Volatility Risk for Regime Switching Models[1]

Adam W. Kolkiewicz
Ken Seng Tan

Department of Statistics and Actuarial Science
University of Waterloo
Waterloo, Ontario N2L 3G1
Canada

Abstract

Regime switching models have proven to be well-suited for capturing the time series behavior of many financial variables. In particular, they have become a popular framework for pricing equity-linked insurance products. The success of these models demonstrates that realistic modeling of financial time series must allow for random changes in volatility. In the context of valuation of contingent claims, however, random volatility poses additional challenges when compared with the standard Black-Scholes framework. The main reason is the incompleteness of such models, which implies that contingent claims cannot be hedged perfectly and that a unique identification of the correct risk-neutral measure is not possible. The objective of the paper is to provide tools for managing the volatility risk. First we present a formula for the expected value of a shortfall caused by misspecification of the realized cumulative variance. This, in particular, leads to a closed-form expression for the expected shortfall for any strategy a hedger may use to deal with the stochastic volatility. Next we identify a method of selection of the initial volatility that minimizes the expected shortfall. This strategy is the same as delta hedging based on the cumulative volatility that matches the Black-Scholes model with the stochastic volatility model. We also discuss methods of managing the volatility risk under model uncertainty. In these cases, super-hedging is a possible strategy but it is expensive. The presented results enable a more accurate analysis of the trade-off between the initial cost and the risk of a shortfall.

1 Introduction and Motivation

In this paper we are interested in the application of regime switching models to the problem of pricing and hedging contingent financial instruments. These models have been proposed, for example, for pricing equity-linked life insurance products, which from the financial economic viewpoint can be interpreted as insurance polices with embedded stock market options (Hardy 2001, 2003). The main underlying assumption for these models is that the process

[1]Forthcoming in the *North American Actuarial Journal*, 2004, volume 8(4).

that determines the current regime follows a finite-state Markov chain. As demonstrated in several studies, this specification is flexible enough to offer models capable of explaining many of the observed features of financial time series, particularly the volatility persistence. Kim and Nelson (1999) give a comprehensive exposition of statistical methods for regime switching models as well as many examples of case studies.

Continuous-time versions of such models form a sub-class of stochastic volatility models, in which instantaneous volatility of the risky asset is allowed to follow its own stochastic process. These models address one of the known imperfections of the Black-Scholes model, which manifests itself through the presence of "smiles" and "skews" in implied volatilities. Reviews of such models can be found, for example, in the work of Taylor (1994) and Shephard (1996). The price we have to pay for this more realistic modelling is that in the context of random volatility contingent claims can no longer be perfectly hedged with the underlying asset and a bond. Although a perfect hedge may be possible by including into the replicating portfolio another derivative security, such procedures are often unsatisfactory due to the higher transactions costs and lesser liquidity associated with trading the second derivative. Therefore, it is important to quantify the volatility risk and determine how to hedge the contingent claim as best as possible using only the underlying asset and a bond.

These problems become even more significant for long-term contingent instruments, such as equity-linked life insurance products, for which mis-pricing and/or lack of efficient hedging strategies may lead to huge losses. It is important to notice that for such contracts the risk of a large shortfall at maturity can be diversified by partitioning the time to maturity into shorter terms and adjusting the writer's position to the market value at the end of each period. Then the distribution of the total shortfall should be close to the normal distribution. To explain this, let us denote by S_t the price of the underlying asset and by $C(S_t, t)$ the market value of the contingent claim at time t. Suppose also that the contract is hedged dynamically with the asset and a bond but, due to the incompleteness of the market, a perfect hedge is not possible. As we explain later in the paper, the value $C(S_t, t)$ can be interpreted as the expected cost of replicating the claim. Therefore, if at the beginning of each sub-period $(t_i, t_{i+1}]$, $i = 0, .., N-1$, a replicating portfolio is created at the cost $C(S_{t_i}, t_i)$, at the end of the same period the expected value of a surplus/shortfall, $R(S_{t_{i+1}}, t_{i+1})$, will be zero. Then the surplus/shortfall for the first l periods of the contract can be represented by the following sum

$$\sum_{i=1}^{l} R(S_{t_i}, t_i). \tag{1}$$

Although the random variables $R(S_{t_i}, t_i)$ are not independent, they have the property that for each i the expectation of $R(S_{t_{i+1}}, t_{i+1})$ conditional on information at time t_i is zero. Hence, the partial sums (1), $l = 1, 2, \dots$, form a martingale, and from the martingale central limit theorem (subject to technical conditions) their distributions converge to a normal with zero mean (e.g., Chow and Teicher 1997). Hence, the variance of the limiting distribution determines the overall risk and can be mitigated only by reducing the size of each surplus/shortfall. This justifies the importance of finding accurate predictions of these values.

In the paper we describe a dynamic hedging strategy that is based on the realized volatility. This method has been introduced by Mykland (2000) in the context of super-replicating strategies, but we propose to use it also for other risk criteria. This is possible because under some assumptions about the terminal payoff function the distribution of the expected surplus/shortfall at maturity can be characterized in a form that is convenient for applications. Consequently, expressions for the probability of a shortfall and the expected shortfall are also available. The latter can be used as a natural measure of the volatility risk. The availability of these expressions enabled us to show that delta hedging based on a particular selection of the initial cumulative variance minimizes the variance of the surplus/shortfall at maturity as well as the expected shortfall. This optimal initial cumulative variance is the one that matches the Black-Scholes price of a derivative with the corresponding price under stochastic volatility.

The main advantage of the method is that it is applicable under quite general conditions on the dynamics of the volatility process. Furthermore, the only feature of a volatility model that is relevant for pricing and hedging is the distribution of the cumulative variance. This quantity is certainly less prone to estimation error than the whole volatility path. In practice, this distribution can be estimated from historical data and/or from market information. We discuss this issue in greater detail in Section 3.

Although the importance of the cumulative variance on the price of a derivative is well known in the literature, this fact is usually discussed under the assumption of independence between the asset and the volatility processes (e.g., Hull and White 1987; Rebonato 1999). In addition, most of the existing results focus on the problem of pricing without addressing the issue of a shortfall risk. One of the exceptions is the paper by El Karoui et al. (1998), but the authors discuss hedging strategies assuming a particular form of the misspecified volatility, which excludes regime switching models.

In the context of a regime switching model, the proposed volatility risk measure is easy to implement. It can be employed to determine the expected size of a shortfall when hedging only with the asset and a bond. In practice, even assuming constant volatility, continuous perfect hedging is not possible. The Black-Scholes framework, however, establishes a benchmark for hedging costs, with respect to which other more realistic hedging strategies can be judged. Similarly, in the context of stochastic volatility models the hedging strategy proposed in the paper represents hedging costs with which other methods can be compared.

Among alternative hedging approaches developed in the literature, the local risk-minimization and the mean-variance hedging seem to be most popular. Both enjoy certain optimality properties; the former minimizes the local risk as measured by the conditional second moment of cost increments while the latter minimizes the global risk over the entire life-time of the contract, as measured by the variance of the hedger's future costs[2].

An undesirable feature of these approaches is that the hedger is penalized equally for losses and gains. In addition, they rely on exact specifications of the volatility model and the correlation between the risky asset and its volatility, which for long-term instruments may be difficult to obtain. Their implementation is also more cumbersome when compared with the method based on realized volatility, since typically simulations must be conducted in order to assess the size of the hedging error. For this problem, the formula that we present

[2]For applications of these methods in the context of stochastic volatility models see, for example, Di Masi et al. (1994); Biagini et al. (2000); Heath et al. (2001), and for hedging equity-linked insurance contracts in discrete-time models, Møller (2001).

for the variance of future costs can be used as an upper bound for the corresponding one in the mean-variance hedging. This follows from the facts that hedging based on the realized volatility and the mean-variance hedging are based on self-financing portfolios and that the latter minimizes the variance of the terminal costs.

There are also other important applications of the results presented in the paper. For example, they allow for a formal risk analysis of hedging strategies with different initial costs. Since there is a trade-off between the required initial cost for setting a replicating portfolio and the expected shortfall, the availability of both values for different strategies allows a hedger to find a solution that best reflects his/her risk tolerance.

Another area of applications is related to the problem of model uncertainty. The number of stochastic volatility models proposed in the literature is large and selection of the one that is worth the time and the expense for development and implementation is not simple. Even in the context of discrete-time regime switching models, one may consider several different specifications of the returns distribution, ranging from a normal distribution with two parameters to mixture distributions with a large number of parameters. In addition, a model may be fitted to historical data using different sampling frequencies. As a result, statistical analysis often may be inconclusive in the identification of a good model.

Another issue is the fact that statistical goodness-of-fit criteria may not be relevant for the problem of risk management. There are two reasons for this. First, statistical models are usually fitted to historical returns, which represent the dynamics under the so-called P-measure. On the other hand, the pricing must be performed under a risk-neutral measure Q, unique determination of which for stochastic volatility models is a challenging problem. Typically the correct measure is the one that is dictated by the market, which in the context of random volatility models can be identified, for example, in terms of the implied volatility surface (e.g., Rebonato 1999). This approach, however, is much more difficult to implement for long-term instruments, since currently there exist no reliable descriptions of the dynamics of such complex objects. Therefore, parsimonious models must be used that would capable of capturing important for pricing and hedging features of the asset price distributions. In this aspect, we demonstrate that for a broad class of models this essential quantity is given by cumulative variation. This suggests, that a correct statistical description of this value should be used as the primary criterion for model selection.

To explain the second reason, let us suppose that the volatility of daily returns follows a two-state regime switching model but a model, with two regimes, is fitted to historical data using monthly returns. For contracts with a lifetime of 20 years, this frequency may be considered fine enough to meet some risk management criteria. Clearly, results of an estimation procedure will depend on the capability of the model to describe correctly the real dynamic. In this case, the cumulative variation follows a multi-state regime switching model, and hence the fitted model will not provide accurate description. The final result will depend on the used statistical criterion, which typically will be the maximization of the likelihood function. However, application of such criteria may not be suitable for identification of volatility dynamic features that are essential for pricing. It is easy to see that in this case the fitted model will not represent correctly the true range of monthly cumulative variances, and this may lead to an underestimation of the total volatility risk. We can gain more insight into this situation only when we understand better the way this risk depends on different elements of statistical models.

The remainder of the paper is organized as follows. The main results, including the characterization of the distribution of the expected surplus/shortfall and the formula for the variance of the expected shortfall, are presented in Section 2. In Section 3 we specialize these results to the case of a regime switching model and discuss some applications. Appendix provides a proof of the result from Section 2.

2 Expected Shortfall for Stochastic Volatility Models

In this section, we first introduce a method of dynamic hedging based on realized volatility and then provide a representation for expected shortfalls for a large class of stochastic volatility models. We also show that dynamic hedging based on implied volatility minimizes the expected shortfall and that it has some robustness properties.

In the sequel we assume that under the objective measure P the price of the underlying asset follows

$$dS_t = \mu_t S_t dt + \sigma_t S_t dB_t, \tag{2}$$

where $B_t, t \geq 0$, is a standard Brownian motion and μ_t denotes an instantaneous expected rate of return. The volatility process σ_t can be deterministic or stochastic. For example, σ_t may be a positive function of a single factor that follows another diffusion process, like in the model considered by Heston (1993) where the variance process v_t, equal to σ_t^2, is the square root mean-reverting process introduced by Cox, Ingersoll, and Ross (1985):

$$dv_t = k[\theta - v_t]dt + \sigma_v \sqrt{v_t} dB_t^v. \tag{3}$$

Here B^v denotes a standard Brownian motion which can be correlated with B. Such a model specification has been motivated by some empirical work in the financial literature which provides evidence of the fact that the volatility is mean reverting (e.g., Scott 1987; Stein and Stein 1991; among others).

Another class of specifications of σ_t is provided by regime switching models in which it is typically assumed that the volatility process is driven by a finite-state Markov process; i.e.,

$$\sigma_t = \sigma_{\rho_t},$$

where ρ_t is a continuous-time Markov chain taking values in the set $\{1, ..., M\}$ and $\sigma_1, ..., \sigma_M$ represent M different values of instantaneous volatility of the asset process. A present state of the latent process ρ_t determines the current "regime".

Our objective is to price and hedge a European type contract whose value at maturity T is given by a nonnegative function g. For this we introduce the following function:

$$W(S, t, \Sigma) := e^{-r(T-t)} E[g(S e^{r(T-t) - \frac{1}{2}\Sigma + \sqrt{\Sigma} Z})],$$

where Z is a standard normal random variable and r is the risk-free interest rate. In the rest of the paper we assume that g is such that $W(S, t, \Sigma)$ is an increasing function of Σ for each S and t. This is true, in particular, when g is a convex function (Hobson 1998).

Function W can be easily interpreted if we assumed that in (2) the volatility was constant, since in this case $W(S_0, 0, \sigma^2 T)$ would give us the Black-Scholes price of the option at time 0. This result can be generalized to the case when the volatility σ_t is deterministic. Then the risk-neutral price of the option would be

$$W(S_0, 0, \int_0^T \sigma_s^2 ds),$$

which follows from the Ito's lemma and the log-normality of S_T. In addition, for deterministic σ_t we would be able to hedge the contract perfectly by taking positions in the underlying asset S_t and the money market account, whose value B_t at time t is given by e^{rt}.

Surprisingly, to hedge the contract perfectly it is not necessary to know exactly the volatility path but only the realized cumulative variation $\int_0^T \sigma_s^2 ds$. We will be able to justify this result by providing an interpretation of $W(S_t, t, \Sigma)$ as time-t value of a certain instrument.

For a positive constant Σ_0, let us define a stochastic process

$$\Sigma_t := \Sigma_0 - \int_0^t \sigma_s^2 ds \tag{4}$$

and a stopping time

$$\tau = \inf\{s : \ \Sigma_s = 0\}.$$

It follows from the definition that the process Σ_t is adapted to the filtration generated by σ_t, which means that at any time t the value of Σ_t can be determined from the actual volatility path. Observe that in a diffusion model with continuous observations of the asset price the current volatility level can be approximated with arbitrary precision by the sum of squared increments.

Consider now another stochastic process whose value at time t, $t \in [0, \tau]$, is

$$V_t := W(S_t, t, \Sigma_t).$$

This process takes only nonnegative values and one may ask whether V_t represents a price of any traded or synthetically created security. The affirmative answer to this question has been provided by Mykland (2000), who studied this process in the context of super-replicating hedging strategies. The author demonstrates that there is a self-financing strategy based on hedging in the underlying asset and the money market account that replicates exactly V_t, $t \in [0, \tau]$, and that the delta, which determines the number of shares held at time t, is equal to $W'_S(S_t, t, \Sigma_t)$. This result may be found surprising as it does not rely on any particular dynamic of the process that governs the behavior of σ_t; to use this hedging method we only need to select an initial value Σ_0, create a portfolio that consists of a pre-specified number of shares of the risky asset and the risk-free asset, and then observe the volatility trajectory.

Let us denote the realized cumulative variation for the time period $[0, T]$ by Σ^R. This value is obviously unknown to a hedger at time 0. Suppose now that the hedger chooses a number, say Σ^H, that is at least equal to Σ^R. This number is used to create a portfolio with the initial value equal to $W(S_0, 0, \Sigma^H)$. By using dynamic hedging with the delta at time t given by $W'_S(S_t, t, \Sigma_t)$, the hedger can guarantee that the value of this portfolio will be equal to V_t at any time t. For such a selection of the initial cumulative variance Σ^H, we have the following relation at maturity

$$V_T \;=\; W(S_T, T, \Sigma_T) \geq W(S_T, T, 0) = g(S_T),$$

which becomes equality when $\Sigma^H = \Sigma^R$ since in this case $\Sigma_T = 0$. Consequently, the difference

$$W(S_T, T, \Sigma_T) - g(S_T) \tag{5}$$

represents a surplus at time T, as the value of replicating portfolio based on Σ^H dominates the value of the option regardless of the terminal price of the underlying asset S_T. The present expected value of (5) is equal to

$$W(S_0, 0, \Sigma^H) - W(S_0, 0, \Sigma^R), \tag{6}$$

and hence this number represents the expected surplus. Note that in the case when the hedger guesses correctly the cumulative variance Σ^R, the option will be replicated exactly (with probability one).

Suppose now that the hedger chooses at time 0 a positive number Σ^H that is smaller than Σ^R, and as before he uses delta hedging to replicate the instrument V_t. In this case $\tau < T$, and at this time we have

$$\begin{aligned} V_\tau &= W(S_\tau, \tau, 0) \\ &\leq W(S_\tau, \tau, \Sigma^R - \int_0^\tau \sigma_s^2 ds). \end{aligned}$$

Now the difference

$$W(S_\tau, \tau, \Sigma^R - \int_0^\tau \sigma_s^2 ds) - W(S_\tau, \tau, 0) \tag{7}$$

represents a shortfall since the values of both portfolios at time τ, given by V_τ and $W(S_\tau, \tau, \Sigma^R - \int_0^\tau \sigma_s^2 ds)$, were obtained through self-financing trading strategies. Hence in the case when $\Sigma^H < \Sigma^R$, the expected value of a shortfall is equal to the discounted expectation of (7), which is

$$W(S_0, 0, \Sigma^R) - W(S_0, 0, \Sigma^H). \tag{8}$$

In the case when $\tau < T$, an external funding is necessary to continue the dynamic hedging. The amount that will be added at time τ, say E_τ, will be related to a new selection of the cumulative variance made by the hedger, Σ_τ^H, through the formula

$$W(S_\tau, \tau, \Sigma_\tau^H) = E_\tau + W(S_\tau, \tau, 0).$$

If

$$W(S_\tau, \tau, \Sigma_\tau^H) \geq W(S_\tau, \tau, \Sigma^R - \int_0^\tau \sigma_s^2 ds), \tag{9}$$

then there will be a surplus at maturity whose expected value at time τ is given by

$$W(S_\tau, \tau, \Sigma^H_\tau) - W(S_\tau, \tau, \Sigma^R - \int_0^\tau \sigma_s^2 ds).$$

If (9) is not satisfied, then the process of adding external funds will have to be repeated, as then the value of the replicating portfolio at τ is not sufficient to hedge dynamically until maturity. Regardless how often this procedure will have to be repeated, however, the difference (8) represents the expected shortfall corresponding to the initial choice of the cumulative variance Σ^H.

Observe that the described hedging strategy will adjust the replicating portfolio to the market value only at times when external funding is necessary. However, a hedger that calculates the delta according to the realized volatility (4) can as well mark the portfolio to the market at other times. It is easy to see that this will not change the above interpretation of expressions (6) and (8). For example, if we set our model under a minimal martingale measure Q then the replicating portfolio can be adjusted, at selected discrete time points, to the values that correspond to a locally risk-minimizing hedging strategy. For regime switching models, this method of hedging has been discussed by Di Masi et al. (1994).

To summarize, expressions (6) and (8) represent an expected surplus and an expected shortfall for selections of the initial cumulative variance being respectively larger and smaller than the realized one. If we can identify the random mechanisms that generate the values of Σ^H and Σ^R, this interpretation of the differences can be employed to quantify the overall expected shortfall.

Let us denote by V^R a random variable that describes the occurrences of the realized cumulative variance. Observe that this variable is completely specified at the time we postulate a dynamic of the volatility paths under a Q-measure. For example, we can adopt model (2) with $\mu_t = r$ and the volatility that either follows (3) or a regime switching model. We shall denote the probability distribution of V^R by π^R. This distribution depends on the current states of the asset and the volatility processes, but to simplify the notation we do not write this explicitly.

Suppose for a moment that the hedger can predict without any error the realized value Σ^R. From the previous analysis it follows that for each Σ^R the initial cost of setting a replicating portfolio will be $W(S_0, 0, \Sigma^R)$, and there will be no shortfalls at maturity. Hence, the average initial cost of hedging with respect to the distribution of Σ^R is

$$P_0 := E[W(S_0, 0, V^R)]. \tag{10}$$

It is easy to verify that this number is the same as the risk-neutral price of the option based on the model (2) under the assumption that the stock price and the volatility σ_t are independent. Here, however, we can interpret it as the cost of hedging when the volatility risk has been eliminated by guessing perfectly the cumulative variance.

Let us denote a random variable that describes the hedger's selection strategy by V^H and the corresponding probability distribution by π^H. This distribution will represent the complete knowledge the hedger has about Σ^R given present values of the asset and volatility processes and hence we may assume that V^R and V^H are independent. The overall cost of hedging using a strategy V^H includes the initial cost plus a shortfall. In the case when

Σ^R could be predicted without any error the average cost would be P_0. This establishes a benchmark for the initial cost of hedging, if the process was repeated and the cumulative variance followed V^R. This suggests that in the case when V^R and V^H are independent, a class of reasonable strategies π^H may be represented by the following set

$$\mathcal{H}_{P_0} := \{\pi^H : E^{\pi^H}[W(S_0, 0, V^H)] = P_0\}.$$

For each distribution from this class the average initial cost will be the same but possible shortfalls will follow different distributions. In order to compare different selection strategies, we need methods of assessing the probability of a possible shortfall as well as its magnitude. The next proposition provides the necessary formulae, justification of which follows directly from the above discussion.

Proposition 2.1 *Suppose that the realized cumulative variance follows a random variable V^R with a probability distribution π^R. Assume also that to hedge dynamically a European option with the terminal payoff $g(S_T)$, an initial replicating portfolio is created at the cost $W(S_0, 0, \Sigma^H)$, where Σ^H is selected according to a random variable V^H whose probability distribution is π^H. The random variables V^R and V^H are assumed to be independent. Then,*

(a) *the probability of a shortfall is equal to*

$$P(V^R > V^H),$$

(b) *the time-zero value of the expected shortfall can be represented as*

$$ES(\pi^H, \pi^R) := E^{\pi^R} E^{\pi^H}[(W(S_0, 0, V^R) - W(S_0, 0, V^H))1_{\{V^R > V^H\}}], \tag{11}$$

where 1_A is an indicator function of an event A.

It is easy to provide some examples of possible selection strategies from $\mathcal{H}_{P_0}$. One such a strategy is to select Σ^H according to the true distribution of Σ^R. It can be described as a strategy that mimics the real mechanism that generates Σ^R.

Another strategy is to select always the same value. In this case, the corresponding probability distribution, denoted by π_{BS}, will put the whole probability mass at a single point, say Σ_{BS}, which by the definition of $\mathcal{H}_{P_0}$ must satisfy the equation

$$W(S_0, 0, \Sigma_{BS}) = P_0. \tag{12}$$

Let us observe that Σ_{BS} chosen according to this method is the same as the volatility that calibrates the Black-Scholes model to model (2) under a Q measure. This however, does not mean that the dynamic hedging strategy described above and the delta hedging based on the Black-Scholes model are equivalent. Firstly, if a hedger decides to use the Black-Scholes model while (2) describes the true dynamic of the asset then although Σ_{BS} leads to the correct cost of setting a replicating portfolio, hedging with this portfolio is consistent with the model only at time 0. Secondly, the two methods handle differently the available information about volatility. While the former keeps track of the value of the realized volatility and uses it to determine the number of shares of the risky asset in the replicating portfolio, the latter ignores this information and calculates the cumulative variance as a proportion of the remaining time to maturity.

The above examples show that we have a non-empty set of hedging strategies with the same initial costs. Therefore a natural question arises as to whether it is possible to identify an optimal, in some sense, strategy π^{opt}. Our discussion in the last section suggests selecting a strategy that minimizes the variability of the surplus/shortfall at the end of each hedging period. In other words, using variance as a criterion the optimal strategy should minimize

$$E^{\pi} E^{\pi^R}[W(S_0, 0, V^R) - W(S_0, 0, V^H)]^2, \tag{13}$$

with respect to $\pi \in \mathcal{H}_{P_0}$, which follows from the fact that all distributions in this class have the same expectations. This approach is valid if we assume that the distribution π^R is known. A more robust approach to this problem would be to search for an optimal selection strategy assuming only that an initial cost P_0 is known. Under such a form of uncertainty of the distribution π^R, an optimal selection strategy should minimize the function

$$\pi \rightarrow \sup_{\pi^R \in \mathcal{H}_{P_0}} E^{\pi} E^{\pi^R}[W(S_0, 0, V^R) - W(S_0, 0, V^H)]^2. \tag{14}$$

From a hedger's viewpoint, the criterion based on variance is not totally satisfactory as it treats surpluses and shortfalls in a symmetric way. To address this, one may search for a strategy from $\mathcal{H}_{P_0}$ that minimizes the expected surplus. As before two versions of the problem can be considered: the first would assume that π^R is known completely and the second would be based on a weaker assumption that the only available information about π^R is that it belongs to $\mathcal{H}_{P_0}$.

Below we show that for all four criteria the optimal strategy is the one that puts the whole probability mass at a single point Σ_{BS} that solves (12). A proof of this result is presented in the appendix.

Proposition 2.2 *Assume that the random variables V^R and V^H are independent.*

(i) *Suppose that the probability distribution π^R is known. Then a strategy that selects the cumulative variance equal to Σ_{BS} with probability one minimizes the variance of a surplus/shortfall (13) and the expected shortfall.*

(ii) *The same strategy minimizes over $\mathcal{H}_{P_0}$ the following robust versions of the above criteria*

$$\sup_{\pi^R \in \mathcal{H}_{P_0}} E^{\pi} E^{\pi^R}[W(S_0, 0, V^R) - W(S_0, 0, V^H)]^2$$

and

$$\sup_{\pi^R \in \mathcal{H}_{P_0}} ES(\pi, \pi^R),$$

provided that these values are finite.

The above result suggests that to minimize the size of a shortfall it suffices to focus on the simple strategy π_{BS}. It is worth emphasizing, however, that the assumption of independence of V^H and V^R is crucial for the optimality of π_{BS}. If not all information relevant for the prediction of Σ^R is reflected in the distribution of V^R, the variables V^H and V^R may be correlated. In that case it is easy to see that to take an advantage of the correlation the support of the distribution π^H must consist of at least two points.

The derived formula for the expected shortfall can be used in practice to quantify and formally manage the volatility risk. The essential ingredient of the analysis is the distribution of the cumulative variance for the time horizon corresponding to the life time of the hedged instrument. For a given model of volatility, this distribution can always be obtained through simulations, but in some cases, for example when σ_t follows of a finite state Markov chain, easier recursive methods are also available.

The presence of an uncertain liability in the future has been studied in many different settings and several strategies have been proposed. In the context of a regime switching model one can easily identify the most conservative approach, which corresponds to the worst case scenario for the cumulative variance. For such models the largest values of the cumulative variance, say Σ_{MAX}, exists. If this value is chosen with probability one to determine the cumulative variance, then it follows from (11) that the expected shortfall is zero and there is no risk of future liability. Such a strategy is called super-replicating, as the terminal value of the replicating portfolio is always at least equal to the hedged instrument.

This result has been proven before and studied by many authors (e.g., El Karoui et al. 1998), but from an economic viewpoint this strategy is not fully satisfactory as it ignores any additional information about particular dynamics of volatility paths. Our results demonstrate that the distribution of the expected shortfall may be available for some models. Therefore, it should be used to balance the trade-off between the initial cost and a shortfall at maturity. This is possible because a large number of new strategies of selecting the initial cumulative variance can be created by assigning probability one to a value of the cumulative variance that is between Σ_{BS} and Σ_{MAX}. For each strategy, both the initial cost and the expected shortfall can be determined. As we reduce the initial cost from the maximal value, which corresponds to Σ_{MAX}, the expected shortfall and the probability of the shortfall will increase. By finding this relationship, a hedger can select a strategy that is optimal according to his/her risk preference.

The explicit formula for the expected shortfall is also helpful in the presence of several candidates for the distribution of Σ^R. Here we briefly outline some of the possible approaches:

1. For hedging volatility risk the essential information is provided by the initial cost P_0 and the expected shortfall. It is possible that some of the candidate models will be quite similar in view of these values. This information may be used to eliminate, for example, models that are more prone to estimation or implementation errors.

2. Suppose that we have several models, each one described by the initial cost and the expected shortfall: $(P_0^{M_1}, ES^{M_1}), ..., (P_0^{M_J}, ES^{M_J})$. If the initial costs were close to each other then a conservative strategy would be to select the model with the highest expected shortfall. The problem becomes more difficult when the initial costs are different, as then a comparison of the models must reflect the risk tolerance of the

hedger. For this purpose one may use, for example, the total expected costs, given by $P_0^{M_i} + ES^{M_i}$, $i = 1, \ldots, J$. Then all the models can be ordered, and the preferable one (e.g., the most conservative) selected. Alternatively, one may set an upper bound for the expected shortfall and then choose the smallest initial cost that would guarantee that regardless of the model the expected shortfall will not exceed this level.

3. In the context of volatility risk management, the existence of different models suggests uncertainty in the identification of the probability distribution of the cumulative variance. Let us denote the candidates by $\pi^{M_1}, .., \pi^{M_J}$. Suppose also that plausibility of each model can be quantified by attaching a number, say $\alpha_i, i = 1, .., J$, between 0 an 1 such that $\alpha_1 + \cdots + \alpha_J = 1$. Then this information about the volatility dynamic can be easily incorporated into a single model by introducing a new distribution of the total cumulative variance in terms of a finite mixture of all distributions:

$$\sum_{i=1}^{J} \pi^{M_i} \alpha_i.$$

The resulting distribution summarizes the overall initial knowledge about the random variable V^R, and as such it may be used to price and hedge a contingent claim. Let us observe that the price P_0 obtained from (10) for this distribution will be consistent with the no-arbitrage assumption, as under this combined measure the discounted price processes of traded securities remain martingales.

3 Volatility risk for regime switching models

In this section, we consider a regime switching model as a special application of the general results obtained in the last section. First we discuss briefly results of statistical analysis related to fitting different specifications of such a model to real time series and present corresponding distributions of cumulative variances over time period of 1 month (or 25 days, assuming 252 trading days per year). Later we employ some of the methods from Section 2 to deal with the trade-off between initial costs and the expected shortfall as well as the problem of model uncertainty.

For the statistical analysis we used daily closing prices of S&P/TSX Composite Index for the time period starting in January 1981 and ending in December 1999. Two models were fitted to the daily log-returns $y_t = \log(S_{t+1}/S_t)$, to which we will refer to as high-frequency data. For simplicity of the exposition, we allowed the regime process to take only two values, which we denote by 1 and 2. In addition, we assumed that in Model 1 the conditional distribution of log-returns was normal

$$y_t|\rho_t = i \ \sim \ N(\mu_i, \sigma_i^2),$$

while in Model 2 it was a Student t-distribution

$$y_t|\rho_t = i \ \sim \ t(\mu_i, c_i, d_i),$$

where μ_i, c_i and d_i denote respectively the mean, the scale and the number of degrees of freedom of the distribution in regime i, $i = 1, 2$. By using the second model we were able to judge the extent to which a regime switching model with normal distribution was capable of explaining daily returns, distribution of which is known to have heavy tails. We should observe that the second model is not inconsistent with the continuous-time specification (2) from Section 2. It is known that Student t-distribution can be characterized as a mixture of normal distribution, when the variance follows a certain continuous distribution. Hence Model 2 can provide a description of situations where the mean variance is switching between two regimes but in each regime the variance is not constant but it fluctuates around its mean. Such a behavior is consistent with the stylized facts observed in the markets.

In the estimation of each model, we allowed the expected returns to depend on the current regime. This additional flexibility enabled us to estimate the variance of returns in the presence of volatility risk premia. For illustration of different applications of the results from Section 2, we have selected a Q measure that has the same variance of log-returns as the one estimated from historical data (see also Hardy 2001). In practice, observations that correspond to P-measure and/or Q-measure may be used to select the martingale component of the model. For a comparison of the behavior of volatilities implicit in option prices and those estimated from stock prices see, for example, Lamoureux and Lastrapes (1993) and Christensen and Hansen (2002).

An analysis of goodness of fit of the two models based on the likelihood ratio test and the Akaike information criterion strongly suggested superiority of the model with t-distribution over the one with the normal distribution (Monroy, 2000). In this case the formal identification of the better model was a straightforward task as the models were embedded. This, however, is not always the case since many of the suggested models for high-frequency data, like the one based on α-stable distributions, hyperbolic distributions or mixture distributions, may not have this property. Thus statistical tools may not provide enough evidence to indicate a model that will be most successful in describing future behavior. From the viewpoint of valuation and hedging of contingent claims, however, we should also use criteria that quantify the involved risk, such as the one discussed in the present paper.

To provide additional examples of plausible distributions of the cumulative variance, we also fitted Models 1 and 2 to logarithms of monthly returns. Since we found little statistical evidence in favor of the t-distribution, we decided to use only the model with normal distribution, which we call Model 3. The fact that monthly log-returns are well described by normal distribution can be attributed to the effect of the central limit theorem. This behavior certainly simplifies the search for models that would describe well statistical properties of low-frequency data. It is not obvious, however, whether models based on such data will also be satisfactory for the purpose of pricing and hedging of contingent claims in continuous time. We can look at this aspect of modeling by comparing the volatility risks of different models.

In Section 2 we demonstrated that for the valuation of contingent claims and an assessment of volatility risk associated with their hedging, we only need to know the distribution of the cumulative variance over a time period that corresponds to the lifetime of the contract. For regime switching models these distributions can be completely determined if we know the transition probabilities

$$p_{ij} = P(\rho_{t+1} = j | \rho_t = i), \quad i, j \in \{1, 2\}.$$

and the variances σ_i^2, $i = 1, 2$, of log-returns of daily data in each regime. The estimated values of these parameters for all three models are summarized below:

	p_{12}	p_{21}	σ_1^2	σ_2^2
Model 1 (daily data, normal distribution)	0.02	0.10	0.000028	0.000234
Model 2 (daily data, Student t-distribution)	0.01	0.03	0.000034	0.000167
Model 3 (monthly data, normal distribution)	0.03	0.19	0.001163	0.005271

As each model specifies uniquely the dynamic of the volatility process σ_t, the same is true about the distribution of the random variable V_K that describes values of the cumulative variance over a time period of length K (in days). For the first two models, this variable takes values in the set

$$\mathcal{W}_K := \{v_l^2 : \ (K - l)\sigma_1^2 + l\sigma_2^2, \quad l \in 0, ..., K\},$$

with the corresponding conditional probabilities given by

$$p_{K,\rho_0}(v_l^2) := P(V_K = v_l^2 | \rho_0) = P(\ \rho_t \text{ visits regime 2 } l \text{ times} | \rho_0), \quad \rho_0 \in \{1, 2\},$$

for $v_l^2 \in \mathcal{W}_K$, and zero otherwise. By defining V_K and R_K as the cumulative variance and total sojourn in regime 2, respectively, over $[0, K]$, the variable V_K can be related to R_K through the following simple formula

$$V_K = (K - R_K)\sigma_1^2 + R_K\sigma_2^2, \tag{15}$$

which implies

$$p_{K,\rho_0}(v_l^2) = P(V_K = v_l^2 | \rho_0) = P(R_K = l | \rho_0), \quad \rho_0 \in \{1, 2\}.$$

Since the probability function of R_K can be easily calculated using a recursive method (e.g., Hardy 2003), the same is true about the distribution of cumulative variance.

The described method furnishes the probability distributions of V_K for the first two models and any integer number K. For Model 3 this distribution is directly estimated from monthly data using the maximum likelihood method. Assuming that $K = 21$ days, histograms of these distributions for Models 1 and 2 are presented in Figure 1. The two graphs in the upper panel are the respective probabilities for regimes 1 and 2 under Model 1 while the two graphs in the lower panel are for regimes 1 and 2 under Model 2. A typical pattern for these distributions is that there is a relatively large probability mass in one value with the remaining probabilities being small. For instance, if the current regime ρ_0 is 1, then there is a spike at v_0^2. This is due to the persistence of the volatility process since $p_{K,1}(v_0^2)$ refers to the probability that regime 1 will be visited K times and this occurs with high probability given the current regime is 1. Similarly, if the current regime is 2 then the probability of visiting the higher volatility regime K times (given by $p_{K,2}(v_K^2)$) tends to be high (but lower than $p_{K,1}(v_0^2)$). Such phenomenon is less pronounced for regime 2 in Model 1 due to the higher estimated transition probability from regime 2 to regime 1. The estimated distribution for Model 3 is similar in characteristics but concentrated at two points only.

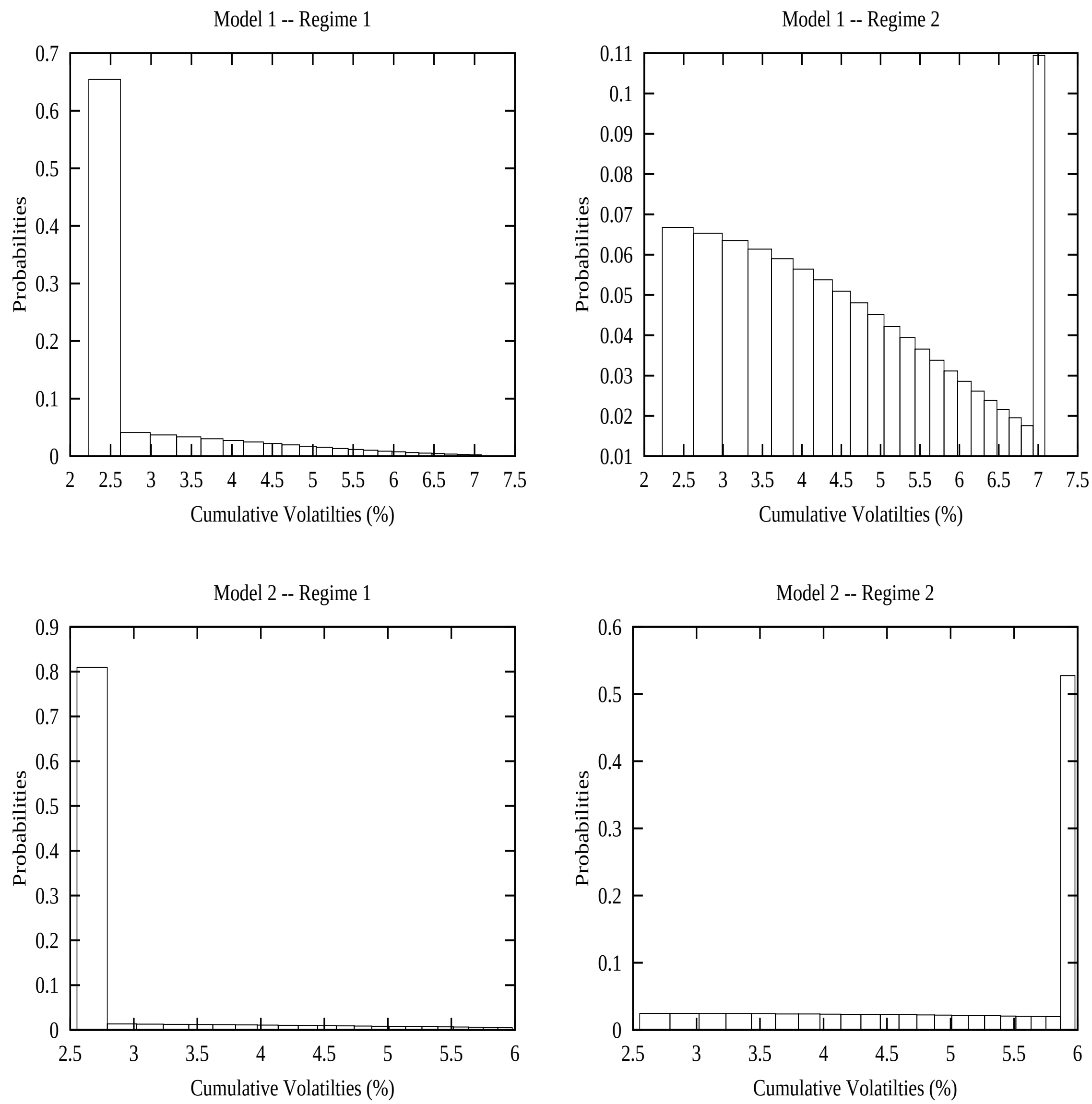


Figure 1: Distribution of Cumulative Volatilities for Models 1 and 2

We now consider the pricing of a European contract under a regime switching model. For this we assume that the current state of the volatility process is known, but in practice ρ_t must be treated as a latent variable and therefore estimated from observable variables. Due to the independence assumption, the distribution of S_T, conditionally on the volatility path, is lognormal with the variance equal to the cumulative variance over the time period $[0, T]$. A discrete-time approximation to this variance is provided by the random variable V_K with values in the set $\mathcal{W}_K$, as defined in (15). Let $BS(g, T, r, \sigma)$ denote the Black-Scholes value of a European contingent claim with payoff function g at maturity T, volatility σ, and the risk-free rate of return r. Then under the regime switching model with initial stock price S_0 and current regime ρ_0, the risk-neutral value of this contract is given by

$$P_{RS}(g; T, S_0, \rho_0) := \sum_{l=0}^{K} BS(g, T, r, v_l) p_{K,\rho_0}(v_l^2). \tag{16}$$

When the cumulative variance is known, possible changes in the value of the underlying asset can be hedged by using dynamic hedging. The incompleteness of the assumed model, however, implies that perfect hedging is not possible. It is therefore important to quantify the volatility risk as well as the potential shortfall for a given hedging strategy.

We now illustrate some of the results from Section 2 using a standard European call option. Let $BS_{\text{call}}(S_0, \sigma)$ denote the Black-Scholes value of a European call option when the initial stock price is S_0 and the cumulative volatility, which is defined as the square-root of the cumulative variance, for the time period $[0, T]$ is σ. Then

$$BS_{\text{call}}(S_0, \sigma) = S_0 \Phi(d_1) - X e^{-rT} \Phi(d_1 - \sigma),$$

where

$$d_1 = \frac{\ln(S_0/X) + rT + \frac{1}{2}\sigma^2}{\sigma},$$

and Φ denotes the cumulative distribution function of a standard normal random variable. Furthermore, it follows from (16) that the risk-neutral value of the option under the regime switching model can be expressed as

$$P_{RS}(S_0, \rho_0) = \sum_{l=0}^{K} BS_{\text{call}}(S_0, v_l) p_{K,\rho_0}(v_l^2), \qquad \rho_0 \in \{1, 2\}. \tag{17}$$

The value of the expected shortfall depends both on the market volatility, as described by the random variable V^R, and the strategy a hedger uses to select his/her initial volatility. This selection will be made according to another random variable, V^H. It was shown in Section 2 that we can compute the expected shortfall and the probability of a shortfall given the present regime ρ_0. Using $SF(\sigma, \rho_0)$ and $p_{SF}(\sigma, \rho_0)$ to denote, respectively, these two quantities of interest, we have

$$SF(\sigma, \rho_0) = \sum_{l=0}^{K} [BS_{\text{call}}(S_0, v_l) - BS_{\text{call}}(S_0, \sigma)] p_{K,\rho_0}(v_l^2) 1_{\{v_l^2 > \sigma^2\}} \tag{18}$$

and

$$p_{SF}(\sigma, \rho_0) = \sum_{l=0}^{K} p_{K,\rho_0}(v_l^2) 1_{\{v_l^2 > \sigma^2\}}. \tag{19}$$

We have also established there that regardless of the market mechanism of selecting the cumulative volatility, the hedger in regime ρ_0 should choose with probability 1 the value $\hat{\sigma}$ that solves

$$BS_{\text{call}}(S_0, \hat{\sigma}) = P_{RS}(S_0, \rho_0). \tag{20}$$

This value can be described as the implied volatility that calibrates the Black-Scholes model to the regime switching model. For illustration, suppose that both the initial stock price and strike price are 100, risk-free interest rate is 5%, and the number of days until the maturity of the option is 21; i.e. $K = 21$ or $T = 1/12$ year. Then, for Model 2 we have

$$\begin{aligned} P_{RS}(100, 1) &= 1.411 \\ P_{RS}(100, 2) &= 2.279. \end{aligned}$$

For regime 1, the implied volatility satisfying (20) is 2.99%, which translates to a shortfall probability of 0.18, expected shortfall of 0.10 and expected shortfall ratio of 7.26%. The expected shortfall ratio is defined as the ratio of expected shortfall to the initial price of the option. Similarly for regime 2, implied volatility is 5.19% with a shortfall probability 0.65, expected shortfall 0.17 and expected shortfall ratio 7.44%.

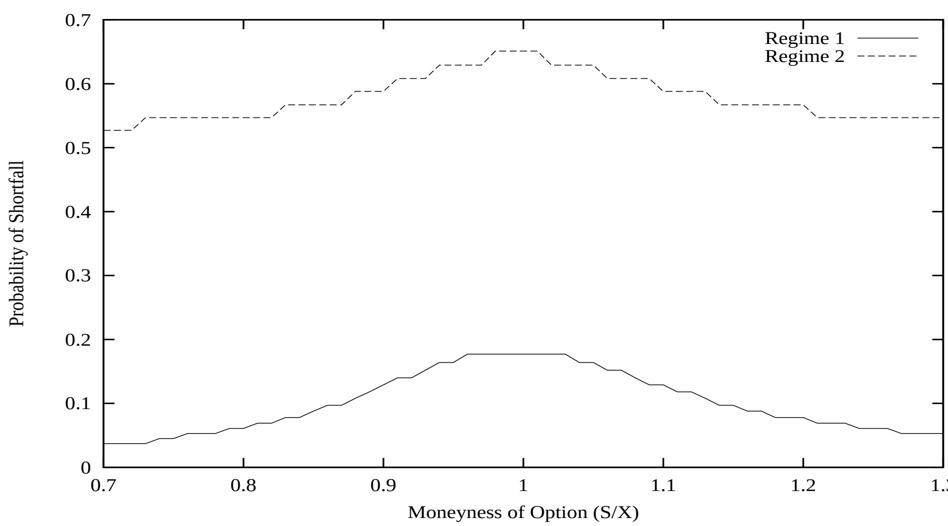


Figure 2: Plot of probability of shortfall $p_{SF}(\hat{\sigma}, \rho_0)$ vs moneyness of option S_0/X for both regimes

Figure 2 depicts the impact that the moneyness of the option has on the probability of a shortfall. The initial stock prices are assumed to lie in the interval $[70, 130]$. An immediate conclusion that can be drawn from this graph is that the shortfall probability is higher for regime 2, indicating that the volatility risk is also considerably higher. Note also that for both regimes, the shortfall probability reaches its peak when the option is at-the-money. This pattern can be explained by the volatility smile effect, which corresponds to the well known fact that the Black-Scholes formula applied to a stochastic volatility model overprices options that are at the money or close to the money, while it underprices options that are deep in or out of the money.

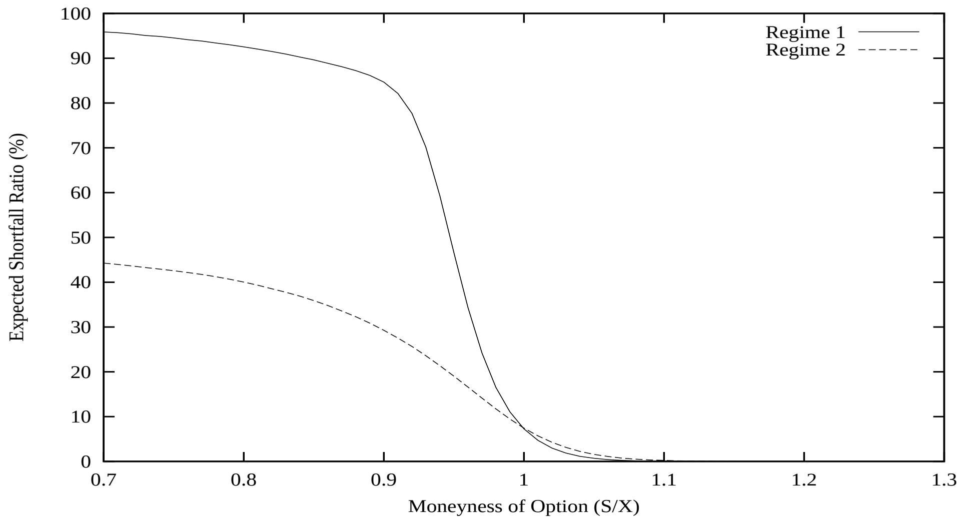


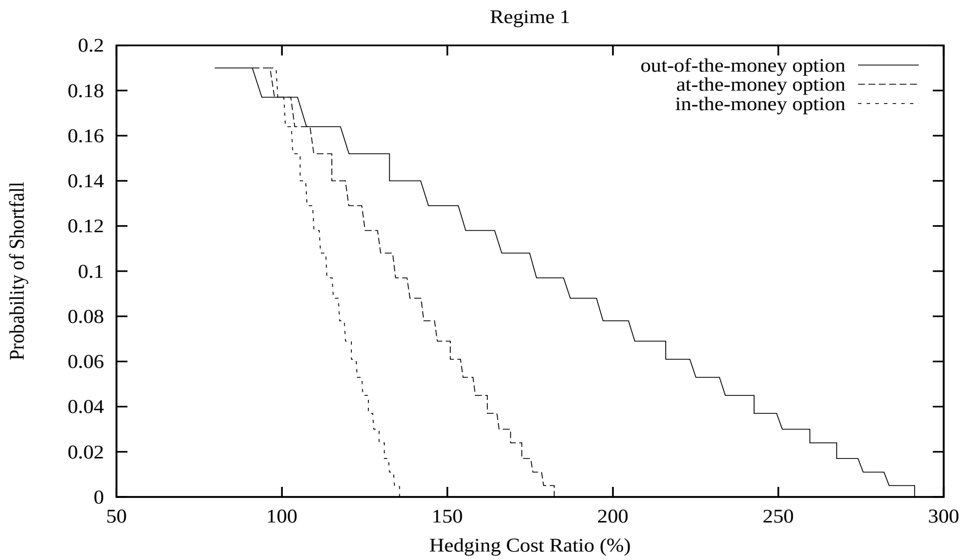
Figure 3: Plot of expected shortfall ratio $SF(\hat{\sigma}, \rho_0)/P_{RS}(S_0, \rho_0)$ vs moneyness of option S_0/X

Figure 3 considers the expected shortfall ratio as a function of moneyness. In terms of the relative proportion, the graph indicates that for out-of-the-money contract, the proportion of shortfall is much higher for the low volatility regime than for the high volatility regime. On the other hand, when the option becomes in-the-money, the high volatility regime has a greater proportion of expected shortfall. It should be pointed out that in terms of the dollar amount, the expected shortfall is greater for the high volatility regime despite the lower proportion for out-of-the money cases.

The preceding analysis has assumed that the hedger selects the initial volatility according to (20). One may then wonder if other levels of the initial cumulative volatility are chosen, what would be the impact on the risk of shortfall as well as the initial cost of hedging? Because of the uncertainty in which the market chooses the volatility path, the most conservative approach is the one where the hedger assumes the worst case scenario. Such approach always leads to a zero shortfall and is commonly known as the super-replication strategy, as discussed in the last section. The downside of this strategy is that it can be very costly which we will illustrate in our examples.

For a regime switching model, the possible values of cumulative volatility are known and lie in the interval $[v_0, v_K]$, regardless of the current state. Hence super-replication strategy corresponds to selecting the largest cumulative volatility v_K when constructing the hedging portfolio. Alternatively, a hedger can choose any value, which we denote as $\tilde{v}$, from the interval $[v_0, v_K]$ and create a corresponding hedging portfolio with cost equal to $BS_{\text{call}}(S_0, \tilde{v})$. Consequently for each chosen level of cumulative volatility $\tilde{v}$, we can study the tradeoff among (i) the cost of setting up the replicating portfolio, (ii) the probability of a shortfall $p_{SF}(\tilde{v}, \rho_0)$, and (iii) the magnitude of expected shortfall $SF(\tilde{v}, \rho_0)$. Note that the shortfall probability and expected shortfall depend on the current state of regime and moneyness of the contract.

Figure 4 shows the relationship between the initial cost of hedging and the probability of shortfall for both regimes 1 and 2. For each regime, we consider three levels of initial stock prices 98, 100 and 102, representing out-of-the-money, at-the-money, and in-the-money



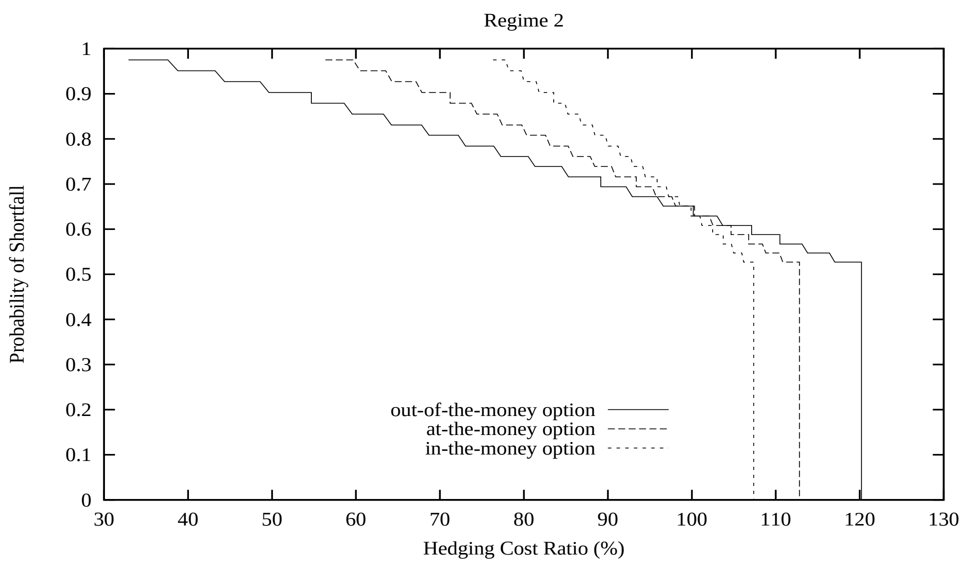
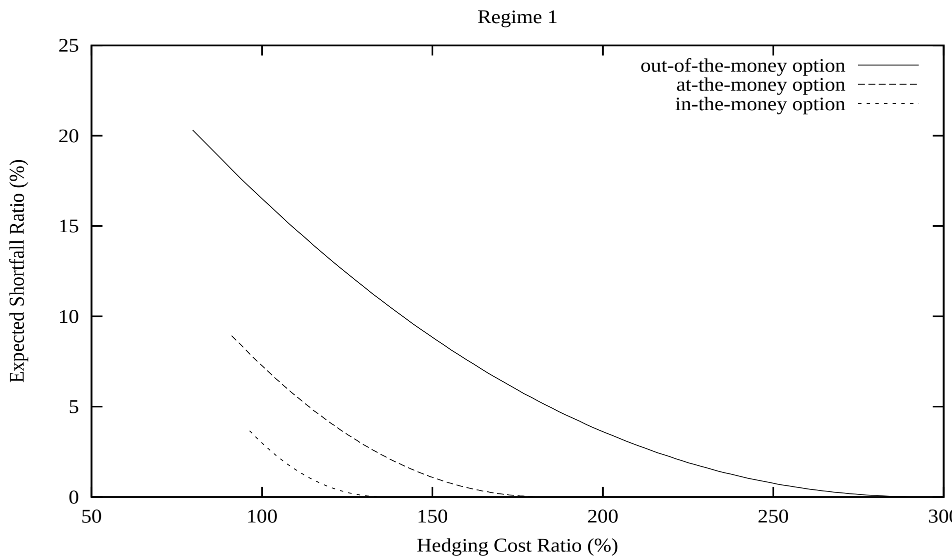


Figure 4: The relationship between Hedging Cost Ratio and Probability of Shortfall



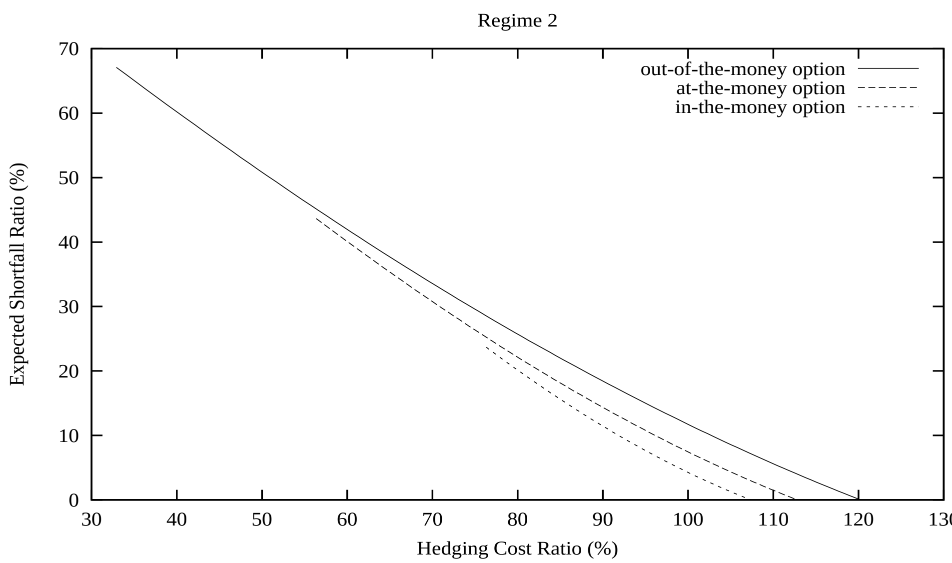


Figure 5: The relationship between Hedging Cost Ratio and Expected Shortfall Ratio

cases. To standardize the comparison across options, the initial hedging cost is reported as the ratio of the cost of the hedging portfolio over the market price of the option; i.e. $BS_{\text{call}}(S_0, \tilde{v})/P_{RS}(S_0, \rho_0)$. The graph indicates that when the initial hedging cost is less than the market value of the option, the in-the-money case has the highest shortfall probability. When the hedging cost ratio is greater than one, the situation is reversed and the out-of-the-money case is then subject to highest shortfall probability. The crossover occurs when the hedging cost coincides with the initial value of the option.

Figure 5 provides a tradeoff analysis between the hedging cost ratio and the expected shortfall ratio. For regime 2, the relationship between the hedging cost ratio and the expected shortfall ratio is close to linear. The reduction of a shortfall can be achieved by adding a proportional amount at time 0. Regime 1, on the other hand, exhibits a stronger non-linearity. For example, to achieve an expected shortfall proportion of 10% for the out-of-the-money case, regime 1 requires 140% of the initial option price while regime 2 only needs 102%. More importantly, if we were to eliminate the risk of shortfall completely, we need to further increase the initial hedging cost by approximately 150% for regime 1 while a mere 18% for regime 2. Consequently, super-replication may not be acceptable when low volatility is the current regime.

We now consider Models 1, 2 and 3 and discuss their implications on hedging. Figure 6 compares the prices of call options for regime 1 and $S_0 \in [98, 102]$ under these models. Clearly, the prices from Model 3 are the most expensive while the prices from Model 2 are the least. Suppose now that for each model we hedge the option according to the corresponding initial cumulative volatility implied from (20). Last section establishes that this strategy is optimal in that it minimizes the variance of a surplus/shortfall and the expected shortfall. Over the range $S_0 \in [80, 120]$, Figure 7 indicates that in general Model 1 subjects the hedger to the highest expected shortfall while Model 3 to the least. A more appropriate comparison is to consider the aggregate cost that takes into account both the initial hedging cost and the expected shortfall. This potentially eliminates the scenario where the small expected shortfall is attributed to the high initial hedging cost (see for example Model 3). Figure 8 confirms that in terms of the overall costs Model 3 is still most costly while Model 2 is the least. However, the aggregate costs for Model 1 are now much closer to Model 3.

The analysis presented above not only provides a systematic approach to the problem of quantifying the volatility risk for each model, but it can also be useful for selecting the appropriate model under model uncertainty. For instance, suppose that a hedger strongly believes that the market will follow one of the three models discussed above. Due to the model uncertainty, a prudent hedging strategy is to select the worst case model, which in our example is Model 3. This ensures that the aggregate cost (initial hedging cost plus expected shortfall) is bounded by this quantity. Such a comparison is possible because the only feature of these models that is relevant for pricing and hedging is the cumulative variance. Therefore, despite the fact that Model 3 has been estimated at a lower sampling frequency than Model 1, it can still be used for hedging as it provides an estimate of this essential quantity. Furthermore, above examples suggest that such a model does not have to underestimate the overall cost of hedging. This issue is important for risk management of long-term financial instruments, for which statistical models are typically estimated from low-frequency data. The above analysis suggest that such an approach does not exclude

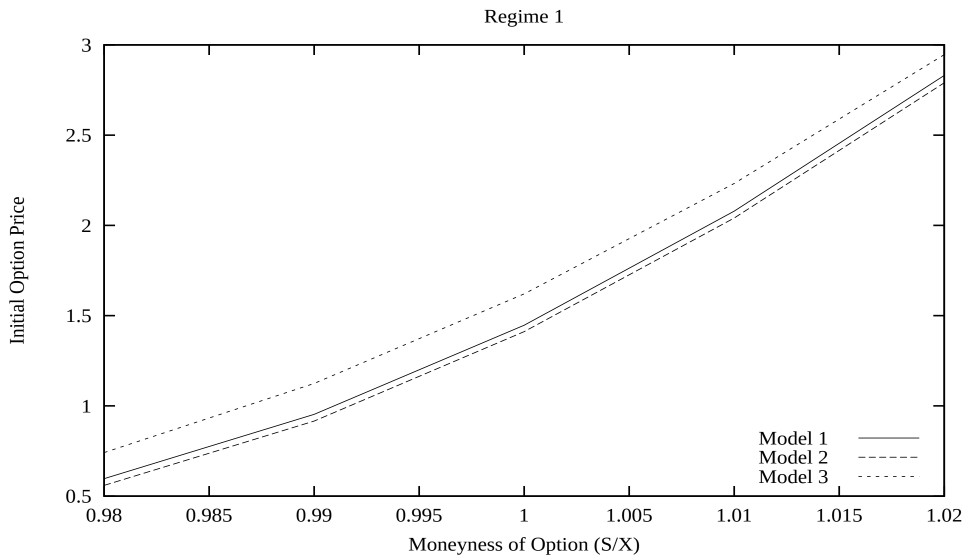


Figure 6: Initial option prices vs moneyness of the options for Models 1, 2 and 3 in regime 1

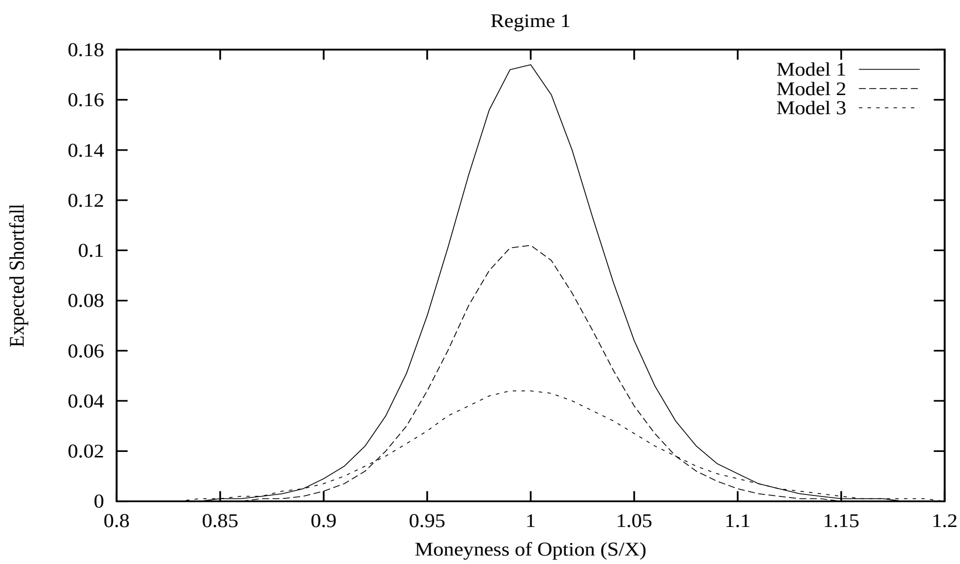


Figure 7: Expected shortfall vs moneyness of the options for Models 1, 2 and 3 in regime 1

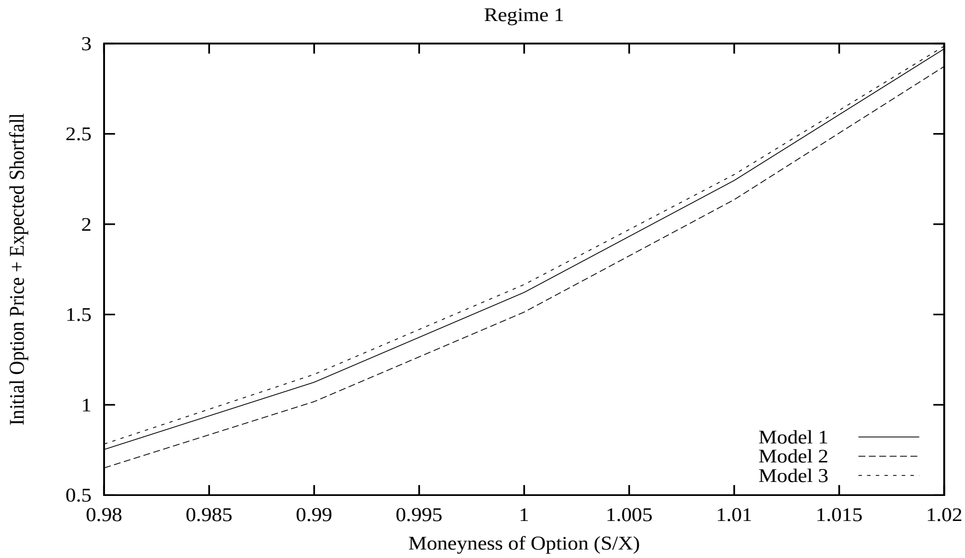


Figure 8: Plot of initial option price plus expected shortfall vs moneyness of the options for Models 1, 2 and 3 in regime 1

hedging strategies based on a more frequent trading.

Appendix

To show that π_{BS} minimizes the variance of the expected surplus/shortfall, first recall that for any square-integrable random variable X the minimum of

$$E[(X-a)^2]$$

over all a is attained for $a = E[X]$. Hence, we have

$$\begin{aligned} E^{\pi}E^{\pi^R}[W(S_0,0,V^R) - W(S_0,0,V^H)]^2 &= E^{\pi}E^{\pi^R}[[W(S_0,0,V^R) - W(S_0,0,V^H)]^2|V^H] \\ &\geq E^{\pi}E^{\pi^R}[[W(S_0,0,V^R) - P_0)]^2|V^H] \\ &= E^{\pi^R}[W(S_0,0,V^R) - P_0)]^2, \end{aligned} \tag{21}$$

where we also used the independence of V^R and V^H. Since the minimal value (21) can be attained by selecting $\pi = \pi_{BS}$, the result follows.

To show that the same strategy minimizes also the expected shortfall, first observe that by the assumed monotonicity of the function $\Sigma \to W(S_0, 0, \Sigma)$ it suffices to prove that the probability distribution concentrated at P_0 minimizes

$$E^{\pi^H}E^{\pi^R}(R-H)^+ \tag{22}$$

over all π^H such that $E^{\pi^H}[H] = P_0$, where x^+ is equal to x if $x \geq 0$ and zero otherwise, and R and H are random variables with probability distributions π^R and π^H, respectively.

Observe that for each y the function

$$x \to (y - x)^+$$

is convex. By linearity of expectation, this implies that the function

$$x \to E^{\pi^R}[(R - x)^+]$$

is also convex. Hence, by Jensen's inequality we have

$$E^{\pi^H} E^{\pi^R}(R - H)^+ \geq E^{\pi^R}(R - P_0)^+.$$

Since the right hand side of this inequality can be attained for π^H concentrated at P_0, this distribution minimizes (22), and consequently this shows that π_{BS} minimizes the expected shortfall.

In order to avoid technicalities, in the proof of part (ii) we assume that supports of all distributions in $\mathcal{H}_{P_0}$ are jointly bounded. Since these distributions describe changes in the cumulative variance, we believe that this assumption is reasonable. Let us equip $\mathcal{H}_{P_0}$ with the Lévy metric and consider two mappings

$$\pi \to E^{\pi}[X^2] \qquad \text{and } \pi \to E^{\pi}[X^+], \tag{23}$$

where π is a probability distribution and X is a random variable that follows this distribution. It can be shown that under this metric the functions (23) are continuous and the class $\mathcal{H}_{P_0}$ is compact. This implies that the suprema

$$\sup_{\pi^R \in \mathcal{H}_{P_0}} E^{\pi} E^{\pi^R}[(W(S_0, 0, V^R) - W(S_0, 0, V^H))^2] \tag{24}$$

and

$$\sup_{\pi^R \in \mathcal{H}_{P_0}} E^{\pi} E^{\pi^R}[(W(S_0, 0, V^R) - W(S_0, 0, V^H)^+] \tag{25}$$

are attained at distributions that must belong to $\mathcal{H}_{P_0}$. Since the optimal solution that we obtained in part (i) does not depend on π^R, it also minimizes (24) and (25).

References

[1] Biagini, F., P. Guasoni, and M. Pratelli (2000). Mean-variance hedging for stochastic volatility models. *Mathematical Finance* 10: 109-123.

[2] Chow, Y.S. and H. Teicher (1997). *Probability Theory. Independence, Interchangeability, Martingales*, 3rd ed. New York: Springer.

[3] Christensen, B.J. and C. Hansen (2002). New Evidence on the implied-realized volatility relation. *The European Journal of Finance* 8(2): 187-205.

[4] Cox, J.C., Ingersoll, J.E., and Ross, S.A. (1985) "A theory of the term structure of interest rates," *Econometrica* 53: 385-408.

[5] Di Masi, G.B., Y.M. Kabanov, and J. Runggaldier (1994). Mean-variance hedging of options on stocks with Markov volatilities. *Theory of Probability and Applications* 39: 172-182.

[6] El Karoui, N., M. Jeanblanc-Picque, and S. E. Shreve (1998). Robustness of the Black and Scholes formula. *Mathematical Finance* 8: 93-126.

[7] Heath, D., E. Platen, and M. Schweizer (2001). A comparison of two quadratic approaches to hedging in incomplete markets. *Mathematical Finance* 11(4): 385-413.

[8] Hardy, M. R. (2001). A regime switching model of long term stock returns. *North American Actuarial Journal* 5(2): 41-53.

[9] Hardy, M. (2003). *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance.* New Jersey: Wiley.

[10] Heston, S. (1993). A closed form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies* 6: 327-343.

[11] Hobson, D. G. (1998). Volatility misspecification, option pricing and super-replication via coupling. *The Annals of Applied Probability* 8: 193-205.

[12] Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42: 281-300.

[13] Kim, C.J, and Nelson, C.R. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications.* MIT Press: Cambridge, Massachusetts.

[14] Lamoureux, C.G. and Lastrapes, W.D. (1993). Forecasting stock-return variance: toward an understanding of stochastic implied volatilities. *Review of Financial Studies* 6: 293-326.

[15] Møller, T. (2001) Hedging equity-linked life insurance contracts. *North American Actuarial Journal* 5(2): 79-95.

[16] Monroy, P.M. (2000). Regime switching models with t-distribution. *Master Essay*, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo.

[17] Mykland, P.A. (2000). Conservative delta hedging. *The Annals of Applied Probability* 10: 664-683.

[18] Rebonato, R. (1999). *Volatility and Correlation in the Pricing of Equity, FX and Interest-Rate Options.* Chichester: Wiley.

[19] Scott, L. (1987). Option pricing when the variance changes randomly: theory, estimation, and an application, *Journal of Financial and Quantitative Analysis* 22: 419-438.

[20] Shephard, N. (1996). Statistical aspect of ARCH and stochastic volatility, in *Time Series Models in Econometrics, Finance and Other Fields*, eds. D.R.Cox, D.V. Hinkley, and O. Barndorff-Nielsen. Chapman & Hall, London, 1-67.

[21] Stein, E., and Stein, J. (1991). Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies* 4: 727-752.

[22] Taylor, S.J. (1994). Modeling stochastic volatility. *Mathematical Finance* 4: 183-204.