ϵ also appears in the v dynamics, it must be independent of T. Hence, our model (95) makes the strong empirical prediction that variance swap rates of different maturities have the same (lognormal) volatility. This feature is not shared by the more general solution of the last section.

As for the statistical process for v, we know that the volatility is the same under $\mathbb{P}$ and $\mathbb{Q}$. In contrast, the drift under the statistical measure $\mathbb{P}$ is:

$$E^{\mathbb{P}}[dv_t|v_t, t] = [p(t)v_t + \lambda(v_t, t)]dt. \tag{96}$$

Suppose that we specify the following quadratic form for the market price of v risk:

$$\lambda(v, t) = \lambda_0(t) + \lambda_1(t)v + \lambda_2(t)v^2, \quad v \geq 0, t \in [0, T]. \tag{97}$$

Then from (96) and (97), the statistical drift of v is

$$E^{\mathbb{P}}[dv_t|v_t, t] = \{\lambda_0(t) + \kappa(t)v_t[\theta(t) - v_t]\}dt, \tag{98}$$

where $\kappa(t) \equiv -\lambda_2(t)$ and $\theta(t) \equiv -\frac{p(t)+\lambda_1(t)}{\lambda_2(t)}$. We assume that $\lambda_2(t) < 0$ and $\lambda_1(t) > -p(t)$ so that $\kappa(t) > 0$ and $\theta(t) > 0$. Now recall that zero is a natural boundary for the risk-neutral process for v in (61). Thus a process which starts at zero stays there forever. However, if the statistical drift is given by (98), then a process which starts at zero moves away from zero unless:

$$\lambda_0(t) = 0. \tag{99}$$

Since $\mathbb{P}$ and $\mathbb{Q}$ must be equivalent probability measures, we set $\lambda_0(t) = 0$ and hence the statistical process for v simplifies to:

$$dv_t = \kappa(t)v_t[\theta(t) - v_t]dt + \epsilon v_t^{3/2} dW_t, \quad t \in [0, T]. \tag{100}$$

Hence, at each $t \in [0, T]$, the process is mean-reverting towards $\theta(t) > 0$ with a speed of mean-reversion $\kappa(t)v_t > 0$. As a result, we refer to the process described in (100) as the mean-reverting 3/2 process.

Finally, we turn to the statistical dynamics for the variance swap rate w. From (97) and (99), the absolute drift compensation at time t for exposure to $\epsilon v_t^{3/2} dW_t$

is $\lambda_1(t)v_t + \lambda_2(t)v_t^2$. Since the martingale component of dw_t is just $\frac{w_t}{v_t}$ times the martingale component of dv_t it follows that:

$$\pi_t^w w_t = \frac{w_t}{v_t}[\lambda_1(t)v_t + \lambda_2(t)v_t^2], \quad t \in [0, T]. \tag{101}$$

Simplifying implies that $w's$ relative risk premium is affine in v:

$$\pi_t^w = \lambda_1(t) + \lambda_2(t)v_t, \quad t \in [0, T]. \tag{102}$$

Restricting dynamics to the model specified in this section, (12) and (94) imply that the valuation function $\Pi(F, w)$ for a path-independent claim satisfies the following second order linear elliptic PDE:

$$\frac{F^2}{2}\frac{\partial^2\Pi}{\partial F^2}(F, w) + \rho\epsilon w F\frac{\partial^2\Pi}{\partial w\partial F}(F, w) + \frac{\epsilon^2 w^2}{2}\frac{\partial^2\Pi}{\partial w^2}(F, w) - \frac{\partial\Pi}{\partial w}(F, w) = 0, \tag{103}$$

on the domain $F > 0$, $w > 0$. We further require that $\Pi(F, w)$ satisfy the boundary conditions:

$$\Pi(F, 0) = f(F), \quad F > 0, \tag{104}$$

$$\Pi(0, w) = f(0), \quad w > 0, \tag{105}$$

and satisfy growth conditions as F and w become infinite.

Note that this valuation model is extremely parsimonious. The forward price of an option can be numerically valued once one knows its strike price, the futures price of the underlying asset, the initial variance swap rate of the same maturity, and the two parameters ρ and ϵ. The average level of the implied volatility smile of maturity T is positively related to the observable variance swap rate w_t. The average slope of the smile is positively related to the correlation parameter ρ. The average curvature of the smile is positively related to the volvol parameter ϵ.

One can use finite differences or finite elements to numerically solve the above boundary value problem. One can also use Monte Carlo simulation of just the variance swap rate over an infinite horizon as discussed in Sect. 6. Alternatively, one can use the Romano Touzi trick to just simulate v over a finite horizon using (61). The choice of whether to simulate w or v depends on how close w is to the origin.

Recall that the last section derived the joint Fourier Laplace transform of X_T and $\langle X\rangle_T$ under the quadratic drift 3/2 process for v. Of course, this formula can be applied to our current setting of proportional risk-neutral drift simply by setting $q = 0$. Furthermore, we can substitute out v and t for w to instead relate this transform and its derived quantities to the more observable variance swap rate. When calibrating an option pricing model, it is useful if call values can be efficiently expressed as a function of strike price. Since the characteristic function of $\ln F_T$ is a known analytic function of $\ln F$, w, and t, the results of Carr and Madan (1999) imply that the fast Fourier transform (FFT) can be used to efficiently obtain call values as a function of log strike. As shown in Chourdakis (2004), the fractional FFT can be used to speed up results even further.

10 Pricing and hedging volatility derivatives

10.1 Pricing and hedging derivatives on realized variance

Before variance swaps became active, swaps were traded over-the-counter on realized volatility. Since volatility is the square root of realized variance, a volatility swap is a derivative security on realized variance. Recently, options on realized variance and options on realized volatility also became available.

To value derivatives on realized variance, recall the assumptions made on the futures price F and the variance swap rate w under the statistical measure $\mathbb{P}$:

$$\frac{dF_t}{F_t} = \mu_t dt + \sqrt{v_t} dB_t, \tag{106}$$

$$dw_t = (\pi_t^w w_t - v_t)dt + g(w_t)\sqrt{v_t} dW_t, \quad t \in [0, T], \tag{107}$$

where:

$$dB_t dW_t = \rho dt, \quad t \in [0, T]. \tag{108}$$

It turns out that to value derivatives on realized variance, only the w dynamics are relevant. Hence for the rest of the paper, we can let the futures price jump and we need no assumption on the nature of the correlation between the two Brownian motions B and W.

Let $X_t \equiv \ln(F_t/F_0)$ be the log price relative and let $\langle X\rangle_t = \int_0^t v_s ds$ be its quadratic variation. We assume that the payoff on the volatility derivative is some known function of just $\langle X\rangle_T$. Let $\mathcal{V}(w, q) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}$ be a $C^{2,1}$ function and let $\mathcal{V}_t \equiv \mathcal{V}(w_t, \langle X\rangle_t)$ be a continuous stochastic process. Itô's formula implies that:

$$\mathcal{V}(w_T, \langle X\rangle_T) = \mathcal{V}(w_0, 0) + \int_0^T \frac{\partial}{\partial w}\mathcal{V}(w_t, \langle X\rangle_t)dw_t + \int_0^T \frac{\partial}{\partial q}\mathcal{V}(w_t, \langle X\rangle_t)d\langle X\rangle_t + \int_0^T \frac{g^2(w_t)}{2}\frac{\partial^2 \mathcal{V}}{\partial w^2}(w_t, \langle X\rangle_t)v_t dt. \tag{109}$$

Suppose that we add and subtract $\int_0^T \frac{\partial}{\partial w}\mathcal{V}(F_t, w_t)v_t dt$. Then $\mathcal{V}(w_T, \langle X\rangle_T)$:

$$= \mathcal{V}(w_0, 0) + \int_0^T \frac{\partial}{\partial w}\mathcal{V}(w_t, \langle X\rangle_t)(dw_t + v_t dt) + \int_0^T \left[\frac{\partial}{\partial q}\mathcal{V}(w_t, \langle X\rangle_t) + \frac{g^2(w_t)}{2}\frac{\partial^2}{\partial w^2}\mathcal{V}(w_t, \langle X\rangle_t) - \frac{\partial}{\partial w}\mathcal{V}(w_t, \langle X\rangle_t)\right] v_t dt. \tag{110}$$

Suppose that the function $\mathcal{V}(F, w)$ solves the following second order linear parabolic PDE:

$$\frac{\partial \mathcal{V}}{\partial q}(w, q) + \frac{g^2(w)}{2}\frac{\partial^2 \mathcal{V}}{\partial w^2}(w, q) - \frac{\partial \mathcal{V}}{\partial w}(w, q) = 0, \quad w > 0, q > 0. \tag{111}$$

Further suppose that:

$$\mathcal{V}(0, q) = f(q), \quad q > 0. \tag{112}$$

Since $w_T = 0$, substitution of (111) and (112) in (110) implies:

$$f(\langle X \rangle_T) = \mathcal{V}(w_0, 0) + \int_0^T \frac{\partial}{\partial w}\mathcal{V}(w_t, \langle X \rangle_t)(dw_t + v_t dt). \tag{113}$$

Thus the payoff on the volatility derivative can be exactly replicated if one charges $\mathcal{V}(w_0, 0)$ initially and holds $\frac{\partial}{\partial w}\mathcal{V}(w_t, \langle X \rangle_t)$ variance swaps at each $t \in [0, T]$.

As in the last section, time is irrelevant because the realized volatility derivative matures at the same time as the variance swap. The boundary value problem (111) and (112) has the same computational complexity as the one arising under a standard stochastic volatility model. However, once again we do not need to know the dynamics of the instantaneous variance or know the market price of volatility risk.

If finite differences are used to numerically determine the value function $\mathcal{V}(w, q)$ on a grid, then one usually supplies 3 boundary conditions. To illustrate, suppose we wish to numerically determine the value of a put on realized variance. Then one boundary condition is:

$$\mathcal{V}(0, q) = \left(K - \frac{q}{T}\right)^+, \quad q > 0. \tag{114}$$

As w gets large, it will take a lot of realized quadratic variation to bring it down to zero and hence:

$$\lim_{w \uparrow \infty} \mathcal{V}(0, q) = 0, \quad q > 0. \tag{115}$$

Finally, since quadratic variation can only increase, the final boundary condition is:

$$\mathcal{V}(w, KT) = 0, \quad w > 0. \tag{116}$$

The PDE can be solved on the domain $q \in [0, KT]$, $w > 0$.

When we specialize to the dynamics of the last section, the valuation PDE simplifies to:

$$\frac{\partial \mathcal{V}}{\partial q}(w, q) + \frac{\epsilon^2 w^2}{2}\frac{\partial^2 \mathcal{V}}{\partial w^2}(w, q) - \frac{\partial \mathcal{V}}{\partial w}(w, q) = 0, \quad w > 0, q > 0. \tag{117}$$

If one wants a supporting SV model where the risk-neutral process for v is a diffusion with maturity independent coefficients, then one should also assume a continuous futures price process, constant correlation, and the quadratic drift 3/2 process (61) for instantaneous variance.

It is also possible to write down the fundamental PDE operating for a European-style claim when the terminal payoff depends on both the spot price and the realized variance of returns. For example, the payoff could be the realized Sharpe ratio. In this case, we need to assume the continuous dynamics in (106) to (108). There are three independent variables F, w, and q in the PDE and it also has the same form as the PDE arising in a stochastic volatility model. Since we are not yet aware of any contracts with this complexity, we leave the details to the reader.

10.2 Pricing and hedging derivatives on the variance swap rate

The CBOE recently revamped the definition of its well known volatility index known as the VIX. The new VIX is an average of two synthetic variance swap rates with the weights deterministically varying over time in such a way that the time to maturity is fixed at 30 days. Once a month, only one synthetic variance swap rate is used to calculate the VIX and by design, this date is also the maturity date for VIX futures, which began trading in March 2004. It turns out that VIX futures at time t are valued by determining the risk-neutral mean of $\sqrt{w_S}$, where $S \in (t, T)$ is the futures maturity date. Hence, VIX futures is an example of a derivative security written on a variance swap rate.

To value derivatives of this kind, let s_t be the value at time $t \in [0, S]$ of the variance swap rate for maturity S:

$$s_t = E^{\mathbb{Q}} \int_t^S v_u du, \quad t \in [0, S]. \tag{118}$$

Working under $\mathbb{P}$, we assume that the dynamics of this variance swap rate are given by:

$$ds_t = (\pi_t^s s_t - v_t)dt + h(s_t)\sqrt{v_t} dB_t, \quad t \in [0, S], \tag{119}$$

where B is a standard Brownian motion under $\mathbb{P}$. Here, π^s is a stochastic process representing compensation for $h(s_t)$ differing from zero. Also, $h(s)$ is a function of just s and is independent of t and v. We continue to assume that the statistical process for w is given by (9), which we repeat here for convenience:

$$dw_t = (\pi_t^w w_t - v_t)dt + g(w_t)\sqrt{v_t} dW_t, \quad t \in [0, S], \tag{120}$$

where $g(w)$ is a function of just w and is independent of t and v. We assume that the correlation between the increments of the two variance swap rates is a function of the two rates:

$$dB_t dW_t = \rho(s_t, w_t)dt, \quad t \in [0, S]. \tag{121}$$

We do not need to specify the process for the futures price F or the correlation that the Brownian motion Z driving it has with B or W.

Let $\mathcal{V}(s,w) : \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}$ be a $C^{2,2}$ function and let $\mathcal{V}_t \equiv \mathcal{V}(s_t, w_t)$ be a continuous stochastic process. Itô's formula implies that:

$$\begin{aligned}\mathcal{V}(s_S, w_S) &= \mathcal{V}(s_0, w_0) + \int_0^S \frac{\partial}{\partial s}\mathcal{V}(s_t, w_t)ds_t + \int_0^S \frac{\partial}{\partial w}\mathcal{V}(s_t, w_t)dw_t \\ &\quad + \int_0^S \left[\frac{h^2(s_t)}{2}\frac{\partial^2 \mathcal{V}}{\partial s^2}(s_t, w_t) + \rho(s_t, w_t)h(s_t)g(w_t)\frac{\partial^2 \mathcal{V}}{\partial s \partial w}(s_t, w_t)\right. \\ &\quad \left. + \frac{g^2(w_t)}{2}\frac{\partial^2 \mathcal{V}}{\partial w^2}(s_t, w_t)\right] v_t dt. \end{aligned} \tag{122}$$

Suppose that we add and subtract $\int_0^S \frac{\partial}{\partial s}\mathcal{V}(s_t, w_t)v_t dt$ and $\int_0^S \frac{\partial}{\partial w}\mathcal{V}(s_t, w_t)v_t dt$. Then $\mathcal{V}(s_S, w_S)$:

$$\begin{aligned}&= \mathcal{V}(s_0, w_0) + \int_0^S \frac{\partial}{\partial s}\mathcal{V}(s_t, w_t)(ds_t + v_t dt) + \int_0^S \frac{\partial}{\partial w}\mathcal{V}(s_t, w_t)(dw_t + v_t dt) \\ &\quad + \int_0^S \left[\frac{h^2(s_t)}{2}\frac{\partial^2}{\partial s^2}\mathcal{V}(s_t, w_t) + \rho(s_t, w_t)h(s_t)g(w_t)\frac{\partial^2 \mathcal{V}}{\partial s \partial w}(s_t, w_t)\right. \\ &\quad \left. + \frac{g^2(w_t)}{2}\frac{\partial^2}{\partial w^2}\mathcal{V}(s_t, w_t) - \frac{\partial}{\partial s}\mathcal{V}(s_t, w_t) - \frac{\partial}{\partial w}\mathcal{V}(s_t, w_t)\right] v_t dt. \end{aligned} \tag{123}$$

Suppose that the function $\mathcal{V}(s,w)$ solves the following second order linear elliptic PDE:

$$\begin{aligned}&\frac{h^2(s)}{2}\frac{\partial^2 \mathcal{V}}{\partial s^2}(s,w) + \rho(s,w)h(s)g(w)\frac{\partial^2 \mathcal{V}}{\partial s \partial w}(s,w) + \frac{g^2(w)}{2}\frac{\partial^2 \mathcal{V}}{\partial w^2}(s,w) \\ &\quad - \frac{\partial \mathcal{V}}{\partial s}(s,w) - \frac{\partial \mathcal{V}}{\partial w}(s,w) = 0, \end{aligned} \tag{124}$$

on the domain $s > 0$, $w > 0$. Further suppose that:

$$\mathcal{V}(0, w) = h(w), \quad w > 0. \tag{125}$$

For example, ignoring normalization constants, the boundary condition for VIX futures is:

$$\mathcal{V}(0, w) = \sqrt{\frac{w}{T - S}}, \quad w > 0. \tag{126}$$

Since $s_S = 0$, substitution of (124) and (125) in (123) implies:

$$h(w_S) = \mathcal{V}(s_0, w_0) + \int_0^S \frac{\partial}{\partial s}\mathcal{V}(s_t, w_t)(ds_t + v_t dt) + \int_0^S \frac{\partial}{\partial w}\mathcal{V}(s_t, w_t)(dw_t + v_t dt). \tag{127}$$

Thus, the payoff on the variance swap derivative can be exactly replicated if one charges $\mathcal{V}(s_0, w_0)$ initially and one holds $\frac{\partial \mathcal{V}}{\partial s}(s_t, w_t)$ variance swaps of maturity S and $\frac{\partial \mathcal{V}}{\partial w}(s_t, w_t)$ variance swaps of maturity T at each $t \in [0, S]$.

Note once again that we do not have to specify the process for v. This robustness survives even if g depends on s and/or h depends on w, but these dependencies seem strained. An example of a supporting stochastic volatility model is when the statistical process for v is the mean-reverting 3/2 process (100). In this case, $W = B$ and hence $\rho(s, w) = 1$. The bivariate statistical process for (s, w) becomes:

$$\begin{aligned} ds_t &= (\pi_t^s s_t - v_t)dt + \epsilon s_t \sqrt{v_t} dW_t, \\ dw_t &= (\pi_t^w w_t - v_t)dt + \epsilon w_t \sqrt{v_t} dW_t, \quad t \in [0, S]. \end{aligned} \tag{128}$$

11 Summary and future research

This paper showed that the payoff to a path-independent claim of maturity T can be replicated by assuming continuous price processes and continuous trading opportunities for the T maturity futures price and the T maturity variance swap rate. Furthermore, so long as the martingale component of the statistical process of the variance swap rate satisfies our SVRH as specified in (5), then one can price the claim relative to these observables without knowledge of the level or statistical process for the instantaneous volatility. We showed how to calibrate the pricing model to a given array of European option prices of all strikes and maturities. Whether or not the model is calibrated to these options, valuation requires solving a boundary value problem containing a second order linear elliptic PDE in just two independent variables. This problem can be numerically solved using finite differences, finite elements, or by Monte Carlo simulation.

Furthermore, we investigated the implications of requiring both SVRH for the variance swap rate and a maturity independent risk-neutral diffusion process for the instantaneous variance rate v. We showed that this latter diffusion must have risk-neutral drift of the form $p(t)v_t + qv_t^2$ and must have a diffusion coefficient proportional to $v_t^{3/2}$. Although this specification was achieved purely from theoretical considerations, it has received a surprisingly large amount of empirical support. We showed that this process has sensible behavior provided that $q < \frac{\epsilon^2}{2}$ and $\rho < \frac{1}{2\epsilon}$. Both of these constraints are automatically met if the v process is mean-reverting and if the leverage effect is present. This v process is flexible in that there is a free function of time $p(t)$ in the risk-neutral drift, which can be used to match a term structure. Although this free function is equivalent to an infinite number of parameters, the proposed v process is still extremely tractable. In particular, we showed that the joint Fourier Laplace transform of returns and their quadratic variation can be derived in closed form. We also derived closed form formulas for the variance swap rate and its volatility. To simplify these latter formulas, we focussed on a special case where the risk-neutral drift for v is linear in v. In this case, the variance swap rate is proportional to v and its volatility is proportional to the instantaneous volatility of the underlying futures. Although this simple risk-neutral v process does not accomodate mean-reversion to a positive level, we showed that the corresponding statistical v process can accomodate this desirable property, provided that the speed of mean-reversion is proportional to the level of the process.

Future research should explore the extension of these results to other path-dependent options such as barrier and Asian options. As it is likely that numerical solution will be required to value these exotic options, it is worth noting that the elimination of time to maturity as an independent variable in the PDE speeds up valuation by an order of magnitude. The extension to finite-lived American options requires adding time as a state variable unless the exercise boundary can be expressed as a function of w.

We note that it is straightforward numerically to allow the correlation between returns and increments in the variance swap rate to be a function of the futures price and the variance swap rate. Since this correlation also describes the correlation between returns and increments in the instantaneous variance, the dependence of this correlation on the variance swap rate can be debated.

We could also consider the application of our result to other kinds of problems. The process y is a bridge process which stays positive prior to T and vanishes right at T. Perhaps, this process could be used to describe the dynamics of a finite-lived annuity under stochastic interest rates. Alternatively, it could be used to model the difference between the face value and present value of a zero coupon bond.

For a fixed maturity T, we have shown that is possible to model a scalar bridge process with stationary dynamics. One can try to extend these results to a multivariate setting. If these results are extended to modeling a continuum, one could evolve a cross section of equal maturity objects. For example, one could evolve the strike structure of option prices by separate modelling of each option's intrinsic and time value. Note that the strike structure of time value is just the strike structure of expected local time. Alternatively, one could model the evolution of a conditional characteristic function by modelling the required convexity correction at each level of the Fourier argument. In the first case, the randomness induced perturbation to intrinsic value is additive and positive, while in the second case it is multiplicative and negative. As it is desirable to impose cross sectional restrictions in either case, these extensions are best left for future research.

Acknowledgements We thank Claudio Albanese, Leif Andersen, Gurdip Bakshi, Bruno Dupire, Paul Feehan, Pat Hagan, Jeong Hyun Lee, Alireza Javaheri, Dilip Madan, Chris Potter, Steve Shreve, Arun Verma, Anna Wen, Liuren Wu, and Ernest Zhu for their comments. They are not responsible for any errors.

Appendix 1: Proof of Theorem 1

This appendix proves that

$$\frac{\partial^n}{\partial F^n}\Pi(F,w) = E^{\hat{Q}_n}\left[f^{(n)}(\hat{F}_\tau)e^{\frac{n(n-1)}{2}\tau}|\hat{F}_0 = F, \hat{w}_0 = w\right], \quad n = 0, 1, 2\ldots, \tag{129}$$

where under the measure $\hat{Q}_n$, the process $\hat{F}_t$ solves the SDE:

$$\frac{d\hat{F}_t}{\hat{F}_t} = ndt + \rho dZ_{1t}^{(n)} + \sqrt{1-\rho^2}dZ_{2t}^{(n)}, \quad t > 0, \tag{130}$$

and the process $\hat{w}_t$ solves the SDE:

$$d\hat{w}_t = [n\rho g(\hat{w}_t) - 1]dt + g(\hat{w}_t)dZ_{1t}^{(n)}, \quad t > 0, \tag{131}$$

where $Z_1^{(n)}$ and $Z_2^{(n)}$ are independent standard Brownian motions under the probability measure $\hat{Q}_n$ and τ is the first passage time of $\hat{w}$ to the origin.

The proof is by induction. For $n = 0$, the result holds by the Feynman Kac Theorem. Now suppose that (129)–(131) holds for some fixed n. To show that it also holds for $n + 1$, differentiate (129) w.r.t. F:

$$\frac{\partial^{n+1}}{\partial F^{n+1}}\Pi(F, w) = E^{\hat{Q}_n}\left[f^{(n+1)}(\hat{F}_\tau)\frac{\partial \hat{F}_\tau}{\partial F}e^{\frac{n(n-1)}{2}\tau}|\hat{F}_0 = F, \hat{w}_0 = w\right],$$
$$n = 0, 1, 2 \ldots, \tag{132}$$

by the chain rule. From (40):

$$\frac{\partial \hat{F}_\tau}{\partial F} = e^{n\tau}M_\tau, \tag{133}$$

where:

$$M_t \equiv e^{-\frac{1}{2}t+\rho Z_{1,t}^{(n)}+\sqrt{1-\rho^2}dZ_{2,t}^{(n)}}, \quad t > 0 \tag{134}$$

is a positive $\hat{Q}_n$ martingale with mean one. Substituting (133) in (129) and using M_τ to change measures from $\hat{Q}_n$ to $\hat{Q}_{n+1}$ implies:

$$\frac{\partial^{n+1}}{\partial F^{n+1}}\Pi(F, w) = E^{\hat{Q}_{n+1}}\left[f^{(n+1)}(\hat{F}_\tau)e^{\frac{n(n+1)}{2}\tau}|\hat{F}_0 = F, \hat{w}_0 = w\right], \quad n = 0, 1, 2 \ldots, \tag{135}$$

where under the measure $\hat{Q}_{n+1}$, Girsanov's theorem implies that the process $\hat{F}_t$ solves the SDE:

$$\begin{aligned}\frac{d\hat{F}_t}{\hat{F}_t} &= ndt + \frac{1}{\hat{F}_t M_t}d\langle \hat{F}, M\rangle_t + \rho dZ_{1t}^{(n+1)} + \sqrt{1-\rho^2}dZ_{2t}^{(n+1)}\\ &= (n+1)dt + \rho dZ_{1t}^{(n+1)} + \sqrt{1-\rho^2}dZ_{2t}^{(n+1)}, \quad t > 0,\end{aligned} \tag{136}$$

and the process $\hat{w}_t$ solves the SDE:

$$d\hat{w}_t = [n\rho g(\hat{w}_t) - 1]dt + \frac{1}{M_t} d\langle \hat{w}, M\rangle_t + g(\hat{w}_t) dZ_{1t}^{(n+1)}$$
$$= [(n+1)\rho g(\hat{w}_t) - 1]dt + g(\hat{w}_t) dZ_{1t}^{(n+1)}, \quad t > 0, \tag{137}$$

where $Z_1^{(n+1)}$ and $Z_2^{(n+1)}$ are independent standard Brownian motions under the probability measure $\hat{Q}_{n+1}$.

Appendix 2: Proof of Theorem 2

Recall the parabolic PDE governing the variance swap rate w:

$$\frac{\partial}{\partial t} w(t, v; T) + \frac{b^2(t, v)v}{2} \frac{\partial^2}{\partial v^2} w(t, v; T) + a(t, v) \frac{\partial}{\partial v} w(t, v; T) + v = 0, \tag{138}$$

on the domain $v > 0$, $t \in [0, T]$, and the terminal condition:

$$w(T, v; T) = 0, \quad v > 0. \tag{139}$$

Differentiating (139) w.r.t. v implies:

$$\frac{\partial}{\partial v} w(T, v; T) = 0, \quad v > 0, \tag{140}$$

and differentiating again implies:

$$\frac{\partial^2}{\partial v^2} w(T, v; T) = 0, \quad v > 0. \tag{141}$$

Evaluating the PDE (138) at $t = T$ and substituting in (140) and (141) implies:

$$\frac{\partial}{\partial t} w(T, v; T) = -v, \quad v > 0. \tag{142}$$

Differentiating (142) w.r.t. v implies:

$$\frac{\partial^2}{\partial t \partial v} w(T, v; T) = -1, \quad v > 0. \tag{143}$$

Now, recall from (57) that we want:

$$b(t, v) = \frac{g(w(t, v; T); T)}{\frac{\partial}{\partial v} w(t, v; T)}, \quad t \in [0, T], v > 0. \tag{144}$$

If we try to evaluate this expression at time T, then in order for b to be finite, we must have:

$$g(0; T) = 0, \quad T > 0, \tag{145}$$

from (139) and (140). Assuming that $g(0) = 0$, L'Hospital's rule implies:

$$b(T, v) = \frac{\partial}{\partial w} g(0; T) \frac{w_t(T, v; T)}{w_{tv}(T, v; T)} = \frac{\partial}{\partial w} g(0; T) v, \quad v > 0 \tag{146}$$

from (142) and (143). However, since b is independent of the specific maturity T, it must have the same form prior to T as it has at T, i.e.:

$$b(t, v) = \frac{\partial}{\partial w} g(0; T) v, \quad t \in [0, T], \tag{147}$$

where:

$$\frac{\partial}{\partial w} g(0; T) = \epsilon, \tag{148}$$

with ϵ independent of t and T.

Having determined the required form of the risk-neutral diffusion coefficient of v, we now turn to the problem of determining the required form of the risk-neutral drift of v. By definition, $w_t(T)$ is just the risk-neutral expected value of the area under the path of $\{v_u, u \in [t, T]\}$. Hence by a coupling argument, the function $w(t, v; T) : [0, T) \times \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ is increasing in v at each $t \in [0, T)$, $T > 0$. This allows us to define an inverse function $v(t, w; T) : [0, T) \times \mathbb{R}^+ \times \mathbb{R}^+ \mapsto \mathbb{R}^+$ with the property that:

$$v(t, w(t, v; T; T) = v, \quad t \in [0, T], T > 0. \tag{149}$$

Since $w(t, 0; T) = 0$, we have:

$$v(t, 0; T) = 0, \quad t \in [0, T), T > 0. \tag{150}$$

Since $\frac{\partial}{\partial v} w(t, v; T) = \frac{1}{\frac{\partial}{\partial w} v(t, w; T)}$ on $t \in [0, T)$, $w \geq 0$, $v \geq 0$, $T > 0$, (57) can be rewritten as:

$$\frac{\frac{\partial}{\partial w} v(t, w; T)}{v(t, w; T)} = \frac{\epsilon}{g(w; T)}, \quad t \in [0, T), w \geq 0, T > 0. \tag{151}$$

Integrating w.r.t. w:

$$\ln v(t, w; T) = G(w; T) + y(t; T), \quad t \in [0, T), w \geq 0, T > 0, \tag{152}$$

where for each fixed $T > 0$, $G(w;T)$ is an anti-derivative of $\frac{\epsilon}{g(w;T)}$ considered as a function of w, i.e.:

$$G(w;T) \equiv \int_{-w}^{1} \frac{\epsilon}{g(z;T)} dz, \quad w \geq 0, T > 0, \tag{153}$$

and $y(t;T)$ is the arbitrary constant of integration. Evaluating (152) at $w = 0$ implies that:

$$G(0;T) = -\infty, \quad T > 0, \tag{154}$$

from (150).

Exponentiating (152) implies:

$$v(t,w;T) = \frac{e^{G(w;T)}}{x(t;T)}, \quad t \in [0,T), w \geq 0, T > 0, \tag{155}$$

where $x(t;T) \equiv e^{-y(t;T)}$. Now $v(t,w;T)$ solves the following nonlinear PDE:

$$\frac{\partial}{\partial t} v(t,w;T) + \frac{g^2(w;T)}{2} v(t,w;T) \frac{\partial^2}{\partial w^2} v(t,w;T) - v(t,w;T) \frac{\partial}{\partial w} v(t,w;T) \\ = a(t, v(t,w;T)), \tag{156}$$

on the domain $t \in [0,T)$, $w \geq 0$, $T > 0$. Differentiating (155) w.r.t. t implies:

$$\begin{aligned} \frac{\partial}{\partial t} v(t,w;T) &= -\frac{\frac{\partial}{\partial t} x(t;T)}{x^2(t;T)} e^{G(w;T)} \\ &= -\frac{\frac{\partial}{\partial t} x(t;T)}{x(t;T)} v(t,w;T), \quad t \in [0,T), w \geq 0, T > 0, \end{aligned} \tag{157}$$

from (155). Differentiating (155) w.r.t. w instead implies:

$$\begin{aligned} \frac{\partial}{\partial w} v(t,w;T) &= v(w,t;T) \frac{\partial}{\partial w} G(w;T) \\ &= v(t,w;T) \frac{\epsilon}{g(w;T)}, \quad t \in [0,T), w \geq 0, T > 0, \end{aligned} \tag{158}$$

from (153). Differentiating w.r.t. w one more time implies:

$$\begin{aligned} \frac{\partial^2}{\partial w^2} v(w,t;T) &= \frac{\partial}{\partial w} v(t,w;T) \frac{\epsilon}{g(w;T)} - v(t,w;T) \frac{\epsilon}{g^2(w;T)} \frac{\partial}{\partial w} g(w;T) \\ &= v(t,w;T) \left[\frac{\epsilon}{g(w;T)} \right]^2 - v(t,w;T) \frac{\epsilon}{g^2(w;T)} \frac{\partial}{\partial w} g(w;T), \end{aligned}$$

on $t \in [0, T)$, $w \geq 0$, $T > 0$ from (158). Multiplying by $g^2(w; T)$ implies that:

$$g^2(w; T)\frac{\partial^2}{\partial w^2}v(t, w; T) = \epsilon v(t, w; T)\left[\epsilon - \frac{\partial}{\partial w}g(w; T)\right],$$
$$t \in [0, T), w \geq 0, T > 0. \tag{159}$$

Substituting (157), (158), and (159) in (156) implies:

$$a(t, v(t, w; T)) = -\frac{\frac{\partial}{\partial t}x(t; T)}{x(t; T)}v(t, w; T) + \frac{\epsilon\left[\epsilon - \frac{\partial}{\partial w}g(w; T)\right]}{2}v^2(t, w; T)$$
$$-v^2(t, w; T)\frac{\partial}{\partial w}G(w; T), \tag{160}$$

on $t \in [0, T)$, $w \geq 0$, $T > 0$. Substituting in (153) and dividing by $v(t, w; T)$:

$$\begin{aligned}\frac{a(t, v(t, w; T))}{v(t, w; T)} &= -\frac{\frac{\partial}{\partial t}x(t; T)}{x(t; T)} + \varepsilon[\varepsilon - g_w(w; T)]\frac{v(t, w; T)}{2} - \frac{\varepsilon}{g(w; T)}v(t, w; T)\\ &= -\frac{\frac{\partial}{\partial t}x(t; T)}{x(t; T)} + \varepsilon[\varepsilon - g_w(w; T)]\frac{v(t, w; T)}{2} - \frac{1}{x(t; T)}\frac{\varepsilon e^{G(w;T)}}{g(w; T)},\end{aligned} \tag{161}$$

from (155).

To analyze further, we need a more explicit form for the function G. Equations (145) and (148) imply that we can write:

$$g(w; T) = \varepsilon w g_1(w; T), \quad w > 0, T > 0, \tag{162}$$

where $g_1(w; T)$ is a smooth function satisfying:

$$g_1(0; T) = 1, \quad T > 0. \tag{163}$$

Substituting (162) in (153) implies:

$$\begin{aligned}G(w; T) &= -\int_w^1 \frac{1}{y g_1(y; T)}dy\\ &= -\int_w^1 \frac{1}{y}dy + \int_w^1 \frac{g_1(y; T) - 1}{y g_1(y; T)}dy.\end{aligned} \tag{164}$$

But now the function:

$$k_w(y; T) \equiv \frac{g_1(y; T) - 1}{y g_1(y; T)} \tag{165}$$

is a bounded function of y in the interval $[0, 1]$ since $g_1(0) = 1$. The boundedness implies that $k_w(y; T)$ is also an integrable function of y in the interval $[0, 1]$, and hence (164) implies that:

$$G(w;T) = \ln w + k(w;T), \quad w > 0, T > 0, \tag{166}$$

where $k(w;T)$ is some smooth function. Exponentiating (166) implies:

$$e^{G(w;T)} = we^{k(w;T)}, \quad w > 0, T > 0. \tag{167}$$

Consider the factor multiplying $\frac{1}{x(t;T)}$ in the last term in (161). Now, (162) and (167) imply that:

$$\lim_{w\downarrow 0} \frac{\varepsilon e^{G(w;T)}}{g(w;T)} = \lim_{w\downarrow 0} \frac{\varepsilon w e^{k(w;T)}}{g(w;T)} = \lim_{w\downarrow 0} \frac{e^{k(w;T)}}{g_1(w;T)} = e^{k(0;T)}, \tag{168}$$

from (163).

As $w \downarrow 0$, substituting (148) and (168) in (161) implies that:

$$\lim_{w\downarrow 0} \frac{a(t, v(t,w;T))}{v(t,w;T)} = -\frac{\frac{\partial}{\partial t}x(t;T)}{x(t;T)} - \frac{e^{k(0;T)}}{x(t;T)}. \tag{169}$$

Now, the MIDH requires that the function $a(t,v)$ is independent of T, so suppose that we choose $x(t;T)$ so that:

$$-\frac{\frac{\partial}{\partial t}x(t;T)}{x(t;T)} - \frac{e^{k(0;T)}}{x(t;T)} = p(t), \tag{170}$$

where $p(t)$ is independent of T. Then $x(t;T)$ solves the ordinary differential equation:

$$\frac{\partial}{\partial t}x(t;T) + x(t;T)p(t) + e^{k(0;T)} = 0, \quad t \in [0,T]. \tag{171}$$

To obtain a terminal condition for x, let us assume that $g(w;T) > 0$ for $w > 0$, $T > 0$. Since $\epsilon > 0$, it follows that $\frac{\epsilon}{g(w;T)} > 0$ for $w > 0$, $T > 0$. Hence, (153) implies that $G(w;T)$ is increasing in w for each $T > 0$. It follows that for each T, the inverse function G^{-1} exists. From (154):

$$G^{-1}(-\infty;T) = 0, \quad T > 0. \tag{172}$$

Solving (155) for w implies:

$$w(t,v;T) = G^{-1}(\ln(x(t;T)v);T), \quad t \in [0,T), v \geq 0, T > 0. \tag{173}$$

Evaluating (173) at $t = T$ and using (54) and (172) implies that our desired terminal condition is $x(T; T) = 0$. Substituting (170) in (161) implies:

$$\frac{a(t, v(t, w; T))}{v(t, w; T)} = p(t) + \left(\frac{e^{k(0;T)}}{x(t; T)} - \frac{1}{x(t; T)} \frac{\varepsilon e^{G(w;T)}}{g(w; T)} \right) + \varepsilon(\varepsilon - g_w(w; T)) \frac{v(t, w; T)}{2}. \tag{174}$$

Dividing by $v(t, w; T)$ implies:

$$\begin{aligned} \frac{a(t, v(t, w; T))}{v^2(t, w; T)} &= \frac{p(t)}{v(t, w; T)} + \frac{e^{k(0;T)}}{x(t; T)v(t, w; T)} - \frac{1}{x(t; T)v(t, w; T)} \frac{\varepsilon e^{G(w;T)}}{g(w; T)} \\ &\quad + \frac{\varepsilon}{2}(\varepsilon - g_w(w; T)) \\ &= \frac{p(t)}{v(t, w; T)} + \frac{e^{k(0;T)}}{e^{G(w;T)}} - \frac{1}{e^{G(w;T)}} \frac{\varepsilon e^{G(w;T)}}{g(w; T)} + \frac{\varepsilon}{2}(\varepsilon - g_w(w; T)), \end{aligned} \tag{175}$$

from (155).

Now letting $t \uparrow T$ and $w \downarrow 0$, we will have $g(w; T) \downarrow 0$ from (145) and $g_w(w; T) \rightarrow \varepsilon$ by (148). Therefore:

$$\begin{aligned} \frac{a(T, v(T, 0; T))}{v^2(T, 0; T)} &= \frac{p(T)}{v(T, 0; T)} + \lim_{w \downarrow 0} \left[\frac{e^{k(0;T)}}{e^{G(w;T)}} - \frac{\varepsilon e^{G(w;T)}}{e^{G(w;T)} g(w; T)} \right] \\ &= \frac{p(T)}{v(T, 0; T)} + \lim_{w \downarrow 0} \left[\frac{e^{k(0;T)} - e^{k(w;T)}/g_1(w; T)}{w e^{k(w;T)}} \right], \end{aligned} \tag{176}$$

from (167) and (162). Let:

$$g_2(w; T) \equiv 1/g_1(w; T). \tag{177}$$

Hence, the last term in (176) can be written as:

$$\begin{aligned} \lim_{w \downarrow 0} \frac{e^{k(0;T)} - e^{k(w;T)}/g_1(w; T)}{w e^{k(w;T)}} &= \lim_{w \downarrow 0} \frac{e^{k(0;T)} - e^{k(w;T)} g_2(w; T)}{w e^{k(w;T)}} \\ &= \lim_{w \downarrow 0} \frac{e^{k(0;T)} - e^{k(w;T)}}{w e^{k(w;T)}} + \frac{1 - g_2(w; T)}{w}. \end{aligned} \tag{178}$$

Substituting (163) in (177) implies that:

$$\lim_{w \downarrow 0} g_2(w; T) = 1. \tag{179}$$

Substituting (179) in (178) implies that:

$$\lim_{w\downarrow 0}\frac{e^{k(0;T)}-e^{k(w;T)}/g_1(w;T)}{we^{k(w;T)}}=-k'(0;T)-g_2'(0;T)\equiv q,\quad T>0,\quad (180)$$

where the required maturity independence of a implies that q is independent of T and t. Substituting (180) and (149) in (176):

$$a(T,v)=p(T)v+qv^2,\quad v>0,T>0. \tag{181}$$

However, the required maturity independence of a implies that for $t\in[0,T]$, the risk-neutral drift $a(t,v)$ has the same form, i.e.:

$$a(t,v)=p(t)v+qv^2,\quad t\in[0,T],v>0. \tag{182}$$

This completes the proof of Theorem 2.

Appendix 3: Proof of Theorem 3

Let $\mathbb{Q}$ denote risk-neutral measure and let the futures price process be continuous:

$$dF_t=\sqrt{v_t}F_t d\tilde{Z}_t,\quad t\in[0,T], \tag{183}$$

where v_t is the instantaneous variance at time t and $\tilde{Z}$ is a $\mathbb{Q}$ standard Brownian motion. We assume that the risk-neutral process for instantaneous variance is:

$$dv_t=\left[p(t)v_t+qv_t^2\right]dt+\epsilon v_t^{3/2}d\tilde{W}_t,\quad t\in[0,T], \tag{184}$$

where $\tilde{W}$ is a $\mathbb{Q}$ standard Brownian motion. Here, $p(t)$ is an arbitrary function of time, q is an arbitrary real scalar, and ϵ is a positive constant. The increments in the two $\mathbb{Q}$ standard Brownian motions have arbitrary constant correlation $\rho\in[-1,1]$ at time $t\in[0,T]$:

$$d\tilde{Z}_t d\tilde{W}_t=\rho dt,\quad t\in[0,T]. \tag{185}$$

Let $X_t\equiv\ln\left(\frac{F_t}{F_0}\right)$ be the log price relative. Then by applying Itô's formula to (183):

$$dX_t=-\frac{v_t}{2}dt+\sqrt{v_t}d\tilde{Z}_t, \tag{186}$$

The quadratic variation of X is given by:

$$\langle X\rangle_t=\int_0^t v_s ds. \tag{187}$$

Let:

$$\mathcal{J}(u,s) \equiv E^{\mathbb{Q}}\left[e^{iux - s(\langle X\rangle_T - \langle X\rangle_t)}\right] |X_t = x, v_t = v] \tag{188}$$

be the joint conditional Fourier Laplace transform of X_T and $\langle X\rangle_T - \langle X\rangle_t$ with arguments u and s defined on some domain $\mathcal{D}$ in $\mathbb{C}^2$ where the expectation exists. This domain always includes the union of $u \in \mathbb{R}$ and s defined on a half plane in $\mathbb{C}$ with a sufficiently large real part.

Let $J(x, v, t)$ denote the joint Fourier Laplace transform considered as a function of the backwards variables x, v, and t. Standard results imply that J solves the following second order linear parabolic PDE:

$$\begin{aligned} &\frac{\partial J}{\partial t}(x,v,t) + \frac{v}{2}\frac{\partial^2 J}{\partial x^2}(x,v,t) + \rho\epsilon v^2 \frac{\partial^2 J}{\partial x \partial v}(x,v,t) + \frac{\epsilon^2 v^3}{2}\frac{\partial^2 J}{\partial v^2}(x,v,t) \\ &\quad -\frac{v}{2}\frac{\partial J}{\partial x}(x,v,t) - \left[p(t)v + qv^2\right]\frac{\partial J}{\partial v}(x,v,t) = svJ(x,v,t), \end{aligned} \tag{189}$$

on $x \in \mathbb{R}, v > 0, t \in [0, T]$. The function J also satisfies the following terminal condition:

$$J(x,v,T) = e^{iux}, \quad x \in \mathcal{D}, v > 0. \tag{190}$$

Since the origin is an aborbing boundary for v, the function $J(x, v, t)$ also obeys the lower boundary condition:

$$\lim_{v\downarrow 0} J(x,v,t) = e^{iux}, \quad x \in \mathcal{D}, t \in [0,T]. \tag{191}$$

By the Riemann Lebesgue lemma[3](see Champeney (1987), p. 23), the function $J(x, v, t)$ also obeys the upper boundary condition:

$$\lim_{v\uparrow \infty} J(x,v,t) = 0, \quad x \in \mathcal{D}, t \in [0,T]. \tag{192}$$

Finally, we require that:

$$\lim_{x\to\pm\infty} |J(x,v,t)| \le 1. \tag{193}$$

Notice that the coefficients in (189) are independent of x. Now suppose that we guess that:

$$J(x,v,t) = e^{iux} L(t,v), \tag{194}$$

[3] This lemma is also known as Mercer's theorem.

where $L(t, v)$ depends on u and s implicitly. Substituting this guess in (189) implies that $L(t, v)$ solves the linear second order PDE:

$$\frac{\partial}{\partial t}L(t,v)+\frac{\epsilon^2 v^3}{2}\frac{\partial^2}{\partial v^2}L(t,v)+\left[p(t)v+\tilde{q}v^2\right]\frac{\partial}{\partial v}L(t,v)-\lambda v L(t,v)=0,$$
$$v>0, t\in[0,T], \tag{195}$$

where:

$$\tilde{q}\equiv q+\rho\epsilon i u \tag{196}$$

and:

$$\lambda\equiv s+\frac{iu}{2}+\frac{u^2}{2}. \tag{197}$$

From (190), the function $L(t, v)$ also satisfies the following terminal condition:

$$L(T,v)=1,\quad v>0. \tag{198}$$

The lower boundary condition corresponding to (190) and (191) is:

$$L(t,0)=1,\quad t\in[0,T]. \tag{199}$$

The upper boundary condition corresponding to (192) is:

$$\lim_{v\uparrow\infty}L(t,v)=0,\quad t\in[0,T]. \tag{200}$$

Now suppose that we guess that L depends on t and v only through some particular intervening variable y. Specifically, we guess:

$$L(t,v)=\ell(y), \tag{201}$$

where:

$$y\equiv\int_t^T e^{\int_t^{t'}p(u)du}dt'\times v. \tag{202}$$

Then, differentiating (201) w.r.t. t:

$$\frac{\partial}{\partial t}L(t,v)=\ell'(y)[-v-p(t)y]. \tag{203}$$

Differentiating (201) w.r.t. v instead:

$$\frac{\partial}{\partial v}L(t,v)=\ell'(y)\int_t^T e^{\int_t^{t'}p(u)du}dt'. \tag{204}$$

Differentiating (204) w.r.t. v:

$$\frac{\partial^2}{\partial v^2} L(t, v) = \ell''(y) \left(\int_t^T e^{\int_t^{t'} p(u)du} dt' \right)^2. \tag{205}$$

Substituting (203), (204), and (205) in (195) implies that:

$$\ell'(y)[-v - p(t)y] + \frac{\epsilon^2 v^3}{2} \ell''(y) \left(\int_t^T e^{\int_t^{t'} p(u)du} dt' \right)^2 \\ +[p(t)v + \tilde{q}v^2]\ell'(y) \int_t^T e^{\int_t^{t'} p(u)du} dt' = \lambda v \ell(y), \tag{206}$$

for $y > 0, t \in [0, T]$. From (202), the terms involving $p(t)$ cancel. Cancelling these terms and multiplying through by $\int_t^T e^{\int_t^{t'} p(u)du} dt'$, (202) implies:

$$- \ell'(y)y + \frac{\epsilon^2 y^3}{2} \ell''(y) + \tilde{q} y^2 \ell'(y) - \lambda y \ell(y) = 0, \quad y > 0. \tag{207}$$

Finally, dividing by y implies that $\ell(y)$ solves the following linear second order ODE:

$$\frac{\epsilon^2 y^2}{2} \ell''(y) + (\tilde{q} y - 1)\ell'(y) - \lambda \ell(y) = 0, \quad y > 0. \tag{208}$$

To satisfy the terminal condition (198) and the boundary condition (199), we require:

$$\ell(0) = 1. \tag{209}$$

To satisfy the upper boundary condition (200), we require:

$$\lim_{y \uparrow \infty} \ell(y) = 0. \tag{210}$$

We now show that we can do a change of dependent and independent variables which converts (208) into a confluent hypergeometric equation. In particular, let:

$$h(z) \equiv z^{-\alpha} \ell(y) \tag{211}$$

be the new dependent variable, where:

$$z \equiv \frac{\beta}{y} \tag{212}$$

is the new independent variable, with the constants α and β to be determined.

From (211):

$$\ell(y) = z^{\alpha} h(z). \tag{213}$$

Differentiating (213) w.r.t. w:

$$\ell'(y) = -\alpha z^{\alpha-1}\frac{\beta}{y^2}h(z) - z^\alpha h'(z)\frac{\beta}{y^2}$$
$$= -\frac{\alpha}{\beta}z^{\alpha+1}h(z) - \frac{z^{\alpha+2}}{\beta}h'(z), \tag{214}$$

from (212). Hence, the middle term in (208) is:

$$(\tilde{q}y-1)\ell'(w) = \left(\frac{1}{\beta} - \frac{\tilde{q}}{z}\right)\left[\alpha z^{\alpha+1}h(z) + z^{\alpha+2}h'(z)\right]. \tag{215}$$

Differentiating (4) w.r.t. y:

$$\ell''(y) = \frac{\alpha(\alpha+1)}{y^2}z^\alpha h(z) + \frac{\alpha z^{\alpha+1}}{y^2}h'(z) + \frac{(\alpha+2)z^{\alpha+1}}{y^2}h'(z) + \frac{z^{\alpha+2}}{y^2}h''(z). \tag{216}$$

Hence, the first term in (208) is:

$$\frac{\epsilon^2 y^2}{2}\ell''(y) = \frac{\epsilon^2\alpha(\alpha+1)}{2}z^\alpha h(z) + \epsilon^2(\alpha+1)z^{\alpha+1}h'(z) + \epsilon^2\frac{z^{\alpha+2}}{2}h''(z). \tag{217}$$

Substituting (213),(215), and (217) into the ODE (208) implies that:

$$\epsilon^2\frac{z^{\alpha+2}}{2}h''(z) + \left\{[\epsilon^2(\alpha+1) - \tilde{q}]z^{\alpha+1} + \frac{z^{\alpha+2}}{\beta}\right\}h'(z)$$
$$+ \left\{\left[\frac{\epsilon^2\alpha(\alpha+1)}{2} - \tilde{q}\alpha - \lambda\right]z^\alpha + \frac{\alpha}{\beta}z^{\alpha+1}\right\}h(z) = 0. \tag{218}$$

Suppose that we require that α solve the quadratic equation:

$$\frac{\epsilon^2\alpha(\alpha+1)}{2} - \tilde{q}\alpha - \lambda = 0. \tag{219}$$

Note that λ is complex so α is as well. Further suppose that we multiply (218) by $\frac{2}{\epsilon^2 z^{\alpha+1}}$ to make the coefficient of $h''(z)$ equal to z.

$$zh''(z) + \left\{2\left[\alpha + 1 - \frac{\tilde{q}}{\epsilon^2}\right] + \frac{2}{\epsilon^2\beta}z\right\}h'(z) + \frac{2\alpha}{\epsilon^2\beta}h(z) = 0. \tag{220}$$

Now suppose that we choose:

$$\beta = -\frac{2}{\epsilon^2} \tag{221}$$

to make the coefficient of z in the middle term equal to -1. Hence, from (212) and (221), the change of independent variable is:

$$z \equiv -\frac{2}{\epsilon^2 y}, \quad y > 0. \tag{222}$$

Note that β and z are real.

Substituting (221) in (220) implies that it simplifies to:

$$zh''(z) + \left\{2\left[\alpha + 1 - \frac{\tilde{q}}{\epsilon^2}\right] - z\right\} h'(z) - \alpha h(z) = 0. \tag{223}$$

Letting:

$$\gamma \equiv 2\left[\alpha + 1 - \frac{\tilde{q}}{\epsilon^2}\right] \tag{224}$$

implies that (223) further simplifies into the following confluent hypergeometric equation:

$$zh''(z) + (\gamma - z)h'(z) - \alpha h(z) = 0. \tag{225}$$

It is well known that the general solution to (226) is:

$$h(z) = AM(\alpha; \gamma; z) + BU(\alpha; \gamma; z), \tag{226}$$

where:

$$M(\alpha; \gamma; z) \equiv \sum_{n=0}^{\infty} \frac{(\alpha)_n}{(\gamma)_n} \frac{z^n}{n!} \tag{227}$$

and:

$$U(\alpha; \gamma; z) \equiv \frac{\pi}{\sin(\pi\gamma)} \left\{\frac{M(\alpha; \gamma; z)}{\Gamma(1 + \alpha - \gamma)\Gamma(\gamma)} - z^{1-\gamma} \frac{M(1 + \alpha - \gamma; 2 - \gamma; z)}{\Gamma(\alpha)\Gamma(2 - \gamma)}\right\} \tag{228}$$

are two linearly independent solutions to the confluent hypergeometric equation (225). The M function is not defined when γ is zero or a negative integer, which from (224) poses little restriction in our case. See Chap. 13 in Abramowitz and Stegun (1964) for an extensive compilation of the properties of these functions.

From (213), the general solution to (208) is:

$$\begin{aligned}\ell(w) &= Az^{\alpha} M(\alpha; \gamma; z) + Bz^{\alpha} U(\alpha; \gamma; z) \\ &= A\left(\frac{-2}{\epsilon^2 y}\right)^{\alpha} M\left(\alpha; \gamma; \frac{-2}{\epsilon^2 y}\right) + B\left(\frac{-2}{\epsilon^2 y}\right)^{\alpha} U\left(\alpha; \gamma; \frac{-2}{\epsilon^2 y}\right),\end{aligned} \tag{229}$$

from (222).

The constants A and B in (229) are determined by the boundary conditions (209) and (210). We guess that we can meet these boundary conditions by taking $B = 0$ and choosing the right root of (219). The quadratic equation (219) determining α can be written as:

$$\frac{\alpha^2}{2} + \left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right)\alpha - \frac{\lambda}{\epsilon^2} = 0. \tag{230}$$

The two roots are:

$$\alpha = -\left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right) \pm \sqrt{\left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right)^2 + 2\frac{\lambda}{\epsilon^2}}. \tag{231}$$

To make progress, we henceforth assume that the argument u is real and the argument s is real and positive. From (196), this implies that $\tilde{q}$ has a nonzero real part and a nonzero imaginary part.

We now prove that we can always choose the sign in (231) so that the real part of α is positive. Take:

$$a + ib = \sqrt{\left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right)^2 + 2\frac{\lambda}{\epsilon^2}}, \tag{232}$$

where a and b are both real.

We then have:

$$\begin{aligned} a^2 - b^2 &= \mathrm{Re}\left\{\left(\frac{1}{2} - \frac{\tilde{q}}{\epsilon^2}\right)^2 + 2\frac{\lambda}{\epsilon^2}\right\} \\ &= \left(\frac{1}{2} - \frac{q}{\epsilon^2}\right)^2 - \frac{\rho^2 u^2}{\epsilon^2} + 2\frac{s}{\epsilon^2} + \frac{u^2}{\epsilon^2}, \end{aligned}$$

from (196) and (197). Since $|\rho| \leq 1$ and $s > 0$, we have:

$$a^2 = b^2 + \left(\frac{1}{2} - \frac{q}{\epsilon^2}\right)^2 + \frac{2s}{\epsilon^2} + \frac{u^2}{\epsilon^2}(1 - \rho^2) \geq \left(\frac{1}{2} - \frac{q}{\epsilon^2}\right)^2. \tag{233}$$

Therefore we may always choose:

$$a > \frac{1}{2} - \frac{q}{\epsilon^2} > 0, \tag{234}$$

such that $a + ib$ is the root and also:

$$\mathrm{Re}\,\alpha = a - \left(\frac{1}{2} - \frac{q}{\epsilon^2}\right) > 0. \tag{235}$$

This proves the result.

As $z \uparrow 0$, the function M in (229) approaches 1. Since the real part of α is positive, the prefactor $\left(\frac{-2}{\epsilon^2 y}\right)^{\alpha} = z^{\alpha}$ in (229) vanishes as $y \uparrow \infty$. This is precisely the behavior we want for ℓ as $y \uparrow \infty$ and $z \uparrow 0$.

As $y \downarrow 0$, the last argument in M approaches negative infinity. From (Abramowitz and Stegun 1964), (13.1.5), as $z \downarrow -\infty$, the function $M(\alpha;\gamma;z)$ behaves like $\frac{\Gamma(\gamma)}{\Gamma(\gamma-\alpha)}(-z)^{-\alpha}$. If we absorb the -1 in the prefactor of M into the constant A, then the behavior of M in z is cancelled by the pre-factor. As our objective is to have $\ell = 1$ as $y \downarrow 0$, we set:

$$A = \frac{\Gamma(\gamma-\alpha)}{\Gamma(\gamma)(-1)^{\alpha}}. \tag{236}$$

Substituting (236) and $B = 0$ into (229) implies that our final answer for the function ℓ is:

$$\ell(y) = \frac{\Gamma(\gamma-\alpha)}{\Gamma(\gamma)}\left(\frac{2}{\epsilon^2 y}\right)^{\alpha} M\left(\alpha;\gamma;\frac{-2}{\epsilon^2 y}\right), \tag{237}$$

where the confluent hypergeometric function M is defined in (227),

$$\alpha = -\left(\frac{1}{2}-\frac{\tilde{q}}{\epsilon^2}\right) + \sqrt{\left(\frac{1}{2}-\frac{\tilde{q}}{\epsilon^2}\right)^2 + 2\frac{\lambda}{\epsilon^2}},$$

$$\gamma \equiv 2\left[\alpha + 1 - \frac{\tilde{q}}{\epsilon^2}\right],$$

and where:

$$\tilde{q} = q + \rho\epsilon i u \quad \lambda = s + \frac{iu}{2} + \frac{u^2}{2}.$$

From (194) and (202), the function $L(t,v)$ is given by:

$$L(t,v) = \frac{\Gamma(\gamma-\alpha)}{\Gamma(\gamma)}\left(\frac{2}{\epsilon^2 y(t,v)}\right)^{\alpha} M\left(\alpha;\gamma;\frac{-2}{\epsilon^2 y(t,v)}\right), \tag{238}$$

where:

$$y(t,v) \equiv \int_t^T e^{\int_t^{t'} p(u)du} dt' \times v. \tag{239}$$

Finally, from (194), the joint Fourier Laplace transform is given by:

$$J(x,v,t) = e^{iux}\frac{\Gamma(\gamma-\alpha)}{\Gamma(\gamma)}\left(\frac{2}{\epsilon^2 y(t,v)}\right)^{\alpha} M\left(\alpha;\gamma;\frac{-2}{\epsilon^2 y(t,v)}\right). \tag{240}$$

The dependence of J on u and s occurs through α and γ, which in turn depend on u and s through $\tilde{q}$ and λ.

Appendix 4: Proof of Theorem 4

In this appendix, we derive a nonlinear ODE for the function $g(w)$. We also derive and solve a linear ODE for a function h relating w to y. This allows us to solve the nonlinear ODE for g.

Recall from (174) that:

$$\begin{aligned} a(t,v) &= p(t)v + \left(\frac{e^{k(0;T)}}{x(t;T)} - \frac{\varepsilon e^{G(w;T)}}{x(t;T)g(w;T)}\right)v + \frac{\varepsilon}{2}[\varepsilon - g_w(w;T)]v^2 \\ &= p(t)v + \left\{\frac{e^{k(0;T)}}{x(t;T)v} - \frac{\varepsilon e^{G(w;T)}}{x(t;T)vg(w;T)} + \frac{\varepsilon}{2}[\varepsilon - g_w(w;T)]\right\}v^2 \\ &= p(t)v + \left\{\frac{e^{k(0;T)}}{e^{G(w;T)}} - \frac{\varepsilon}{g(w;T)} + \frac{\varepsilon}{2}[\varepsilon - g_w(w;T)]\right\}v^2 \end{aligned} \tag{241}$$

If we set the expression:

$$\frac{e^{k(0;T)}}{e^{G(w;T)}} - \frac{\varepsilon}{g(w;T)} + \frac{\varepsilon}{2}[\varepsilon - g_w(w;T)] = q, \tag{242}$$

then the risk-neutral drift $a(t,v)$ would become a quadratic function of v.

Note that differentiating (242) w.r.t. w yields a second order ODE for the function g. To determine this ODE, rewrite (242) as:

$$e^{k(0;T)}e^{-G(w;T)} = q + \frac{\varepsilon}{g(w;T)} - \frac{\varepsilon}{2}[\varepsilon - g_w(w;T)]. \tag{243}$$

Differentiating w.r.t. w implies:

$$-e^{k(0;T)}e^{-G(w;T)}\frac{\varepsilon}{g(w;T)} = -\frac{\varepsilon g_w(w;T)}{g^2(w)} + \frac{\varepsilon}{2}g_{ww}(w;T), \tag{244}$$

from (153). Dividing out $-\varepsilon$ implies:

$$e^{k(0;T)}e^{-G(w;T)}\frac{1}{g(w;T)} = \frac{g_w(w;T)}{g^2(w)} - \frac{1}{2}g_{ww}(w;T), \tag{245}$$

Substituting (243) into (245) implies:

$$\left\{q + \frac{\varepsilon}{g(w;T)} - \frac{\varepsilon}{2}[\varepsilon - g_w(w;T)]\right\}\frac{1}{g(w;T)} = \frac{g_w(w;T)}{g^2(w;T)} - \frac{1}{2}g_{ww}(w;T). \tag{246}$$

Re-arranging (246) leads to (82). Note that (246) is a second order nonlinear ODE for $g(w)$ which can always be numerically solved subject to the initial conditions $g(0; T) = 0$ and $g_w(0; T) = \epsilon$.

To find an analytic solution for g, first recall that $w(t, v; T)$ solves the PDE:

$$\frac{\partial}{\partial t}w(t, v; T) + \frac{\varepsilon^2 v^3}{2}w_{vv}(t, v; T) + [p(t)v + qv^2]w_v(t, v; T) + v = 0,$$
$$t \in [0, T], v > 0, T > 0. \tag{247}$$

However (173) implies that there exists a function h that relates w to $y \equiv x(t; T)v$, i.e. the function h is defined by:

$$w = h(y; T) \equiv G^{-1}(\ln(y); T), \quad y > 0. \tag{248}$$

Substituting (248) in (247) implies that the function $h(y; T)$ satisfies:

$$h_y(y; T)x'(t; T)v + \frac{\varepsilon^2 v^3}{2}h_{yy}(y; T)x^2(t; T)$$
$$+h_y(y; T)x(t; T)[p(t)v + qv^2] + v = 0. \tag{249}$$

Now recall from (171) that the function $x(t; T)$ satisfies the ODE:

$$\frac{\partial}{\partial t}x(t; T) + p(t)x(t; T) + e^{k(0;T)} = 0. \tag{250}$$

As $k(0, T)$ is just some arbitrary finite quantity at each T, we will set $k(0, T) = 0$ and sacrifice some of our freedom in finding a solution. Hence, we require that the function $x(t; T)$ satisfies the ODE:

$$\frac{\partial}{\partial t}x(t; T) + p(t)x(t; T) + 1 = 0. \tag{251}$$

With this restriction, the solution of (250) subject to $x(T; T) = 0$ is:

$$x(t; T) = \int_t^T e^{\int_t^{t'} p(u)du} dt', \quad t \in [0, T], T > 0. \tag{252}$$

Using (251), we can reduce the ODE (249) for $h(y)$ to:

$$\frac{\varepsilon^2}{2}h_{yy}(y; T)y^2 + h_y(y; T)(qy - 1) + 1 = 0, \quad y \geq 0. \tag{253}$$

Evaluating at $y = 0$ and assuming that $h_{yy}(0; T)$ and $h_y(0; T)$ are both bounded implies that:

$$h_y(0; T) = 1. \tag{254}$$

Since $x(T;T)=0$ and $w(T,v;T)=0$, (248) implies that a second initial condition on h is $h(0;T)=0$. To find a solution to the ODE (253), we define:

$$h'(y;T)\equiv h_y(y;T). \tag{255}$$

The second order ODE (253) for h implies that $h'_y(y;T)$ satisfies the following first order ODE:

$$\frac{\varepsilon^2 y^2}{2}h'_y(y;T)+(qy-1)h'(y;T)+1=0, \quad y\geq 0. \tag{256}$$

Dividing by $\frac{\varepsilon^2 y^2}{2}$:

$$h'_y(y;T)+\left(\frac{2q}{\epsilon^2 y}-\frac{2}{\epsilon^2 y^2}\right)h'(y;T)+\frac{2}{\epsilon^2 y^2}=0, \quad y\geq 0. \tag{257}$$

Multiplying both sides of (257) by $y^{\frac{2q}{\varepsilon^2}}e^{\frac{2}{\varepsilon^2 y}}$ implies that:

$$\frac{\partial}{\partial y}\left[y^{\frac{2q}{\varepsilon^2}}e^{\frac{2}{\varepsilon^2 y}}h'(y;T)\right]=-\frac{2}{\varepsilon^2}y^{\frac{2q}{\varepsilon^2}-2}e^{\frac{2}{\varepsilon^2 y}}. \tag{258}$$

Hence, integrating both sides w.r.t. y implies:

$$y^{\frac{2q}{\varepsilon^2}}e^{\frac{2}{\varepsilon^2 y}}h'(y;T)=C(T)+\int_y^\infty \frac{2}{\varepsilon^2}u^{\frac{2q}{\varepsilon^2}-2}e^{\frac{2}{\varepsilon^2 u}}du, \tag{259}$$

where $C(T)$ is the constant of integration. Notice that:

$$\int_y^\infty \frac{2}{\varepsilon^2}u^{\frac{2q}{\varepsilon^2}-2}e^{\frac{2}{\varepsilon^2 u}}du<\infty, \tag{260}$$

since we required that $\frac{2q}{\varepsilon^2}<1$ to avoid explosions.

Thus, the desired function h relating w to y is given by:

$$h(y;T)=\int_0^y e^{-\frac{2}{\varepsilon^2 z}}z^{-\frac{2q}{\varepsilon^2}}\left(C(T)+\int_z^\infty \frac{2}{\varepsilon^2}e^{\frac{2}{\varepsilon^2 u}}u^{\frac{2q}{\varepsilon^2}-2}du\right)dz. \tag{261}$$

We now will prove that $C(T)\equiv 0$. We first derive the asymptotic behavior as $y\to+\infty$ of the function $h(y)$ in (261). Splitting h into the sum of two terms, we have:

$$\begin{aligned}h(y;T)&=C(T)\int_0^y e^{-\frac{2}{\varepsilon^2 z}}z^{-\frac{2q}{\varepsilon^2}}\\&\quad+\int_0^y z^{-\frac{2q}{\varepsilon^2}}e^{-\frac{2}{\varepsilon^2 z}}\int_z^\infty \frac{2}{\varepsilon^2}u^{\frac{2q}{\varepsilon^2}-2}e^{\frac{2}{\varepsilon^2 u}}du\,dz.\end{aligned} \tag{262}$$

The following lemma shows the asymptotic behavior for the first term in (262):

Lemma 1

$$C(T)\int_0^y z^{-\frac{2q}{\varepsilon^2}} e^{-\frac{2}{\varepsilon^2 z}} dz \sim \frac{C(T)}{1-\frac{2q}{\varepsilon^2}} y^{1-\frac{2q}{\varepsilon^2}}, \tag{263}$$

as $y \to +\infty$.

Proof The proof is straightforward. We only need to calculate the limit:

$$\begin{aligned} \lim_{y\to+\infty} \frac{\int_0^y z^{-\frac{2q}{\varepsilon^2}} e^{-\frac{2}{\varepsilon^2 z}} dz}{y^{1-\frac{2q}{\varepsilon^2}}} &= \lim_{y\to+\infty} \frac{y^{-\frac{2q}{\varepsilon^2}} e^{-\frac{2}{\varepsilon^2 y}}}{\left(1-\frac{2q}{\varepsilon^2}\right) y^{1-\frac{2q}{\varepsilon^2}}} \\ &= \lim_{y\to+\infty} \frac{1}{1-\frac{2q}{\varepsilon^2}} e^{-\frac{2}{\varepsilon^2 y}} \\ &= \frac{1}{1-\frac{2q}{\varepsilon^2}}. \end{aligned} \tag{264}$$

□

The asymptotic behavior for the second term in (262) is given by the following lemma:

Lemma 2 *We have the following asymptotic behavior:*

$$\int_0^y e^{-\frac{2}{\varepsilon^2 z}} z^{-\frac{2q}{\varepsilon^2}} \left(\int_z^{+\infty} \frac{2}{\varepsilon^2} u^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 u}} du\right) dz \sim \frac{1}{\frac{\varepsilon^2}{2}-q} \ln y, \tag{265}$$

as $y \to +\infty$.

Proof Once again, we can calculate the following limit:

$$\begin{aligned} &\lim_{y\to+\infty} \frac{\int_0^y e^{-\frac{2}{\varepsilon^2 z}} z^{-\frac{2q}{\varepsilon^2}} \left(\int_z^{+\infty} \frac{2}{\varepsilon^2} u^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 u}} du\right) dz}{\ln y} \\ &= \lim_{y\to+\infty} y^{-\frac{2q}{\varepsilon^2}} \frac{e^{-\frac{2}{\varepsilon^2 y}} \left(\int_y^{+\infty} \frac{2}{\varepsilon^2} u^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 u}} du\right)}{\frac{1}{y}} \\ &= \lim_{y\to+\infty} \frac{\int_y^{+\infty} \frac{2}{\varepsilon^2} u^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 u}} du}{y^{\frac{2q}{\varepsilon^2}-1}} \end{aligned}$$

$$= \lim_{y \to +\infty} \frac{\frac{2}{\varepsilon^2} y^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 y}}}{\left(1 - \frac{2q}{\varepsilon^2}\right) y^{\frac{2q}{\varepsilon^2}-1}}$$

$$= \frac{1}{\frac{\varepsilon^2}{2} - q}. \tag{266}$$

□

Combining the results of the last two lemmas, we have just proved that:

Lemma 3

$$h(y) \sim \frac{C(T)}{1 - \frac{2q}{\varepsilon^2}} y^{1-\frac{2q}{\varepsilon^2}} + \frac{1}{\frac{\varepsilon^2}{2} - q} \ln y, \tag{267}$$

as $y \to +\infty$.

We now prove that in fact $C(T) = 0$. For this purpose, consider the Laplace transform:

$$L(t, v; s) = E^{\mathbb{Q}} \left(e^{-s \int_t^T v_s \, ds} \middle| v_t = v \right), \tag{268}$$

where $s \in \mathbb{R}^+$. Setting $u = 0$ in (240), the solution is:

$$L(t, v; s) = \ell \left(\int_t^T e^{\int_t^{t'} p(u) du} dt' \times v; s \right), \tag{269}$$

where the function ℓ is given by:

$$\ell(y; s) = \frac{\Gamma(\gamma - \alpha)}{\Gamma(\gamma)} \left(\frac{2}{\varepsilon^2 y} \right)^{\alpha} M \left(\alpha, \gamma; -\frac{2}{\varepsilon^2 y} \right), \tag{270}$$

where α is given by:

$$\alpha = -\left(\frac{1}{2} - \frac{q}{\varepsilon^2} \right) + \sqrt{\left(\frac{1}{2} - \frac{q}{\varepsilon^2} \right)^2 + \frac{2s}{\varepsilon^2}} \tag{271}$$

and γ is given by:

$$\gamma = 2 \left[\alpha + 1 - \frac{q}{\varepsilon^2} \right]$$

$$= 1 + 2 \sqrt{\left(\frac{1}{2} - \frac{q}{\varepsilon^2} \right)^2 + \frac{2s}{\varepsilon^2}}, \tag{272}$$

from (271). Now (271) and (272) imply that:

$$\gamma - \alpha = \frac{3}{2} - \frac{q}{\varepsilon^2} + \sqrt{\left(\frac{1}{2} - \frac{q}{\varepsilon^2}\right)^2 + \frac{2s}{\varepsilon^2}}. \tag{273}$$

We can verify that:

$$\begin{aligned}\gamma - \alpha &> \frac{3}{2} - \frac{q}{\varepsilon^2} + \frac{1}{2} - \frac{q}{\varepsilon^2} \\ &= 2 - \frac{2q}{\varepsilon^2} \\ &> 1\end{aligned} \tag{274}$$

by our usual assumption. Also note that in our current setting both α and γ are positive real numbers. As shown in Abramowitz and Stegun (1964), the confluent hypergeometric function $M(a, b; x)$ hs the following integral representation:

$$M(a, b; x) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 e^{xt} t^{a-1} (1-t)^{b-a-1} dt. \tag{275}$$

Substituting (275) in (270) implies:

$$\ell(y; s) = \frac{1}{\Gamma(\alpha)} \left(\frac{2}{\varepsilon^2 y}\right)^{\alpha} \int_0^1 e^{-\frac{2}{\varepsilon^2 y} t} t^{\alpha-1} (1-t)^{\gamma-\alpha-1} dt. \tag{276}$$

The definition of the Laplace transform suggests that $s \to 0$, we have $\ell(y; s) \to 1$. This is indeed the case here because of the following:

Lemma 4 *We have the limit:*

$$\lim_{a \to 0} \frac{\int_0^1 e^{xt} t^{a-1} (1-t)^{b-1} dt}{\Gamma(a)} = 1, \tag{277}$$

for any real number x and $b > 0$.

Proof As $a \to 0$,

$$\Gamma(a) \equiv \int_0^{+\infty} t^{a-1} e^{-t}\, dt \to +\infty \tag{278}$$

because $1/t$ is not integrable around the point 0. On the other hand:

$$\lim_{a \to 0} \int_0^1 e^{xt} t^{a-1} (1-t)^{b-a-1} dt = \int_0^1 e^{xt} t^{-1} (1-t)^{b-1} dt = \infty \tag{279}$$

for the same reason. Hence, a determination of the limiting value of the ratio requires more careful analysis. Using integration by parts, it is well known that the gamma function satisfies the following recursion:

$$\alpha\Gamma(a) = \Gamma(a+1). \tag{280}$$

On the other hand:

$$\begin{aligned} a\int_0^1 e^{xt}t^{a-1}(1-t)^{b-1}dt &= \int_0^1 e^{xt}(1-t)^{b-1}d(t^a) \\ &= e^{xt}(1-t)^{b-1}\Big|_0^1 - \int_0^1 t^a d(e^{xt}(1-t)^{b-1}) \\ &\text{and integrating by parts implies:} \\ &= -\int_0^1 t^a d(e^{xt}(1-t)^{b-1}). \end{aligned} \tag{281}$$

Therefore:

$$\begin{aligned} \frac{\int_0^1 e^{xt}t^{a-1}(1-t)^{b-1}dt}{\Gamma(a)} &= \frac{a\int_0^1 e^{xt}t^{a-1}(1-t)^{b-1}dt}{a\Gamma(a)} \\ &= \frac{-\int_0^1 t^a d(e^{xt}(1-t)^{b-1})}{\Gamma(a+1)}. \end{aligned} \tag{282}$$

Now taking the limit, we will have:

$$\begin{aligned} \lim_{a\to 0}\frac{\int_0^1 e^{xt}t^{a-1}(1-t)^{b-1}dt}{\Gamma(a)} &= \frac{-\int_0^1 d(e^{xt}(1-t)^{b-1})}{\Gamma(1)} \\ &= -e^{xt}(1-t)^{b-1}\Big|_0^1 \\ &= 1. \end{aligned} \tag{283}$$

□

First, (271) implies that:

$$\lim_{s\to 0}\alpha = 0. \tag{284}$$

Second:

$$\lim_{\alpha\to 0}\left(\frac{2}{\varepsilon^2 y}\right)^{\alpha} = 1, \tag{285}$$

for nonzero y. Third:

$$\lim_{\alpha\to 0}\frac{1}{\Gamma(\alpha)}\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}t^{\alpha-1}(1-t)^{\gamma-\alpha-1}dt = 1. \tag{286}$$

These three results along with (270) together imply that:

$$\lim_{s\to 0} \ell(y;s) = 1. \tag{287}$$

Now since:

$$\frac{\partial L(t,v;s)}{\partial s}\Big|_{s=0} = E\left(-\int_t^T v_s ds \Big| v_t = v\right), \tag{288}$$

we have:

$$h(y) = -\frac{\partial \ell(y;s)}{\partial s}\Big|_{s=0}. \tag{289}$$

We now use this relation to derive the asymptotic behavior of the function $h(y)$. This behavior will allow us to determine the constant $C(T)$. In particular, we prove the following lemma:

Lemma 5 *As $y \to +\infty$, we have the following asymptotic behavior:*

$$-\frac{\partial h(y;s)}{\partial s}\Big|_{s=0} \sim \frac{1}{\frac{\varepsilon^2}{2} - q} \ln y. \tag{290}$$

Proof First, differentiating (271) w.r.t s implies that:

$$\frac{\partial \alpha}{\partial s} = \frac{1}{\varepsilon^2} \frac{1}{\sqrt{\left(\frac{1}{2} - \frac{q}{\varepsilon^2}\right)^2 + \frac{2s}{\varepsilon^2}}} \tag{291}$$

hence, setting $s = 0$, we see that:

$$\frac{\partial \alpha}{\partial s}\Big|_{s=0} = \frac{1}{\frac{\varepsilon^2}{2} - q}. \tag{292}$$

We now just differentiate with respect to the parameter α instead of s. There will be two major terms after taking the derivatives:

$$\begin{aligned}\frac{\partial h(y;s)}{\partial s} &= \frac{1}{\Gamma(\alpha)}\left(\frac{2}{\varepsilon^2 y}\right)^{\alpha} \ln\left(\frac{2}{\varepsilon^2 y}\right) \int_0^1 e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1} dt \\ &\quad + \left(\frac{2}{\varepsilon^2 y}\right)^{\alpha} \frac{d}{d\alpha}\frac{1}{\Gamma(\alpha)}\left(\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1} dt\right). \end{aligned} \tag{293}$$

The first term is easy to deal with. Letting $\alpha \to 0$, we have:

$$\begin{aligned}
&\lim_{\alpha\to 0} \tfrac{1}{\Gamma(\alpha)} \left(\frac{2}{\varepsilon^2 y}\right)^{\alpha} \ln\left(\frac{2}{\varepsilon^2 y}\right) \int_0^1 e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1}dt \\
&\quad = \ln\left(\frac{2}{\varepsilon^2 y}\right) \lim_{\alpha\to 0} h(y;s) \\
&\quad = \ln\left(\frac{2}{\varepsilon^2 y}\right). \qquad (294)
\end{aligned}$$

To deal with the second term, we use the identity:

$$\alpha\Gamma(\alpha) = \Gamma(\alpha+1). \qquad (295)$$

Hence:

$$\begin{aligned}
&\alpha \int_0^1 e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1}dt \\
&= \int_0^1 e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1}d(t^{\alpha}) \text{ and integrating by parts:} \\
&= e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1}t^{\alpha}\Big|_0^1 - \int_0^1 t^{\alpha} d\left(e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{\gamma-\alpha-1}\right) \\
&= -\int_0^1 t^{\alpha} d\left(e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1}\right). \qquad (296)
\end{aligned}$$

We have used the fact that $\gamma - \alpha - 1 > 0$. Therefore:

$$\begin{aligned}
&\tfrac{1}{\Gamma(\alpha)} \int_0^1 e^{-\frac{2}{\varepsilon^2 y}t} t^{\alpha-1}(1-t)^{\gamma-\alpha-1}dt \\
&= -\frac{1}{\Gamma(\alpha+1)} \int_0^1 t^{\alpha} d\left(e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{\gamma-\alpha-1}\right) \\
&= \frac{\int_0^1 \left(t^{\alpha}\frac{2}{\varepsilon^2 y}e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{\gamma-\alpha-1} + (\gamma-\alpha-1)e^{-\frac{2}{\varepsilon^2 y}t}t^{\alpha}(1-t)^{\gamma-\alpha-2}\right)dt}{\Gamma(1+\alpha)}. \qquad (297)
\end{aligned}$$

We now take the derivative w.r.t. α using the quotient rule and noticing that:

$$\frac{d(\gamma-\alpha)}{d\alpha} = 1, \qquad (298)$$

we have:

$$\begin{aligned}
&\frac{d}{d\alpha}\frac{\int_0^1 t^\alpha \frac{2}{\varepsilon^2 y}e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{\gamma-\alpha-1}dt}{\Gamma(\alpha+1)}\Big|_{\alpha=0}\\
&=\left(\frac{2}{\varepsilon^2 y}\right)\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{1-\frac{2q}{\varepsilon^2}}(\ln t+\ln(1-t))\,dt\\
&\quad-\Gamma'(1)\int_0^1\left(\frac{2}{\varepsilon^2 y}\right)e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{2-\frac{2q}{\varepsilon^2}}dt, \qquad (299)
\end{aligned}$$

while the second term:

$$\begin{aligned}
&\frac{d}{d\alpha}\frac{\int_0^1(\gamma-\alpha-1)e^{-\frac{2}{\varepsilon^2 y}t}t^\alpha(1-t)^{\gamma-\alpha-2}dt}{\Gamma(\alpha+1)}\Big|_{\alpha=0}\\
&=\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{-\frac{2q}{\varepsilon^2}}\,dt\\
&\quad+\left(1-\frac{2q}{\varepsilon^2}\right)\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{-\frac{2q}{\varepsilon^2}}(\ln t+\ln(1-t))\,dt\\
&\quad-\Gamma'(1)\left(2-\frac{2q}{\varepsilon^2}\right)\int e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{-\frac{2q}{\varepsilon^2}}\,dt. \qquad (300)
\end{aligned}$$

In the end, we have:

$$\begin{aligned}
&\frac{\partial h(y;s)}{\partial s}\Big|_{s=0}\\
&=\ln\left(\frac{2}{\varepsilon^2 y}\right)+\left(\frac{2}{\varepsilon^2 y}\right)\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{1-\frac{2q}{\varepsilon^2}}(\ln t+\ln(1-t))\,dt\\
&\quad-\Gamma'(1)\int_0^1\left(\frac{2}{\varepsilon^2 y}\right)e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{2-\frac{2q}{\varepsilon^2}}dt\\
&\quad+\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{-\frac{2q}{\varepsilon^2}}\,dt\\
&\quad+\left(1-\frac{2q}{\varepsilon^2}\right)\int_0^1 e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{-\frac{2q}{\varepsilon^2}}(\ln t+\ln(1-t))\,dt\\
&\quad-\Gamma'(1)\left(2-\frac{2q}{\varepsilon^2}\right)\int e^{-\frac{2}{\varepsilon^2 y}t}(1-t)^{-\frac{2q}{\varepsilon^2}}\,dt. \qquad (301)
\end{aligned}$$

However, every integral in this identity is a finite number since it is easy to check that all the integrals are proper. Moreover, except for the very first term, as $y \to +\infty$, all of the other terms approach a constant. Therefore:

$$-\left.\frac{\partial h(y;s)}{\partial s}\right|_{s=0} \sim \frac{1}{\frac{\varepsilon^2}{2}-q}\ln y, \tag{302}$$

as $y \to +\infty$ and this proves the lemma. □

Finally, we have assembled enough results to prove that the solution to the variance swap rate is given by:

$$h(y) = \frac{2}{\varepsilon^2}\int_0^y \left(z^{-\frac{2q}{\varepsilon^2}} e^{-\frac{2}{\varepsilon^2 z}} \int_z^\infty u^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 u}} du\right) dz. \tag{303}$$

To see this, note that we have just proved that:

$$\lim_{s\to 0} -\frac{\partial \ell(y;s)}{\partial s} \sim \frac{1}{\frac{\varepsilon^2}{2}-q}\ln y \tag{304}$$

as $y \to +\infty$. Therefore:

$$h(y) \sim \frac{1}{\frac{\varepsilon^2}{2}-q}\ln y. \tag{305}$$

This is the asymptotic behavior that we have been looking for. Now from (267), we also have the asymptotic behavior:

$$h(y) \sim C(T)y^{1-\frac{2q}{\varepsilon^2}} + \frac{1}{\frac{\varepsilon^2}{2}-q}\ln y. \tag{306}$$

Comparing (305) and (306), we must have that $C(T) = 0$. Hence, we have proved that:

$$w(t,v) = h\left(\int_t^T e^{\int_t^{t'} p(u)du} dt' \times v\right), \tag{307}$$

where $h(y)$ is given by:

$$h(y) = \frac{2}{\varepsilon^2}\int_0^y \left(z^{-\frac{2q}{\varepsilon^2}} e^{-\frac{2}{\varepsilon^2 z}} \int_z^\infty u^{\frac{2q}{\varepsilon^2}-2} e^{\frac{2}{\varepsilon^2 u}} du\right) dz. \tag{308}$$

This completes the proof that (78) holds.

Given that we know the function h, (248) implies that:

$$G(h(y)) = \ln y, \quad y > 0. \tag{309}$$

Differentiating (153) w.r.t. w implies:

$$G_w(w) = \frac{\varepsilon}{g(w)}, \quad w > 0. \tag{310}$$

Solving (310) for g and substituting in (309) implies:

$$g(w) = \frac{\varepsilon}{G_w(w)} = \varepsilon h^{-1}(w) h_y(h^{-1}(w)), \quad w > 0. \tag{311}$$

Since the function h is known, (311) gives our candidate for the function g governing the volatility of w.

References

Abramowitz, M., & Stegun, I. (1964). *Handbook of mathematical functions*. New York: Dover Publications, Inc.

Ahn, D., & Gao, B. (1999). A parametric nonlinear model of term structure dynamics. *Review of Financial Studies, 12*, 721–762.

Andersen, L., & Piterbarg, V. (2004). Moment explosions in stochastic volatility models. Bank of America working paper.

Andersen, L., & Brotherton-Ratcliffe, R. (1996). Exact exotics. *Risk, 9*, 85–89.

Andreasen, J. (2001). Credit explosives. Bank of America working paper.

Bakshi, G., Ju, N., & Yang, H. (2004). Estimation of continuous time models with an application to equity volatility. University of Maryland working paper.

Baldi, P., Caramellino, L., & Iovino, M. G. (1999). Pricing general barrier options: A numerical approach using sharp large deviations. *Mathematical Finance, 9*, 293–322.

Beaglehole, D., Dybvig, P., & Zhou, G. (1997). Going to extremes: Correcting simulation bias in exotic option valuation. *Financial Analysts Journal, 53*, 62–68.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy, 81*, 637–659.

Carr, P. (2001). Deriving derivatives of derivative securities. *Journal of Computational Finance, 4*(2), 5–30.

Carr, P., & Madan, D. (1999). Option pricing and the fast Fourier transform. *Journal of Computational Finance, 2*(4), 61–73.

Chacko, G., & Viceira, L. (1999). Spectral GMM estimation of continuous-time processes. Harvard University working paper.

Champeney, D. (1987). *A handbook of Fourier theorems*. New York: Cambridge University Press.

Chourdakis, K. (2004). Option pricing using the fractional FFT. Queen Mary, University of London working paper.

Cox, J. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. Stanford University working paper.

Cox, J., Ingersoll, J., & Ross, S. (1980). An analysis of variable rate loan contracts. *The Journal of Finance, 35*(2), 389–403.

Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica, 53*(2).

Duanmu, Z. (2004). Rational pricing of options on realized volatility—The black scholes way. Merrill Lynch working paper.

Dupire, B. (1992). *Arbitrage pricing with stochastic volatility*. Societe Generale Division Options, Paris.

El Babsiri, M., & Noel, G. (1998). Simulating path-dependent options: A new approach. *Journal of Derivatives, 6*(2), 65–83.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation. *Econometrica, 66*, 77–105.

Heston, S. (1997). A simple new formula for options with stochastic volatility. Washington University of St. Louis working paper.

Ishida, I., & Engle, R. (2002). Modelling variance of variance: The square root, the affine, and the CEV GARCH models. NYU working paper.

Javaheri, A. (2004). The volatility process: A study of stock market dynamics via parametric stochastic volatility models and a comparison to the information embedded in option prices. Ph.D. dissertation.

Jones, C. (2003). The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics, 116*, 118–224.

Lewis, A. (2000). *Option valuation under stochastic volatility: With mathematica code*. Finance Press.
Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science, 4*, 141–183.
Neuberger, A. (1990). Volatility trading. London Business School working paper.
Poteshman, A. (1998). Estimating a general stochastic variance model from option prices. University of Chicago working paper.
Potter, C. W. (2004). Complete stochastic volatility models with variance swaps. Oxford University working paper.
Romano, M., & Touzi, N. (1997). Contingent claims and market completeness in a stochastic volatility model. *Mathematical Finance, 7*, 399–410.
Spencer, P. (2003). Coupon bond valuation with a non-affine discount yield model. Department of Economics, University of York working paper.