Robust Replication of Volatility Derivatives

Peter Carr* and Roger Lee†

This version: October 4, 2007

Abstract

We develop model-free trading strategies to replicate volatility derivatives – contracts which pay functions of the realized variance of [the returns on] an underlying price process. The replicating portfolios consist of the underlying shares and vanilla options, traded dynamically in quantities specified by our formulas, which are exact and robust across all underlying price processes satisfying an independence assumption on the instantaneous volatility. In case the independence condition does not hold, our schemes are moreover immunized, to first order, against the presence of correlation. Explicit model-free valuation formulas follow from the replication strategies. Examples include volatility swaps and variance options.

*Bloomberg LP and Courant Institute. pcarr4@bloomberg.com
†University of Chicago. rogerlee@uchicago.edu

1 Introduction

We define the realized variance of the returns on a positive underlying price S from time 0 to time T to be the quadratic variation of $\log S$ at time T. If S has an instantaneous volatility process σ_t, then realized variance equals integrated variance, meaning the time integral of σ_t^2. In practice, contracts written on realized variance typically define it discretely as the sample variance of daily or weekly log returns. Following the custom in the derivatives literature, we study the (continuously-sampled) quadratic variation / integrated variance, leaving tests of discrete sampling for future research, such as [10].

Realized variance can be traded by means of a variance swap, a contract which pays at time T the difference between realized variance and an agreed fixed leg. The variance swap has become a leading tool – perhaps *the* leading tool – for portfolio managers to trade variance. As reported in the Financial Times [16] in 2006,

> Volatility is becoming an asset class in its own right. A range of structured derivative products, particularly those known as variance swaps, are now the preferred route for many hedge fund managers and proprietary traders to make bets on market volatility.

According to some estimates [1], the daily trading volume in equity index variance swaps reached USD 4–5 million vega notional in 2006. On an annual basis, this corresponds to payments of more than USD 1 billion, per percentage point of volatility.

From a dealer's perspective, the variance swap admits replication by a T-expiry log contract (which decomposes into static positions in calls and puts on S), together with dynamic trading in S, as shown in Neuberger [19], Dupire [13], Carr-Madan [11], Derman et al [12], and Britten-Jones/Neuberger [8]. Perfect replication requires frictionless markets and continuity of the price process, but does not require the dynamics of instantaneous volatility to be specified. In that sense, the result is model-free. Well-known to practitioners, the variance swap's model-free replicating portfolio became in 2003 the basis for how the Chicago Board Option Exchange (CBOE) calculates the VIX index, an indicator of the risk-neutral market expectation of short-term volatility. Implementation issues arising from data limitations (particularly the need to interpolate and extrapolate from a limited number of strikes) are documented and addressed in Jiang-Tian [18].

1.1 Volatility derivatives

More generally, *volatility derivatives*, which pay *functions* of realized variance, are of interest to portfolio managers who desire non-linear exposure to variance. Important examples include calls and puts on realized variance; and volatility swaps (popular especially in foreign exchange markets) which pay *realized volatility*, the square root of realized variance.

In contrast to the variance swap's replicability by a log contract, general functions of variance present greater hedging difficulties to the dealer. In theory, if one specifies the dynamics of instantaneous volatility as a one-dimensional diffusion, then one can replicate a volatility derivative by trading the underlier and one option. Such simple stochastic volatility models are, however, misspecified according to empirical evidence, such as difficulties in fitting the observed cross-section of option prices, and pricing errors out-of-sample, as documented in Bakshi-Cao-Chen [3] and Bates [4]. Moreover, even if one could find a well-specified model, further error can arise in trying to calibrate or estimate the model's parameters, not directly observable from options prices.

Equity derivatives dealers have struggled with these issues. According to a 2003 article [20] in RiskNews,

> While variance swaps - where the underlying is volatility squared - can be perfectly replicated under classical derivatives pricing theory, this has not generally been thought to be possible with volatility swaps. So while a few equity derivatives desks are comfortable with taking on the risk associated with dealing volatility swaps, many are not.

A 2006 Financial Times article [16] quotes a derivatives trader:

> Variance is easier to hedge. Volatility can be a nightmare.

We challenge this conventional wisdom, by developing model-free trading strategies to price and replicate volatility derivatives.

1.2 Our approach

We prove that general functions of variance, including volatility swaps, do admit valuation and replication using portfolios of the underlying shares and European options, dynamically traded according to strategies valid across all underlying dynamics specified in Section 2. In practice, trading in options incurs transaction costs, but the costs can be mitigated when hedging a portfolio of volatility contracts, because any offsetting trades need not actually be conducted.

Our approach has the following benefits.

First, we take a non-parametric *model-free* approach, in the sense that we do not specify the dynamics or estimate the parameters of instantaneous volatility. We handle non-Markovian and jumpy volatility processes just as easily as diffusive volatility. By establishing pricing and hedging strategies valid across a whole class of models, we avoid the risk of misspecification and miscalibration present in any one model. Specifically, our replication strategy is robust across all underlying price processes whose instantaneous volatility satisfies an independence assumption. Moreover, in case the independence condition does not hold, we immunize our schemes, to first order, against the presence of correlation; thus we can price approximately under dynamics which generate implied volatility skews – without relying on an particular model of instantaneous volatility.

Second, we find *exact formulas* for values of various classes of contracts on realized volatility $\langle X\rangle_T$, in terms of observable values of contracts on price S_T. The typical pricing result in this paper takes the form of an equality of risk-neutral expectations

$$\mathbb{E}h(\langle X\rangle_T) = \mathbb{E}G(S_T),$$

where we find formulas for G, given various classes of payoff functions h, including the square root function which defines the volatility swap. The left-hand side is the value of the desired *volatility* contract. The right-hand side is the value of a claim on a function of *price* S_T, and is therefore model-independently observable from the values of Europeans. Thus our exact formula for the volatility contract value is not in terms of the parameters of any model, but rather in terms of prices observable, in principle, in the vanilla options market.

Third, we cover not just valuation but also *replication*, by proving explicit trading strategies which enforce the valuation results. The holdings in our replicating portfolios are rebalanced dynamically, but the quantity to hold, at each time, depends only on contemporaneously observable *prices*, not on the parameters of any model.

2 Assumptions

Fix an arbitrary time horizon $T > 0$. Assume for simplicity zero interest rates on a risk-free asset B with price 1 at all times, which we will call the *bond* or *cash*.

On a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ satisfying the usual conditions, assume there exists a equivalent probability measure $\mathbb{P}$ such that S solves

$$dS_t = \sigma_t S_t dW_t, \qquad S_0 > 0 \qquad \text{(Assumption W)}$$

for some $(\mathcal{F}_t, \mathbb{P})$-Brownian motion W_t and some measurable $\mathcal{F}_t$-adapted process σ_t which satisfy

$$\int_0^T \sigma_t^2 dt \text{ is bounded by some } m \in \mathbb{R} \qquad \text{(Assumption B)}$$

and

$$\sigma \text{ and } W \text{ are independent} \qquad \text{(Assumption I)}$$

and such that $\mathbb{P}$ is a risk-neutral pricing measure in the following sense: for all $p \in \mathbb{C}$ and $t \leq T$, a claim paying the real part of S_T^p at time T has time-t price equal to the real part of $\mathbb{E}_t S_T^p$, where $\mathbb{E}_t$ denotes $\mathcal{F}_t$-conditional $\mathbb{P}$-expectation; and likewise for the imaginary parts. Denote the logarithmic returns process by

$$X_t := \log(S_t/S_0) \tag{2.1}$$

and write $\langle X \rangle$ for the quadratic variation of X. Under assumption (I),

$$\langle X \rangle_t = \int_0^t \sigma_u^2 du. \tag{2.2}$$

Unless otherwise stated, the assumptions (B, W, I) are in effect throughout the paper.

Remark 2.1. We will relax assumption (I), by finding results robust – in a sense to be defined in section 4 – to violations of (I). The independence assumption implies that implied volatility skews are symmetric – contrary to typical implied volatility skews in equity markets, which slope downward. Therefore robustness to correlation has practical importance.

Remark 2.2. We drop assumption (B) in section 8.

Remark 2.3. We drop assumption (W) in a separate paper, by introducing jumps in the price process and also by introducing an S-dependent "local volatility" multiplier. In particular, we allow asymmetries in the jump distribution and the local volatility function, which can generate asymmetric volatility skews.

Remark 2.4. We need not and will not work under the actual physical probability measure P. All expectations are with respect to risk-neutral measure $\mathbb{P}$.

This paper concerns the *replication* and arbitrage-free *valuation* of functions of future variance, not the forecasting/prediction of functions of future variance. Our typical result, of the form

$$\mathbb{E}h(\langle X\rangle_T) = \mathbb{E}G(S_T),$$

does not claim that the price of a $G(S_T)$ payoff is an "unbiased forecast" of the realized-variance-contingent random variable $h(\langle X\rangle_T)$; indeed we make no claims about the physical expectation of $h(\langle X\rangle_T)$. What we do show is that the price of the $G(S_T)$ payoff equals the *value* of a claim that pays $h(\langle X\rangle_T)$, because the latter claim replicates the $G(S_T)$ payoff. The concept of perfect replication is measure-invariant; because we will show that the replication occurs with risk-neutral probability one, the replication must also occur with physical probability one, as P and $\mathbb{P}$ agree on all events of probability 0 or 1. Thus, given the availability of the appropriate European-style contracts as hedging instruments, the variance payoff is dynamically spanned with probability 1, and valuation follows, by absence of arbitrage.

The irrelevance of physical expectations (for this paper's valuation and replication purposes) renders also irrelevant the mapping between risk-neutral expectations and physical expectations. Thus we have no need of any assumptions about the market's volatility risk premia (or indeed any other market risk premia) which mediate between the risk-neutral and the physical. In particular, our results are valid without regard to the market's risk preferences, and without regard to whether volatility risk is priced or unpriced.

Remark 2.5. While this paper concerns valuation and replication, a direction for future research is to investigate whether our developments improve the forecasting of [functions of] variance.

Such a research path would parallel how the fundamental variance swap theorem (that the log contract correctly *values* variance, by replication) gave rise to the conjecture that the log contract value may have ability to *predict* variance or functions thereof. For empirical analysis of that conjecture, see the recent work of Andersen-Frederiksen-Staal [2] and the references therein.

Likewise, the contracts which we develop to *value general functions* of variance (by replication) may have the ability to predict functions of variance, and thus to serve as alternatives to the log contract which values plain variance (and which does not value functions of variance, such as volatility).

3 Variance swap

A *variance swap* pays $\langle X\rangle_T$ minus an agreed fixed amount, which we take to be zero unless otherwise specified.

Replication of a variance swap does not require assumption (I). As shown in Neuberger [19], Dupire [13], Carr-Madan [11], and Derman et al [12], Itô's rule implies

$$X_T = \log(S_T/S_0) = \int_0^T \frac{1}{S_t}dS_t - \frac{1}{2}\int_0^T \left(\frac{-1}{S_t^2}\right)\sigma_t^2 S_t^2 dt.$$

so

$$\langle X\rangle_T = -2X_T + \int_0^T \frac{2}{S_t}dS_t.$$

Therefore the following self-financing strategy replicates the $\langle X\rangle_T$ payoff. At each time $t \le T$ hold

1	*log contract*, which pays $-2\log(S_T/S_0)$
$\frac{2}{S_t}$	shares
$\int_0^t \frac{2}{S_u}dS_u - 2$	cash

which is a static position in a European-style claim plus a dynamic position in shares.

In particular, the variance swap's time-0 value equals the price of the log contract.

Remark 3.1. To create a claim on the quadratic variation $[S]$ of price (instead of log price), continuity is not needed. Indeed, provided only that S is a local martingale – without assuming (B, W, I) – we have

$$[S]_T = S_T^2 - S_0^2 - 2\int_0^T S_{t-}dS_t$$

(with the convention that $[S]_0 = 0$). Thus a European-style claim on S_T^2 together with cash and a dynamically traded position of $-2S_{t-}$ shares at each time t replicates the "arithmetic" variance swap.

Remark 3.2. By Breeden-Litzenberger [6] and Carr-Madan [11], the log contract, and more generally a claim on a function $G(S_T)$, can be synthesized if we have bonds and T-expiry puts and calls at all strikes. Specifically, if $G : \mathbb{R}_+ \to \mathbb{R}$ is a difference of convex functions, then for any $\kappa \in \mathbb{R}_+$ we have for all $x \in \mathbb{R}_+$ the representation

$$G(x) = G(\kappa) + G'(\kappa)(x-\kappa) + \int_{K\ge\kappa} G''(K)(x-K)^+ dK + \int_{0<K<\kappa} G''(K)(K-x)^+ dK$$

where G' denotes the left-derivative, and G'' the second derivative, which exists as a signed measure.

4 Robustness to correlation

The typical pricing result in this paper has the following form. Given a desired function h of variance, we find a formula for a function G of price, such that

$$\mathbb{E}h(\langle X\rangle_T) = \mathbb{E}G(S_T) \tag{4.1}$$

Indeed, we will find an *infinite family* of G such that the equality (4.1) holds for all processes S satisfying assumptions (B, W, I). From among this family, we will eventually choose a G such that $\mathbb{E}G(S_T)$ is insensitive to correlation between price and volatility shocks. Thus we gain robustness, in the sense that (4.1) still holds approximately, even if assumption (I) does not hold.

To quantify the effect of correlation, Theorem 4.1 will give a mixing formula that (without assuming independence) expresses the price of any European-style payoff as the expectation of the Black-Scholes formula for that payoff, evaluated at a randomized stock price and random volatility. We then examine the mixing formula's sensitivity to correlation.

First we define what is meant by the Black-Scholes formula for a payoff.

Let $t \leq T$. Let $\mathcal{B}$ denote the Borel sets of $\mathbb{R}_+$ and let $m\mathcal{F}_t$ denote the set of $\mathcal{F}_t$-measurable random variables. Consider a time-t-contracted *payoff function*, by which we mean a

$$F : \mathbb{R}_+ \times \Omega \to \mathbb{R}, \qquad F \text{ is } (\mathcal{B} \otimes \mathcal{F}_t)\text{-measurable}, \tag{4.2}$$

and a random probability distribution P on $\mathbb{R}_+$, such that for all $B \in \mathcal{B}$, we have $\mathsf{P}(B) \in m\mathcal{F}_t$.

Think of F as a function which maps S_T to a European-style payout; for example, an ATM call would have $F(\mathsf{S}) = (\mathsf{S} - S_t)^+$. Think of P as an $\mathcal{F}_t$-conditional risk-neutral distribution that may be imputed to S_T/S_t – not necessarily the *true* $\mathcal{F}_t$-conditional risk-neutral distribution of S_T/S_t. The ω-dependence of F (and P) allows payoffs constructed at time t (and conditional distributions observed at time t) to depend on information in $\mathcal{F}_t$. Our notation may suppress this ω-dependence; for example, $F(\mathsf{S}) = (\mathsf{S} - S_t)^+$ is shorthand for $F(\mathsf{S}, \omega) = (\mathsf{S} - S_t(\omega))^+$.

Given F and P, define the *valuation function* $F^{\mathsf{P}} : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ by

$$F^{\mathsf{P}}(\mathsf{s}, \omega) := \int F(\mathsf{s}y, \omega)\mathsf{P}(dy, \omega).$$

Thus F^{P} maps S_t to the time-t value of $F(S_T)$ induced by P. In particular, define the *Black-Scholes formula*

$$F^{BS}(\mathsf{s}, \sigma, \omega) := F^{\mathsf{P}_\sigma}(\mathsf{s}, \omega)$$

where P_σ denotes the lognormal distribution with parameters $(-\sigma^2/2,\ \sigma^2)$. Note that the valuations F^{P} and F^{BS} are defined as functions of *today's* price (where "today" means the valuation date), unlike the payoff F which is defined as a function of *expiration* price. Notationally, we make a distinction: the placeholder for today's price is s, whereas the placeholder for expiration price is S. Again, our notation may suppress the ω-dependence.

To prove the mixing formula, we give the same argument as Romano-Touzi [22] and Willard [24], but in a more general setting where we do not assume that instantaneous volatility follows a 1-factor diffusion.

Theorem 4.1. *Without assuming* (I)*, let*

$$dS_t = \sigma_t S_t \sqrt{1 - \rho^2} dW_{1t} + \sigma_t S_t \rho dW_{2t}$$

where $|\rho| \leq 1$*, and* W_1 *and* W_2 *are* $\mathcal{F}_t$*-Brownian motions, and* σ *and* W_2 *are adapted to some filtration* $\mathcal{H}_t \subseteq \mathcal{F}_t$*, where* $\mathcal{H}_T$ *and* $\mathcal{F}_T^{W_1}$ *are independent. Then*

$$\mathbb{E}_t F(S_T) = \mathbb{E}_t F^{BS}(S_t M_{t,T}(\rho),\ \bar{\sigma}_{t,T}\sqrt{1-\rho^2}), \tag{4.3}$$

where

$$M_{t,T}(\rho) := \exp\Big(-\frac{\rho^2}{2}\int_t^T \sigma_u^2 du + \rho \int_t^T \sigma_u W_{2u} \Big) \tag{4.4}$$

and $\bar{\sigma}_{t,T} := (\int_t^T \sigma_u^2 du)^{1/2}$*.*

Remark 4.2. This setting includes the standard stochastic volatility models with correlation, which take the form

$$\begin{aligned} dS_t &= \sigma_t S_t dW_{0t} \\ d\sigma_t &= \alpha(\sigma_t)dt + \beta(\sigma_t)dW_{2t}, \end{aligned}$$

where W_2 and $W_0 := \sqrt{1-\rho^2}W_1 + \rho W_2$ have correlation ρ.

Our setting also allows more general dynamics; for example, σ can have jumps independent of W_1.

Remark 4.3. Expanding (4.3) in a formal Taylor series about $\rho = 0$,

$$\begin{aligned} \mathbb{E}_t F(S_T) &= \mathbb{E}_t F^{BS}(S_t M_{t,T}(\rho),\ \bar{\sigma}_{t,T}\sqrt{1-\rho^2}) \\ &\approx \mathbb{E}_t F^{BS}(S_t, \bar{\sigma}_{t,T}) + \rho S_t \mathbb{E}_t\Big[\frac{\partial F^{BS}}{\partial \mathsf{s}}(S_t, \bar{\sigma}_{t,T})\int_t^T \sigma_u dW_{2u}\Big] + O(\rho^2) \end{aligned}$$

Viewed at time t, the Black delta $\partial F^{BS}/\partial \mathsf{s}$ can be random, because $\bar{\sigma}_{t,T}$ is random. Suppose, however, that F has the property that $\partial F^{BS}/\partial \mathsf{s}$ is invariant to its second argument. Then the $\partial F^{BS}/\partial \mathsf{s}$ comes out of the expectation and the ρ term vanishes:

$$\mathbb{E}_t F(S_T) \approx \mathbb{E}_t F^{BS}(S_t, \bar{\sigma}_{t,T}) + O(\rho^2),$$

so the F claim is first-order *correlation neutral*. We favor synthetic volatility claims which satisfy this property, because of the robustness (indeed, the first-order *invariance*) of the claim's valuation in the presence of correlation.

This motivates the following definition.

Definition 4.4. Let $t < T$. We say that a payoff function F is [first-order] ρ*-neutral* at time t if there exists $c \in m\mathcal{F}_t$ such that

$$\frac{\partial F^{\mathsf{P}_\sigma}}{\partial \mathsf{s}}(S_t) \equiv \frac{\partial F^{BS}}{\partial \mathsf{s}}(S_t, \sigma) = c \qquad \text{for } \textit{all} \text{ constants } \sigma \geq 0,$$

almost surely. In other words, "the payoff's Black-Scholes delta is constant across all volatility parameters".

The ρ-neutralization concept of making portfolio values insensitive to correlation improves the robustness of our valuation schemes. To improve also the robustness of our *replication* schemes, we will favor trading strategies whose portfolio *delta* is zero even in the presence of correlation. Hence we define also a concept of Δ-neutrality.

Definition 4.5. We say that F at time t is Δ-*neutral* with respect to a distribution P, if

$$\frac{\partial F^{\mathsf{P}}}{\partial s}(S_t) = 0$$

almost surely. If we simply say Δ-*neutral*, without specifying the distribution, then the distribution is understood to be P_{true}, which denotes the conditional $\mathbb{P}$-distribution of S_T/S_t given $\mathcal{F}_t$.

Remark 4.6. A payoff's Δ-neutrality with respect to a particular P does not imply ρ-neutrality; but Δ-neutrality with respect to *all lognormal* P does imply ρ-neutrality. Conversely, a payoff's ρ-neutrality, even with $c = 0$, does not imply Δ-neutrality with respect to arbitrary P; but it does imply Δ-neutrality with respect to a class of P that includes all lognormal distributions.

Any payoff may be made Δ-neutral by adding an appropriately chosen affine function $\alpha S_T + \beta$ (where $\alpha, \beta \in m\mathcal{F}_t$). In contrast, adding such affine functions has no effect on whether or not a payoff is ρ-neutral.

Definition 4.7. Consider a *trading strategy* which holds at each time $t < T$ a portfolio of claims whose combined time-T payout is $F_t(S_T)$, where F_t is a payoff function in the sense of (4.2). We say that the trading strategy is [first-order] (ρ, Δ)-*neutral* if for each $t < T$, the payoff function F_t is ρ-neutral and Δ-neutral.

5 Exponentials

Consider an exponential variance claim which pays $e^{\lambda\langle X\rangle_T}$ for some constant λ.

5.1 Basic replication

We introduce first a "correlation-sensitive" replication strategy for the exponential of variance, relying on the independence assumption (I). In the next section, we will improve this to a "correlation-robust" strategy, which is robust to first-order against violations of the independence assumption.

Theorem 5.1 (Basic pricing of exponentials)**.** *For each $\lambda \in \mathbb{C}$ and $t \le T$,*

$$\mathbb{E}_t e^{\lambda\langle X\rangle_T} = e^{\lambda\langle X\rangle_t}\mathbb{E}_t(S_T/S_t)^{1/2\pm\sqrt{1/4+2\lambda}}. \tag{5.1}$$

In particular, for $t = 0$,

$$\mathbb{E}_0 e^{\lambda\langle X\rangle_T} = \mathbb{E}_0(S_T/S_0)^{1/2\pm\sqrt{1/4+2\lambda}}. \tag{5.2}$$

Theorem 5.2 (Basic replication of exponentials)**.** *Let $\lambda \in \mathbb{R}$. If $p := 1/2 \pm \sqrt{1/4 + 2\lambda} \in \mathbb{R}$ then the payoff $e^{\lambda\langle X\rangle_T}$ admits replication by the self-financing strategy*

$$\begin{aligned} N_t &\quad \textit{claims on } S_T^p \\ -pN_tP_{t-}/S_t &\quad \textit{shares} \\ pN_tP_{t-} &\quad \textit{bonds} \end{aligned} \tag{5.3}$$

where $N_t := e^{\lambda\langle X\rangle_t}/S_t^p$ and $P_t := \mathbb{E}_t S_T^p$.

Remark 5.3. Although S and N are continuous, P is merely cadlag and may jump. Nonetheless, we are free to replace the predictable process P_{t-} with the adapted process P_t everywhere in the statement and proof of Theorem 5.2, because the relevant integrators (S, B, N) are continuous, so the distinction between P_t and P_{t-} is immaterial. Thus we have proved that the strategy

$$\begin{aligned} N_t &\quad \text{claims on } S_T^p \\ -pN_tP_t/S_t &\quad \text{shares} \\ pN_tP_t &\quad \text{bonds} \end{aligned} \tag{5.4}$$

replicates $e^{\lambda\langle X\rangle_T}$. Henceforth we follow the standard practice of allowing trading strategies and stochastic integrands with respect to continuous price processes to be one-side-continuous adapted processes, as in (5.4).

Remark 5.4. If futures are available as hedging instruments, then they can replace the shares and cash; the strategy to replicate the payoff $e^{\lambda\langle X\rangle_T}$ becomes

$$\begin{aligned} N_t &\quad \text{claims on } S_T^p \\ -pN_tP_t/S_t &\quad \text{futures} \end{aligned}$$

by similar reasoning.

Remark 5.5. For complex λ and p, and complex $\alpha = \alpha(\lambda)$,

$$\mathrm{Re}(\alpha N_T P_T) = \mathrm{Re}(\alpha P_0) + \int_0^T \mathrm{Re}(\alpha N_t) d\mathrm{Re}(P_t) - \int_0^T \mathrm{Im}(\alpha N_t) d\mathrm{Im}(P_t) - \int_0^T \frac{\mathrm{Re}(p\alpha P_t N_t)}{S_t} dS_t$$

so we can replicate $\mathrm{Re}(\alpha e^{\lambda\langle X\rangle_T})$ by trading cosine and sine claims. Specifically, at time t, hold

$$\begin{aligned} \mathrm{Re}(\alpha N_t) &\quad \text{claims on } \mathrm{Re}(e^{pX_T}) \\ -\mathrm{Im}(\alpha N_t) &\quad \text{claims on } \mathrm{Im}(e^{pX_T}) \\ -\mathrm{Re}(p\alpha N_t P_t)/S_t &\quad \text{shares} \\ \mathrm{Re}(p\alpha N_t P_t) &\quad \text{bonds.} \end{aligned}$$

Remark 5.6. Under assumption (I), Theorem 5.2 shows that the trading strategy works perfectly; moreover the strategy is Δ-neutral, because the combined payoff function of the time-t holdings is

$$F(\mathsf{S}) := N_t(\mathsf{S}/S_t)^p - \frac{pN_tP_t}{S_t}(\mathsf{S} - S_t),$$

which satisfies

$$\frac{\partial F^{\mathsf{P}_{\text{true}}}}{\partial \mathsf{s}}(S_t) = N_t \mathbb{E}_t\left(\frac{\partial}{\partial \mathsf{s}}(\mathsf{s} S_T/S_t^2)^p\right)\Big|_{\mathsf{s}=S_t} - \frac{pN_tP_t}{S_t} = 0.$$

Thus the share position $-pN_tP_t/S_t$ can be interpreted as a delta-hedge of the option position consisting of N_t claims on S_T^p. This makes sense, from the standpoint that if we want to create a purely volatility-dependent payoff, then we do not want to have net exposure to directional risk, hence we delta-neutralize.

Of course, this observation is neither necessary nor sufficient to prove the validity of our hedging strategy (for that purpose the Theorem 5.2 proof speaks for itself); but it can help us to understand and implement the strategy.

The strategy is not ρ-neutral. We remove this correlation-sensitivity in the next section.

Remark 5.7. From Carr-Lee [9], it follows that assumption (I) implies a general form of put-call symmetry: for *any* time-t-contracted payoff function f such that $f(S_T/S_t)$ is integrable,

$$\mathbb{E}_t f\left(\frac{S_T}{S_t}\right) = \mathbb{E}_t\left[\frac{S_T}{S_t} f\left(\frac{S_t}{S_T}\right)\right]. \tag{5.5}$$

Combining Theorem 5.1 and (5.5), we have a large *family* of European-style payoffs which correctly price the variance payoff: For all such f,

$$\mathbb{E}_t e^{\lambda\langle X\rangle_T} = e^{\lambda\langle X\rangle_t}\mathbb{E}_t\left[(S_T/S_t)^{1/2+\sqrt{1/4+2\lambda}} + f(S_T/S_t) - \frac{S_T}{S_t} f(S_t/S_T)\right]. \tag{5.6}$$

In particular, choosing $f(\mathsf{S}) := \alpha \mathsf{S}^{1/2-\sqrt{1/4+2\lambda}} + \beta \mathsf{S}$ for $\alpha, \beta \in m\mathcal{F}_t$ yields the sub-family of identities

$$\mathbb{E}_t e^{\lambda\langle X\rangle_T} = e^{\lambda\langle X\rangle_t}\mathbb{E}_t\left[(1-\alpha)(S_T/S_t)^{1/2+\sqrt{1/4+2\lambda}} + \alpha(S_T/S_t)^{1/2-\sqrt{1/4+2\lambda}} + \beta(S_T/S_t - 1)\right] \tag{5.7}$$

under (I). In the next section we choose α and β in such a way as to achieve (ρ, Δ)-neutrality.

5.2 Correlation-robust replication of exponentials

Define

$$\begin{aligned} \theta_\pm(\lambda) &:= \frac{1}{2} \mp \frac{1}{2\sqrt{1+8\lambda}} \\ p_\pm(\lambda) &:= \frac{1}{2} \pm \frac{1}{2}\sqrt{1+8\lambda} \end{aligned} \tag{5.8}$$

Theorem 5.8 (Correlation-robust pricing of exponentials)**.** *Let $t \le T$. For any $\lambda \in \mathbb{C}$,*

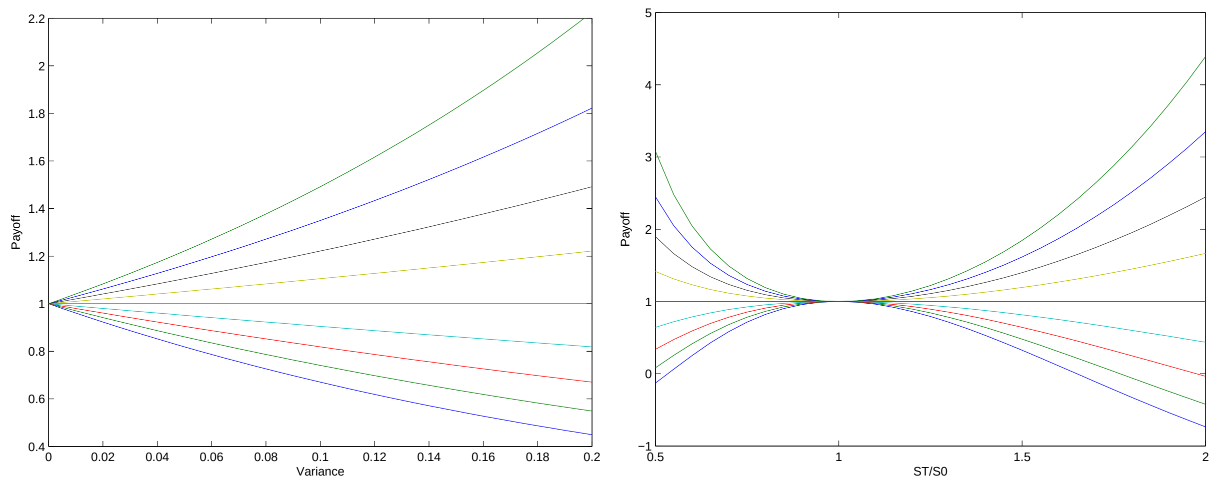
$$\mathbb{E}_t e^{\lambda\langle X\rangle_T} = \mathbb{E}_t G_{\text{exp}}(S_T, S_t, \langle X\rangle_t; \lambda). \tag{5.9}$$

where

$$G_{\text{exp}}(\mathsf{S}, \mathsf{u}, q; \lambda) := e^{\lambda q}\left[\theta_+(\mathsf{S}/\mathsf{u})^{p_+} + \theta_-(\mathsf{S}/\mathsf{u})^{p_-}\right] \tag{5.10}$$

For each t, the payoff function $F(\mathsf{S}) := G_{\text{exp}}(\mathsf{S}, S_t, \langle X\rangle_t; \lambda)$ is ρ-neutral.

Figure 5.1: Exponential variance claims $e^{\lambda\langle X\rangle_T}$ on the left, and their European-style synthetic counterparts $G_{\exp}(S_T; S_0; \langle X\rangle_0; \lambda)$ on the right, for $\langle X\rangle_0 = 0$ and $\lambda \in \{-4, -3, \ldots, 3, 4\}$.



Remark 5.9. Therefore the relationship

$$\mathbb{E}_t e^{\lambda\langle X\rangle_T} = e^{\lambda\langle X\rangle_t}\mathbb{E}_t\Big[\theta_+(S_T/S_t)^{p_+} + \theta_-(S_T/S_t)^{p_-}\Big] \tag{5.11}$$

holds exactly under independence (I), and is first-order *robust* to the presence of correlation. Figure 5.1 plots the payoff functions appearing in the left and right-hand sides.

Theorem 5.10 (Correlation-robust replication of exponentials)**.** *Define* $p_\pm$ *and* $\theta_\pm$ *as in* (5.8)*. Let*

$$N_t^\pm := e^{\lambda\langle X\rangle_t}/S_t^{p_\pm}$$
$$P_t^\pm := \mathbb{E}_t S_T^{p_\pm}$$

If $\lambda \in \mathbb{R}$ *and* $p_\pm \in \mathbb{R}$*, then the strategy*

$$\begin{array}{ll} \theta_+ N_t^+ & \textit{claims on} \quad S_T^{p_+} \\ \theta_- N_t^- & \textit{claims on} \quad S_T^{p_-} \\ -\delta_t & \textit{shares} \\ \delta_t S_t & \textit{cash} \end{array}$$

replicates the payoff $e^{\lambda\langle X\rangle_T}$*, where*

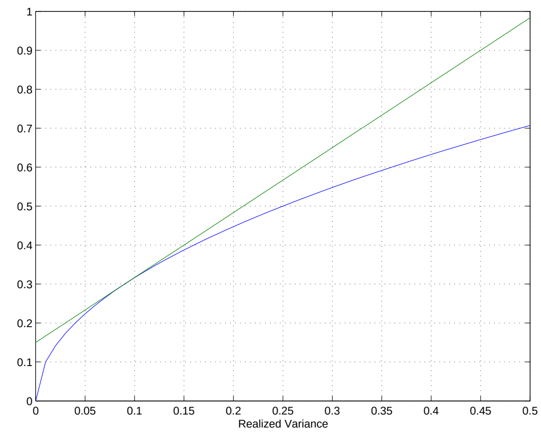
$$\delta_t := \frac{\theta_+ p_+ N_t^+ P_t^+ + \theta_- p_- N_t^- P_t^-}{S_t}.$$

If (I) *holds then* $\delta_t = 0$*. Regardless of whether* (I) *holds, the strategy is* (ρ, Δ)*-neutral.*

6 Volatility swap

A *volatility swap* pays $\sqrt{\langle X\rangle_T}$ minus some agreed fixed amount, which we take to be 0 unless otherwise specified.

Figure 6.1: The volatility swap payoff admits model-independent superreplication



6.1 Bounds and approximations

For $F_{\text{atmc}}(S) := (S - S_0)^+$, a direct computation shows that

$$F^{BS}_{\text{atmc}}(S_0, \sigma) = S_0(N(\sigma/2) - N(-\sigma/2))$$

which is strictly increasing and concave in σ.

Define the unannualized at-the-money *implied volatility* IV_0 as the unique solution to

$$F^{BS}_{\text{atmc}}(S_0, \text{IV}_0) = \mathbb{E}_0 F_{\text{atmc}}(S_T). \tag{6.1}$$

Let VAR_0 and VOL_0 denote respectively the time-0 values of the variance swap (quoted as a volatility) and the volatility swap.

$$\text{VAR}_0 := \sqrt{\mathbb{E}_0 \langle X \rangle_T} \tag{6.2}$$

$$\text{VOL}_0 := \mathbb{E}_0 \sqrt{\langle X \rangle_T}. \tag{6.3}$$

These values are model-independently determined by prices of Europeans, according to sections 3 and 6.2 respectively. In particular, VAR_0 equals the square root of the value of the log contract; VAR_0 is what the VIX attempts to approximate, and is sometimes described as the *model-free implied volatility.*

Theorem 6.1. *We have the following observable lower and upper bounds on* VOL_0

$$\frac{\sqrt{2\pi}}{S_0}\mathbb{E}_0(S_T - S_0)^+ \underset{(1)}{\leq} \text{IV}_0 \underset{(2)}{\leq} \text{VOL}_0 \underset{(3)}{\leq} \text{VAR}_0 = \sqrt{-2\mathbb{E}_0 \log(S_T/S_0)}.$$

Inequalities (1) and (3) do not assume (I).

Remark 6.2. If variance and volatility swap rates fail to respect (3), then an arbitrage profit can be locked in, model independently. The volatility swap admits robust superreplication, as follows. Then a portfolio of VAR_0 in cash, together with $1/(2\,\mathrm{VAR}_0)$ variance swaps each of which pays $\langle X\rangle_T - \mathrm{VAR}_0^2$, superreplicates the $\langle X\rangle_T^{1/2}$ payoff. Essentially this portfolio enforces Jensen's inequality, by constructing the appropriate tangent, as shown in Figure 6.1.

If VOL_0 fails to respect (2), then arbitrage profit can be locked in, using the trading strategy of the next section.

In Remarks 6.3 and 6.4, we include some approximations, mainly to provide reference points and context for our theory. We emphasize that we do not actually advocate the use of these two approximations, because our theory is more powerful and robust, in ways described in Remark 6.5.

Remark 6.3. Although $F^{BS}_{\mathrm{atmc}}(S_0,\cdot)$ is concave, it is nearly linear – indeed, linear to a *second* order approximation near 0, because its second derivative vanishes at 0. Thus the inequality in (A.1) is an approximate equality (as shown by Feinstein [14] and Poteshman [21]); and the inequality in (A.2) is an approximate equality (as shown by Brenner and Subrahmanyam [7]). Therefore, the Theorem 6.1 lower bounds are indeed *approximately equal* to the volatility swap rate:

$$\mathrm{VOL}_0 \approx \frac{\sqrt{2\pi}}{S_0}\mathbb{E}_0(S_T - S_0)^+ \approx \mathrm{IV}_0 \tag{6.4}$$

where the first $\approx$ assumes (I), but the second does not.

This does not show that ATM implied volatility is approximately an "unbiased predictor" of future realized volatility, but rather that ATM implied volatility approximately *prices correctly* a claim on future realized volatility, where correctness means absence of arbitrage against Europeans; the $\mathbb{E}$ is with respect to a pricing measure, not the physical measure.

Remark 6.4. Under assumption (I), the approximation (6.4) can be refined, to the following simple approximation using ATM implied and the variance swap rate:

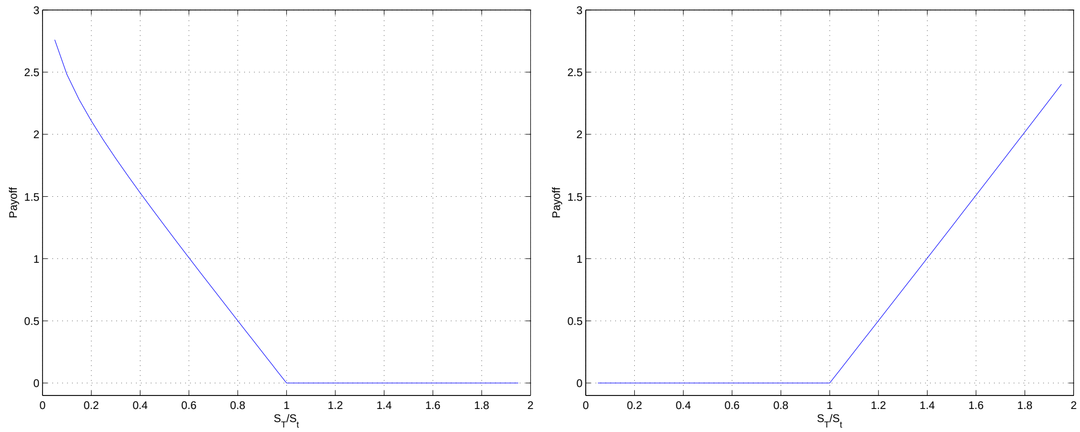
$$\mathrm{VOL}_0 \approx \mathrm{IV}_0\left(1 + \frac{\mathrm{VAR}_0^2 - \mathrm{IV}_0^2}{8 + 2\,\mathrm{IV}_0^2}\right). \tag{6.5}$$

Remark 6.5. We do not endorse the approximations (6.4) and (6.5). They do not establish a replication strategy for realized volatility (or functions thereof), they do not apply at times after inception, they do not price general volatility derivatives, and they do not suggest what to do in the presence of correlation. Our theory does all of the above. Regarding the last point in particular, Section 6.5 will illustrate the correlation-robustness of our approach, compared to the naive approximation (6.4).

6.2 Basic / correlation-sensitive replication

We introduce first a basic "correlation-sensitive" valuation strategy for the volatility swap, relying on the independence assumption (I). In the next section, we will improve this to a "correlation-robust" strategy, which is immunized, to first order, against the presence of correlation.

Figure 6.2: Basic replication of a volatility swap: European-style payoffs $g_-(S_T)$ and $g_+(S_T)$



For our correlation-robust strategy, we will give a full treatment, including seasoned volatility swaps at times $t > 0$, and including the replication argument. For our basic strategy, however, we restrict our coverage to the valuation of volatility swaps at inception $t = 0$, because we do not advocate the basic strategy; for the basic case we include only enough material to draw some connections with other representations/approximations, in Remarks 6.7 and 6.13 and Section 6.5.

Theorem 6.6 (Basic / correlation-sensitive synthetic volatility swap)**.** *We have*

$$\mathbb{E}_0\sqrt{\langle X\rangle_T} = \mathbb{E}_0 g_\pm(S_T/S_0)$$

where

$$g_\pm(x) := \frac{1}{2\sqrt{\pi}}\int_0^\infty \frac{e^{(1/2\pm1/2)\log x} - \mathrm{Re}[e^{(1/2\pm\sqrt{1/4-2z})\log x}]}{z^{3/2}}dz \tag{6.6}$$

In particular, the convergence of the integral is guaranteed.

Remark 6.7. Figure 6.2 plots the functions $g_\pm$. Note their strong resemblance to $\sqrt{2\pi}/S_0$ at-the-money puts and calls, respectively. Our result is consistent with the naive approximation (6.4), but as discussed in Remark 6.5, our theory has implications far beyond the naive approximations.

We call a claim on $g_+(S_T/S_0)$ the *basic / correlation-sensitive synthetic volatility swap.*

6.3 Correlation-robust replication

Theorem 6.8 (Correlation-robust synthetic volatility swap)**.** *For all* $t \in [0, T]$,

$$\mathbb{E}_t\sqrt{\langle X\rangle_T} = \mathbb{E}_t G_{\mathrm{svs}}(S_T, S_t, \langle X\rangle_t) \tag{6.7}$$

where

$$G_{\mathrm{svs}}(\mathsf{S}, \mathsf{u}, q) := \frac{1}{2\sqrt{\pi}}\int_0^\infty \theta_+\frac{1-e^{-zq}(\mathsf{S}/\mathsf{u})^{p_+}}{z^{3/2}} + \theta_-\frac{1-e^{-zq}(\mathsf{S}/\mathsf{u})^{p_-}}{z^{3/2}}dz. \tag{6.8}$$

$$\theta_\pm := \theta_\pm(-z) := \frac{1}{2} \mp \frac{1}{2\sqrt{1-8z}} \qquad p_\pm := p_\pm(-z) := \frac{1}{2} \pm \frac{1}{2}\sqrt{1-8z} \tag{6.9}$$

In particular, the convergence and integrability of G_{svs} are guaranteed.

For each t, the payoff function $F(\mathsf{S}) := G_{\text{svs}}(\mathsf{S}, S_t, \langle X\rangle_t)$ is ρ-neutral.

Remark 6.9. We call a claim on $G_{\text{svs}}(S_T, S_t, \langle X\rangle_t)$ the time-t correlation-robust *synthetic volatility swap.*

Note that the correlation-robust synthetic volatility swap is *not* simply a linear combination of the put-like and call-like payoffs (6.6), because the linear combinations are taken *inside* the z-integral, and the weights $\theta_\pm$ depend on z.

It does resemble a straddle, but its arms are curved, not straight. Indeed, the three arguments of the payoff function $G_{\text{svs}}(\mathsf{S}, \mathsf{u}, q)$ have the following interpretation: S stands for the terminal share price; u represents the "strike" of the curved straddle; and q controls the "curvature" of the curved straddle. Theorem 6.8 shows that the "strike" should be chosen at-the-money and that the "curvature" should be depend on how much variance has been already accumulated.

At inception, the correlation-robust synthetic volatility swap may be written concisely in terms of Bessel functions.

Corollary 6.10 (Payoff of newly-issued synthetic volatility swap: Bessel formula)**.** *Let I_ν denote the modified Bessel function of order ν. Then*

$$\mathbb{E}_0\sqrt{\langle X\rangle_T} = \mathbb{E}_0\psi(S_T) \tag{6.10}$$

where $\psi(\mathsf{S}) := \phi(\log(\mathsf{S}/S_0))$ where

$$\phi(x) := \sqrt{\frac{\pi}{2}}e^{x/2}\Big|xI_0(x/2) - xI_1(x/2)\Big|. \tag{6.11}$$

The payoff is ρ-neutral.

Instead of expressing the synthetic volatility swap as a payoff function, we may express it as a mixture of put and call payoffs. We treat separately the case of a newly-issued volatility swap and the case of a seasoned volatility swap.

Corollary 6.11 (Put/call decomposition of newly-issued synthetic volatility swap: Bessel formula)**.** *The initial ($\langle X\rangle_t = 0$) correlation-robust synthetic volatility swap decomposes into the payoffs of*

$$\begin{aligned} \sqrt{\pi/2}/S_0 &\quad \textit{straddles at strike } K = S_0 \\ \sqrt{\frac{\pi}{8K^3S_0}}\Big[I_1(\log\sqrt{K/S_0}) - I_0(\log\sqrt{K/S_0})\Big]dK &\quad \textit{calls at strikes } K > S_0 \\ \sqrt{\frac{\pi}{8K^3S_0}}\Big[I_0(\log\sqrt{K/S_0}) - I_1(\log\sqrt{K/S_0})\Big]dK &\quad \textit{puts at strikes } K < S_0 \end{aligned} \tag{6.12}$$

Corollary 6.12 (Put/call decomposition of seasoned synthetic volatility swap)**.** *The seasoned* $(\langle X\rangle_t > 0)$ *correlation-robust synthetic volatility swap decomposes into the payoffs of*

$$\begin{aligned} \frac{dK}{\sqrt{\pi}} \int_0^\infty \frac{e^{-z\langle X\rangle_t}}{K^2 z^{1/2}} \big[\theta_+ (K/S_t)^{p_+} + \theta_- (K/S_t)^{p_-}\big] dz \quad & \textit{calls at strikes } K > S_t, \textit{ puts at } K < S_t \\ \langle X\rangle_t^{1/2} \quad & \textit{cash} \end{aligned} \tag{6.13}$$

together with a zero-cost delta-hedge.

Remark 6.13. Friz-Gatheral [15] found one Bessel representation of the Carr-Lee *basic* synthetic volatility swap (6.6). In contrast, we find two Bessel representations of the Carr-Lee *correlation-robust* volatility swap, in Corollaries 6.10 (Bessel formula for payoff) and 6.11 (Bessel formula for put/call decomposition).

Theorem 6.14 (Synthetic volatility swap replicates the volatility swap)**.** *Holding at each time t a delta-hedged claim on $G_{\rm svs}(S_T, S_t, \langle X\rangle_t)$ replicates the volatility swap. In other words:*

Choose an arbitrary constant $\kappa > 0$ as a put/call separator. For $K \in (0, \kappa)$ let $P_t(K)$ be the time-t value of a K-strike T-expiry binary put. For $K \geq \kappa$ let $P_t(K)$ be the time-t value of a K-strike T-expiry binary call.

Let the time-t binary option holdings be given by the signed measure φ_t defined by the density function $K \mapsto \pm\partial G_{\rm svs}/\partial \mathsf{S}(K; S_t, \langle X\rangle_t)$ on the domain $K \in (0, \infty)$, where the $+$ and $-$ correspond to $K > \kappa$ and $K < \kappa$ respectively.

Then the strategy of holding at each time t

$$\begin{aligned} \varphi_t \quad & \textit{options} \\ G_{\rm svs}(\kappa, S_t, \langle X\rangle_t) + \delta_t S_t \quad & \textit{bonds} \\ -\delta_t \quad & \textit{shares} \end{aligned} \tag{6.14}$$

replicates the payoff $\sqrt{\langle X\rangle_T}$, where

$$\delta_t := \mathbb{E}_t \frac{\partial}{\partial \mathsf{s}}\bigg|_{\mathsf{s}=S_t} G_{\rm svs}(\mathsf{s} S_T/S_t, S_t, \langle X\rangle_t) = -\mathbb{E}_t \frac{\partial G_{\rm svs}}{\partial \mathsf{u}}(S_T, S_t, \langle X\rangle_t) \tag{6.15}$$

is observable from the time-t prices of T-expiry options.

If (I) *holds, then $\delta_t = 0$. Regardless of whether* (I) *holds, the strategy is (ρ, Δ)-neutral.*

6.4 Evolution of the synthetic volatility swap

As variance accumulates during the life of the synthetic volatility swap, its payoff profile evolves. Theorem 6.8 makes this precise, but here let us give some intuition.

The initial payoff resembles a straddle struck at-the-money. The dynamics of the payoff depend on two factors. First, as the spot moves, the "strike" of the "straddle" floats to stay at-the-money. Second, as quadratic variation (an increasing process) accumulates, the "straddle" smooths out, losing its kink; indeed, only when $\langle X\rangle_t = 0$ does the kink literally exist.

Figure 6.3: Evolution of the synthetic volatility swap: $\langle X \rangle_t = 0.0$, and $\langle X \rangle_t = 0.2$

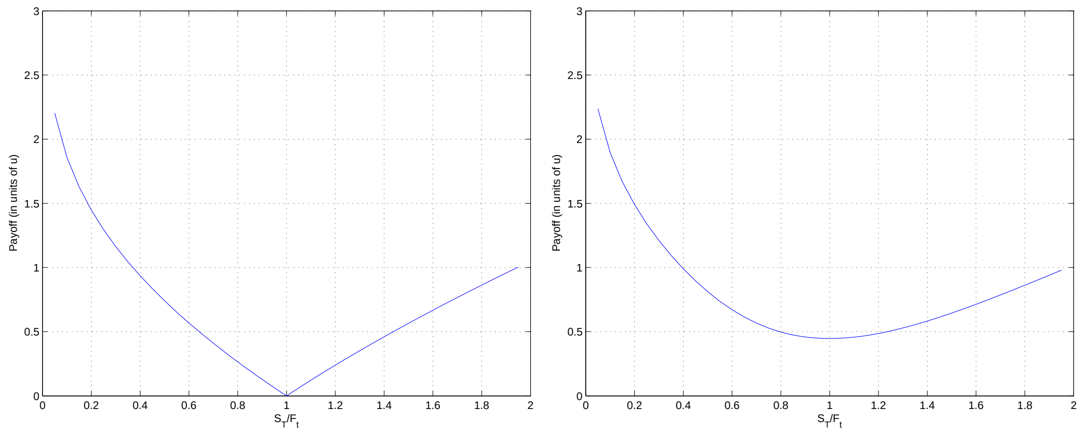


Figure 6.4: Evolution of the synthetic volatility swap: $\langle X \rangle_t = 0.4$, and $\langle X \rangle_t = 0.6$

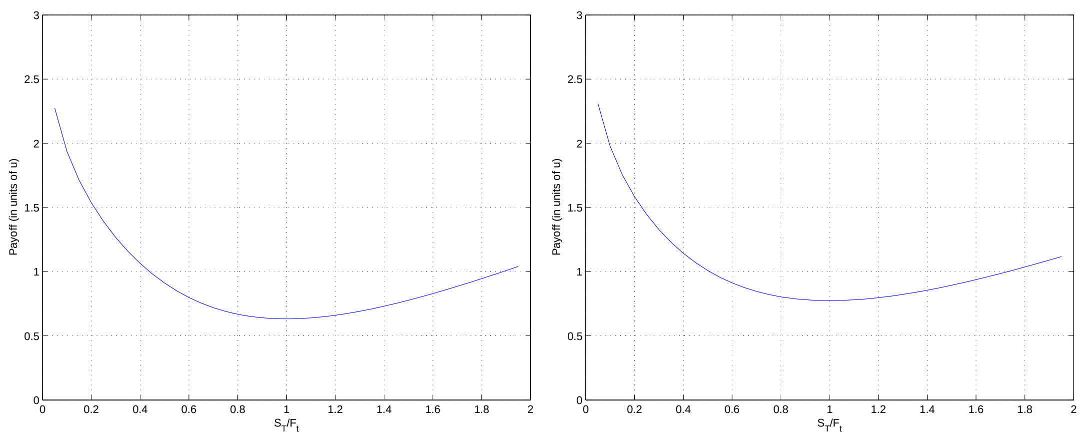
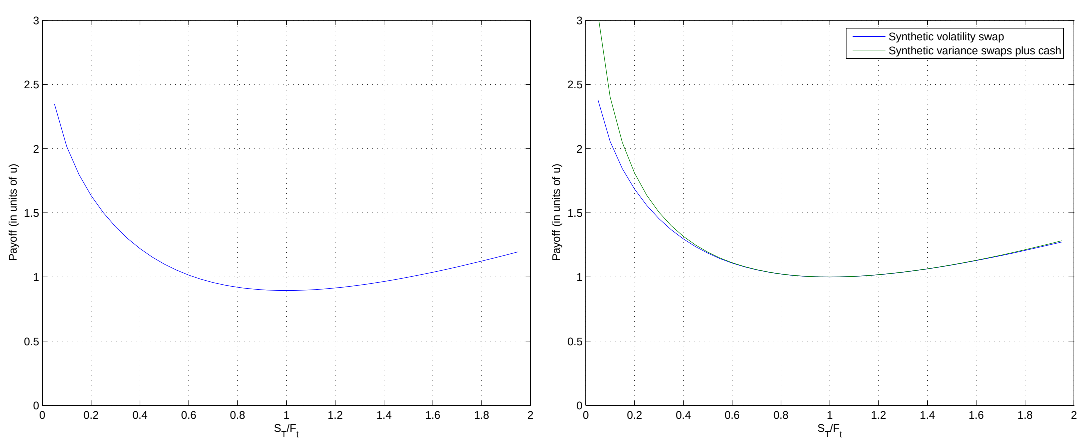


Figure 6.5: Evolution of the SVS: $\langle X \rangle_t = 0.8$, and $\langle X \rangle_t = 1.0$ compared to the log approximation



We can, moreover, understand the shape which the payoff approaches as it smooths out. Suppose that at time t the "running variance" $\langle X\rangle_t$ is large compared to the (random) remaining variance $R_T := \langle X\rangle_T - \langle X\rangle_t$. The volatility contract pays

$$\sqrt{\langle X\rangle_T} = \sqrt{\langle X\rangle_t + R_T} \approx \sqrt{\langle X\rangle_t} + \frac{1}{2\sqrt{\langle X\rangle_t}} R_T,$$

but from variance swap theory, we have

$$\mathbb{E}_t R_T = \mathbb{E}_t\big[-2\log(S_T/S_t) + 2(S_T/S_t - 1)\big].$$

So at time t, the synthetic volatility swap's time-T payoff should resemble

$$\sqrt{\langle X\rangle_t} + \frac{1}{\sqrt{\langle X\rangle_t}}(S_T/S_t - 1 - \log(S_T/S_t))$$

for $S_T/S_t \approx 1$. In other words, the synthetic volatility swap should *evolve toward a hedged log contract*, which is confirmed in Figure 6.5.

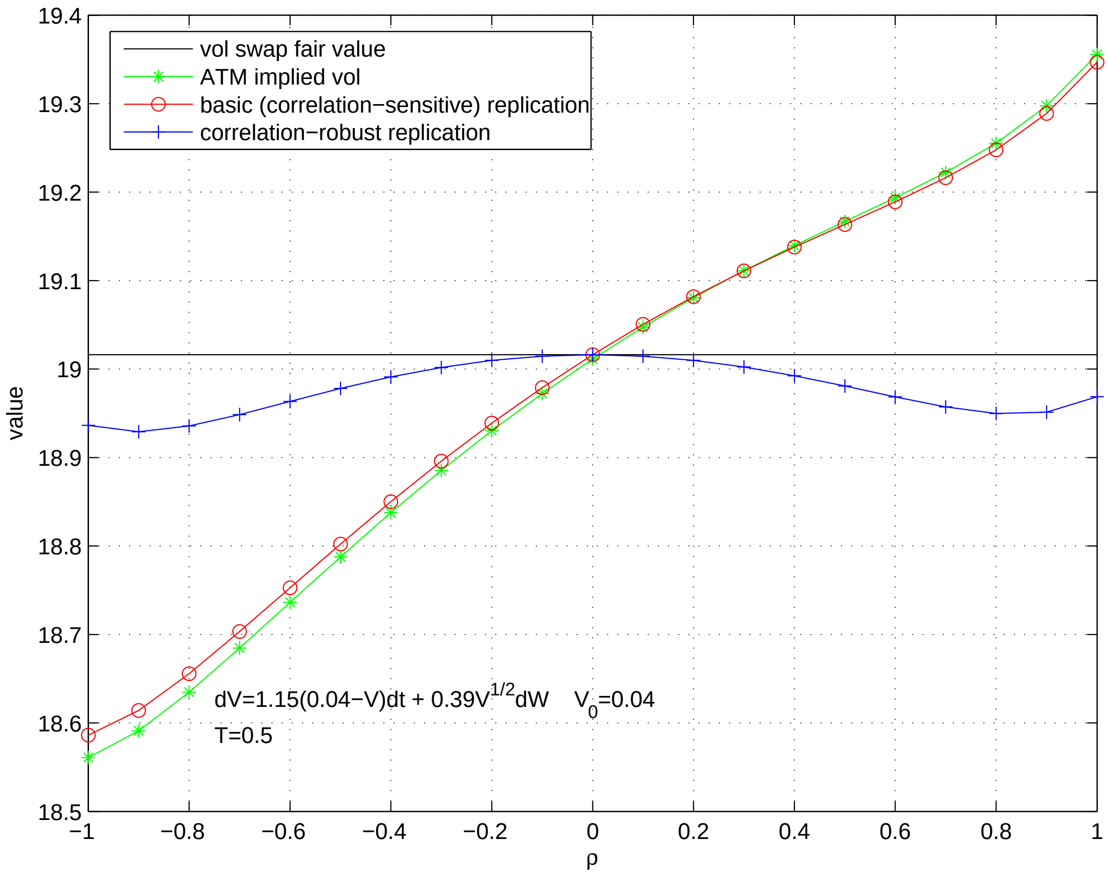
Intuitively, as variance accumulates, the square-root function of Figure 6.1 loses convexity, and the linear-in-variance upper bound (which can be enforced by a superreplicating position in log contracts) becomes a better approximation to the volatility swap payoff.

6.5 Accuracy of the ρ-neutral synthetic volatility swap

Figure 6.6 shows how closely the time-0 ρ-neutral synthetic volatility swap price approximates the true volatility swap fair value, under Heston dynamics with parameters from Bakshi-Cao-Chen [3]. For comparison, we plot also the ATM implied volatility, and the basic (correlation-sensitive) hedge portfolio price.

As approximations of the true volatility swap value, our correlation-robust synthetic volatility swap outperforms ATM implied volatility and outperforms our basic (correlation-sensitive) replication – across essentially all correlation assumptions. In the case $\rho = 0$, our methods are (as promised) exact and the implied volatility approximation is nearly exact; but more importantly, in the empirically relevant case of $\rho \neq 0$, our correlation-robust synthetic volatility swap's relative "flatness" with respect to ρ results in its greater accuracy. This illustrates why, in equity markets, we do not recommend methods or approximations which rely on assumption (I) without the robustness provided by the ρ-neutralization present in our correlation-robust synthetic volatility swap.

Figure 6.6: Heston dynamics: Volatility swap valuations as functions of correlation



We comment on each curve in greater detail.

The volatility swap fair value (which we denote by VOL_0 as in Section 6.1) equals the expectation of realized volatility. It is determined by the distribution of realized variance $\int V_t dt$, which is determined entirely by the given dynamics

$$dV_t = 1.15(0.04 - V_t)dt + 0.39V_t^{1/2}dW_t, \quad V_0 = 0.04$$

of the instantaneous variance $V_t = \sigma_t^2$. Hence the correlation ρ is irrelevant to VOL_0, which therefore plots as a horizontal line.

The basic (correlation-sensitive) synthetic volatility swap payoff is approximately the payoff of $\sqrt{2\pi}/S_0$ calls, as noted in Remark 6.7. Therefore its value and the ATM Black-Scholes implied volatility are nearly equal, due to (6.4). The plots confirm this across the full range of ρ. More importantly, the plots confirm that VOL_0 is well-approximated by these two values *if* $\rho = 0$, but due to the correlation-sensitivity of ATM implied volatility and the basic synthetic volatility swap, both values underestimate VOL_0 by more than 40 basis points, for certain values of ρ.

Our correlation-robust synthetic volatility swap, as promised, has exactly the same value as VOL_0 if $\rho = 0$. Furthermore, its value is, as intended by its design, ρ-invariant to first-order, at $\rho = 0$. While there is no guarantee that this flatness will extend to ρ far from 0, the numerical evidence here is that indeed the ρ-neutrality results in accuracy gains across the entire range of ρ.

Finally we comment on a benchmark not plotted in the figure. The variance swap value (which equals the log-contract value) is 0.04; and its square root (which we denote by VAR_0 as in Section 6.1, and which the VIX seeks to approximate) is 0.20, regardless of ρ. Therefore, a plot of VAR_0 would be a horizontal line at 20.0 volatility points, far above the upper boundary of Figure 6.6. (In particular, consistently with Theorem 6.1, VAR_0 exceeds VOL_0.)

To summarize, in this example the best approximation of VOL_0 is given by our correlation-robust synthetic volatility swap value, and the worst (due to Jensen's inequality) is given by the VIX-style quantity VAR_0. The other approximations – ATM implied volatility and our basic (correlation-sensitive) volatility swap – are accurate *if* assumption (I) holds.

Remark 6.15. Figure 6.6 can be regarded as a numerical comparison of two notions of model-free implied volatility (MFIV). As defined in the "VIX-style," MFIV is typically understood to mean VAR_0, the square root of the variance swap (or log contract) value. Here we have introduced the correlation-robust *synthetic volatility swap*, whose value can be regarded as an alternative notion of MFIV, which we can describe as "SVS-style" MFIV.

The SVS-style MFIV is arguably both more and less aptly called "model-free implied volatility" than the VIX-style MFIV is. The SVS notion is arguably more apt, due to the word "volatility" in MFIV: the synthetic volatility swap's value, we have shown, gives truly the risk-neutral expectation of realized *volatility* – a value distinct from (and by Jensen's inequality smaller than) the VIX-style definition which gives the square root of expected realized *variance*. The SVS notion is arguably less apt, due to the words "model free" in MFIV: the square of the VIX-style MFIV gives expected variance for essentially all continuous price processes, whereas the SVS-style MFIV gives exactly the expected volatility across for a smaller set of models (those satisfying (I)). Nonetheless, even for the $\rho \neq 0$ dynamics in Figure 6.6 which violate (I), expected volatility VOL_0 is approximated much more accurately by the SVS-style MFIV than by the VIX-style MFIV.

7 Pricing other volatility derivatives

Using exponential variance payoffs, we can synthesize general variance payoffs.

7.1 Fractional power payoffs

Our volatility swap formula is the $r = 1/2$ case of the following generalization to powers in $(0, 1)$.

Theorem 7.1. *For $0 < r < 1$,*

$$\mathbb{E}_t\langle X\rangle_T^r = \mathbb{E}_t G_{\text{pow}(r)}(S_T, S_t, \langle X\rangle_t)$$

where

$$G_{\text{pow}(r)}(\mathsf{S}, \mathsf{u}, q) := \frac{r}{\Gamma(1-r)} \int_0^\infty \theta_+ \frac{1 - e^{-zq}(\mathsf{S}/\mathsf{u})^{p_+}}{z^{r+1}} + \theta_- \frac{1 - e^{-zq}(\mathsf{S}/\mathsf{u})^{p_-}}{z^{r+1}} dz \tag{7.1}$$

$$\theta_\pm := \theta_\pm(-z) := \frac{1}{2} \mp \frac{1}{2\sqrt{1-8z}} \qquad p_\pm := p_\pm(-z) := \frac{1}{2} \pm \frac{1}{2}\sqrt{1-8z} \tag{7.2}$$

For each t, the payoff function $\mathsf{S} \mapsto G_{\text{pow}(r)}(\mathsf{S}, S_t, \langle X\rangle_t)$ is ρ-neutral.

For arbitrary negative powers, we have the following formula for "inverse variance" claims.

Theorem 7.2. *For any $r > 0$ and any ε such that $\langle X\rangle_t + \varepsilon > 0$,*

$$\mathbb{E}_t(\langle X\rangle_T + \varepsilon)^{-r} = \mathbb{E}_t G_{\text{pow}(-r)}(S_T, S_t, \langle X\rangle_t + \varepsilon)$$

where

$$G_{\text{pow}(-r)}(\mathsf{S}, \mathsf{u}, q) := \frac{1}{r\Gamma(r)} \int_0^\infty (\theta_+(\mathsf{S}/\mathsf{u})^{p_+} + \theta_-(\mathsf{S}/\mathsf{u})^{p_-})e^{-z^{1/r}q} dz$$

$$\theta_\pm := \theta_\pm(-z^{1/r}) := \frac{1}{2} \mp \frac{1}{2\sqrt{1-8z^{1/r}}} \qquad p_\pm := p_\pm(-z^{1/r}) := \frac{1}{2} \pm \sqrt{1/4 - 2z^{1/r}}.$$

For each t, the payoff function $F(\mathsf{S}) := G_{\text{pow}(-r)}(\mathsf{S}, S_t, \langle X\rangle_t)$ is ρ-neutral.

7.2 Integer power payoffs

We obtain integer powers of variance by differentiating, in λ, the exponential of $\lambda\langle X\rangle_T$.

Theorem 7.3. *For each positive integer n,*

$$\mathbb{E}_t\langle X\rangle_T^n = \mathbb{E}_t G_{\text{pow}(n)}(S_T, S_t, \langle X\rangle_t)$$

where

$$G_{\text{pow}(n)}(\mathsf{S}, \mathsf{u}, q) := \partial_\lambda^n G_{\exp}(\mathsf{S}, \mathsf{u}, q, \lambda)\big|_{\lambda=0} \tag{7.3}$$

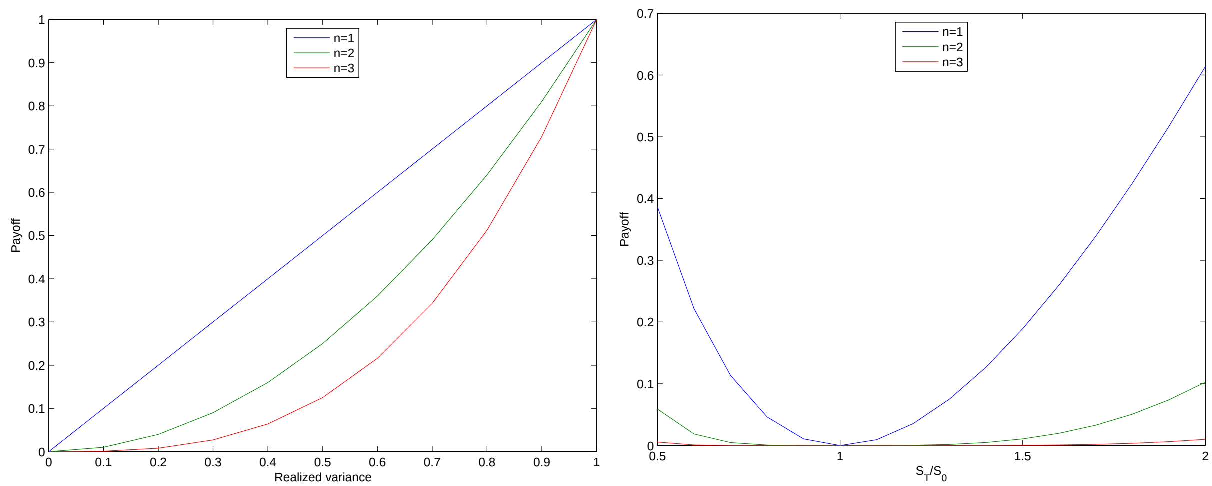
with $G_{\exp}$ defined in (5.10). For each t, the payoff function $F(\mathsf{S}) := G_{\text{pow}(n)}(\mathsf{S}, S_t, \langle X\rangle_t)$ is ρ-neutral.

In particular, we have the following time-0 valuations for $n = 1, 2, 3$:

$$\mathbb{E}_0\langle X\rangle_T = \mathbb{E}_0(-2X_T + 2e^{X_T} - 2)$$
$$\mathbb{E}_0\langle X\rangle_T^2 = \mathbb{E}_0(4X_T^2 + 16X_T + 8X_Te^{X_T} - 24e^{X_T} + 24)$$
$$\mathbb{E}_0\langle X\rangle_T^3 = \mathbb{E}_0(-8X_T^3 + 24X_T^2e^{X_T} - 72X_T^2 - 192X_Te^{X_T} - 288X_T + 480e^{X_T} - 480)$$

Note that $n = 1$ gives the usual valuation of the variance swap using a hedged log contract.

Figure 7.1: Polynomial variance claims $\langle X\rangle_T^n$ on the left, and their European-style synthetic counterparts $G_{\mathrm{pow}(n)}(S_T, S_0, \langle X\rangle_0)$ on the right, for $n = 1, 2, 3$ and $\langle X\rangle_0 = 0$



7.3 Payoffs whose transforms decay exponentially

In sections 7.3 to 7.5 we make use of exponential variance payoffs as *basis* functions, to span a space of general variance payoffs.

Theorem 7.4. *Assume the continuous payoff function $h : \mathbb{R} \to \mathbb{R}$ has bilateral Laplace transform*

$$H(z) = \int_{-\infty}^{\infty} e^{-zq} h(q) dq, \qquad \mathrm{Re}(z) > A, \tag{7.4}$$

such that $|H(\alpha + \beta i)| = O(e^{-|\beta|\mu})$ as $|\beta| \to \infty$ for some $\alpha > A$ and some $\mu > m/2$. Then

$$\mathbb{E}_t h(\langle X\rangle_T) = \mathbb{E}_t G_h(S_T, S_t, \langle X\rangle_t)$$

where

$$G_h(\mathsf{S}, \mathsf{u}, q) := \frac{1}{2\pi i}\int_{\alpha-\infty i}^{\alpha+\infty i} H(z) e^{zq}[\theta_+(\mathsf{S}/\mathsf{u})^{p_+} + \theta_-(\mathsf{S}/\mathsf{u})^{p_-}]dz \tag{7.5}$$

$$\theta_\pm := \theta_\pm(z) := \frac{1}{2} \mp \frac{1}{2\sqrt{1+8z}} \qquad p_\pm := p_\pm(z) := \frac{1}{2} \pm \sqrt{1/4 + 2z}.$$

In particular, the convergence and finite expectation of G_h are guaranteed.

For each t, the payoff function $\mathsf{S} \mapsto G_h(\mathsf{S}, S_t, \langle X\rangle_t)$ is ρ-neutral.

Remark 7.5. Recall the heuristic that the smoother a function, the more rapid the decay of its transform. For payoff functions h which are not sufficiently smooth (including call and put payoffs), the transform H will not decay rapidly enough to satisfy the assumption of Theorem 7.4.

For payoff functions h well-behaved enough to satisfy the stated assumptions, Theorem 7.4 guarantees that the volatility contract can be priced identically to our "synthetic" volatility contract with payoff $G_h(S_T, S_t, \langle X\rangle_t)$, defined by the convergent integral in (7.5). Although this payoff G_h

may be oscillatory in S_T, Theorem 7.4 guarantees that the payoff has a well-defined price, in the sense that the payoff's positive and negative components each have finite expectation.

Nonetheless, implementation difficulties can arise if these finite-priced components are very large and/or concentrated at illiquid strikes, which can occur for volatility contracts h whose replicating price-contracts G_h have payoff profiles with too much variation. In such cases, regularization of the payoff profile can be accomplished by using a finite set of basis functions (such as in Section 7.5); alternatively, distributional inference can be conducted using a finite set of pricing benchmarks (such as in Friz-Gatheral [15]).

7.4 Payoffs whose transforms are integrable

If instead of having exponential decay, the payoff's transform is merely integrable, then our usual pricing formulas of the form $\mathbb{E}h(\langle X\rangle_T) = \mathbb{E}G(S_T)$ may not be available by the Laplace transform method. Nonetheless, the prices of claims on S_T do still determine the price of the $h(\langle X\rangle_T)$ contract.

Theorem 7.6. *Assume the continuous payoff function $h : \mathbb{R} \to \mathbb{R}$ has bilateral Laplace transform H, defined in* (7.4)*, and integrable along* $\mathrm{Re}(z) = \alpha$ *for some $\alpha > A$. Then*

$$\mathbb{E}_t h(\langle X\rangle_T) = \frac{1}{2\pi i}\int_{\alpha-\infty i}^{\alpha+\infty i} H(z)e^{z\langle X\rangle_t}\mathbb{E}_t[\theta_+(S_T/S_t)^{p_+} + \theta_-(S_T/S_t)^{p_-}]dz \tag{7.6}$$

where

$$\theta_\pm := \theta_\pm(z) := \frac{1}{2} \mp \frac{1}{2\sqrt{1+8z}} \qquad p_\pm := p_\pm(z) := \frac{1}{2} \pm \sqrt{1/4+2z}.$$

In particular, the convergence of the integral is guaranteed.

In the case of a variance call, defined by $h(q) = (q-K)^+$, we have $H(z) = e^{-zK}/z^2$ for all $\mathrm{Re}(z) > 0$. For all $\alpha > 0$, therefore, (7.6) exists and gives the variance call price.

Remark 7.7. Despite its generality, Theorem 7.6 has a practical drawback, relative to the earlier results. To price a variance contract exactly using Theorem 7.6 requires the valuation of infinitely *many* different functions of S_T (one for each z). In contrast, using Theorems 6.8, 7.1, 7.2, 7.3, 7.4, to price one variance contract exactly requires the valuation of only one *individual* function of S_T.

If, instead of demanding closed form formulas, we accept [a sequence of] approximate prices which converge to the exact price, then a general class of variance contracts can be priced using [a sequence of] *individual* functions of S_T. That is the subject of the next section.

7.5 General payoffs continuous on $[0,\infty]$

Let $C[0,\infty]$ denote the set of continuous $h : [0,\infty) \to \mathbb{R}$ such that $h(\infty) := \lim_{q\to\infty} h(q)$ exists in $\mathbb{R}$.

For example, a *variance put* payoff $h(q) = (K-q)^+$ is in $C[0,\infty]$. This section gives two ways to determine prices of general payoffs in $C[0,\infty]$. Moreover, by put-call parity (in the sense that a variance call equals a variance put plus a variance swap), a variance call can also be priced by the methods of this section.

In this section let $h \in C[0,\infty]$ and let $c > 0$ be an arbitrary constant.

Theorem 7.8 (Prices as limits of uniform approximations' prices)**.** *Define $h^* : [0,1] \to \mathbb{R}$ by $h^*(0) := h(\infty)$ and $h^*(x) := h(-(1/c)\log x)$ for $x > 0$. For integers $n \geq k \geq 0$, let*

$$b_{n,k} := \sum_{j=0}^{k} h^*(j/n)\binom{n}{k}\binom{k}{j}(-1)^{k-j}. \tag{7.7}$$

Then

$$\mathbb{E}_t h(\langle X\rangle_T) = \lim_{n\to\infty} \mathbb{E}_t \sum_{k=0}^{n} b_{n,k} e^{-ck\langle X\rangle_t}[\theta_+(S_T/S_t)^{p_+} + \theta_-(S_T/S_t)^{p_-}], \tag{7.8}$$

where

$$\theta_\pm := \frac{1}{2} \mp \frac{1}{2\sqrt{1-8ck}} \qquad p_\pm := \frac{1}{2} \pm \sqrt{1/4 - 2ck}. \tag{7.9}$$

In particular, the existence of the limit is guaranteed.

Theorem 7.9 (Prices as limits of L^2 projections' prices)**.** *Let μ be a finite measure on $[0,\infty)$. Let*

$$a_{n,n}e^{-cnq} + a_{n,n-1}e^{-c(n-1)q} + \cdots + a_{n,0} =: A_n(q)$$

be the $L^2(\mu)$ projection of h onto $\operatorname{span}\{1, e^{-cq}, \ldots, e^{-cnq}\}$*. Let P denote the $\mathbb{P}$-distribution of $\langle X\rangle_T$, conditional on $\mathcal{F}_t$. Assume P is absolutely continuous with respect to μ and $d\mathsf{P}/d\mu \in L^2(\mu)$. Then*

$$\mathbb{E}_t h(\langle X\rangle_T) = \lim_{n\to\infty} \mathbb{E}_t \sum_{k=0}^{n} a_{n,k} e^{-ck\langle X\rangle_t}[\theta_+(S_T/S_t)^{p_+} + \theta_-(S_T/S_t)^{p_-}] \tag{7.10}$$

where

$$\theta_\pm := \frac{1}{2} \mp \frac{1}{2\sqrt{1-8ck}} \qquad p_\pm := \frac{1}{2} \pm \sqrt{1/4 - 2ck}. \tag{7.11}$$

In particular, the existence of the limit is guaranteed.

Remark 7.10. For each n, the $a_{n,k}$ $(k = 0, \ldots, n)$ are given by the solution to the linear system

$$\sum_{k=0}^{n} a_{n,k}\langle e^{-cjq}, e^{-ckq}\rangle = \langle h(q), e^{-cjq}\rangle, \qquad j = 0, \ldots, n \tag{7.12}$$

of normal equations, where $\langle \alpha(q), \beta(q)\rangle := \int_0^\infty \alpha(q)\beta(q)\, d\mu(q)$. In practice, one can compute $a_{n,k}$ as the coefficients in a weighted least squares regression of the $h(q)$ function on the regressors $\{q \mapsto e^{-ckq} : k = 0, \ldots, n\}$, with weights given by the measure μ.

For example, consider the variance put payoff $h(\langle X\rangle_T) = (0.04 - \langle X\rangle_T)^+$ with expiry $T = 1$. Under the Heston variance dynamics specified in Figure 6.6 with $\rho = 0$, let us compare the put's true time-0 value $\mathbb{E}h(\langle X\rangle_T)$ against the sequence of European prices in the RHS of (7.10). For example, let $c = 0.5$, and let μ be the lognormal distribution whose parameters are consistent with the prices of T-expiry variance and volatility swaps (which are observable from Europeans, by Theorems 6.8 and 7.3). We compute:

$\mathbb{E}A_3(\langle X\rangle_T)$	$\mathbb{E}A_4(\langle X\rangle_T)$	$\mathbb{E}A_5(\langle X\rangle_T)$	$\cdots$	$\mathbb{E}h(\langle X\rangle_T)$
0.01108	0.01133	0.01147	$\cdots$	0.01149

(7.13)

Here small values of n have sufficed to produce an accurate approximation of $\mathbb{E}h(\langle X\rangle_T)$.

Remark 7.11. In principle, each A_n and B_n function admits perfect pricing by Europeans, via (7.8) and (7.10) respectively; in practice, the convergence benefits of incrementing n must be considered in the context of whether the available European options data (which may have noisy or missing observations) can provide sufficient resolution.

Remark 7.12. Each A_n and B_n function is a linear combination of exponentials, hence admits perfect *replication* by Europeans, according to Theorem 5.10. Consequently, by the explicit uniform approximation (A.11), any variance payoff continuous on $[0, \infty]$ can be replicated to within an arbitrarily small uniform error.

8 Extension to unbounded quadratic variation

Here we show how to drop the assumption (B) that $\langle X \rangle_T \leqslant m$ for some constant m.

It could be argued that this section is mainly of theoretical interest, because in practice a bound of, say, $m = 10^{10}T$ may be an acceptable assumption for an equity index. Theoretically, however, models such as Heston do violate (B).

Theorem 8.1. *Assume the measurable functions h and G satisfy*

$$\mathbb{E}h(\langle X \rangle_T) = \mathbb{E}G(S_T) \tag{8.1}$$

for all S which satisfy (B, W, I)*.*

Assume that h is bounded or *that h is nonnegative and increasing.*

Assume that G has a decomposition $G = G_1 - G_2$, where $G_{1,2}$ are convex and $\mathbb{E}G_{1,2}(S_T) < \infty$.

Then (8.1) *holds, more generally, for all S which satisfy* (W) *and* (I) *and $\mathbb{E}\langle X \rangle_T < \infty$.*

Remark 8.2. The finiteness of $\mathbb{E}h(\langle X \rangle_T)$ is a conclusion, not an assumption.

Remark 8.3. The assumptions on G are very mild, in the following sense: They are satisfied by any payoff function which can be represented as a mixture of calls and puts at all strikes, such that the long and short positions have finite values.

Corollary 8.4. *Theorems 5.1 and 5.8 on exponential variance valuation, and Theorems 6.6 and 6.8 on volatility swap valuation, and Theorems 7.1 7.2, 7.3, on valuation of fractional and integer powers of variance, hold* without *assuming* (B) *– provided that the long and short positions in call and puts in the replicating portfolios have finite values.*

9 Conclusion

Contracts on realized variance allow investors to tailor their exposure to volatility risk, but derivatives dealers have faced difficulties in pricing and hedging such contracts. We find robust solutions by deriving explicit model-free formulas to value realized variance contracts in terms of vanilla option prices – not in terms of the parameters of any model. The formulas are exact under an independence condition, and they are first-order robust to the presence of correlation. We enforce these valuations by replicating the variance payoffs using explicit trading strategies in vanilla options and the underlying shares.

Future research can build on the dynamics we study and the risks we replicate. This paper lays the groundwork for research to add jumps and local volatility to the price dynamics; and this paper contributes to a broad program, including [9], which aims to use European options – which pay functions of S_T alone – to extract information about risks dependent on the entire *path* of S, and to hedge those risks robustly.

A Proofs

Proof of Theorem 4.1. We have

$$\begin{aligned} dX_t &= -\frac{1}{2}\sigma_t^2 dt + \sigma_t\sqrt{1-\rho^2}dW_{1t} + \sigma_t\rho dW_{2t} \\ &= -\frac{1-\rho^2}{2}\sigma_t^2 dt + \sigma_t\sqrt{1-\rho^2}dW_{1t} - \frac{\rho^2}{2}\sigma_t^2 + \sigma_t\rho dW_{2t} \end{aligned}$$

So conditional on $\mathcal{H}_T \vee \mathcal{F}_t$,

$$X_T \sim \text{Normal}\Big(X_t + \log M_{t,T}(\rho) - \bar{\sigma}_{t,T}^2\frac{1-\rho^2}{2},\ \bar{\sigma}_{t,T}\sqrt{1-\rho^2}\Big).$$

Hence the time-t price of the $F(S_T)$ claim is

$$\mathbb{E}_t F(S_T) = \mathbb{E}_t(\mathbb{E}_t(F(S_T)|\mathcal{H}_T)) = \mathbb{E}_t F^{BS}(S_t M_{t,T}(\rho),\ \bar{\sigma}_{t,T}\sqrt{1-\rho^2})$$

as desired. □

Proof of Theorem 5.1. We apply a more general version of Hull-White's [17] conditioning argument. Conditional on $\mathcal{F}_T^\sigma$, the W is still a Brownian motion, by independence. So conditional on $\mathcal{F}_t \vee \mathcal{F}_T^\sigma$,

$$X_T - X_t = \int_t^T \sigma_u dW_u - \frac{1}{2}(\langle X\rangle_T - \langle X\rangle_t) \quad \sim \quad \text{Normal}\Big(-\frac{\langle X\rangle_T - \langle X\rangle_t}{2},\ \langle X\rangle_T - \langle X\rangle_t\Big).$$

For each $p \in \mathbb{C}$, therefore,

$$\begin{aligned} \mathbb{E}_t e^{p(X_T - X_t)} &= \mathbb{E}_t\Big[\mathbb{E}_t(e^{p(X_T-X_t)}|\mathcal{F}_T^\sigma)\Big] \\ &= \mathbb{E}_t\Big[e^{\mathbb{E}_t(pX_T - pX_t|\mathcal{F}_T^\sigma) + \text{Var}_t(pX_T - pX_t|\mathcal{F}_T^\sigma)/2}\Big] \\ &= \mathbb{E}_t\Big[e^{(p^2/2 - p/2)(\langle X\rangle_T - \langle X\rangle_t)}\Big] = \mathbb{E}_t e^{\lambda(\langle X\rangle_T - \langle X\rangle_t)}, \end{aligned}$$

where $\lambda = p^2/2 - p/2$. Equivalently, $p = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2\lambda}$. □

Proof of Theorem 5.2. Our trading strategy at each time t has value $N_t P_t - (pN_t P_{t-}/S_t)S_t + pN_t P_{t-} = N_t P_t$, and in particular it has at time T the desired terminal value $N_T P_T = e^{\lambda\langle X\rangle_T}$. To prove that it self-finances, we have

$$\begin{aligned} d(N_t P_t) &= N_t dP_t + P_{t-}dN_t + d[P,N]_t \\ &= N_t dP_t + P_{t-}\Big(\frac{-pN_t}{S_t}dS_t\Big) + dA_t \end{aligned}$$

where A has finite variation. The continuity of S implies the continuity of N, hence $[P, N]$, hence A. Moreover, A is a local martingale because N_tP_t ($= \mathbb{E}_t e^{\lambda\langle X\rangle_T}$ by Theorem 5.1) and the stochastic integrals with respect to P and S are all local martingales. Therefore dA vanishes.

Moreover, because $dB = 0$, we have

$$d(N_tP_t) = N_t dP_t - (pN_tP_{t-}/S_t)dS_t + pN_tP_{t-}dB_t$$

which proves self-financing. □

Proof of Theorem 5.8. The weights $\theta_\pm$ have the properties that $\theta_+ + \theta_- = 1$ and $\theta_+p_+ + \theta_-p_- = 0$.

The first property, together with Remark 5.7, implies (5.9).

To see that the second property implies ρ-neutrality, let ϕ_v be the lognormal density with parameters $(-v/2, v)$. Then

$$\begin{aligned}\frac{\partial F^{\mathsf{P}_v}}{\partial \mathsf{s}}(S_t) &= e^{\lambda\langle X\rangle_t}\frac{\partial}{\partial \mathsf{s}}\bigg|_{\mathsf{s}=S_t}\int_0^\infty \Big[\theta_+(\mathsf{s}/S_t)^{p_+}y^{p_+} + \theta_-(\mathsf{s}/S_t)^{p_-}y^{p_-}\Big]\phi_v(y)dy \\ &= e^{\lambda\langle X\rangle_t}\int_0^\infty\left(\theta_+\frac{p_+}{S_t}y^{p_+} + \theta_-\frac{p_-}{S_t}y^{p_-}\right)\phi_v(y)dy = e^{\lambda\langle X\rangle_t}\frac{\theta_+p_+ + \theta_-p_-}{S_t}\int_0^\infty y^{p_+}\phi_v(y)dy = 0\end{aligned}$$

using the equality of integrals of $y^{p_+}\phi_v(y)$ and $y^{p_-}\phi_v(y)$. □

Proof of Theorem 5.10. The strategy is a linear combination of the two strategies $(+, -)$ specified in Theorem 5.2, with constant weights θ_+ and θ_- which sum to 1. Each strategy self-finances and replicates $e^{\lambda\langle X\rangle_T}$, so the combination does also. Moreover, $\delta_t = 0$ because $N_t^+P_t^+ = N_t^-P_t^-$ and $\theta_+p_+ + \theta_-p_- = 0$.

Regardless of (I), we have Δ-neutrality because the time-t holdings have combined payoff function (in the sense of (4.2))

$$F(\mathsf{S}) := G_{\exp}(\mathsf{S}, S_t, \langle X\rangle_t; \lambda) - \delta_t(\mathsf{S} - S_t),$$

which satisfies

$$\frac{\partial F^{\mathsf{P}_{\text{true}}}}{\partial \mathsf{s}}(S_t) = e^{\lambda\langle X\rangle_t}\mathbb{E}_t\left(\frac{\partial}{\partial \mathsf{s}}\Big[\theta_+(\mathsf{s}S_T/S_t^2)^{p_+} + \theta_-(\mathsf{s}S_T/S_t^2)^{p_-}\Big]\right)\bigg|_{\mathsf{s}=S_t} - \delta_t = 0.$$

Finally, Theorem 5.8 implies the ρ-neutrality condition. □

Proof of Theorem 6.1. The upper bound (3) is known (Britten-Jones/Neuberger [8]) to hold, by Jensen's inequality.

For (2), we have by Theorem 4.1 and the concavity of F^{BS}_{atmc},

$$F^{BS}_{\text{atmc}}(S_0, \text{IV}_0) = \mathbb{E}_0 F_{\text{atmc}}(S_T) = \mathbb{E}_0 F^{BS}_{\text{atmc}}(S_0, \bar\sigma_{0,T}) \le F^{BS}_{\text{atmc}}(S_0, \mathbb{E}_0\bar\sigma_{0,T}). \tag{A.1}$$

By the monotonicity of F^{BS}_{atmc}, therefore, $\text{IV}_0 \le \mathbb{E}_0\bar\sigma_{0,T}$. For (1),

$$\frac{\sqrt{2\pi}}{S_0}\mathbb{E}_0(S_T - S_0)^+ = \frac{\sqrt{2\pi}}{S_0}F^{BS}_{\text{atmc}}(S_0, \text{IV}_0) \le \frac{\sqrt{2\pi}}{S_0}\frac{S_0\,\text{IV}_0}{\sqrt{2\pi}} = \text{IV}_0 \tag{A.2}$$

because concavity implies that $F^{BS}_{\text{atmc}}(S_0, \cdot)$ lies everywhere below its tangent at 0. □

Proof of Theorem 6.6. Of the $\pm$, we prove the $+$ equation; the $-$ equation is similar.

The square root function has the integral representation

$$\sqrt{q} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{1-e^{-zq}}{z^{3/2}} dz \qquad q \geq 0;$$

as shown in sources such as Schürger [23]. So

$$\begin{aligned}
\mathbb{E}_0\sqrt{\langle X\rangle_T} &= \frac{1}{2\sqrt{\pi}}\mathbb{E}_0 \int_0^\infty \frac{1-e^{-z\langle X\rangle_T}}{z^{3/2}} dz \\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty \mathbb{E}_0 \frac{1-e^{-z\langle X\rangle_T}}{z^{3/2}} dz \\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty \mathbb{E}_0 \frac{1-e^{(1/2-\sqrt{1/4-2z})X_T}}{z^{3/2}} dz \\
&= \frac{1}{2\sqrt{\pi}}\mathbb{E}_0 \int_0^\infty \frac{1-e^{(1/2-\sqrt{1/4-2z})X_T}}{z^{3/2}} dz.
\end{aligned}$$

and take real parts. The first application of Fubini is justified by $|1-e^{-z\langle X\rangle_T}| < 1-e^{-zm}$. The second application of Fubini is justified by $\mathbb{E}_0|1-e^{(1/2-\sqrt{1/4-2z})X_T}| = O(1)$ as $z\to\infty$; and on the other hand for z sufficiently small,

$$\begin{aligned}
(\mathbb{E}_0|1-e^{(1/2-\sqrt{1/4-2z})X_T}|)^2 &\leq \mathbb{E}_0(|1-e^{(1/2-\sqrt{1/4-2z})X_T}|^2) \\
&= \mathbb{E}_0(1-2e^{(1/2-\sqrt{1/4-2z})X_T} + e^{(1/2-\sqrt{1/4-2z})2X_T}) \\
&= 1 - 2\mathbb{E}_0 e^{-z\langle X\rangle_T} + \mathbb{E}_0 e^{(\frac{1-8z-\sqrt{1-8z}}{2})\langle X\rangle_T} \\
&= 1-2(1-zf'(0)+O(z^2)) + 1 - 2zf'(0) + O(z^2) \\
&= O(z^2) \qquad \text{as } z\to 0
\end{aligned}$$

using the analyticity of the moment generating function $f(\xi) := e^{\xi\langle X\rangle_T}$, which follows from (B). ☐

Proof of Theorem 6.8. For arbitrary $\mathcal{F}_t$-measurable $q\geq 0$ we have

$$\begin{aligned}
\mathbb{E}_t\sqrt{\langle X\rangle_T - \langle X\rangle_t + q} &= \frac{1}{2\sqrt{\pi}}\mathbb{E}_t \int_0^\infty \frac{1-e^{-z(\langle X\rangle_T-\langle X\rangle_t+q)}}{z^{3/2}} dz && \text{(A.3)}\\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty (\theta_+ + \theta_-) \frac{1-\mathbb{E}_t e^{-z(\langle X\rangle_T-\langle X\rangle_t+q)}}{z^{3/2}} dz && \text{(A.4)}\\
&= \frac{1}{2\sqrt{\pi}} \int_0^\infty \sum_\pm \theta_\pm \frac{1-e^{-zq}\mathbb{E}_t e^{p_\pm(X_T-X_t)}}{z^{3/2}} dz && \text{(A.5)}\\
&= \frac{1}{2\sqrt{\pi}}\mathbb{E}_t \int_0^\infty \sum_\pm \theta_\pm \frac{1-e^{-zq}e^{p_\pm(X_T-X_t)}}{z^{3/2}} dz && \text{(A.6)}
\end{aligned}$$

Taking $q := \langle X\rangle_t$ yields the result.

The first application of Fubini (A.4) is justified by $|1-e^{-z(\langle X\rangle_T+q)}| < 1-e^{-z(m+q)}$. The second application of Fubini (A.6) is justified by

$$A_\pm := (\mathbb{E}_t|1-e^{-qz+(1/2\pm\sqrt{1/4-2z})(X_T-X_t)}|)^2 \leq \mathbb{E}_t(|(1-e^{-qz+(1/2\pm\sqrt{1/4-2z})(X_T-X_t)}|^2) \qquad \text{(A.7)}$$

which is $O(1)$ as $z \to \infty$, hence

$$\mathbb{E}_t \frac{|\theta_\pm(1 - e^{-qz+(1/2\pm\sqrt{1/4-2z})(X_T - X_t)})|}{z^{3/2}} = O(z^{-3/2}) \qquad z \to \infty.$$

On the other hand, for z sufficiently small, the term in the absolute values in (A.7) is real, so

$$\begin{aligned} A_\pm &\le \mathbb{E}_t\big[1 - 2e^{-qz+(1/2\pm\sqrt{1/4-2z})(X_T-X_t)} + e^{-2qz+2(1/2\pm\sqrt{1/4-2z})(X_T-X_t)}\big] \\ &= 1 - 2e^{-qz}\mathbb{E}_t e^{-z(\langle X\rangle_T - \langle X\rangle_t)} + e^{-2qz}\mathbb{E}_t e^{(\frac{1-8z\pm\sqrt{1-8z}}{2})(\langle X\rangle_T - \langle X\rangle_t)} \end{aligned}$$

hence as $z \to 0$,

$$A_+ = O(1) \tag{A.8}$$

$$A_- = 1 - 2(1 - zf'(0) - qz + O(z^2)) + 1 - 2zf'(0) - 2qz + O(z^2) = O(z^2) \tag{A.9}$$

using analyticity of moment generating function $f(\xi) := e^{\xi\langle X\rangle_T}$, which follows from (B). Combining this with $\theta_- = O(1)$ and $\theta_+ = O(z)$ as $z \to 0$, we have

$$\mathbb{E}_t \frac{|\theta_\pm(1 - e^{-qz+(1/2\pm\sqrt{1/4-2z})(X_T - X_t)})|}{z^{3/2}} = \frac{|\theta_\pm| A_\pm^{1/2}}{z^{3/2}} = O(z^{-1/2}) \qquad z \to 0,$$

which allows the interchange in (A.6).

To establish ρ-neutrality, let ϕ_v be the lognormal density with parameters $(-v/2, v)$. Then

$$\begin{aligned} \frac{\partial F^{\mathsf{P}_v}}{\partial \mathsf{s}}(S_t) &= \frac{1}{2\sqrt{\pi}} \frac{\partial}{\partial \mathsf{s}}\bigg|_{\mathsf{s}=S_t} \int_0^\infty \int_0^\infty \left[\theta_+ \frac{1 - e^{-z\langle X\rangle_t}(\mathsf{s}y/S_t)^{p_+}}{z^{3/2}} + \theta_- \frac{1 - e^{-z\langle X\rangle_t}(\mathsf{s}y/S_t)^{p_-}}{z^{3/2}} dz\right] \phi_v(y) dy \\ &= \frac{1}{2\sqrt{\pi}} \int_0^\infty \int_0^\infty \frac{-e^{-z\langle X\rangle_t}(\theta_+ p_+ y^{p_+} + \theta_- p_- y^{p_-})}{S_t z^{3/2}} \phi_v(y) dz dy = 0 \end{aligned}$$

using the equality of integrals of $y^{p_+}\phi_v(y)$ and $y^{p_-}\phi_v(y)$, and the identity $\theta_+ p_+ + \theta_- p_- = 0$. □

Proof of Theorem 6.10. By a Mathematica computation,

$$\frac{1}{2\sqrt{\pi}} \int_0^\infty \theta_+ \frac{1 - e^{p_+ X_T}}{z^{3/2}} + \theta_- \frac{1 - e^{p_- X_T}}{z^{3/2}} dz = \sqrt{\frac{\pi}{2}} e^{X_T/2} |X_T I_0(X_T/2) - X_T I_1(X_T/2)|.$$

The result now follows from Theorem 6.8. □

Proof of Theorem 6.11. From (6.11), compute $\psi''(K)$, and apply Remark 3.2. □

Proof of Theorem 6.12. From (6.8), compute $\partial^2 G_{\text{svs}}/\partial \mathsf{S}^2(K, S_t, \langle X\rangle_t)$, and apply Remark 3.2. □

Proof of Theorem 6.14. For background in measure-valued trading strategies, see [5]. The trading strategy at each time t has value

$$\begin{aligned} V_t &= \int P_t(K)\varphi_t(dK) + G_{\text{svs}}(\kappa, S_t, \langle X\rangle_t) \\ &= \int P_t(K)(-1)^{\mathbb{I}_{K<\kappa}} \frac{\partial G_{\text{svs}}}{\partial \mathsf{S}}(K; S_t, \langle X\rangle_t) dK + G_{\text{svs}}(\kappa, S_t, \langle X\rangle_t) \\ &= \mathbb{E}_t G_{\text{svs}}(S_T, S_t, \langle X\rangle_t) = \mathbb{E}_t \sqrt{\langle X\rangle_T} \end{aligned}$$

by Theorem 6.8. In particular it has at time $t = T$ the desired terminal value.

To prove that it self-finances, we have

$$\begin{aligned}
dV_t &= \varphi_t dP_t + \left[\int_0^\infty P_t(K)(-1)^{\mathbb{I}_{K<\kappa}} \frac{\partial^2 G_{\text{svs}}}{\partial S \partial u}(K, S_t, \langle X\rangle_t) dK\right] dS_t + d\tilde{A}_t + dG_{\text{svs}}(\kappa, S_t, \langle X\rangle_t) \\
&= \varphi_t dP_t + \left[\int_0^\infty P_t(K)(-1)^{\mathbb{I}_{K<\kappa}} \frac{\partial^2 G_{\text{svs}}}{\partial S \partial u}(K, S_t, \langle X\rangle_t) dK + \frac{\partial G_{\text{svs}}}{\partial u}(\kappa, S_t, \langle X\rangle_t)\right] dS_t + dA_t \\
&= \varphi_t dP_t + \left[\mathbb{E}_t \frac{\partial G_{\text{svs}}}{\partial u}(S_T, S_t, \langle X\rangle_t)\right] dS_t + dA_t
\end{aligned}$$

where $\tilde{A}$ and A denote time-continuous finite-variation processes. Moreover, A is a local martingale because $\varphi_t P_t$ and the integrals with respect to P and S are local martingales. Therefore dA vanishes. Because $dB = 0$, we have

$$dL_t = \varphi_t dP_t - \delta_t dS_t + [G_{\text{svs}}(\kappa, S_t, \langle X\rangle_t) + \delta_t S_t] dB_t,$$

which is the self-financing condition.

That $\delta_t = 0$ follows from $\mathbb{E}_t(S_T/S_t)^{p+} = \mathbb{E}_t(S_T/S_t)^{p-}$ and $\theta_+ p_+ + \theta_- p_- = 0$.

Regardless of (I), we have Δ-neutrality because the time-t holdings have combined payoff function (in the sense of (4.2))

$$F(\mathsf{S}) := G_{\text{svs}}(\mathsf{S}, S_t, \langle X\rangle_t; \lambda) - \delta_t(\mathsf{S} - S_t),$$

which satisfies

$$\frac{\partial F^{\mathsf{P}_{\text{true}}}}{\partial \mathsf{s}}(S_t) = \mathbb{E}_t \frac{\partial}{\partial \mathsf{s}}\bigg|_{\mathsf{s}=S_t} G_{\text{svs}}(\mathsf{s} S_T/S_t, S_t, \langle X\rangle_t) - \delta_t = 0.$$

Moreover, the second representation of δ in (6.15) follows from

$$\frac{\partial}{\partial \mathsf{s}}\bigg|_{\mathsf{s}=S_t} G_{\text{svs}}(\mathsf{s} S_T/S_t, S_t, \langle X\rangle_t) = \frac{S_T}{S_t}\frac{\partial G_{\text{svs}}}{\partial \mathsf{S}}(S_T, S_t, \langle X\rangle_t) = -\frac{\partial G_{\text{svs}}}{\partial \mathsf{u}}(S_T, S_t, \langle X\rangle_t).$$

The ρ-neutrality is proved in Theorem 6.8. □

Proof of Theorem 7.1. Using the identity [23]

$$q^r = \frac{r}{\Gamma(1-r)} \int_0^\infty \frac{1 - e^{-zq}}{z^{r+1}} dz \qquad\qquad 0 < r < 1,\ q \geq 0$$

follow the proof of Theorem 6.8. □

Proof of Theorem 7.2. Using the identity [23]

$$q^{-r} = \frac{1}{r\Gamma(r)} \int_0^\infty e^{-z^{1/r} q} dz, \qquad\qquad r > 0,\ q > 0$$

we have

$$\begin{aligned}\mathbb{E}_t(\langle X\rangle_T+\varepsilon)^{-r} &= \frac{1}{r\Gamma(r)}\mathbb{E}_t\int_0^\infty e^{-z^{1/r}(\langle X\rangle_T+\varepsilon)}dz \\ &= \frac{1}{r\Gamma(r)}\int_0^\infty e^{-z^{1/r}(\langle X\rangle_t+\varepsilon)}\mathbb{E}_t e^{-z^{1/r}(\langle X\rangle_T-\langle X\rangle_t)}dz \\ &= \frac{1}{r\Gamma(r)}\mathbb{E}_t\int_0^\infty (\theta_+ e^{p_+(X_T-X_t)}+\theta_- e^{p_-(X_T-X_t)})e^{-z^{1/r}(\langle X\rangle_t+\varepsilon)}dz\end{aligned}$$

where the two applications of Fubini are justified by (B)and $|e^{(1/2\pm\sqrt{1/4-2z^{1/r}})(X_T-X_t)}|\leq 1$ respectively.

To establish ρ-neutrality, let ϕ_v be the lognormal density with parameters $(-v/2, v)$. Then

$$\begin{aligned}\frac{\partial F^{\mathsf{P}_v}}{\partial \mathsf{s}}(S_t) &= \frac{1}{r\Gamma(r)}\frac{\partial}{\partial \mathsf{s}}\Big|_{\mathsf{s}=S_t}\int_0^\infty\int_0^\infty \Big[\theta_+(\mathsf{s}y/S_t)^{p_+}+\theta_-(\mathsf{s}y/S_t)^{p_-}dz\Big]e^{-z^{1/r}\langle X\rangle_t}\phi_v(y)dy \\ &= \frac{1}{r\Gamma(r)S_t}\int_0^\infty\int_0^\infty e^{-z^{1/r}\langle X\rangle_t}(\theta_+p_+y^{p_+}+\theta_-p_-y^{p_-})\phi_v(y)dzdy = 0\end{aligned}$$

using the equality of integrals of $y^{p_+}\phi_v(y)$ and $y^{p_-}\phi_v(y)$, and the identity $\theta_+p_++\theta_-p_-=0$. ☐

Proof of Theorem 7.3. Take the nth derivative of (5.9) with respect to λ, and evaluate at $\lambda=0$:

$$\mathbb{E}_t\partial_\lambda^n e^{\lambda\langle X\rangle_T}\big|_{\lambda=0} = \mathbb{E}_t\partial_\lambda^n G_{\exp}(S_T, S_t, \langle X\rangle_t;\lambda)\big|_{\lambda=0} \tag{A.10}$$

Differentiation through the expectations is justified by the boundedness of $\langle X\rangle_T$ and the analyticity of the moment generating function of X_T.

To establish ρ-neutrality, let ϕ_v be the lognormal density with parameters $(-v/2, v)$. Then

$$\begin{aligned}\frac{\partial F^{\mathsf{P}_v}}{\partial \mathsf{s}}(S_t) &= \frac{\partial}{\partial \mathsf{s}}\Big|_{\mathsf{s}=S_t} G_{\mathrm{pow}(n)}(\mathsf{s}y/S_t, S_t, \langle X\rangle_t)\phi_v(y)dy \\ &= \frac{\partial^n}{\partial\lambda^n}\Big|_{\lambda=0}\frac{\partial}{\partial \mathsf{s}}\Big|_{\mathsf{s}=S_t} G_{\exp}(\mathsf{s}y/S_t, S_t, \langle X\rangle_t, \lambda)\phi_v(y)dy = 0\end{aligned}$$

by the ρ-neutrality of $G_{\exp}$. ☐

Proof of Theorem 7.4. Inverting the Laplace transform,

$$h(q) = \frac{1}{2\pi i}\int_{\alpha-\infty i}^{\alpha+\infty i} H(z)e^{zq}dz$$

Therefore

$$\begin{aligned}\mathbb{E}_t h(\langle X\rangle_T) &= \frac{1}{2\pi i}\mathbb{E}_t\int_{\alpha-\infty i}^{\alpha+\infty i} H(z)e^{z\langle X\rangle_T}dz = \frac{1}{2\pi i}\int_{\alpha-\infty i}^{\alpha+\infty i} H(z)\mathbb{E}_t e^{z\langle X\rangle_T}dz \\ &= \frac{1}{2\pi i}\int_{\alpha-\infty i}^{\alpha+\infty i} H(z)e^{z\langle X\rangle_t}\mathbb{E}_t[\theta_+e^{p_+(X_T-X_t)}+\theta_-e^{p_-(X_T-X_t)}]dz \\ &= \mathbb{E}_t G_h(S_T, S_t, \langle X\rangle_t)\end{aligned}$$

where the two applications of Fubini (and, in particular, the convergence of the integral in (7.5)) are justified respectively by assumption (B) and by

$$\begin{aligned}\mathbb{E}_t|e^{p_\pm(X_T-X_t)}| = \mathbb{E}_t e^{\mathrm{Re}(1/2\pm\sqrt{1/4+2(\alpha+\beta i)})(X_T-X_t)} &= \mathbb{E}_t e^{(1/2\pm\sqrt{|\beta|}+O(|\beta|^{-1/2}))(X_T-X_t)} \\ &= \mathbb{E}_t e^{(|\beta|/2+O(1))(\langle X\rangle_T-\langle X\rangle_t)} = O(e^{|\beta|m/2})\end{aligned}$$

and $|H(z)e^{z\langle X\rangle_t}\theta_\pm(z)| = O(e^{-|\beta|\mu})$ as $|\beta|\to\infty$.

Proof of ρ-neutrality is by calculation similar to the proof of Theorem 6.8. □

Proof of Theorem 7.6. Follow the proof of Theorem 7.4, stopping before the second application of Fubini. □

Proof of Theorem 7.8. The nth Bernstein approximation for h^* is defined by

$$B_n(x) := b_{n,n}x^n + b_{n,n-1}x^{n-1} + \cdots + b_{n,0}$$

and satisfies $h^*(x) = \lim_{n\to\infty} B_n(x)$ uniformly in $x\in[0,1]$. Therefore

$$h(q) = \lim_{n\to\infty} B_n(e^{-cq}) \tag{A.11}$$

uniformly in $q\in[0,\infty)$. Hence

$$\mathbb{E}_t h(\langle X\rangle_T) = \lim_{n\to\infty}\mathbb{E}_t B_n(e^{-c\langle X\rangle_T}) = \lim_{n\to\infty}\mathbb{E}_t\sum_{k=0}^n b_{n,k}e^{-ck\langle X\rangle_t}[\theta_+ e^{p_+(X_T-X_t)} + \theta_- e^{p_-(X_T-X_t)}]$$

as claimed. □

Proof of Theorem 7.9. The span of the polynomials $\{1, x, x^2, \ldots\}$ is dense in $C[0,1]$ with respect to the uniform norm. By the transformation $q = -(1/c)\log x$, the span of exponential functions $\{1, e^{-cq}, e^{-2cq}, \ldots\}$ is dense in $C[0,\infty]$ with respect to the uniform norm, hence dense in $C[0,\infty]$ with respect to the $L^2(\mu)$ norm. Then

$$h = \lim_{n\to\infty} A_n \tag{A.12}$$

in the $L^2(\mu)$ sense, hence

$$\begin{aligned}\Big(\mathbb{E}\big[h(\langle X\rangle_T) - A_n(\langle X\rangle_T)\big]\Big)^2 &= \Big(\int \frac{dP}{d\mu}\big[h(\langle X\rangle_T) - A_n(\langle X\rangle_T)\big]d\mu\Big)^2 \\ &\le \int\Big(\frac{dP}{d\mu}\Big)^2 d\mu \int \big[h(\langle X\rangle_T) - A_n(\langle X\rangle_T)\big]^2 d\mu \longrightarrow 0\end{aligned}$$

as $n\to\infty$. Thus

$$\mathbb{E}_t h(\langle X\rangle_T) = \lim_{n\to\infty}\mathbb{E}_t A_n(\langle X\rangle_T) = \lim_{n\to\infty}\mathbb{E}_t\sum_{k=0}^n a_{n,k}e^{-ck\langle X\rangle_t}[\theta_+ e^{p_+(X_T-X_t)} + \theta_- e^{p_-(X_T-X_t)}] \tag{A.13}$$

as desired. □

Proof of Theorem 8.1. For each positive integer m, define the process $\sigma_t^m := \sigma_t \mathbb{I}(\langle X \rangle_t \le m)$.

Define the process S_t^m by $dS_t^m = \sigma_t^m S_t^m dW_t$. Let $X_t^m := \log(S_t^m)$.

Then S^m satisfies (B), so

$$\mathbb{E}h(\langle X^m \rangle_T) = \mathbb{E}G(S_T^m). \tag{A.14}$$

Now let $m \to \infty$. The LHS approaches $\mathbb{E}h(\langle X \rangle_T)$ because $\langle X^m \rangle_T \to \langle X \rangle_T$ almost surely, and either monotone convergence or dominated convergence applies.

It remains to show that the RHS of (A.14) approaches $\mathbb{E}G(S_T)$. Note that there exist constants α and β and convex nonnegative functions G_+ and G_- such that $G_\pm(S_0) = 0$ and $\mathbb{E}G_\pm(S_T) < \infty$ and

$$G(S) = G_+(S) - G_-(S) + \alpha S + \beta \qquad \text{for all } S \ge 0.$$

We need only show that $\mathbb{E}G_+(S_T^m) \to \mathbb{E}G_+(S_T)$; convergence proofs for the other terms are then trivial. It suffices to show that the family

$$\{G_+(S_T^m) : \ m \ge 1\}$$

is uniformly integrable. Since $\mathbb{E}G_+(S_T) < \infty$, it is enough to show that for all m and all $A > 0$,

$$\mathbb{E}G_+(S_T^m)\mathbb{I}(G_+(S_T^m) > A) \le \mathbb{E}G_+(S_T)\mathbb{I}(G_+(S_T) > A).$$

By the convexity of G_+, there exist $a, b \in [0, \infty]$ such that

$$\mathbb{I}(G_+(S) > A) = \mathbb{I}(S < S_0 - a) + \mathbb{I}(S > S_0 + b) \qquad \text{for all } S > 0$$

Moreover, the function

$$U(S) := G_+(S)\mathbb{I}(G_+(S) > A) - \frac{A}{b}(S - S_0)\mathbb{I}(S > S_0 + b) - \frac{A}{a}(S_0 - S)\mathbb{I}(S < S_0 - a)$$

is convex. We have

$$\begin{aligned}
&\mathbb{E}G_+(S_T^m)\mathbb{I}(G_+(S_T^m) > A) \\
&= \mathbb{E}\left[\frac{A}{a}(S_0 - S_T^m)\mathbb{I}(S_T^m < S_0 - a) + \frac{A}{b}(S_T^m - S_0)\mathbb{I}(S_T^m > S_0 + b) + U(S_T^m)\right] \\
&= \mathbb{E}\left[\frac{A}{a}\mathbb{E}[(S_0 - S_T^m)\mathbb{I}(S_T^m < S_0 - a)|\langle X \rangle_T] + \frac{A}{b}\mathbb{E}[(S_T^m - S_0)\mathbb{I}(S_T^m > S_0 + b)|\langle X \rangle_T] + \mathbb{E}[U(S_T^m)|\langle X \rangle_T]\right] \\
&\le \mathbb{E}\left[\frac{A}{a}\mathbb{E}[(S_0 - S_T)\mathbb{I}(S_T < S_0 - a)|\langle X \rangle_T] + \frac{A}{b}\mathbb{E}[(S_T - S_0)\mathbb{I}(S_T > S_0 + b)|\langle X \rangle_T] + \mathbb{E}[U(S_T)|\langle X \rangle_T]\right] \\
&= \mathbb{E}G_+(S_T)\mathbb{I}(G_+(S_T) > A)
\end{aligned}$$

where the inequality holds because if Z has a mean-S_0 lognormal distribution with $\operatorname{Var}\log(Z) = \sigma^2$, then each of the quantities $\mathbb{E}[(S_0 - Z)\mathbb{I}(Z < S_0 - a)]$, $\mathbb{E}[(Z - S_0)\mathbb{I}(Z > S_0 + b)]$, and $\mathbb{E}U(Z)$ is an increasing function of σ. For the first two expectations, this comes from direct calculation; for the third expectation, it follows from convexity and Jensen's inequality. $\square$

References

[1] Vexed by variance. *Risk*, August 2006.

[2] Torben G. Andersen, Per Frederiksen, and Arne D. Staal. The information content of realized volatility forecasts. Northwestern University, Nordea Bank, and Lehman Brothers, 2007.

[3] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52:2003–2049, 1997.

[4] David Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94:181–238, 2000.

[5] Tomas Björk, Giovanni Di Masi, Yuri Kabanov, and Wolfgang Runggaldier. Towards a general theory of bond markets. *Finance And Stochastics*, 1:141–174, 1997.

[6] David Breeden and Robert Litzenberger. Prices of state contingent claims implicit in options prices. *Journal of Business*, 51:621–651, 1978.

[7] Menachem Brenner and Marti Subrahmanyam. A simple formula to compute the implied standard deviation. *Financial Analysts Journal*, 44(5):80–83, 1988.

[8] Mark Britten-Jones and Anthony Neuberger. Option prices, implied price processes, and stochastic volatility. *Journal of Finance*, 55(2):839–866, 2000.

[9] Peter Carr and Roger Lee. Put-call symmetry: Extensions and applications. *Mathematical Finance*, 2007. Forthcoming.

[10] Peter Carr and Roger Lee. Realized volatility and variance: Options via swaps. *Risk*, 20(5):76–83, 2007.

[11] Peter Carr and Dilip Madan. Towards a theory of volatility trading. In R. Jarrow, editor, *Volatility*, pages 417–427. Risk Publications, 1998.

[12] Emanuel Derman, Kresimir Demeterfi, Michael Kamal, and Joseph Zou. A guide to volatility and variance swaps. *Journal of Derivatives*, 6(4):9–32, 1999.

[13] Bruno Dupire. Arbitrage pricing with stochastic volatility. Société Générale, 1992.

[14] Steven P. Feinstein. The Black-Scholes formula is nearly linear in sigma for at-the-money options: Therefore implied volatilities from at-the-money options are virtually unbiased. Federal Reserve Bank of Atlanta, 1989.

[15] Peter Friz and Jim Gatheral. Valuation of volatility derivatives as an inverse problem. *Quantitative Finance*, 5(6):531–542, 2005.

[16] Anuj Gangahar. Volatility becomes an asset class. *Financial Times*, May 23, 2006.

[17] John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, June 1987.

[18] George J. Jiang and Yisong S. Tian. The model-free implied volatility and its information content. *Review of Financial Studies*, 18:1305–1342.

[19] Anthony Neuberger. The log contract. *Journal of Portfolio Management*, 20(2):74–80, 1994.

[20] Navroz Patel. *RiskNews*, June 12, 2003.

[21] Allen Poteshman. Forecasting future volatility from option prices. University of Illinois at Urbana-Champaign, 2000.

[22] Eric Renault and Nizar Touzi. Option hedging and implied volatilities in a stochastic volatility model. *Mathematical Finance*, 6(3):279–302, 1996.

[23] Klaus Schürger. Laplace transforms and suprema of stochastic processes. University of Bonn, 2002.

[24] Gregory A. Willard. Calculating prices and sensitivities for path-independent derivative securities in multifactor models. *Journal of Derivatives*, pages 45–61, Fall 1997.