

Robust Stochastic Discount Factors

Phelim Boyle

Wilfrid Laurier University

Shui Feng

McMaster University

Weidong Tian

University of Waterloo

Tan Wang

University of British Columbia and CCFR

When the market is incomplete, a new non-redundant derivative security cannot be priced by no-arbitrage arguments alone. Moreover, there will be a multiplicity of stochastic discount factors and each of them may give a different price for the new derivative security. This paper develops an approach to the selection of a stochastic discount factor for pricing a new derivative security. The approach is based on the idea that the price of a derivative security should not vary too much when the payoff of the primitive security is slightly perturbed, i.e., the price of the derivative should be robust to model misspecification. The paper develops two metrics of robustness. The first is based on robustness in expectation. The second is based on robustness in probability and draws on tools from the theory of large deviations. We show that in a stochastic volatility model, the two metrics yield analytically tractable bounds for the derivative price, as the underlying stochastic volatility model is perturbed. The bounds can be readily used for numerical examination of the sensitivity of the price of the derivative to model misspecification. (*JEL* G11, G12)

1. Introduction

Derivative pricing in incomplete markets is an interesting and important problem in financial economics. When markets are complete and there is no arbitrage, there exists a unique stochastic discount factor. In this case, every

We are very grateful to the editor, Yacine Aït-Sahalia, and an anonymous referee for several constructive and insightful suggestions on how to improve the paper. We would like to acknowledge comments from George M. Constantinides, Jérôme Detemple, Jun Liu, Jiang Wang, Hong Yan, and participants at presentations given at McGill University, the University of Montreal, the University of British Columbia, the University of Florida, the University of Western Ontario, and the Third World Congress of the Bachelier Finance Society in Chicago. We are responsible for any errors. The authors thank the Natural Sciences and Engineering Research Council and the Social Sciences and Humanities Research Council for their support. Address correspondence to Weidong Tian, Department of Statistics & Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1; e-mail: wdtian@uwaterloo.ca.

security has a unique price. When markets are incomplete, however, the absence of arbitrage no longer guarantees a unique stochastic discount factor. There will be a multiplicity of admissible stochastic discount factors and this means that the pricing of a new security becomes a difficult problem.¹ In this paper, we develop an approach to this type of problem in the context of pricing a new derivative security. The basic idea is to develop a concept of robustness by requiring that the price of the derivative is not too sensitive to perturbations in the payoffs on the underlying asset.

The pricing of a new asset in incomplete markets often comes down to the selection of a particular stochastic discount factor. The existing literature offers several approaches. One approach is to postulate a representative investor and value the new security from the perspective of this investor. The preference of this representative investor is used to identify a particular stochastic discount factor. Lucas (1978); Cox, Ingersoll, and Ross (1985); and the large literature that follows are all examples of this representative agent approach. Since the maximization of a concave (expected utility) function is the dual of the minimization of the corresponding convex function, the second approach is to identify a unique stochastic discount factor by the minimization of a convex function (e.g., Rogers, 1994, and MacKay and Prisman, 1997). This problem is isomorphic to finding the Q measure that is closest to the objective probability measure P in the sense that it minimizes a measure of entropy (Stutzer, 1996). A third approach is to narrow the range of stochastic discount factors. It has been proposed by Bizid, Jouini, and Koehl (1999). They assume that the prices of a set of primitive securities are known and that a set of non-redundant derivative securities completes the market. Their paper exploits the properties of the equilibrium to constrain the range of stochastic discount factors. Jouini and Napp (1999) have extended this approach to a continuous time framework.

This paper offers an alternative approach to the selection of stochastic discount factors. The basic idea is as follows. Although all admissible stochastic discount factors price the existing primitive securities consistently, they typically give different prices for a new non-redundant derivative security. If, in addition, there is potential model misspecification, each stochastic discount factor gives a range of prices for the new security, one for each potentially misspecified model. These price ranges can typically be represented as intervals on the real line and they will differ in size, some smaller than others. The stochastic discount factor that gives the tightest (narrowest) range of prices is robust in the sense that under this stochastic discount factor, the price of the new security is the least sensitive to potential model misspecification. Our approach is to select the robust stochastic discount factor and use this stochastic discount factor to price the new derivative security.

Our approach uses two different metrics to capture the notion of robustness. Both start with perturbations of the existing primitive security and consider

¹ An admissible stochastic discount factor is one that prices the existing assets correctly.

the difference in the prices of the new derivative security based on different perturbations of the primitive security. Our first metric of robustness is obtained by finding the minimal range of the pricing difference as one varies through all admissible stochastic discount factors. If there are closed-form expressions for the price of the new derivative security under alternative perturbations of the primitive security, then the evaluation of the pricing difference is straightforward. Otherwise, one has to rely on the fundamental theorem of asset pricing to evaluate the pricing difference by calculating the difference in expectations under a risk-neutral probability.

Our second metric of robustness, instead of looking at the pricing difference today, looks first at the conditional pricing difference given an information set. Then, for each stochastic discount factor, it aggregates back to unconditional pricing difference by evaluating the risk-neutral probability of large conditional pricing difference under that stochastic discount factor. If the probability is significant, then under that stochastic discount factor, the price of the new derivative security is sensitive to potential model misspecification. If, for some stochastic discount factor, the probability is small, in other words, if the range of the probabilities as one varies through all possible perturbations of the primitive security is small, then the price of the new derivative security is robust under that stochastic discount factor.

Our paper builds on ideas from several streams of the literature. The first one is the incomplete market literature briefly reviewed above. The second one is the model misspecification literature. In a series of papers, Hansen, Sargent and their co-authors advocated a robust control framework in which the decision-maker views the model as an approximation.² Maenhout (2004); Uppal and Wang (2003); and Garlappi, Uppal, and Wang (2007) used this framework to study the impact of potential model misspecification on individual portfolio choice. These papers develop decision rules that not only work for the underlying model, but also perform reasonably well if there is some form of model misspecification. In general, these decision rules concentrate on the worst-case scenarios. The third literature that we draw our ideas from is the literature on potential misspecification of the stochastic discount factors and its implications for asset pricing. Prominent contributions include Hansen and Jagannathan (1997); and Hodrick and Zhang (2001), who studied stochastic discount factor misspecification errors. The issue that they study is, if an asset pricing model cannot price assets correctly, i.e., if the asset pricing model/stochastic discount factor is misspecified, how should one measure its misspecification? They propose to measure the degree of misspecification by the least square distance between the stochastic discount factor and the set of arbitrage-free stochastic discount factors. In duality terms, this leads to a quadratic distance that captures the maximum pricing errors across all assets. Bernardo and Ledoit (2000); and

² See Anderson, Hansen, and Sargent (2003); Cagetti et al. (2002); Hansen and Sargent (1995); and Hansen and Sargent (2001).

Cochrane and Saá-Requejo (2000) defined a (relative) distance metric between two stochastic discount factors in terms of the extreme values of the ratios of the stochastic discount factors across states. Their distance metrics also capture the worst-case error across the contingent claims.³ The focus of these studies is not on finding a stochastic discount factor that minimizes the sensitivity of asset prices to model misspecifications.

The approach proposed in this paper for selecting stochastic discount factors differs from the Hansen and Jagannathan (1997) approach. Our approach focuses on the choice of the stochastic discount factor when there is potential model misspecification of the payoffs of the assets, based on a given derivative security. Therefore, the choice of the stochastic discount factor depends on the given security. In the Hansen-Jagannathan approach, all possible securities are considered, and therefore, their robust stochastic discount factor is based on worst-case scenarios. Also the Hansen-Jagannathan approach is based on a one-period discrete model, whereas we employ a continuous time framework. The Hansen-Jagannathan approach is perhaps better motivated from economic considerations, since it provides a clear link between asset pricing and stochastic discount factors. In our paper, the linkage is less stressed because the pricing of a derivative security is at least partially determined by the underlying. So we can take a reduced form approach by taking the prices of the underlying as given, and hence, we do not have to specify the asset pricing model. The main advantage of our approach is that it considers robustness of stochastic discount factors and model misspecification in a unified framework.

While our approach can be applied more generally, we have illustrated it in the context of stochastic volatility models, since this provides a natural application to showcase the approach. Stochastic volatility models are of great interest in finance. Surprisingly, however, the literature has typically ignored model misspecification and focused on obtaining a price of a new derivative security (e.g., Heston, 1993). Frey and Sin, (1999) have analyzed the pricing bound of a standard European call option in a Black-Scholes-Merton world with stochastic volatility. They permit a wide range of values for the market price of volatility risk. Frey and Sin demonstrate that the bounds on the price of the call that represent the supremum and the infimum of the no-arbitrage prices correspond to Merton's no-arbitrage bounds. If the current asset price is S and the strike price is K , these bounds are given by $\max(S - K, 0)$ and S , respectively. However, this result does not give any guidance as to how to restrict the range of stochastic discount factors to narrow the pricing bounds. Levy (1985) uses a stochastic dominance argument to bound the option prices. In the current paper, following Frey and Sin, (1999), we illustrate our approach using a stochastic volatility model with a range of values of the market price of risk. In addition, we allow the model to be potentially misspecified. We derive

³ See also Bondarenko (2003); Cerney (2003); Hansen and Jagannathan (1981); Hansen and Jagannathan (1997); Ross (2005); and Stutzer (1995).

the pricing error under each market price of risk. The robust stochastic discount factor is then obtained by solving a minimization problem.

In summary, the aim of our paper is to provide one possible framework for dealing with two important aspects of stochastic discount factors. The first is related to their lack of uniqueness in an incomplete market. The second is related to possible model misspecification. Our notion of robustness provides procedures for selecting stochastic discount factors that are robust to model misspecification in the context of pricing a given derivative security.

The rest of the paper is organized as follows. Section 2 presents a simple one-period model to explain and motivate our approach. This section captures the essence of our approach and may be of interest in its own right. Section 3 describes a family of continuous time stochastic volatility models. We explain how to construct a benchmark model for a given market price of risk. Section 4 discusses the robustness in expectation approach to robust derivative pricing. Section 5 examines the robustness in probability approach to robust derivative pricing using a large deviations technique. Section 6 contains numerical analysis to illustrate the two approaches. Section 7 concludes. Proofs are collected in the appendices.

2. Discrete Time Setting

In this section, we illustrate the basic idea behind our approach in a single period setting.⁴ The basic idea is as follows. Since the market is incomplete, stochastic discount factors are not unique. There is, however, a stochastic discount factor under which the price of the derivative is the least sensitive to perturbations to its payoffs. Then, under this stochastic discount factor, the price of the derivative security is said to be robust to potential model misspecifications.⁵ We formalize this notion of robustness and then illustrate it with two simple numerical examples.

2.1 Robust derivative pricing

We assume a one-period frictionless financial market and that the uncertainty is described by the physical probability measure P . We assume that the number of traded assets is less than the number of states at time one, so that the market is incomplete. There is a set of stochastic discount factors m_α , such that the price of the payoff vector p is given by $E_P[m_\alpha p]$.

For simplicity, we consider a market in which there exists just a single risky primitive asset and a risk-free bond. The payoff of the risky asset is described by a model p . Since it is only a model, agents in the economy have concerns

⁴ We thank the referee for suggesting the approach used in this section.

⁵ We formalize model misspecifications as perturbations to the stochastic process of the price of the underlying asset. The potentially misspecified models are those that cannot be distinguished based on existing data and econometric technique.

that the payoff p may be misspecified and that the actual payoff is a payoff p_0 plus a perturbation/noise $z_{\alpha,\epsilon}$ in a family of possible perturbations indexed by $\alpha \in A$ and $\epsilon \in \mathcal{E}$. Here, we have allowed the perturbation to depend on α . The reason for that will be explained momentarily. We assume that $\mathcal{E}$ is a segment in the real line including zero and that for all α , $\lim_{\epsilon \rightarrow 0} z_{\alpha,\epsilon} = 0$. We call p_0 the benchmark model or the benchmark payoff and $p_0 + z_{\alpha,\epsilon}$ the perturbed model or the perturbed payoff. We require that

$$p = p_0 + z_{\alpha,\epsilon} \quad (1)$$

for some $(\alpha, \epsilon) \in A \times \mathcal{E}$ and that for all $\alpha \in A$ and $\epsilon \in \mathcal{E}$,

$$E^\alpha[p_0 + z_{\alpha,\epsilon}] = E_P[m_\alpha(p_0 + z_{\alpha,\epsilon})] = E_P[m_\alpha p] = E^\alpha[p] \quad (2)$$

or, equivalently, for all α and ϵ ,

$$E^\alpha[z_{\alpha,\epsilon}] = 0. \quad (3)$$

This condition means that since the prices of basic securities are market observable, the perturbed payoff has to be consistently priced by the same set of stochastic discount factors. Equation (3) makes clear that not every perturbation is admissible. By allowing for its dependence on α , we can accommodate a wider range of perturbations.

It may seem most natural to take p as the benchmark payoff and to consider perturbation around p . However, since p may be a misspecification of the actual payoff, there is no economic reason that it has to be the benchmark. To make economic sense, though, the benchmark has to be close to p . In fact, there is no particular reason that the benchmark payoff p_0 has to be fixed. It may (see the example in Section 2.2) depend on the market price of risk. Thus, we extend Equations (1) and (2) to

$$p = p_0(\alpha) + z_{\alpha,\epsilon} \quad (4)$$

for some $(\alpha, \epsilon) \in A \times \mathcal{E}$ and that for all $\alpha \in A$ and $\epsilon \in \mathcal{E}$,

$$E^\alpha[p_0(\alpha) + z_{\alpha,\epsilon}] = E^\alpha[p_0(\alpha)]. \quad (5)$$

We are interested in pricing a derivative security that is not spanned by the set of existing marketed securities. Suppose that the payoff of the derivative is $x = \phi(p)$. Since the $m_{\alpha,s}$ are not unique, neither is the price of the derivative. The focus of our study will be on how to choose a stochastic discount factor so that $E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon})]$ is robust, i.e., the least sensitive to ϵ .

There are two approaches to formalizing that notion of robustness. One is in terms of prices. That is, we will find the stochastic discount factor m_{α^*} , so that the pricing error band in the derivative security,

$$\{E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon_1})] - E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon_2})] : \epsilon_1, \epsilon_2 \in \mathcal{E}\}, \quad (6)$$

is the smallest. The price of the derivative security $\phi(p)$ is then $E^{\alpha^*}[\phi(p)]$.

The motivation of this approach is straightforward. Given a stochastic discount factor indexed by α , $E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_i})]$, $i = 1, 2$, is the price of the derivative security under model indexed by ϵ_i . The larger the difference of prices under the two models, the bigger the impact of model misspecification is on the price of the derivative security. The stochastic discount factor that makes the band in Equation (6) the smallest gives a price of the derivative security that is the least sensitive, and hence, robust to model misspecification.

While Equation (6) has clear intuitive appeal, it is often analytically more advantageous to evaluate the following band,

$$\{|E^\alpha[\phi(p_0(\alpha))] - E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon})]| : \epsilon \in \mathcal{E}\}. \quad (7)$$

Clearly, this band is smaller than Equation (6). However,

$$\begin{aligned} & |E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_1})] - E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_2})]| \\ & \leq |E^\alpha[\phi(p_0(\alpha))] - E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_1})]| + |E^\alpha[\phi(p_0(\alpha))] \\ & \quad - E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_2})]|. \end{aligned}$$

By a judicious choice of a benchmark model, the band (7) is not only easier to estimate than (6), but also a good approximation of the band (6). For this reason, we will focus on Equation (7) instead of Equation (6). Since the approach is based on difference in expected values, we call it the robustness in expectation approach.

The second approach is based on a band in probabilities,

$$\{P^\alpha(\{|E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_1})|\mathcal{F}] - E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon_2})|\mathcal{F}]| > \delta\}) : \epsilon_1, \epsilon_2 \in \mathcal{E}\} \quad (8)$$

for a given $\delta > 0$ and an information set $\mathcal{F}$. Here, δ is a measure of closeness. The role of $\mathcal{F}$ will be explained shortly. The objective is to find the stochastic discount factor m_{α^*} , such that the band is the smallest.

The intuition behind this approach is best explained in the context of option pricing in a world with stochastic volatility. Assume a setting as in Hull and White (1987), where the underlying stock price follows a geometric Brownian motion with a stochastic volatility. As Hull and White show, if the average variance over the sample path of the stochastic variance is known, then the option price depends only on the current stock price and the average variance, i.e., $C(S_0, \sqrt{V})$, where V is the average variance. In other words, $E^\alpha[\max\{S_T - K, 0\}|V] = C(S_0, \sqrt{V})$. Moreover, given a market price of risk indexed by α , the price of the option under stochastic volatility is given by $E^\alpha[\max\{S_T - K, 0\}] = E^\alpha[C(S_0, \sqrt{V})]$. Now suppose that the stochastic volatility process cannot be estimated precisely, so that the price of the option under any particular potentially misspecified stochastic volatility model is $E^\alpha[C(S_0, \sqrt{V(\epsilon)})]$. In this context, if we take $\mathcal{F}$ as the information on the average variance, then $E^\alpha[\phi(p_0(\alpha) + z_{\alpha, \epsilon})|\mathcal{F}]$ in Equation (8) is $C(S_0, \sqrt{V(\epsilon)})$, which is the price of the option under the stochastic discount factor indexed

by α and conditional knowing the average variance that comes from a potentially misspecified stochastic volatility model indexed by ϵ . Next, consider the probability in Equation (8), which in current context is

$$P^\alpha(\{|C(S_0, \sqrt{V(\epsilon_1)}) - C(S_0, \sqrt{V(\epsilon_2)})| > \delta\}), \quad (9)$$

where ϵ_1 and ϵ_2 are two arbitrary indices in $\mathcal{E}$. Intuitively, if this probability is large, then the option prices under the two stochastic volatility models are very different. Thus, if the probability band in Equation (8) is large, the option price is sensitive to potential model misspecification. Put differently, if a stochastic discount factor can be chosen, such that this probability band is the smallest, then under that stochastic discount factor, the option price is the least sensitive, and hence, robust to potential model misspecifications.

While the discussion above is in the context of call option pricing under a stochastic volatility model, the idea applies more generally. In fact, the price y of any contingent claim Y , under stochastic discount factor index by α , satisfies

$$E^\alpha[Y] = E^\alpha[E^\alpha[Y|\mathcal{F}]],$$

where $\mathcal{F}$ is an arbitrary information set, and $E^\alpha[Y|\mathcal{F}]$ is the price of the contingent claim conditional on the information. If the description of the payoff of the contingent claim Y is subject to potential model misspecification, then the difference in conditional price of the contingent claim can be used as a measure of the sensitivity of the price of the contingent claim to model misspecification. Equation (8) is a way of formalizing a measure of such sensitivity.

By exactly the same reason as for the robustness in expectation approach, it is often analytically advantageous to examine the band

$$\{P^\alpha(\{|E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon})|\mathcal{F}] - E^\alpha[\phi(p_0(\alpha))]| > \delta\}) : \epsilon \in \mathcal{E}\} \quad (10)$$

instead of Equation (8). The following observation provides the connection between Equations (8) and (10). For any two models indexed by ϵ_1 and ϵ_2 , respectively,

$$\begin{aligned} &P^\alpha(\{|E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon_1})|\mathcal{F}] - E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon_2})|\mathcal{F}]| > 2\delta\}) \\ &\leq P^\alpha(\{|E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon_1})|\mathcal{F}] - E^\alpha[\phi(p_0(\alpha))]| > \delta\}) \\ &+ P^\alpha(\{|E^\alpha[\phi(p_0(\alpha) + z_{\alpha,\epsilon_2})|\mathcal{F}] - E^\alpha[\phi(p_0(\alpha))]| > \delta\}). \end{aligned} \quad (11)$$

As in the robustness in expectation approach, the choice of the benchmark model is crucial to make the estimation of the band (10) analytically tractable and to make it a good approximation of Equation (8). As we will show, for stochastic volatility models, the Black-Scholes model is a good benchmark.⁶ Since the approach involves evaluating the probability of pricing differences, we call it the robustness in probability approach.

⁶ See Section 6.

Before moving on to the numerical examples, we note that the two approaches could be supplemented by other economic considerations. In practice, we would not have to consider all possible stochastic discount factors if we could remove some unreasonable ones. For instance, the good-deal approach developed by Bernardo and Ledoit (2000); Cochrane and Saá-Requejo (2000); and Ross (2005) suggests that some stochastic discount factors should be precluded, and the upper bound and lower bound using this restricted set can narrow the no-arbitrage bounds.

2.2 Numerical examples

We now discuss two numerical examples that demonstrate the main ideas behind the two approaches. To focus on the basic ideas, we have made these examples as simple as possible. Both examples are based on the same market structure. We have an incomplete market with a basic risky asset and a risk-free asset. We use a call option to illustrate the notion of robustness. The first example illustrates the notion of robustness under the first approach, while the second example does that under the second approach. It is convenient to use different benchmarks in the two examples.

There are four states at time one, denoted by $(\omega_1, \omega_2, \omega_3, \omega_4)$. The probabilities of these states are $(1/3, 1/3, 1/6, 1/6)$ under the physical probability measure P . The payoffs of the risky and risk-free assets at the end of the period are as follows:

State	Payoff on risky asset	Payoff on risk-free asset
ω_1	8	1
ω_2	6	1
ω_3	3	1
ω_4	3	1

The current price of the risky asset is 4. The current price of the risk-free asset is one, so that the risk-free rate is zero. This market is incomplete and the stochastic discount factor is not unique. It is straightforward to verify that the stochastic discount factors are given by

$$m_\alpha := \left[3 \left(\frac{3}{2}\alpha - 1 \right), 3 \left(2 - \frac{5}{2}\alpha \right), 3\alpha, 3\alpha \right],$$

where $\alpha \in (2/3, 4/5)$.

In this market, the no-arbitrage price of a general contingent claim with payoff vector $[x(\omega_1), x(\omega_2), x(\omega_3), x(\omega_4)]'$ belongs to the set

$$\{E^\alpha[x] : \alpha \in (2/3, 4/5)\}.$$

2.2.1 First example. Suppose now that we have a call option with a strike price of $K = 7$ written on the risky asset. The no-arbitrage price of the call, which is given by $\frac{3}{2}\alpha - 1$, falls in the set $(0, 1/5)$. If the agents in the economy

are not absolutely sure of the payoff vector of the risky asset, they may express their concern about the uncertainty by considering a family of possible payoff vectors given by

$$p(\alpha, \epsilon) \equiv p + z_{\alpha, \epsilon} = \left(8 + \left(\frac{5}{2}\alpha - 2 \right) \epsilon, 6 + \left(\frac{3}{2}\alpha - 1 \right) \epsilon, 3, 3 \right) \quad (12)$$

for $\epsilon \in [0, 1]$ with p as the benchmark payoff. In this case, the benchmark is of the type described in Equations (1)–(3). For any α , the pricing band of the call is given by

$$\left| E^\alpha [\max\{p(\alpha, \epsilon) - 7, 0\}] - \left(\frac{3}{2}\alpha - 1 \right) \right| \in \left[0, \left(\frac{3}{2}\alpha - 1 \right) \left(2 - \frac{5}{2}\alpha \right) \right].$$

Clearly, for α close to $2/3$ or $4/5$, the pricing band is the smallest, while it is the widest when $\alpha = 11/15$. In fact, at $\alpha \approx 4/5$, the price of the call is approximately $1/5$ regardless of ϵ . Similarly, at $\alpha \approx 2/3$, the price of the call is approximately 0 regardless of ϵ . In this sense, the price of the call, evaluated with stochastic discount factors m_α with α just above $2/3$ or just below $4/5$, is robust to model misspecification.

2.2.2 Second Example. We now turn to our second example. This example is of necessity a little more complex, since we want it to illustrate a more sophisticated notion: robustness in probability. To illustrate this second approach, we assume that the agents view the payoffs 8 and 6 in states ω_1 and ω_2 , respectively, as being very close so that the agents may view the payoff vector $p = (8, 6, 3, 3)$ as the result of a small perturbation of the benchmark payoff vector whose first and second components are identical, i.e., a perturbation of a binomial tree. Consider the family of perturbations of the benchmark payoff given by

$$p(\alpha, \epsilon) = \left(8 + \left(\frac{5\alpha - 4}{1 - \alpha} \right) (1 - \epsilon), 6 + \left(\frac{3\alpha - 2}{1 - \alpha} \right) (1 - \epsilon), 3, 3 \right) \quad (13)$$

for $\epsilon \in [0, 1]$. When $\epsilon = 1$, $p(\alpha, 1) = p$, which is the payoff of the risky asset. When $\epsilon = 0$, $p(\alpha, 0) = \left(\frac{4-3\alpha}{1-\alpha}, \frac{4-3\alpha}{1-\alpha}, 3, 3 \right)$, which has the same first and second components, required of the benchmark payoff vectors. Note that the benchmark in this case is a function of α , and thus, is of the type described in Equations (4) and (5). Under the benchmark payoffs, the price of the option is

$$C(\alpha, 0) \equiv E^\alpha [\max\{p(\alpha, 0) - K, 0\}] = 4 - 3\alpha - K(1 - \alpha),$$

for $3 < K \leq (4 - 3\alpha)/(1 - \alpha)$.

Let $\mathcal{F}$ denote the information set generated by the events $\{\omega_1, \omega_3\}$ and $\{\omega_2, \omega_4\}$. Let $C(\alpha, \epsilon)$ denote the option price conditional on the events. Then,

$$C(\alpha, \epsilon) = E^\alpha [\max\{p(\alpha, \epsilon) - K, 0\} | \mathcal{F}],$$

and the option price at time 0 is given by $C_0(\alpha, \epsilon) = E^\alpha [C(\alpha, \epsilon)]$.

As described earlier, the second approach to robustness is based on the following bands of probabilities,

$$\{P^\alpha(|C(\alpha, \epsilon) - C(\alpha, 0)| > \delta) : \epsilon \in \mathcal{E}\}. \quad (14)$$

indexed by α . If δ is small and if for some α , the probability in the expression is also the smallest for all $\epsilon \in \mathcal{E}$, then (a) $E^\alpha[C(\alpha, \epsilon)]$ will be very close to $C(\alpha, 0)$ and (b) at such α , the value of the option is robust to model misspecification. Note that $C(\alpha, 0)$ is essentially priced by a binomial model. In other words, it is priced in a complete market.

For the example with $K = 6$,

$$C(\alpha, 0) = 3\alpha - 2.$$

The conditional option prices are given by

$$C(\alpha, \epsilon) = (3\alpha - 2) + \left[\left(1 + (1 - \epsilon) \frac{5\alpha - 4}{2(1 - \alpha)} \right) \frac{1}{2\alpha - 1} - 1 \right] (3\alpha - 2)$$

in the event $\{\omega_1, \omega_3\}$, and

$$C(\alpha, \epsilon) = (3\alpha - 2) + \left[(1 - \epsilon) \frac{4 - 5\alpha}{4(1 - \alpha)^2} - 1 \right] (3\alpha - 2)$$

in the event $\{\omega_2, \omega_4\}$. Thus, the probability band in Equation (8) is the smallest when $\alpha \rightarrow 2/3$.

It should be stressed that the sole purpose of these two numerical examples is to introduce the two concepts of robustness and our focus is on simplicity rather than generality. We use different benchmarks in each case and this framework is not rich enough to support an integrated discussion of the two notions of robustness. However, when we discuss the continuous time application, we will be able to compare the two notions of robustness in a unified framework.

3. Stochastic Volatility Models

We now turn to the continuous time case. The approach is similar to that used in the previous section, but there are differences. Some of the differences arise from the more complex structure of continuous time models, which means that there are more technical issues to deal with. In this paper, we focus on stochastic volatility models since stochastic volatility models constitute an important class of incomplete market models and they are convenient for illustrating the approaches.

The layout of this section is as follows. We first discuss two classes of benchmark models. Then we describe a class of stochastic volatility models that are perturbations of the benchmark models. Finally, we discuss the connection between the stochastic volatility models and their corresponding benchmarks

and explain how the market price of risk plays a critical role here in establishing this connection.

3.1 Benchmark models

We assume that there is a single risky asset and a risk-free asset. The price dynamics of the risky asset follow the basic Black-Scholes assumptions,

$$\frac{dS(t)}{S(t)} = \mu dt + \sqrt{v(t)} dW_1(t), \quad S(0) = 1, \quad (15)$$

where W_1 is a one-dimensional Brownian motion. We also assume that risk-free rate is constant.

The square of the volatility, $v(t)$, in the description of the stock price will be the main focus of the rest of this section. Different specifications of it will lead to different models. It turns out that all the benchmark models and perturbed models differ only in the specification of the variance process, which we will now turn to.

3.1.1 Deterministic volatility model—Black-Scholes model. We now describe the first class of models we will use later as benchmarks. These correspond to the Black-Scholes model with deterministic volatility. We assume that the square of the volatility, $v(t)$, of the risky asset is a deterministic function of time and that it evolves toward some long run level,

$$dv(t) = (\kappa_1 - \kappa_2 v(t))dt, \quad v_0 = \sigma_0^2. \quad (16)$$

We assume $\kappa_1 \geq 0$ and $\kappa_2 \geq 0$, where κ_2 is the speed of reversion and $\bar{v} = \kappa_1/\kappa_2$ is the long-run target level. The solution for $v(t)$ is

$$v(t) = \begin{cases} \sigma_0^2 e^{-\kappa_2 t} + \frac{\kappa_1}{\kappa_2} (1 - e^{-\kappa_2 t}) & \text{if } \kappa_2 > 0 \\ \sigma_0^2 + \kappa_1 t & \text{if } \kappa_2 = 0, \end{cases}$$

and we note that $v(t)$ is always positive.

Note that as we vary κ_1 and κ_2 , we will obtain different models and we will use this fact to obtain a range of benchmark models.⁷ We will see later how each benchmark model can be associated with a different market price of risk.

3.1.2 Stochastic volatility model. We now move to stochastic volatility models. Everything is the same as in the preceding section except that the square of the volatility $v(t)$ is stochastic and obeys the following diffusion:

$$dv(t) = (\kappa_1 - \kappa_2 v(t))dt + \sigma dW_2(t), \quad v_0 = \sigma_0^2, \quad (17)$$

⁷ We will assume that the speed of reversion is positive later, even though all results are still hold from the technical point of view.

where W_2 is a one-dimensional Brownian motion that is independent of W_1 . Here, κ_1 , κ_2 , and $\sigma > 0$ are model parameters. The instantaneous variance is $|v(t)|$ and the instantaneous volatility is $\sqrt{|v(t)|}$. The random noise $W_2(t)$ makes the market incomplete because the two sources of risk cannot be hedged with the risky asset and the risk-free asset.

Before proceeding, a few comments about our specification of the process $v(t)$ may be helpful. In our model, $v(t)$ is mean-reverting. This differs from Stein and Stein (1991), where the instantaneous volatility is mean-reverting. Our volatility specification will simplify the comparison between the stochastic volatility model and the benchmark model. Also note that in the asset price dynamics, we used the absolute value of $v(t)$. This is because when $v(t)$ follows an Ornstein Uhlenbeck process, Equation (17) can produce negative values for $v(t)$. Note that we can write Equation (17) as

$$dv(t) = \kappa_2 \left(\frac{\kappa_1}{\kappa_2} - v(t) \right) dt + \sigma dW_2(t), \quad v_0 = \sigma_0^2.$$

For some purposes the mean-reversion speed, κ_2 , and the long-run mean-reversion target rate, $\bar{v} = \kappa_1/\kappa_2$, represent a more intuitive parametrization than the original pair, (κ_1, κ_2) .

The solution for $v(t)$ of Equation (17) is

$$v(t) = \sigma_0^2 e^{-\kappa_2 t} + \int_0^t \kappa_1 e^{-\kappa_2(t-s)} ds + \int_0^t \sigma e^{-\kappa_2(t-s)} dW_2(s),$$

which has a normal distribution with unconditional mean and unconditional variance given by

$$\begin{aligned} E[v(t)] &= \sigma_0^2 e^{-\kappa_2 t} + \int_0^t \kappa_1 e^{-\kappa_2(t-s)} ds, \\ \text{Var}[v(t)] &= \int_0^t \sigma^2 e^{-2\kappa_2(t-s)} ds. \end{aligned}$$

As the *volatility of v_t* , σ has a critical impact on the model. When $\sigma = 0$, model (17) reduces to the Black-Scholes model (16). If σ becomes small enough, this model will become closer to the Black-Scholes model. However, for any small positive σ , this model still has a stochastic volatility component, and the no-arbitrage bound (for a call option) is the Merton bound (Frey and Sin, 1999).⁸ Hence, making σ smaller does not eliminate this effect. On the other hand, if one fixes a stochastic discount factor, the smaller σ becomes, the smaller the difference between the price of the option in the stochastic volatility market and in the deterministic volatility market. When $\sigma \rightarrow 0$, the option price in the stochastic volatility model approaches the Black-Scholes value. Hence, making σ smaller makes the two markets closer in some sense. This behavior

⁸ Merton's bounds correspond to the options' intrinsic value and the underlying asset itself.

arises from the presence of the volatility of volatility. Our approach provides a helpful way of looking at this phenomenon.

In the stochastic volatility model, there are infinitely many stochastic discount factors. Equivalently, there are infinitely many choices of the market price of (volatility) risk. We assume a linear market price of risk,⁹

$$\lambda_{(a,b)} \equiv \lambda = a + bv(t), \quad (18)$$

where both a and b are real numbers. The corresponding stochastic discount factor is given by

$$\eta_T^{(a,b)} = \exp \left[- \int_0^T r_t dt - \int_0^T \frac{(\mu - r)^2}{v(t)} dt - \frac{1}{2} \int_0^T (a + bv(t))^2 dt - \int_0^T \frac{\mu - r}{\sqrt{v(t)}} dW_1(t) - \int_0^T (a + bv(t)) dW_2(t) \right]. \quad (19)$$

There exists another two-dimensional Brownian motion, $W^{(a,b)}(t) \equiv (W_1^{(a,b)}(t), W_2^{(a,b)}(t))$, such that the risky asset's dynamics under this measure satisfy the following stochastic differential equations:

$$\begin{aligned} \frac{dS(t)}{S(t)} &= r dt + \sqrt{|v(t)|} dW_1^{(a,b)}(t), \quad S(0) = 1, \\ dv(t) &= [\kappa_1 - \kappa_2 v(t) - \sigma(a + bv(t))] dt + \sigma dW_2^{(a,b)}(t). \end{aligned}$$

The relationship between the two Brownian motions is given by the Girsanov transformation

$$W_1^{(a,b)} = W_1(t) + \int_0^t \frac{\mu - r}{\sqrt{v(s)}} ds; \quad W_2^{(a,b)}(t) = W_2(t) + \int_0^t (a + bv(s)) ds.$$

In the following, we frame the discussion in terms of the market price of risk rather than the corresponding stochastic discount factor.

3.2 Perturbations to the benchmark models

We now describe the perturbed models. Fix a market price of risk. Consider the following family of stochastic volatility models indexed by (a, b) and ϵ , under the probability measure $P^{(a,b)}$,

$$\frac{dS(t)}{S(t)} = r dt + \sqrt{|v(t)|} dW_1^{(a,b)}(t), \quad S(0) = 1, \quad (20)$$

$$dv(t) = [\kappa_1 - \kappa_2 v(t) - \sigma(a + bv(t))] dt + \epsilon \sigma dW_2^{(a,b)}(t). \quad (21)$$

⁹ The choice of linear specification has analytical advantages in this setting. In fact, all linear-type market price of risk generates Merton's bounds; see Frey and Sin (1999).

It turns out that we can use this family of stochastic volatility models as perturbed versions of the two classes of benchmark models described in Sections 3.1.1 and 3.1.2 by suitable choice of the range of ϵ . For example, if $\epsilon \in (1 - \bar{\epsilon}, 1 + \bar{\epsilon})$ for some $\bar{\epsilon} > 0$, then the family can be viewed as a perturbed model of the stochastic volatility model in Section 3.1.2. Indeed, when $\epsilon = 1$, we get the stochastic volatility model in Section 3.1.2. If, however, $\epsilon \in (0, 1]$, then the family can be viewed as a perturbed model of the deterministic volatility model in Section 3.1.1.

It is useful to introduce

$$\kappa_1^{(a,b)} = \kappa_1 - \sigma a; \quad \kappa_2^{(a,b)} = \kappa_2 + \sigma b. \quad (22)$$

With this notation, the system (20)–(21) can be expressed in terms of $\kappa_1^{(a,b)}$ and $\kappa_2^{(a,b)}$ and (a, b) contributes only through $\kappa_1^{(a,b)}$ and $\kappa_2^{(a,b)}$. Since there is a one-to-one mapping between (a, b) and $(\kappa_1^{(a,b)}, \kappa_2^{(a,b)})$, for the rest of this paper, we will parameterize the system by $(\kappa_1^{(a,b)}, \kappa_2^{(a,b)}, \epsilon)$ instead of (a, b, ϵ) . Furthermore, we will suppress the superscript (a, b) except when clarity requires otherwise. Then the system (20)–(21) can be expressed as

$$\frac{dS(t)}{S(t)} = r dt + \sqrt{|v(t; \kappa_1, \kappa_2, \epsilon)|} dW_1^{(\kappa_1, \kappa_2)}(t), \quad S(0) = 1, \quad (23)$$

$$dv(t; \kappa_1, \kappa_2, \epsilon) = [\kappa_1 - \kappa_2 v(t; \kappa_1, \kappa_2, \epsilon)] dt + \epsilon \sigma dW_2^{(\kappa_1, \kappa_2)}(t). \quad (24)$$

4. Robustness in Expectation Approach

We have described the benchmark models and the perturbed models. We now present the robustness in expectation approach to robust derivative pricing.¹⁰ For illustrative purpose, we use a standard European call option as the security to be priced.

For reasons that will be explained in Section 5.1, we will focus on the deterministic volatility benchmark. Under the benchmark model, the stock price is given by

$$\frac{dS(t)}{S(t)} = r dt + \sqrt{v(t; \kappa_1, \kappa_2)} dW_1^{(\kappa_1, \kappa_2)}(t), \quad S(0) = 1, \quad (25)$$

$$dv(t; \kappa_1, \kappa_2) = [\kappa_1 - \kappa_2 v(t; \kappa_1, \kappa_2)] dt, \quad v(0; \kappa_1, \kappa_2) = \sigma_0^2, \quad (26)$$

where $\kappa_1 > 0$ and $\kappa_2 > 0$.

Since $v(t; \kappa_1, \kappa_2)$ is deterministic, we can use no-arbitrage arguments to obtain a *unique* price for any derivative written on the underlying asset. For

¹⁰ Because we make use of a Taylor series expansion of the difference in the two option prices to capture convergence speed, this approach is similar to an asymptotic approach that was first used in Hull and White (1987), and extended by Johnson and Sircar (2002) in the context of contingent claim pricing and hedging. However, in these papers, the market price of risk is either fixed or not specified.

example, we can obtain an explicit Black-Scholes type formula for the price of a standard call option. It is well known that the path integral of the instantaneous variance,¹¹

$$V(\kappa_1, \kappa_2, 0) = \int_0^T v(t; \kappa_1, \kappa_2) dt, \quad (27)$$

plays a key role in option pricing in this model. The price of a call option with strike price K and time to maturity T on the risky asset is

$$C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) = S(0) \Phi(d_1) - K e^{-rT} \Phi(d_2), \quad (28)$$

where

$$d_1 = \frac{\ln(S(0)/K e^{-rT}) + \frac{1}{2} V(\kappa_1, \kappa_2, 0)}{\sqrt{V(\kappa_1, \kappa_2, 0)}}, \quad d_2 \equiv d_1 - \sqrt{V(\kappa_1, \kappa_2, 0)}$$

and $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Under the stochastic volatility dynamics given by Equation (24), let

$$V(\kappa_1, \kappa_2, \varepsilon) = \int_0^T |v(t; \kappa_1, \kappa_2, \varepsilon)| dt. \quad (29)$$

The price of the call option is equal to¹²

$$E^{(\kappa_1, \kappa_2)} \left[C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) \right], \quad (30)$$

where $E^{(\kappa_1, \kappa_2)}$ is the expectation under the risk-neutral probability indexed by the market price of risk parameter (κ_1, κ_2) .

We approximate the difference

$$E^{(\kappa_1, \kappa_2)} \left[C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) \right] - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right), \quad (31)$$

by a second-order Taylor expansion

$$\begin{aligned} & E^{(\kappa_1, \kappa_2)} \left[C'_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} - \sqrt{V(\kappa_1, \kappa_2, 0)} \right) \right] \\ & + \frac{1}{2} E^{(\kappa_1, \kappa_2)} \left[C''_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} - \sqrt{V(\kappa_1, \kappa_2, 0)} \right)^2 \right], \end{aligned} \quad (32)$$

¹¹ We have taken mild liberties with the notation here to anticipate a third argument. Strictly speaking, we should define the integrated variance as $V(\kappa_1, \kappa_2)$.

¹² Hull and White (1987) derive the formula under the assumption that the variance follows a geometric Brownian motion. The formula, however, holds also under our assumption on the variance.

where C'_{BS} and C''_{BS} denote the first, and second-order, derivatives of function $C_{BS}(V)$. The appendix shows that when $\varepsilon\sigma$ is small,

$$E \left[\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} - \sqrt{V(\kappa_1, \kappa_2, 0)} \right] = \frac{\varepsilon\sigma}{2\sqrt{V(\kappa_1, \kappa_2, 0)}} E[y] \\ - \frac{\varepsilon^2\sigma^2}{8(V(\kappa_1, \kappa_2, 0))^{3/2}} E[y^2] + o(\varepsilon^2\sigma^2),$$

where

$$y = \int_0^T \frac{1 - e^{-\kappa_2(T-u)}}{\kappa_2} dW_2^{(\kappa_1, \kappa_2)}(u).$$

Since $E[y] = 0$, the second-order approximation (32) is approximately linear in $\varepsilon^2\sigma^2$.

In the robustness in expectation approach to robust pricing, the band of Equation (31) or (6) in single period case is to be minimized by choice of the market price of risk. As $\varepsilon\sigma$ is typically small, one can look instead at the second-order approximation of the band. The following proposition reports such an approximation.

Proposition 1. *Define*

$$F(\kappa_1, \kappa_2) = \lim_{\varepsilon\sigma \downarrow 0} \frac{E^{(\kappa_1, \kappa_2)} [C_{BS}(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)})] - C_{BS}(\sqrt{V(\kappa_1, \kappa_2, 0)})}{\varepsilon^2\sigma^2}. \quad (33)$$

Then

$$F(\kappa_1, \kappa_2) = \left| \frac{C''_{BS}(\sqrt{V(\kappa_1, \kappa_2, 0)})}{8V(\kappa_1, \kappa_2, 0)} - \frac{C'_{BS}(\sqrt{V(\kappa_1, \kappa_2, 0)})}{8(V(\kappa_1, \kappa_2, 0))^{3/2}} \right| \\ \times \int_0^T \left[\frac{1 - e^{-\kappa_2(T-u)}}{\kappa_2} \right]^2 du. \quad (34)$$

The second-order approximation of the band in Equation (31) is then $F(\kappa_1, \kappa_2)$ multiplied by the square of the length of $\sigma\mathcal{E}$.

The function $F(\kappa_1, \kappa_2)$, thus, captures the rate of convergence of the option price from the stochastic volatility models to the deterministic benchmark model. This function depends on the model parameters κ_1 and κ_2 . We will explore this dependence in more detail in Section 6. Here, we just trace the contours of this dependence. The integral on the right-hand side of Equation (34) depends only on κ_2 and the option's time to maturity. The higher the mean-reversion rate κ_2 is, the smaller is this integral component. On the other hand, the middle components that involve Black-Scholes pricing functions

have a complicated relationship with (κ_1, κ_2) . We will illustrate the properties of $F(\kappa_1, \kappa_2)$ numerically in Section 6.

5. Robustness in Probability Approach

In this section, we discuss the robustness in probability approach to robust derivative pricing. For a reason similar to the robustness in expectation case, it is difficult to employ a stochastic volatility model as a benchmark with the robustness in probability approach. In Section 5.1, we will provide a brief discussion of the difficulty. In this section, we will focus only on the case of a deterministic benchmark model.

To better link what is to come to Equation (10), we note that the price of the call option is

$$E^{(\kappa_1, \kappa_2)}[\max\{S_T - K, 0\}] = E^{(\kappa_1, \kappa_2)}\left[C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)}\right)\right],$$

and

$$C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)}\right) = E^{(\kappa_1, \kappa_2)}[\max\{S_T - K, 0\}|\mathcal{F}],$$

where $\mathcal{F}$ is the information of average variance, $V(\kappa_1, \kappa_2, \varepsilon)/T$, on the variance sample path.

With this link in the background, we will study the convergence in probability of the option price under stochastic volatility models to the price under the Black-Scholes model:

$$\lim_{\varepsilon \rightarrow 0} \text{Prob}^{(\kappa_1, \kappa_2)}\left(\left|C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)}\right) - C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, 0)}\right)\right| \geq \delta\right) = 0. \quad (35)$$

We will provide a bound on the probability in Equation (35). Once we know the bound, the bound for the pricing difference between any stochastic volatility models (indexed by ε_1 and ε_2 , respectively) can be estimated by the following inequality,

$$\begin{aligned} & \text{Prob}^{(\kappa_1, \kappa_2)}\left(\left|C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon_1)}\right) - C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon_2)}\right)\right| \geq \delta\right) \\ & \leq \text{Prob}^{(\kappa_1, \kappa_2)}\left(\left|C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon_1)}\right) - C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, 0)}\right)\right| \geq \delta/2\right) \\ & + \text{Prob}^{(\kappa_1, \kappa_2)}\left(\left|C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon_2)}\right) - C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, 0)}\right)\right| \geq \delta/2\right). \end{aligned}$$

The motivation for studying Equation (35) is given in Section 2. While the motivation is different from that for the robustness in expectation approach, we can still gain insight by examining how the two approaches differ. For that purpose, we divide the price difference

$$E^{(\kappa_1, \kappa_2)}\left[C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)}\right)\right] - C_{BS}\left(\sqrt{V(\kappa_1, \kappa_2, 0)}\right) \quad (36)$$

into two terms based on a partition of the sample paths $V(\kappa_1, \kappa_2, \varepsilon)$ for a given arbitrary real number $\delta > 0$. The two terms are

$$E^{(\kappa_1, \kappa_2)} \left[\left(C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \right) \times 1_{\{|V(\kappa_1, \kappa_2, \varepsilon) - V(\kappa_1, \kappa_2, 0)| < \delta\}} \right], \quad (37)$$

and

$$E^{(\kappa_1, \kappa_2)} \left[\left(C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \right) \times 1_{\{|V(\kappa_1, \kappa_2, \varepsilon) - V(\kappa_1, \kappa_2, 0)| \geq \delta\}} \right], \quad (38)$$

where 1_A is the indicator function for the set A . The first term, Equation (37), is based on the set of paths of $V(\kappa_1, \kappa_2, \varepsilon)$ such that $|V(\kappa_1, \kappa_2, \varepsilon) - V(\kappa_1, \kappa_2, 0)| < \delta$, while the second term, Equation (38), is based on the set of paths $V(\kappa_1, \kappa_2, \varepsilon)$ such that $|V(\kappa_1, \kappa_2, \varepsilon) - V(\kappa_1, \kappa_2, 0)| \geq \delta$. As the function $C_{BS}(V)$ is continuous in V , the first term, Equation (37), will be close to zero for small δ . The second term, Equation (38), is based on the set of scenarios for which $|V(\kappa_1, \kappa_2, \varepsilon) - V(\kappa_1, \kappa_2, 0)| \geq \delta$ and is hence called the large deviation term. While the convergence in Equation (35) implies

$$\lim_{\varepsilon \sigma \rightarrow 0} \text{Prob}^{(\kappa_1, \kappa_2)} (|V(\kappa_1, \kappa_2, \varepsilon) - V(\kappa_1, \kappa_2, 0)| \geq \delta) = 0,$$

it does not imply the convergence of Equation (38).¹³ However, the probability that

$$\left| C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \right| \geq \delta$$

can be estimated precisely, as shown in Proposition 2 below.

Proposition 2. *For all $\delta \in (0, \delta_0(\kappa_1, \kappa_2)]$ where $\delta_0(\kappa_1, \kappa_2) = \sqrt{V(\kappa_1, \kappa_2, 0)}$, and for all $0 < c < 1$, we have*

$$\begin{aligned} & \ln P^{(\kappa_1, \kappa_2)} \left[\left| C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \right| \geq \delta \right] \\ & \leq -\frac{c}{\varepsilon^2 \sigma^2} J(\kappa_1, \kappa_2, \delta), \end{aligned} \quad (39)$$

as $\varepsilon \sigma \downarrow 0$, where

$$J(\kappa_1, \kappa_2, \delta) = J_1(\kappa_2) \left(2\sqrt{V(\kappa_1, \kappa_2, 0)}\delta - \delta^2 \right)^2, \quad \delta \leq \delta_0(\kappa_1, \kappa_2), \quad (40)$$

$$J_1(x) = \frac{x^3(e^{2xT} - 1)}{(e^{\frac{x}{2}T} - e^{-\frac{x}{2}T})^4}. \quad (41)$$

¹³ The convergence in expectation in Equation (36) does not necessarily imply that in Equation (35) either.

An immediate implication of Proposition 2 is that the probability band

$$\left\{ P^{(\kappa_1, \kappa_2)} \left[\left| C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right) \right| \geq \delta \right] : \varepsilon \in \mathcal{E} \right\}$$

is bounded by $\exp \left(-\frac{c}{\varepsilon^2 \sigma^2} J(\kappa_1, \kappa_2, \delta) \right)$ at least for $\mathcal{E}$ that is small. Thus, in first-order approximation, $\exp \left(-\frac{c}{\varepsilon^2 \sigma^2} J(\kappa_1, \kappa_2, \delta) \right)$ provides the probability band we are interested in. Proposition 2 also implies that the probability of large deviation of $C_{BS}(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)})$ from $C_{BS}(\sqrt{V(\kappa_1, \kappa_2, 0)})$ decays to zero exponentially as $\varepsilon \sigma$ tends to zero.¹⁴ The $J(\kappa_1, \kappa_2, \delta)$ term represents the convergence speed for this case. Thus, at least to a first-order approximation, the robust market price of risk is the one that maximizes $J(\kappa_1, \kappa_2, \delta)$.

In addition to providing an approximate for the probability band, Proposition 2 can potentially be useful for risk management. For risk managers, scenarios where

$$\left| C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon_1)} \right) - C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon_2)} \right) \right| \geq \delta,$$

that is, scenarios in which the prices of the option under two alternative stochastic volatility models deviate significantly from each other, may be of particular interest. Proposition 2 provides a bound of the risk-neutral probability of these scenarios when the distribution of $V(\kappa_1, \kappa_2, \varepsilon)$ is given by Equation (29).

Proposition 2 is derived using the tools of large deviation theory.¹⁵ To describe the idea of large deviations, consider the case where a sequence of random variables converges to a constant when n is large. From the law of large numbers (LLN), we know that with probability approaching one, the sequence will stay in a small neighborhood of the constant. Because of randomness, large deviations from the constant are still possible even though it becomes more and more unlikely when n approaches infinity. The theory of large deviations is concerned with the likelihood of the unlikely deviations from the limit. Typically a large deviation theorem contains a statement of the form

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln P_n(C) \leq -I(C), \quad (42)$$

where C is a closed set. Here, $I(\cdot)$ is called the rate function because it describes the rate at which the probability $P_n(C)$ on the left-hand side of the above expression converges (exponentially) to zero. However, a general theorem of large deviation can only claim the existence of such a rate function and in general it is difficult to get a closed-form expression for $I(\cdot)$. Proposition 2 provides a closed-form expression for $I(\cdot)$, which makes it possible for us to

¹⁴ Since $C_{BS}(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)})$ converges to $C_{BS}(\sqrt{V(\kappa_1, \kappa_2, 0)})$ in probability, any difference of the two bigger than $\delta > 0$ is viewed as an rare event, as $\varepsilon \downarrow 0$, and hence, a large deviation.

¹⁵ We give a brief discussion of large deviations in Appendix A as well as some references.

numerically examine the probability band that is of interest to us and to choose the market price of risk that gives the robust option price.

In this connection, the use of a large deviation *principle* for bounding probability of (large) deviations, the constant c in Proposition 2 deserves a comment. As a large deviation principle typically takes a limit form, one cannot simply drop the limit in Equation (42) and obtain

$$P_n(C) \leq \exp(-nI(C)),$$

although inequalities of this kind are what we want. So we make use of the fact that the rate function $I(\cdot)$ is strictly positive and then obtain the inequality on the probability bound in Proposition 2 for small $\varepsilon\sigma$ by introducing a strictly positive constant c that is less than one.

We have couched the analysis in terms of a call option. To finish this subsection, we illustrate how our two approaches could be used for any general European style derivative as well. We use the deterministic benchmark models and the robustness in probability approach to illustrate the major points. The idea is the same for the robustness in expectation approach. Assume first that x is a general European-style derivative. Given a choice of market price of risk, the price of x in the class of misspecified models is written as an expectation in terms of the joint distribution of the path $\{v(t; \kappa_1, \kappa_2, \varepsilon)\}$ and the price path $\{S_t\}$. In the corresponding benchmark model, the price of x is expressed as an expectation of the path $\{S_t\}$, and by replacing $\{v(t; \kappa_1, \kappa_2, \varepsilon)\}$ by the path $\{v(t; \kappa_1, \kappa_2, 0)\}$. Therefore, the convergence in terms of probability can be reduced to the properties of the rate function on a set of paths $\{v(t; \kappa_1, \kappa_2, \varepsilon)\}$. As stated in Appendix A, for any *reasonable* set, the rate function can be obtained from large deviations techniques. Hence, our approach goes through for any European-style derivatives. For other derivatives such as barrier-style or American-style, the proposed approach should be combined with some further techniques in pricing these derivatives, and then reduced to a rate function problem.¹⁶

5.1 Discussion

In this section, we discuss briefly why it is difficult to use a stochastic benchmark model in the robustness in probability approach as well as in the robustness in expectation approach.

We follow the same notations as above. Assume that

$$\lim_{\varepsilon \rightarrow 1} E^{(\kappa_1, \kappa_2)} \left[C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)} \right) \right] = E^{(\kappa_1, \kappa_2)} \left[C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 1)} \right) \right],$$

¹⁶ For instance, given an American-style derivative x , we can use a recent result from Bank and Karouri (2004). Bank and Karouri (2004) showed that the American option price can be written as an expectation in terms of the distribution of state variables. Because of the generality of the large deviation principle technique, our approach also works for American-style options.

it is tempting to think that one can formulate a large deviation principle of the following form:

$$\limsup_{\varepsilon \rightarrow 1} (\varepsilon - 1)^2 \ln P^{(\kappa_1, \kappa_2)}(\{|C_{BS}(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)}) - E^{(\kappa_1, \kappa_2)}[C_{BS}(\sqrt{V(\kappa_1, \kappa_2, 1)})]| > \delta\}) \leq -I(\kappa_1, \kappa_2, \delta).$$

It is actually impossible to formulate such a large deviation principle. The reason is roughly as follows. Large deviation principle of the form we have used so far is based on the convergence of sample paths of $v(t)$ to a limit path. That is, as $\varepsilon \rightarrow 0$, the sample path of Equation (24) converges to a deterministic path. If we were to derive a large deviation principle of the form given above by using a similar argument, we would have to take the sample paths of Equation (17), or Equation (24) with $\varepsilon = 1$ as the limit. Because the sample paths of Equation (17) are random, there is not a single path to converge to, which makes it very difficult to estimate the probability of a large deviation by the sample paths of Equation (24) from those of Equation (17). Therefore, it would be very difficult to obtain a rate function.

For the robustness in expectation approach, the difficulty with a stochastic volatility benchmark comes from exactly the same source. For Proposition 1, the proof relies on a large deviation principle to deal with the situation where $S(0) = K e^{-rT}$. When the benchmark is a stochastic volatility model, that situation is difficult to deal with. Hence, we are not able to provide a first-order approximation of the error band.

It is for these reasons that our study has focused on deterministic benchmark only.

6. Numerical Analysis

In this section, we will use numerical examples to illustrate how to select market price of risk parameters in our framework. We use deterministic benchmark models to compare both approaches. As will be seen in Section 6.2, since the volatility of volatility is small, using deterministic models as a benchmark gives a tight pricing band. The fact that the volatility of the instantaneous variance is small motivates our decision to use deterministic benchmarks to study the misspecified stochastic volatility models. Recall that in Section 4, we introduced the function $F(\kappa_1, \kappa_2)$. This function was derived in Proposition 1 from the robustness in expectation approach. In Section 5, we introduced the function $J(\kappa_1, \kappa_2, \delta)$ based on the robustness of probability approach. These functions will now be used to examine how close the misspecified stochastic volatility model is to the benchmark model. Under each approach, the market price of risk with the smallest $F(\kappa_1, \kappa_2)$ or the largest $J(\kappa_1, \kappa_2, \delta)$, respectively, gives us a robust choice of the corresponding stochastic discount factor.

Table 1
Model Parameters

a	b	θ	$\bar{v}$	κ_2
0.800	4.000	0.100	0.04153	9.88904
0.800	4.000	0.200	0.04514	9.88909
0.800	4.000	0.300	0.05119	9.88916
0.800	8.000	0.100	0.01128	9.88864
0.800	8.000	0.200	0.01331	9.88866
0.800	8.000	0.300	0.01668	9.88871
0.800	12.000	0.100	0.00282	9.88852
0.800	12.000	0.200	0.00419	9.88854
0.800	12.000	0.300	0.00647	9.88857
1.600	4.000	0.100	0.11444	9.89000
1.600	4.000	0.200	0.11832	9.89005
1.600	4.000	0.300	0.12486	9.89012
1.600	8.000	0.100	0.04101	9.88904
1.600	8.000	0.200	0.04306	9.88906
1.600	8.000	0.300	0.04648	9.88910
1.600	12.000	0.100	0.01741	9.88872
1.600	12.000	0.200	0.01878	9.88874
1.600	12.000	0.300	0.02107	9.88877
2.400	4.000	0.100	0.22216	9.89141
2.400	4.000	0.200	0.22676	9.89146
2.400	4.000	0.300	0.23453	9.89154
2.400	8.000	0.100	0.08709	9.88965
2.400	8.000	0.200	0.08922	9.88967
2.400	8.000	0.300	0.09278	9.88971
2.400	12.000	0.100	0.04079	9.88903
2.400	12.000	0.200	0.04218	9.88905
2.400	12.000	0.300	0.04449	9.88908

This table reports the calibration of the parameters for Equation (17). The three columns on the left side of this table are parameters from Stein and Stein (1991). The two columns on the right side of this table are parameters in our model calibrated to the Stein and Stein estimates by moment matching. Since all the moments of a normally distributed random variable are determined by the mean and variance of the random variable, there are really only two parameters that we can identify by matching moments. So we fix σ (for instance, $\sigma = 1\%$) and match moments by varying $\bar{v}$ and κ_2 , where $\bar{v} = \kappa_1/\kappa_2$. In the process of calibrating the parameters, we observed that the sum of squares of the differences is very flat in κ_2 . It is essentially $\bar{v}$ that does all the adjustment to match the moments. The initial volatility $\sigma_0 = 0.2$ and the terminal time $T = 0.25$ are used in the calibration.

6.1 Model parameters

The ranges of the parameters in our numerical analysis are based on empirical studies of stochastic volatility in the literature. Stein and Stein (1991) assume that the instantaneous volatility σ_t satisfies

$$d\sigma_t = (\alpha - \beta\sigma_t)dt + \theta dW_2(t).$$

Based on the estimates in Stein (1989); Merville and Pieptea (1989); and Stein and Stein (1991), the parameter θ ranges from 0.15 to 0.30, α ranges from 0.8 to 5.2, and β ranges from 4 to 16. The initial volatility σ_0 ranges from 20 to 40%. We use this information to calibrate the parameters in our model by a simple moment matching method. Specifically, we match $E[v(T)^i]$ with $E[\sigma_T^{2i}]$, $i = 1$ and 2, respectively. We assume $T = 0.25$. Table 1 displays the range of our

parameters κ_1 and κ_2 that are consistent with the empirical parameters in the empirical studies cited above.

In calibration, we choose $\sigma = 1\%$ and used the moment matching conditions to calibrate $\bar{v}$ and κ_2 . The reason is that our stochastic variance $v(T)$ is normally distributed and that the moment matching conditions can only pin down two parameters. We experimented with other values of $\sigma \in (0, 1\%)$ and found that the calibrated $(\bar{v}, \kappa_2)$ were not very sensitive to σ . To further support our choice of σ , we compare it with the empirical findings in Heston, (1993) as well. In Heston, (1993), v_t satisfies the CIR process, and the diffusion term of the variance is $\theta\sqrt{v_t}$. According to Table 1 of Heston (1993), θ is around 10%, and the current volatility is 10%. Hence, the volatility of the instantaneous variance is approximately 1%.

There is little guidance from the literature on the range of plausible values for market prices of risk parameters (κ_1, κ_2) .¹⁷ Hull and White (1987) assumed that the market price of risk was zero. Frey and Sin, (1999) considered all possible values for the market prices of risk. In our numerical examples, we considered a reasonably wide range for the parameters.

6.2 Robustness in probability approach

In this subsection, we consider the robustness in probability approach. We first perform some transformations of the variables that make the presentation easier. Then we state a proposition that summarizes some useful comparative statics. Finally, we present and discuss some numerical results.

Recall that $J(\kappa_1, \kappa_2, \delta)$ measures the rate of convergence in probability (Proposition 2), given three parameters κ_1 , κ_2 , and δ . Thus, our numerical study will focus mostly on the function $J(\kappa_1, \kappa_2, \delta)$. It will be useful to introduce a parameter η and a function H ,

$$\eta = \frac{\delta}{H(\kappa_1, \kappa_2)}, \quad H(\kappa_1, \kappa_2) = C_{BS} \left(\sqrt{V(\kappa_1, \kappa_2, 0)} \right). \quad (43)$$

The parameter is used to express the percentage price difference. Formula (39) can then be rewritten as

$$\begin{aligned} P^{(\kappa_1, \kappa_2)} \left(\left| \frac{C_{BS}(\sqrt{V(\kappa_1, \kappa_2, \epsilon)}) - H(\kappa_1, \kappa_2)}{H(\kappa_1, \kappa_2)} \right| \geq \eta \right) \\ \leq \exp \left\{ - \frac{c J(\kappa_1, \kappa_2, \eta H(\kappa_1, \kappa_2))}{\sigma^2 \epsilon^2} \right\}, \end{aligned} \quad (44)$$

¹⁷ We have called (κ_1, κ_2) the market price of risk parameter. In fact, the market price of risk parameter is (a, b) . However, there is a one-to-one correspondence between (κ_1, κ_2) and (a, b) given by Equation (22). So for simplicity of terminology, in the rest of the paper, we will call (κ_1, κ_2) , and equivalently $(\bar{v}, \kappa_2)$, the market price of risk parameter.

for all $0 < c < 1$ and all $\eta \leq \eta_0$, where

$$\eta_0 = \frac{\delta_0}{H(\kappa_1, \kappa_2)}$$

and δ_0 is defined in Proposition 2. In the large deviation literature, $\frac{1}{\sigma^2 \varepsilon^2} J(\kappa_1, \kappa_2, \eta H(\kappa_1, \kappa_2))$ is called the rate function because it characterizes the rate at which the probability on the left-hand side of Equation (44) converges to zero. With an abuse of terminology, we call J the rate function. The first component of the exponent on the right-hand side of formula (44) is $\frac{1}{\sigma^2 \varepsilon^2}$. Holding other parameters constant, the smaller σ , the larger the rate of convergence. This makes intuitive sense. Moreover, it appears as a product with ε in formula (44), and $\varepsilon \sigma$ represents the volatility of the variance in the misspecified stochastic volatility models. As the variance term $(\varepsilon \sigma)^2$ decreases, the exponent on the right-hand side of Equation (44) increases and the probability bound decreases. This means that the price in the misspecified stochastic volatility model becomes closer to that in the benchmark model.

The following proposition records some properties of the rate function J , which will be useful for our numerical analysis and for potential applications.

Proposition 3.

- (1) $J(\kappa_1, \kappa_2, \eta H(\kappa_1, \kappa_2))$ is increasing in η over the range $(0, \eta_0(\kappa_1, \kappa_2))$.
- (2) $J(\kappa_1, \kappa_2, \eta H(\kappa_1, \kappa_2))$ is a decreasing function of the strike price K .
- (3) For $\eta > 0$ and $\kappa_2 > 0$, $J(\bar{v}\kappa_2, \kappa_2, \eta H(\bar{v}\kappa_2, \kappa_2))$ is increasing in $\bar{v}$.
- (4) For $\eta > 0$ and $\bar{v} > \sigma_0^2$, $J(\bar{v}\kappa_2, \kappa_2, \eta H(\bar{v}\kappa_2, \kappa_2))$ is increasing in κ_2 .

Proposition 3(1) is intuitive. As η increases, the probability of a deviation in the call option price greater than $\eta H(\kappa_1, \kappa_2)$ becomes smaller. Proposition 3(2) is evident from the expression of the function $J(\kappa_1, \kappa_2, \delta_0)$. It is well known that $H(\kappa_1, \kappa_2)$ is decreasing in K . Proposition 3(3) is more interesting. It says that, other things being equal, the higher the long-run mean-reversion target $\bar{v}$ of the stochastic process $v(t)$, the larger the rate function J , and hence, the smaller the probability of large deviation from the Black-Scholes call option price. On the other hand, Proposition 3(4) says that holding the long-run mean-reversion target $\bar{v}$ constant, the greater κ_2 , the smaller the probability of large deviations from the Black-Scholes call option price. This is because the larger κ_2 is, other things being equal, the greater the speed of mean reversion in $v(t; \bar{v}\kappa_2, \kappa_2, \varepsilon)$. The less the stochastic variance deviates from its long-run target, the closer the average variance is to its deterministic counterpart, and hence, the closer the option values are.

Propositions 3(3) and 3(4) together imply that the robust market price of risk parameter $(\bar{v}^*, \kappa_2^*)$ must lie on the boundary of the set of market price of risk parameters, as illustrated in Figure 1. In the left graph, the set of market price of risk parameters is the area inside and on the boundary of the square box.

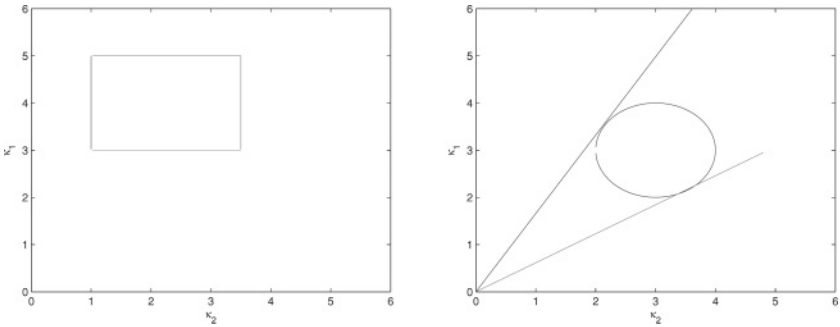


Figure 1
Sets of market price of risk parameters. In the left graph, the set of market price of risk parameters is the area inside, including the square box. In the right graph, it is the area inside, including the circle.

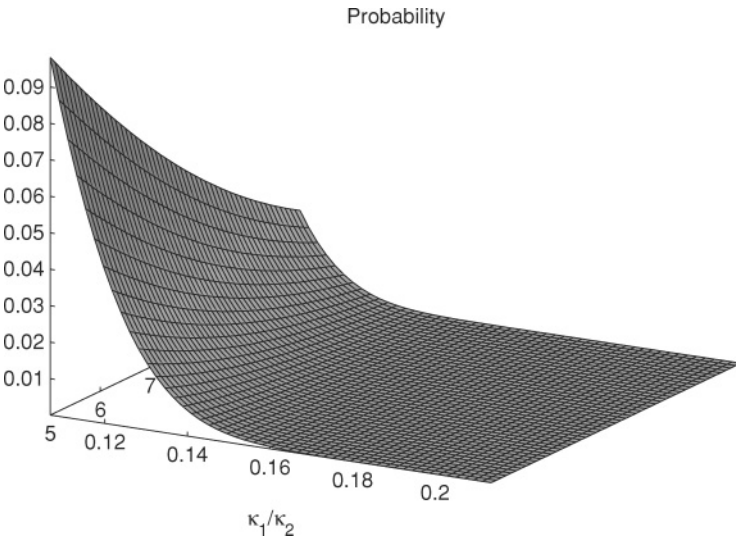


Figure 2
Sensitivity of the probability bound in $\bar{v}$ and κ_2 . This three-dimensional figure shows the sensitivity of the probability bound, $\exp(-\frac{1}{\sigma^2} J(\bar{v}\kappa_2, \kappa_2, \eta H(\bar{v}\kappa_2, \kappa_2)))$, in the long-run mean-reversion target, $\bar{v} = \kappa_1/\kappa_2$, of $v(t; \kappa_1, \kappa_2, 1)$, and the speed of the mean-reversion parameter, κ_2 . The values of the parameters are: $\eta = 5\%$, $T = 3$ months, $\sigma_0 = 0.2$, $\sigma = 0.01$, $r = 3\%$, $S(0) = K = 10$.

The robust market price of risk parameter lies somewhere on the top boundary of the box. In the right graph, the set of market price of risk parameters is the area inside, and on the boundary of the circle. The robust market price of risk parameter lies somewhere on the north east boundary between the two straight lines.

The next set of figures and tables provides numerical illustrations of the magnitude of J and its sensitivity to its underlying parameters. The vertical axis of Figure 2 is the probability bound, and the two horizontal axes are the speed of mean-reversion parameter κ_2 and the long-run mean-reversion

Table 2
Sensitivity of probability bound in η

	1%	2%	3%	4%	η 5%	6%	7%	8%	9%	10%
$(\bar{v}, \kappa_2) = (0.10, 2)$	0.97	0.90	0.78	0.65	0.51	0.38	0.27	0.19	0.12	0.07
$(\bar{v}, \kappa_2) = (0.10, 4)$	0.94	0.78	0.58	0.38	0.22	0.12	0.05	0.02	0.01	0.00
$(\bar{v}, \kappa_2) = (0.10, 6)$	0.89	0.62	0.34	0.15	0.05	0.01	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.10, 8)$	0.80	0.41	0.14	0.03	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.10, 10)$	0.68	0.22	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.20, 2)$	0.95	0.81	0.63	0.44	0.28	0.16	0.08	0.04	0.02	0.01
$(\bar{v}, \kappa_2) = (0.20, 4)$	0.86	0.55	0.26	0.09	0.03	0.01	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.20, 6)$	0.71	0.26	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.20, 8)$	0.51	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.20, 10)$	0.30	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.30, 2)$	0.92	0.71	0.47	0.26	0.12	0.05	0.02	0.01	0.00	0.00
$(\bar{v}, \kappa_2) = (0.30, 4)$	0.76	0.33	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.30, 6)$	0.52	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.30, 8)$	0.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.30, 10)$	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.40, 2)$	0.88	0.61	0.33	0.14	0.05	0.01	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.40, 4)$	0.64	0.17	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.40, 6)$	0.34	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.40, 8)$	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.40, 10)$	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.50, 2)$	0.84	0.50	0.21	0.06	0.01	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.50, 4)$	0.52	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.50, 6)$	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.50, 8)$	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$(\bar{v}, \kappa_2) = (0.50, 10)$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

This table reports the sensitivity of the probability bound, $\exp\{-J(\kappa_1, \kappa_2, \delta)/\varepsilon^2\sigma^2\}$ in Proposition 2 as η is varied. The parameters are $T = 0.25$, $\sigma = 1\%$, $r = 3\%$, $\sigma_0 = 0.2$, $S(0) = K = 1$.

target $\bar{v} = \kappa_1/\kappa_2$. The figure shows that the bound of the probability of large deviation is monotone in $\bar{v}$ and κ_2 , respectively, as described in Proposition 3. Furthermore, the bound is quite sensitive to $\bar{v}$ and κ_2 . In other words, the figure indicates that the bound on the probability of large deviations is sensitive to the market price of risk.

Table 2 illustrates the impact of the percentage error parameter, η (see Equation (43)), on the probability bound. Each row in the table corresponds to a market price of risk parameter $(\bar{v}, \kappa_2)$. Each column corresponds to a percentage error level. For example, when $(\bar{v}, \kappa_2) = (0.20, 6)$ and $\eta = 3\%$, the bound of the probability of a large deviation exceeding 3% is 0.05. The table suggests that the bound is decreasing in η (Proposition 3). This is intuitive because the larger the η , the less likely the event of large deviation exceeding η . The table also suggests that the bound is very sensitive to η . For example, for $(\bar{v}, \kappa_2) = (0.20, 6)$, the bound at 1% level is 0.71, while at 2%, it is 0.26, a 37% drop.

Table 3 illustrates the impact of the moneyness of the call option on the probability bound. Each row in the table corresponds to a market price of risk parameter $(\bar{v}, \kappa_2)$. Each column corresponds to a level of moneyness. The central column is for the at-the-money option. The columns to the left are for out-of-the-money options, while the columns to the right are for in-the-

Table 3
Sensitivity of probability bound in moneyness

	0.90	0.925	0.95	0.975	$S(0)/K$		1.025	1.05	1.0725	1.10
					1.00					
$(\bar{v}, \kappa_2) = (0.10, 2)$	0.998	0.996	0.991	0.984	0.973	0.957	0.936	0.910	0.878	
$(\bar{v}, \kappa_2) = (0.10, 4)$	0.994	0.989	0.979	0.964	0.941	0.910	0.871	0.823	0.768	
$(\bar{v}, \kappa_2) = (0.10, 6)$	0.987	0.975	0.956	0.927	0.886	0.833	0.767	0.692	0.611	
$(\bar{v}, \kappa_2) = (0.10, 8)$	0.974	0.952	0.918	0.868	0.801	0.718	0.623	0.522	0.421	
$(\bar{v}, \kappa_2) = (0.10, 10)$	0.954	0.916	0.860	0.781	0.683	0.569	0.450	0.336	0.236	
$(\bar{v}, \kappa_2) = (0.20, 2)$	0.994	0.989	0.980	0.967	0.949	0.925	0.895	0.858	0.816	
$(\bar{v}, \kappa_2) = (0.20, 4)$	0.976	0.959	0.936	0.903	0.861	0.809	0.749	0.681	0.609	
$(\bar{v}, \kappa_2) = (0.20, 6)$	0.936	0.900	0.851	0.788	0.713	0.628	0.538	0.446	0.359	
$(\bar{v}, \kappa_2) = (0.20, 8)$	0.865	0.799	0.716	0.618	0.512	0.405	0.305	0.217	0.147	
$(\bar{v}, \kappa_2) = (0.20, 10)$	0.756	0.653	0.537	0.414	0.299	0.199	0.122	0.069	0.036	
$(\bar{v}, \kappa_2) = (0.30, 2)$	0.987	0.977	0.963	0.944	0.919	0.886	0.847	0.801	0.751	
$(\bar{v}, \kappa_2) = (0.30, 4)$	0.940	0.909	0.869	0.819	0.758	0.690	0.615	0.537	0.459	
$(\bar{v}, \kappa_2) = (0.30, 6)$	0.840	0.775	0.697	0.610	0.516	0.422	0.332	0.251	0.183	
$(\bar{v}, \kappa_2) = (0.30, 8)$	0.675	0.573	0.464	0.357	0.259	0.177	0.113	0.067	0.038	
$(\bar{v}, \kappa_2) = (0.30, 10)$	0.463	0.342	0.234	0.146	0.083	0.042	0.019	0.008	0.003	
$(\bar{v}, \kappa_2) = (0.40, 2)$	0.976	0.962	0.942	0.915	0.882	0.841	0.794	0.741	0.683	
$(\bar{v}, \kappa_2) = (0.40, 4)$	0.886	0.840	0.783	0.717	0.643	0.564	0.483	0.403	0.329	
$(\bar{v}, \kappa_2) = (0.40, 6)$	0.706	0.619	0.525	0.428	0.336	0.252	0.181	0.124	0.081	
$(\bar{v}, \kappa_2) = (0.40, 8)$	0.455	0.347	0.248	0.166	0.104	0.060	0.032	0.016	0.007	
$(\bar{v}, \kappa_2) = (0.40, 10)$	0.212	0.128	0.070	0.034	0.015	0.006	0.002	0.001	0.000	
$(\bar{v}, \kappa_2) = (0.50, 2)$	0.961	0.941	0.915	0.881	0.840	0.792	0.737	0.678	0.615	
$(\bar{v}, \kappa_2) = (0.50, 4)$	0.817	0.755	0.685	0.607	0.525	0.442	0.363	0.289	0.224	
$(\bar{v}, \kappa_2) = (0.50, 6)$	0.555	0.457	0.362	0.273	0.197	0.135	0.088	0.054	0.032	
$(\bar{v}, \kappa_2) = (0.50, 8)$	0.261	0.176	0.109	0.063	0.033	0.016	0.007	0.003	0.001	
$(\bar{v}, \kappa_2) = (0.50, 10)$	0.071	0.034	0.014	0.005	0.002	0.000	0.000	0.000	0.000	

This table reports the sensitivity of the probability bound, $\exp\{-J(\kappa_1, \kappa_2, \delta)/\varepsilon^2\sigma^2\}$, in Proposition 2 as the moneyness of the option is varied. The parameters are $T = 0.25, \sigma = 1\%, \sigma_0 = 0.2, r = 3\%, S(0) = 1, \eta = 1\%$.

money options. The table suggests that the probability bound is also sensitive to the moneyness of the option. The bound is tighter when the option is in-the-money than when the option is out-of-the-money. The intuition is that when an option becomes more in-the-money it behaves more like the underlying asset. Hence, there is less mispricing, and consequently, the probability of a deviation becomes smaller. Indeed as the strike price K goes to zero, the option becomes the risky asset. The difference of the two option prices in Equation (44) is zero.

6.3 Robustness in expectation approach

In this subsection, we consider the robustness in expectation approach. We summarize our numerical results in a single table.

Assuming that $\epsilon \in [0, 1]$, we examine the pricing error band, $\sigma^2 F(\kappa_1, \kappa_2)$, given in Proposition 1. As it is less meaningful to look at the absolute pricing error band, we divide $\sigma^2 F(\kappa_1, \kappa_2)$ by the option price under the benchmark model, $H(\kappa_1, \kappa_2)$. This provides a percentage pricing error bands.

Table 4 reports an array of percentage pricing error band for varying market price of risk parameters and for varying degrees of moneyness. As the error is often very small, the numbers in the table are $F(\kappa_1, \kappa_2)/H(\kappa_1, \kappa_2)$ multiplied by 1000. For example, for $(\bar{v}, \kappa_2) = (0.05, 1.00)$ and $S(0)/K = 0.9$, $F(\kappa_1, \kappa_2)/H(\kappa_1, \kappa_2) = 1.0646\%$.

Table 4
Pricing error bound

	0.90	0.925	0.95	0.975	$S(0)/K$ 1.00	1.025	1.05	1.0725	1.10
$(\bar{v}, \kappa_2) = (0.05, 1.0)$	10.646	5.826	3.624	2.543	1.863	1.295	0.774	0.318	0.040
$(\bar{v}, \kappa_2) = (0.05, 1.5)$	9.374	5.158	3.223	2.268	1.664	1.161	0.698	0.292	0.029
$(\bar{v}, \kappa_2) = (0.05, 2.0)$	8.299	4.589	2.879	2.032	1.493	1.044	0.631	0.268	0.020
$(\bar{v}, \kappa_2) = (0.05, 2.5)$	7.386	4.102	2.584	1.827	1.345	0.943	0.572	0.246	0.013
$(\bar{v}, \kappa_2) = (0.05, 3.0)$	6.604	3.683	2.327	1.650	1.216	0.854	0.521	0.227	0.007
$(\bar{v}, \kappa_2) = (0.10, 1.0)$	7.083	4.098	2.672	1.930	1.442	1.032	0.653	0.314	0.038
$(\bar{v}, \kappa_2) = (0.10, 1.5)$	5.351	3.177	2.114	1.547	1.167	0.845	0.548	0.280	0.058
$(\bar{v}, \kappa_2) = (0.10, 2.0)$	4.163	2.526	1.710	1.264	0.961	0.704	0.465	0.249	0.067
$(\bar{v}, \kappa_2) = (0.10, 2.5)$	3.318	2.050	1.408	1.050	0.804	0.594	0.399	0.222	0.071
$(\bar{v}, \kappa_2) = (0.10, 3.0)$	2.697	1.693	1.177	0.885	0.681	0.507	0.345	0.197	0.070
$(\bar{v}, \kappa_2) = (0.15, 1.0)$	4.991	3.025	2.047	1.513	1.150	0.841	0.556	0.297	0.080
$(\bar{v}, \kappa_2) = (0.15, 1.5)$	3.392	2.136	1.489	1.121	0.863	0.643	0.439	0.253	0.092
$(\bar{v}, \kappa_2) = (0.15, 2.0)$	2.435	1.580	1.127	0.860	0.670	0.507	0.355	0.215	0.092
$(\bar{v}, \kappa_2) = (0.15, 2.5)$	1.822	1.212	0.880	0.680	0.534	0.409	0.292	0.183	0.088
$(\bar{v}, \kappa_2) = (0.15, 3.0)$	1.408	0.956	0.704	0.549	0.435	0.336	0.244	0.158	0.081
$(\bar{v}, \kappa_2) = (0.20, 1.0)$	3.676	2.317	1.616	1.217	0.938	0.699	0.478	0.275	0.101
$(\bar{v}, \kappa_2) = (0.20, 1.5)$	2.314	1.527	1.102	0.848	0.665	0.507	0.359	0.223	0.103
$(\bar{v}, \kappa_2) = (0.20, 2.0)$	1.574	1.075	0.796	0.622	0.494	0.383	0.279	0.182	0.096
$(\bar{v}, \kappa_2) = (0.20, 2.5)$	1.132	0.794	0.600	0.475	0.381	0.299	0.222	0.151	0.086
$(\bar{v}, \kappa_2) = (0.20, 3.0)$	0.850	0.609	0.467	0.373	0.302	0.239	0.181	0.126	0.076
$(\bar{v}, \kappa_2) = (0.25, 1.0)$	2.804	1.827	1.307	1.000	0.780	0.591	0.415	0.253	0.111
$(\bar{v}, \kappa_2) = (0.25, 1.5)$	1.667	1.142	0.847	0.664	0.528	0.409	0.299	0.196	0.104
$(\bar{v}, \kappa_2) = (0.25, 2.0)$	1.092	0.776	0.591	0.471	0.379	0.300	0.225	0.155	0.091
$(\bar{v}, \kappa_2) = (0.25, 2.5)$	0.765	0.558	0.434	0.350	0.285	0.228	0.175	0.125	0.078
$(\bar{v}, \kappa_2) = (0.25, 3.0)$	0.563	0.420	0.331	0.270	0.222	0.179	0.140	0.102	0.067

This table reports the percentage pricing error bound, $\sigma^2 F(\kappa_1, \kappa_2)/H(\kappa_1, \kappa_2)$, in Proposition 1 multiplied by 1000. The model parameters are $\sigma = 1\%$, $\sigma_0 = 0.2$, $r = 3\%$, $T = 0.25$.

Table 4 displays some interesting patterns. First, looking across each row, we see that for each market price of risk, the percentage pricing error band increases as the option moves out of money. The intuition here is the same as in Table 3. As the option gets more in the money, the option resembles more the underlying stock, and hence, the pricing error becomes smaller. Second, looking down each column, the numbers are not monotone. For example, looking down the last column, for the subset of market price of risk parameters where $\bar{v} = 0.05$, the percentage pricing error band decreases as κ_2 increases. For the subset of market price of risk where $\bar{v} = 0.15$ and $S(0)/K = 1.10$, as κ_2 increases, the percentage pricing error band first increases and then decreases. Third, not only is there no monotonicity in κ_2 , there is no monotonicity in $\bar{v}$ either. Again, looking down the last column, holding $\kappa_2 = 1$, as $\bar{v}$ increases, the percentage pricing error band first goes down and then up.

The pattern that deserves the most attention is perhaps how small the pricing error bands are, as exhibited in Table 4. In the table, we have reported the numbers for a small set of market price of risk parameters. In the calculations that are not reported here, we looked at a much wider range of market price of risk parameters. The percentage pricing error bands are all less than 1%. So far, the calculation is done for $\sigma = 1\%$. For σ that is significantly larger than

1%, the error band gets large, and the large band typically occurs for options that are significantly out-of-the-money.

We have now applied our two approaches to option pricing in a stochastic volatility setting and we have seen the type of results they deliver. The robustness in expectation approach produces a narrow error band. However, we observed that there did not seem to be any clear pattern of behavior with regard to the dependence of the function $F(\kappa_1, \kappa_2)$ on its arguments. Thus makes it a bit difficult to find the robust stochastic discount factor. On the other hand, the robustness in probability approach tends to produce a wider error band. However, in this case, we can obtain a general result for the dependence of our robustness metric on the market price of risk. Specifically, Proposition 3 gives very clear comparative statics regarding the dependence of the function J on its various arguments and also on the strike price of the call option. Indeed, this proposition enables us to show that the most robust market prices of risk will lie on the boundary of the feasible set. Overall, the numerical analysis in this section gives us a quantitative sense of how big the error band under each of the approaches can be and how sensitive the error band is to market price of risk, which we think is of ultimate importance from the standpoint of applications.

7. Conclusions

Derivative security pricing in an incomplete market is an important problem, particularly when agents face model misspecification risk. While the literature has offered some guidance as to how the market price of risk can be selected, it is not completely satisfactory, especially when there is potential model misspecification. In this paper, we develop a framework that enables one to select a market price of risk, such that the price of a new derivative security is robust to potential model misspecification. In the context of a stochastic volatility model, we derive an analytic expression for the bound of the probability that the price of the derivative security under a stochastic volatility model, which may be misspecified, deviates significantly from the price of the derivative under a benchmark model. We also derive the expression that bounds the pricing difference of the derivative security under alternative models. These results are then applied to a calibrated setting to study the bounds numerically. The study suggests that the bounds can be made tighter by an appropriate choice of the market price of risk. Therefore, the price of the derivative security can be made robust to potential model misspecification by an appropriate choice of the market price of risk.

Appendix A: Large Deviation Principle

Large deviation principle is closely related to the strong law of large numbers and the central limit theorem. It is in some respect a stronger statement. Consider an *iid* process X_n with mean μ and unit variance. Let $S_n = \frac{1}{n} \sum_{i=1}^n X_i$

be its sample mean. Then the strong law of large numbers states that S_n converges to μ almost surely. The central limit theorem states that the distribution of $\sqrt{n}(S_n - \mu)$ converges to the standard normal distribution. Both of these standard theorems make a statement about the fact that $S_n - \mu$ converges to zero. However, none of them make an explicit statement about the rate at which $S_n - \mu$ converges to zero. In contrast, large deviation principles are statements about the rate of convergence. For example, one version of the large deviation principle for the *iid* process is that there exists a rate function $I(\delta)$ such that, for any $\delta > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \text{Prob}(|S_n - \mu| \geq \delta) = -I(\delta).$$

The large deviation principle makes two statements about the convergence of $S_n - \mu$. First, any large deviation of S_n from μ has zero probability in the limit. Note that since $S_n - \mu$ converges to zero, $|S_n - \mu| > \delta$ for any $\delta > 0$ is viewed as a large deviation in the limit. Second, the probability of a large deviation converges to zero at the speed given by the exponential on the right-hand side of the above expression. The first statement, similar to the strong law of large numbers and the central limit theorem, is a statement about the convergence of $S_n - \mu$ to zero. The second statement, however, provides the additional information that neither the strong law of large numbers nor the central limit theorem provides. It is this additional information about the speed of convergence that makes the large deviation principle most useful for our purpose of assessing derivative pricing error.

The version of the large deviation principle for our purpose is as follows. The family of laws Q^ε of $v(t, \varepsilon)$ is said to satisfy the large deviation principle with rate functional I if

- (i) for any closed set $F \subset C([0, T], \mathcal{R}^+)$,

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^2 \log Q^\varepsilon(v(t, \varepsilon) \in F) \leq - \inf_{\phi \in F} I(\phi),$$

- (ii) for any open set $G \subset C([0, T], \mathcal{R}^+)$,

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon^2 \log Q^\varepsilon(v(t, \varepsilon) \in G) \geq - \inf_{\phi \in G} I(\phi),$$

- (iii) for any constant $c \geq 0$, the set $\{\phi \in C([0, T], \mathcal{R}^+) : I(\phi) \leq c\}$ is compact.

The rate functional I is usually given in variational form. It can be expressed in an explicit integral form for some models.

As candidates for stochastic volatility models, consider the following family of stochastic processes indexed by $\varepsilon \in [0, 1]$,

$$dv(t) = a(v(t)) + \varepsilon b(v(t)) dW_2(t), \quad v_0 = \sigma_0^2 > 0,$$

where $a(v)$ is a linear function and $b(v)$ belongs to any of the following three subclasses:

- (i) $b(v) = \alpha + \beta v$, where $\alpha, \beta \in R^+ \cup \{0\}$;
- (ii) $b(v) = \alpha\sqrt{v}$, where $\alpha \in R^+$;
- (iii) $b(v) = \alpha\sqrt{v(\beta - v)}$, where $\alpha, \beta \in R^+$.

For subclass (i), the Freidlin-Wintzell theory can be used to get the large deviation principle. For subclasses (ii) and (iii), the large deviation principles are proved in Dawson and Feng (1998); Dawson and Feng (2001); and Feng and Xiong (2002). For an accessible exposure on large deviation principle, we refer the interested reader to Dembo and Zeitouni (2003).

For the stochastic volatility models considered in this paper, $v(t; \kappa_1, \kappa_2, \varepsilon)$ solves the following SDE:

$$dv(t; \kappa_1, \kappa_2, \varepsilon) = (\kappa_1 - \kappa_2 v(t; \kappa_1, \kappa_2, \varepsilon))dt + \varepsilon \sigma dW_2(t).$$

The large deviation principle for $v(t; \kappa_1, \kappa_2, \varepsilon)$ follows directly from the Freidlin-Wentzell theory, with rate functional given by

$$I(\phi) = \begin{cases} \frac{1}{2\sigma^2} \int_0^T (\phi(t) - \kappa_1 + \kappa_2 \phi(t))^2 dt, & \phi \in \mathcal{H} \\ \infty, & \phi \notin \mathcal{H} \end{cases}, \quad (\text{A1})$$

where $\mathcal{H}$ is the set of all absolutely continuous functions in $C([0, T], \mathcal{R}^+)$ starting at σ_0^2 .

Another interesting candidate for stochastic volatility models that satisfies the large deviation principle is $v(t) = X(t)^{\frac{1}{\beta}}$, where $X(t)$ satisfies a square-root process

$$dX(t) = (\kappa - KX(t))dt + \varepsilon \sigma \sqrt{X(t)} dW_2(t).$$

This model can be dealt with by combining subclass (ii) above and the contraction principle. See Liu (2006) for details. Liu's model generalizes many stochastic volatility models in the literature. When $\beta = 1$, it reduces to Heston's model. When $\beta = -1$, it reduces to a model proposed by Chacko and Viceira (2005).

Appendix B: Proof of Propositions

Proof of Proposition 1

To simplify notation, we will suppress the arguments, κ_1 and κ_2 , of $v(t; \kappa_1, \kappa_2, \varepsilon)$. We write

$$v(t, \varepsilon) = v(t, 0) + \varepsilon \sigma \int_0^t e^{-\kappa_2(t-s)} dW_s.$$

Moreover,

$$|v(t, \varepsilon)| = v(t, \varepsilon) - 2v(t, \varepsilon)1_{\{v(t, \varepsilon) \leq 0\}}.$$

Then, by stochastic Fubini's theorem, we have

$$V(\varepsilon) = V(0) + \varepsilon \sigma y - 2\alpha,$$

where

$$y = \int_0^T \frac{1 - e^{-\kappa_2(T-u)}}{\kappa_2} dW(u), \quad \alpha = \int_0^T v(t, \varepsilon) 1_{\{v(t, \varepsilon) \leq 0\}} dt.$$

Recall the Black-Scholes option pricing formula

$$C_{BS}(\sqrt{V}) = S(0)\Phi(d_1) - K e^{-rT} \Phi(d_2),$$

where

$$d_1 = \frac{m}{\sqrt{V}} + \frac{1}{2}\sqrt{V}, \quad d_2 = \frac{m}{\sqrt{V}} - \frac{1}{2}\sqrt{V}, \quad m = \ln\left(\frac{S(0)}{K e^{-rT}}\right).$$

Let

$$h(V) = C_{BS}(\sqrt{V}).$$

Then

$$\begin{aligned} h'(V) &= C'_{BS}(\sqrt{V}) \frac{1}{2\sqrt{V}}, \\ h''(V) &= \frac{C''_{BS}(\sqrt{V})}{4V} - \frac{C'_{BS}(\sqrt{V})}{4V^{3/2}}, \\ h'''(V) &= C'''_{BS}(\sqrt{V}) \frac{1}{8V^{3/2}} - C''_{BS}(\sqrt{V}) \frac{3}{8V^2} + C'_{BS}(\sqrt{V}) \frac{3}{8V^{5/2}}. \end{aligned}$$

Moreover, we have

$$\begin{aligned} C'_{BS}(\sqrt{V}) &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}d_1^2\right], \\ C''_{BS}(\sqrt{V}) &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}d_1^2\right] \left(\frac{m^2}{V^{3/2}} - \frac{V^{1/2}}{4}\right), \\ C'''_{BS}(\sqrt{V}) &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}d_1^2\right] \left\{\left(\frac{V^{1/2}}{4} - \frac{m^2}{V^{3/2}}\right)^2 - \frac{1}{4} - \frac{3m^2}{V^2}\right\}. \end{aligned}$$

Case I. $m \neq 0$. It is readily seen that

$$\lim_{\sigma \rightarrow 0} \frac{C_{BS}^{(k)}(\sigma)}{\sigma^n} = \lim_{\sigma \rightarrow \infty} \frac{C_{BS}^{(k)}(\sigma)}{\sigma^n} = 0,$$

for $k = 1, 2, 3$, and $n \geq 1$. Hence,

$$\lim_{V \rightarrow 0} h'''(V) = \lim_{V \rightarrow \infty} h'''(V) = 0.$$

Since the function $h(V) \in C^\infty((0, \infty))$, there exists a positive real number K , such that

$$|h'''(V)| \leq K, \quad (\text{B1})$$

for all $V > 0$. By Taylor expansion, we have

$$\begin{aligned} h(V(\varepsilon)) &= h(V(0)) + h'(V(0))(V(\varepsilon) - V(0)) + \frac{1}{2}h''(V(0))(V(\varepsilon) - V(0))^2 \\ &\quad + \frac{1}{6}h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3, \end{aligned} \quad (\text{B2})$$

where $|\theta_\varepsilon - V(0)| \leq |V(\varepsilon) - V(0)|$. Taking expectation yields,

$$\begin{aligned} E[h(V(\varepsilon))] &= h(V(0)) + h'(V(0))E[V(\varepsilon) - V(0)] \\ &\quad + \frac{1}{2}h''(V(0))E[(V(\varepsilon) - V(0))^2] + \frac{1}{6}E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3]. \end{aligned}$$

By the Cauchy-Schwartz inequality,

$$E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3]^2 \leq E[h'''(\theta_\varepsilon)^2]E[(V(\varepsilon) - V(0))^6].$$

Since $h'''(V)$ is bounded as shown above,

$$E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3] \leq K\sqrt{E[(V(\varepsilon) - V(0))^6]}.$$

If we can show that

$$E[V(\varepsilon) - V(0)] = o((\varepsilon\sigma)^2) \quad (\text{B3})$$

$$E[(V(\varepsilon) - V(0))^2] = (\varepsilon\sigma)^2 y^2 + o((\varepsilon\sigma)^2) \quad (\text{B4})$$

$$E[(V(\varepsilon) - V(0))^6] = o((\varepsilon\sigma)^4), \quad (\text{B5})$$

then

$$\begin{aligned} E\left[C_{BS}\left(\sqrt{V(\varepsilon)}\right)\middle|\mathcal{F}\right] &= C_{BS}\left(\sqrt{V(0)}\right) \\ &\quad + (\varepsilon\sigma)^2 \left[\frac{C''_{BS}\left(\sqrt{V(0)}\right)}{8V(0)} - \frac{C'_{BS}\left(\sqrt{V(0)}\right)}{8(V(0))^{3/2}} \right] E[y^2] \\ &\quad + o((\varepsilon\sigma)^2), \end{aligned}$$

and the proposition follows for the case where $m \neq 0$.

It remains to show Equations (B3)–(B5). Note that $V(\varepsilon) - V(0) = \varepsilon\sigma y - 2\alpha$. We see that y is normal with mean zero and its variance is given by

$$E[y^2] = \int_0^T \left[\frac{1 - e^{-\kappa_2(T-u)}}{\kappa_2} \right]^2 du.$$

Now we evaluate $E[\alpha]$, $E[\alpha y]$, and $E[\alpha^2]$. We want to prove that $E[\alpha] = o((\varepsilon\sigma)^{2n})$ for any $n \geq 1$. Write

$$v(t, \varepsilon) = v(t, 0) + \varepsilon\sigma M(t)\xi, \quad \xi \sim N(0, 1).$$

Since

$$\begin{aligned} E[v(t, \varepsilon)1_{\{v(t, \varepsilon) \leq 0\}}] &= E\left[(v(t, 0) + \varepsilon\sigma M(t)\xi) 1_{\{\xi \leq -\frac{v(t, 0)}{\varepsilon\sigma M(t)}\}}\right] \\ &= v(t, 0)\Phi\left(-\frac{v(t, 0)}{\varepsilon\sigma M(t)}\right) + \frac{\varepsilon\sigma M(t)}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{v(t, 0)}{\varepsilon\sigma M(t)}\right)^2}, \end{aligned}$$

by L' Hopital rule, $e^{-\frac{1}{2\sigma^2}c} = o((\varepsilon\sigma)^{2n})$ for $c > 0$, $n > 0$. Thus, $E[\alpha] = o((\varepsilon\sigma)^{2n})$. Therefore, formula (B3) is proved.

For $E[\alpha y]$, observe first that $E[\alpha y]$ equals

$$\begin{aligned} E\left[\int_0^T \left(\int_0^t \frac{1 - e^{-\kappa_2(T-u)}}{\kappa_2} dW(u)\right) v(t, \varepsilon) 1_{\{v(t, \varepsilon) \leq 0\}} dt\right] \\ = E\left[\int_0^T N(t) v(t, \varepsilon) 1_{\{v(t, \varepsilon) \leq 0\}} dt\right], \end{aligned}$$

where

$$N(t) = \int_0^t \frac{1 - e^{-\kappa_2(T-u)}}{\kappa_2} dW(u).$$

Thus,

$$E\left[\int_0^T N(t) v(t, \varepsilon) 1_{\{v(t, \varepsilon) \leq 0\}} dt\right] = E\left[\int_0^T E[N(t) | v(t, \varepsilon)] v(t, \varepsilon) 1_{\{v(t, \varepsilon) \leq 0\}} dt\right].$$

Since $N(t)$ is perfectly correlated with $v(t, \varepsilon)$ and they follow a joint normal distribution, then by a similar argument as above, it is readily verified that

$$E\left[\int_0^T N(t) v(t, \varepsilon) 1_{\{v(t, \varepsilon) \leq 0\}} dt\right] = o((\varepsilon\sigma)^{2n}).$$

For $E[\alpha^2]$, observe that it is equal to

$$\int_0^T \int_0^T E[E[v(t, \varepsilon)1_{\{v(t, \varepsilon) \leq 0\}} | v(s, \varepsilon)] v(s, \varepsilon) 1_{\{v(s, \varepsilon) \leq 0\}}] dt ds.$$

Since $v(t, \varepsilon)$ and $v(s, \varepsilon)$ are jointly normal, by a similar argument as above, one can show that $E[v(t, \varepsilon)1_{\{v(t, \varepsilon) \leq 0\}} | v(s, \varepsilon)]$ is $o((\varepsilon\sigma)^{2n})$ and is bounded by a polynomial of $1/(\varepsilon\sigma)$. Thus, $E[\alpha^2]$ is $o((\varepsilon\sigma)^{2n})$.

Since $E[\alpha]$, $E[\alpha y]$, and $E[\alpha^2]$ are all $o((\varepsilon\sigma)^{2n})$, then Equation (B4) is proved. The proof of Equation (B5) is similar and the proposition is proved for Case I.

Case II. $m = 0$. In this case, noting that $S(0) = 1$, we have

$$C_{BS}(\sqrt{V}) = 2\Phi(d_1) - 1,$$

where $d_1 = \sqrt{V}/2$. Thus, we have

$$C_{BS}(\sqrt{V(\varepsilon)}) - C_{BS}(\sqrt{V(0)}) = 2 \left[\Phi \left(\frac{1}{2} \sqrt{V(\varepsilon)} \right) - \Phi \left(\frac{1}{2} \sqrt{V(0)} \right) \right].$$

Let $g(x) := \Phi(\frac{1}{2}\sqrt{x})$, $x > 0$. There exists a $\frac{V(0)}{2} \geq x_0 > 0$ and $K > 0$, such that

$$|g'''(x)| \leq K,$$

for all $x \geq x_0$.

Let

$$\mathcal{A} = \{|V(\varepsilon) - V(0)| \leq V(0) - x_0 := \delta\}$$

and let $\mathcal{A}^c$ denote the complementary set of $\mathcal{A}$. By the Freidlin-Wentzell theorem (Dembo and Zeitouni, 2003), we know that $P(\mathcal{A}^c) = o((\varepsilon\sigma)^2)$.

Consider the Taylor expansion

$$\begin{aligned} E \left[C_{BS} \left(\sqrt{V(\varepsilon)} \right) \right] &= C_{BS} \left(\sqrt{V(0)} \right) + h'(V(0))E[V(\varepsilon) - V(0)] \\ &\quad + \frac{1}{2} h''(V(0))E[(V(\varepsilon) - V(0))^2] \\ &\quad + \frac{1}{6} E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3 : V(\varepsilon) \in \mathcal{A}] \\ &\quad + \frac{1}{6} E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3 : V(\varepsilon) \in \mathcal{A}^c]. \end{aligned}$$

Based on the proof for Case I, it suffices to prove that

$$E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3 : V(\varepsilon) \in \mathcal{A}] = o((\varepsilon\sigma)^2) \quad (\text{B6})$$

$$E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3 : V(\varepsilon) \in \mathcal{A}^c] = o((\varepsilon\sigma)^2). \quad (\text{B7})$$

We first consider Equation (B6). If $V(\varepsilon) \in \mathcal{A}$, then θ_ε satisfies that $|\theta_\varepsilon - V(0)| \leq V(0) - x_0$. Hence, $|\theta_\varepsilon| \geq x_0$ and therefore, for all $V(\varepsilon) \in \mathcal{A}$, $|h'''(\theta_\varepsilon)| \leq 2K$. Therefore,

$$E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3 : V(\varepsilon) \in \mathcal{A}] \leq 2K \sqrt{E[(V(\varepsilon) - V(0))^6 : V(\varepsilon) \in \mathcal{A}]}.$$

But

$$E[(V(\varepsilon) - V(0))^6 : V(\varepsilon) \in \mathcal{A}] \leq E[(V(\varepsilon) - V(0))^6] = o((\varepsilon\sigma)^4).$$

Hence, we have proved Equation (B6).

To prove Equation (B7), we use the Taylor expansion again:

$$\begin{aligned} E[C_{BS}(\sqrt{V(\varepsilon)}) : V(\varepsilon) \in \mathcal{A}^c] \\ &= C_{BS}(\sqrt{V(0)}) + h'(V(0))E[V(\varepsilon) - V(0) : V(\varepsilon) \in \mathcal{A}^c] \\ &\quad + \frac{1}{2}h''(V(0))E[(V(\varepsilon) - V(0))^2 : V(\varepsilon) \in \mathcal{A}^c] \\ &\quad + \frac{1}{6}E[h'''(\theta_\varepsilon)(V(\varepsilon) - V(0))^3 : V(\varepsilon) \in \mathcal{A}^c]. \end{aligned}$$

Because $C_{BS}(x)$ is bounded (since $m = 0$), we have

$$E[C_{BS}(\sqrt{V(\varepsilon)}) - C_{BS}(\sqrt{V(0)}) : V(\varepsilon) \in \mathcal{A}^c] = o((\varepsilon\sigma)^2).$$

Then it suffices to prove that

$$E[V(\varepsilon) - V(0) : V(\varepsilon) \in \mathcal{A}^c] = o((\varepsilon\sigma)^2) \quad (\text{B8})$$

$$E[(V(\varepsilon) - V(0))^2 : V(\varepsilon) \in \mathcal{A}^c] = o((\varepsilon\sigma)^2). \quad (\text{B9})$$

Note that $V(\varepsilon) - V(0) = \varepsilon\sigma y - \alpha$ and by definition $\alpha \leq 0$. Because $E[\alpha] = o((\varepsilon\sigma)^{2n})$, for all $n \geq 0$, we have $E[\alpha : V_\varepsilon \in \mathcal{A}^c] = o((\varepsilon\sigma)^{2n})$, for all $n \geq 0$ (since α is always nonnegative). On the other hand,

$$E[y : V(\varepsilon) \in \mathcal{A}^c] = E[y 1_{\{V(\varepsilon) \in \mathcal{A}^c\}}] \leq \sqrt{E[y^2]}P(\mathcal{A}^c) = o((\varepsilon\sigma)^2).$$

Hence, the formula (B8) is proved. The proof for Equation (B9) is similar.

Proof of Proposition 2

Fix $\delta > 0$, $\kappa_1 \geq 0$, and $\kappa_2 > 0$. Let $v(t; \kappa_1, \kappa_2)$ satisfy

$$dv(t) = (\kappa_1 - \kappa_2 v(t))dt, \quad v(0) = \sigma_\varepsilon^2.$$

Then, $v(t; \kappa_1, \kappa_2) > 0$ and the integral $V(\kappa_1, \kappa_2, 0) = \int_0^T v(t; \kappa_1, \kappa_2)dt > 0$.

Since $C'_{BS}(V) = N'(d_1)$, the event

$$\left\{ \left| C_{BS}(\sqrt{V(\kappa_1, \kappa_2, \varepsilon)}) - C_{BS}(\sqrt{V(\kappa_1, \kappa_2, 0)}) \right| \geq \delta \right\} \quad (\text{B10})$$

is a subset of the event

$$\left\{ \left| \sqrt{V(\kappa_1, \kappa_2, \varepsilon)} - \sqrt{V(\kappa_1, \kappa_2, 0)} \right| \geq \delta \right\}. \quad (\text{B11})$$

Because of this inclusion relation, instead of estimating the probability of the event (B10), we will establish a law of large deviation for the event (B11).

Let A_δ denote the set of all continuous functions $\phi(t)$ on $[0, T]$ with $\phi(0) = \sigma_0^2$ that satisfy

$$\left| \sqrt{\int_0^T |\phi(t)| dt} - \sqrt{\int_0^T v(t; \kappa_1, \kappa_2) dt} \right| \geq \delta. \quad (\text{B12})$$

Then, event (B11) occurs if and only if the sample path of $v(t, \kappa_1, \kappa_2, \varepsilon) \in A_\delta$. Define

$$R(\kappa_1, \kappa_2, \delta) = \inf_{\phi \in A_\delta} I(\phi),$$

where $I(\phi)$ is given by Equation (A1). We will derive below the closed-form expression of $R(\kappa_1, \kappa_2, \delta)$ and show that $J(\kappa_1, \kappa_2, \delta) \leq \sigma^2 R(\kappa_1, \kappa_2, \delta)$ which, when combined with the large deviation principle, implies Proposition 2. The closed-form expression of $R(\kappa_1, \kappa_2, \delta)$ is of interest in its own right.

STEP 1. In this step, we identify all extremal points of $I(y)$, including possible candidates for the minimizer in the definition of $R(\kappa_1, \kappa_2, \delta)$.

Define $F(t, y, y') = (y'(t) - \kappa_1 + \kappa_2 y(t))^2$. By the Freidlin-Wentzell theorem (Dembo and Zeitouni, 2003),

$$I(y) = \frac{1}{2\sigma^2} \int_0^T F(t, y', y) dt = \frac{1}{2\sigma^2} \int_0^T (y'(t) - \kappa_1 + \kappa_2 y(t))^2 dt.$$

By the Euler-Lagrange equation of the variational problem with objective function $F(t, y, y')$, the optimal function $y(t)$ has to satisfy

$$F_y = \frac{d}{dt} F_{y'}. \quad (\text{B13})$$

The solution of $y(t)$ of Equation (B13), with the initial condition $y(0) = \sigma_0^2$, is given by

$$y(t; c_1) = c_1 e^{\kappa_2 t} + \left[\sigma_0^2 - c_1 - \frac{\kappa_1}{\kappa_2} \right] e^{-\kappa_2 t} + \frac{\kappa_1}{\kappa_2}, \quad (\text{B14})$$

where c_1 is a constant parameter, and the function $y(t; c_1)$ is indexed by the parameter c_1 . The benchmark volatility $v(t)$ corresponds to the function $y(t; 0)$ with $c_1 = 0$, i.e., $v(t) = (\sigma_0^2 - \frac{\kappa_1}{\kappa_2}) e^{-\kappa_2 t} + \frac{\kappa_1}{\kappa_2}$.

Let $z(t) = e^{\kappa_2 t} - e^{-\kappa_2 t}$, and

$$\bar{Z} = \int_0^T z(t) dt = \frac{(e^{\frac{\kappa_2}{2} T} - e^{-\frac{\kappa_2}{2} T})^2}{\kappa_2}.$$

Then we have

$$y(t; c_1) = c_1 z(t) + v(t).$$

For this particular $y(t; c_1)$, we have

$$I(y) = \frac{c_1^2}{2\sigma^2} \int_0^T (z'(t) + \kappa_2 z(t))^2 dt = \frac{c_1^2 \kappa_2}{\sigma^2} (e^{2\kappa_2 T} - 1). \quad (\text{B15})$$

The range of c_1 is determined by $y(t; c_1) \in A_\delta$. That is, Equation (B12). Therefore, to calculate $R(\kappa_1, \kappa_2, \delta)$, it suffices to determine

$$\min_{\{y(t; c_1) \in A_\delta\}} |c_1|. \quad (\text{B16})$$

STEP 2. In this step we consider the case that $\delta \geq \sqrt{V}$, and solve problem (B16).

In this case, by Equation (B12),

$$\sqrt{\int_0^T |y(t; c_1)| dt} \geq \delta + \sqrt{V}.$$

Then

$$\int_0^T |y(t; c_1)| dt \geq (\delta + \sqrt{V})^2.$$

Since

$$|y(t; c_1)| \leq |c_1| z(t) + v(t),$$

we have

$$|c_1| \bar{Z} + V \geq (\delta + \sqrt{V})^2.$$

Hence,

$$|c_1| \geq \alpha(\delta) = \frac{1}{\bar{Z}} (\delta^2 + 2\sqrt{V}\delta),$$

and the equality can be achieved by choosing $c_1 = \alpha(\delta)$. Therefore, for $\delta \geq \sqrt{V}$, we have

$$\min_{\{y(t; c_1) \in A_\delta\}} |c_1| = \alpha(\delta). \quad (\text{B17})$$

STEP 3. In this step, we consider the case that $\delta < \sqrt{V}$, and present a lower bound of the problem (B16).

We first consider the case that $c_1 > 0$. In this case, $\int_0^T |y(t; c_1)| dt > V$. The same argument as in Step 2 goes through and Equation (B17) holds as well.

That is,

$$\min_{\{c_1 > 0, y(t; c_1) \in A_\delta\}} |c_1| = \alpha(\delta). \quad (\text{B18})$$

We now consider the case that $c_1 < 0$. In this case, Equation (B12) is equivalent to either

$$\sqrt{\int_0^T |y(t; c_1)| dt} - \sqrt{V} \geq \delta, \quad (\text{B19})$$

or

$$\sqrt{\int_0^T |y(t; c_1)| dt} - \sqrt{V} \leq -\delta. \quad (\text{B20})$$

To proceed, we consider two cases: (i) the function $y(t; c_1)$ is always nonnegative over $[0, T]$; and (ii) the function $y(t; c_1)$ is negative for some $t \in [0, T]$. We consider case (ii) first. Let $A_\delta(1) = \{y(t) \in A_\delta, y(T) < 0\}$, and $A_\delta(2) = A_\delta - A_\delta(1)$. Note that $y(0; c_1) > 0$. Therefore, $y(t; c_1) \in A_\delta(1)$ if and only if there exists a positive $t_0 \in (0, T)$ such that $y(t_0; c_1) = 0$. Since

$$y(t; c_1) = e^{-\kappa_2 t} \left[c_1 e^{2\kappa_2 t} + \frac{\kappa_1}{\kappa_2} e^{\kappa_2 t} + \sigma_0^2 - \frac{\kappa_1}{\kappa_2} - c_1 \right],$$

$e^{\kappa_2 t_0}$ solves the equation

$$c_1 x^2 + \frac{\kappa_1}{\kappa_2} x + \left(\sigma_0^2 - \frac{\kappa_1}{\kappa_2} - c_1 \right) = 0.$$

This equation has two solutions. Because this equation is concave in x and $y(0; c_1) > 0$, then $y(t; c_1) < 0$ for $t > t_0$ and $y(t; c_1) > 0$ for $t \in [0, t_0]$. Then it follows from Equation (B19) that

$$c_1 \bar{Z} + 2 \int_{t_0}^T z(t) (|c_1| - v(t)/z(t)) dt \geq \delta^2 + 2\delta\sqrt{V}.$$

Since

$$|c_1| \bar{Z} \geq c_1 \bar{Z} + 2 \int_{t_0}^T z(t) (|c_1| - v(t)/z(t)) dt,$$

a necessary condition for Equation (B19) is

$$|c_1| \geq \alpha(\delta).$$

Similarly, we have from Equation (B20) that

$$|c_1| \left\{ \bar{Z} - 2 \int_{t_0}^T z(t) dt \right\} \geq 2\delta\sqrt{V} - \delta^2 - 2 \int_{t_0}^T v(t) dt. \quad (\text{B21})$$

Then

$$|c_1| \geq \beta(\delta) + \frac{2}{Z} \int_{t_0}^T z(t) \left(|c_1| - \frac{v(t)}{z(t)} \right) dt \geq \beta(\delta), \quad (\text{B22})$$

where

$$\beta(\delta) = \frac{1}{Z} (2\sqrt{V}\delta - \delta^2),$$

and the inequality in Equation (B22) follows by noting that the function $v(t)/z(t)$ is nonincreasing in t and $v(t_0)/z(t_0) = |c_1|$. It is clear that $\beta(\delta) \leq \alpha(\delta)$. Thus, we have that

$$\min_{\{c_1 < 0, y(t; c_1) \in A_\delta(1)\}} |c_1| \geq \beta(\delta).$$

We now consider case (i), where $y(t; c_1) \geq 0$ for all $t \in [0, T]$. In this case, $|y(t; c_1)| = y(t; c_1) = -|c_1|z(t) + v(t)$. Clearly, Equation (B19) does not hold. We consider Equation (B20), which is equivalent to

$$|c_1| \geq \beta(\delta).$$

Therefore,

$$\min_{\{c_1 < 0, y(t; c_1) \in A_\delta(2)\}} |c_1| \geq \beta(\delta).$$

Putting cases (i) and (ii) together, we have

$$\min_{\{c_1 < 0, y(t; c_1) \in A_\delta\}} |c_1| \geq \beta(\delta). \quad (\text{B23})$$

STEP 4. In this step, we consider the case that $\delta < \sqrt{V}$ and solve the problem (B16).

By Step 3, it suffices to examine the case when $\beta(\delta)$ is achieved. Set

$$h_T = Z \frac{v(T)}{z(T)}.$$

Noting that $\frac{V}{Z} \geq \inf_{t \in [0, T]} \frac{v(t)}{z(t)}$, it follows from the monotonicity of $v(t)/z(t)$ that $V \geq h_T$. Let

$$\delta^* = \sqrt{V} - \sqrt{V - h_T} > 0. \quad (\text{B24})$$

Then δ^* satisfies the following equation of δ :

$$2\sqrt{V}\delta - \delta^2 = h_T. \quad (\text{B25})$$

Case A. $\delta \leq \delta^*$.

In this case, $\beta(\delta) \leq \frac{v(T)}{z(T)}$ and we choose $c_1 = -\beta(\delta)$. We now prove that $y(t; c_1) \in A_\delta$ with this choice of c_1 .

Since $z(t)/v(t)$ is nonincreasing, $\beta(\delta) \leq v(t)/z(t)$ for all $t \in [0, T]$. Thus, $y(t; c_1) \geq 0$ for all $t \in [0, T]$. Then it is readily checked that Equation (B20) holds for $y(t; c_1)$. Hence, we have shown in Case A,

$$\min_{\{y(t; c_1) \in A_\delta\}} |c| = \min_{\{c < 0, y(t; c_1) \in A_\delta\}} |c| = \beta(\delta). \quad (\text{B26})$$

Case B. $\delta \in (\delta^*, \sqrt{V})$.

Let

$$\Gamma(\delta) = \min_{\{y(t; c_1) \in A_\delta\}} |c| = \min \left\{ \min_{\{c < 0, y(t; c_1) \in A_\delta\}} |c|, \alpha(\delta) \right\}, \quad (\text{B27})$$

where the second equality comes from the Equation (B18). Hence by Equation (B23), $\Gamma(\delta) \geq \beta(\delta)$.

STEP 5. We compute $R(\kappa_1, \kappa_2, \delta)$ in this step.

From Equation (B15) and the previous discussion, we have for $\delta \geq \sqrt{V}$,

$$R(\kappa_1, \kappa_2, \delta) = \frac{\alpha^2(\delta)\kappa_2}{\sigma^2} (e^{2\kappa_2 T} - 1); \quad (\text{B28})$$

for $\delta \leq \delta^*$,

$$R(\kappa_1, \kappa_2, \delta) = \frac{\beta^2(\delta)\kappa_2}{\sigma^2} (e^{2\kappa_2 T} - 1); \quad (\text{B29})$$

and for δ in $(\delta^*, \sqrt{V})$,

$$R(\kappa_1, \kappa_2, \delta) = \frac{\Gamma^2(\delta)\kappa_2}{\sigma^2} (e^{2\kappa_2 T} - 1). \quad (\text{B30})$$

Note that there is no closed-form expression of $\Gamma(\delta)$.

STEP 6. We investigate the relation between function $R(\kappa_1, \kappa_2, \delta)$ and function $J(\kappa_1, \kappa_2, \delta)$.

For $\delta \leq \delta^*$, we clearly have $R(\kappa_1, \kappa_2, \delta) = \frac{1}{\sigma^2} J(\kappa_1, \kappa_2, \delta)$. For $\delta \in (\delta_1^*, \sqrt{V}]$, since $\Gamma(\delta) \geq \beta(\delta)$, we have $R(\kappa_1, \kappa_2, \delta) \geq \frac{1}{\sigma^2} J(\kappa_1, \kappa_2, \delta)$. Hence, for all $\delta \in (0, \sqrt{V}]$, we have

$$e^{-\frac{1}{\sigma^2} R(\kappa_1, \kappa_2, \delta)} \leq e^{-\frac{1}{\sigma^2} J(\kappa_1, \kappa_2, \delta)}.$$

This gives us a probability bound, as claimed.

Proof of Proposition 3

To prove the analytical properties of $J(\kappa_1, \kappa_2, \eta H(\kappa_1, \kappa_2))$ as claimed in Proposition 3, we need the following lemma. Note that $\delta_0 = V(\kappa_1, \kappa_2, 0)$. To simplify notation, we will omit the argument 0 and write it as $V(\kappa_1, \kappa_2)$.

Lemma 1. *For $\kappa_1, \kappa_2 > 0$, $V(\kappa_1, \kappa_2)$ is strictly increasing with respect to κ_1 .*

Proof. It is readily verified that

$$V(\kappa_1, \kappa_2) = \frac{\kappa_1}{\kappa_2} T + \left(\sigma_0^2 - \frac{\kappa_1}{\kappa_2} \right) \frac{1 - e^{-\kappa_2 T}}{\kappa_2}.$$

Thus,

$$\frac{\partial V(\kappa_1, \kappa_2)}{\partial \kappa_1} = \frac{T}{\kappa_2} - \frac{1}{\kappa_2^2} (1 - e^{-\kappa_2 T}).$$

It is readily seen by taking the derivative with respect to T that the above expression is positive. The claim of the lemma follows.

Turning to the proof of Proposition 3, claim (1) follows from the fact that

$$\left(2\sqrt{V(\kappa_1, \kappa_2, 0)}\delta - \delta^2 \right)^2$$

is an increasing function of δ for $0 < \delta \leq \delta_0$.

Claim (2) follows from claim (1) and the fact that the Black-Scholes call option price $H(\kappa_1, \kappa_2)$ is a decreasing function of strike price K .

To prove claim (3), fix κ_2 and η . By Lemma 1, $V(\kappa_1, \kappa_2)$ is increasing in κ_1 . This implies that $H(\bar{v}\kappa_2, \kappa_2)$ is increasing in $\bar{v}$ for $\kappa_2 > 0$. Since both $V(\bar{v}\kappa_2, \kappa_2)$ and $H(\bar{v}\kappa_2, \kappa_2)$ are increasing in $\bar{v}$, by the proof of claim (1), $J(\bar{v}\kappa_2, \kappa_2, \eta H(\bar{v}\kappa_2, \kappa_2))$ is increasing in $\bar{v}$.

For claim (4), note first that when $\kappa_1 = \bar{v}\kappa_2$,

$$V(\bar{v}\kappa_2, \kappa_2) = \bar{v}T + (\sigma_0^2 - \bar{v})(1 - e^{-\kappa_2 T})/\kappa_2.$$

Its partial derivative with respect to κ_2 is

$$(\sigma_0^2 - \bar{v}) \frac{(1 + T\kappa_2)e^{-\kappa_2 T} - 1}{\kappa_2^2}.$$

This expression is strictly increasing in T when $\bar{v} > \sigma_0^2$ and takes its minimum 0 at $T = 0$. Thus, $V(\bar{v}\kappa_2, \kappa_2)$ is strictly increasing in κ_2 when $\bar{v} > \sigma_0^2$. Hence, both $V(\bar{v}\kappa_2, \kappa_2)$ and $H(\bar{v}\kappa_2, \kappa_2)$ are increasing in κ_2 when $\bar{v} > \sigma_0^2$. Then, by the same argument as in the proof of claim (3),

$$\left(\frac{\eta H(V(\bar{v}\kappa_2, \kappa_2))}{S(0)} \sqrt{V(\bar{v}\kappa_2, \kappa_2)} - \left(\frac{\eta H(V(\bar{v}\kappa_2, \kappa_2))}{S(0)} \right)^2 \right)^2$$

is increasing in $\bar{v}$ when $\bar{v} > \sigma_0^2$.

To prove claim (4) it remains to show that $J_1(\kappa_2)$ is increasing in κ_2 . Let $b = e^{\kappa_2 T} > 1$. Then

$$J_1(\kappa_2) = \frac{1}{(\ln T)^3} \frac{b^2(b+1)(\ln b)^3}{(b-1)^3}.$$

Ignoring the constant term, the derivative of the expression with respect to b is

$$F(b) = \frac{(\ln b)^2}{(b-1)^4} (3b^3 - 3b - (4b^2 + 2b) \ln b).$$

We wish to show $F(b) > 0$ for $b > 1$. Let

$$f(b) = 3b^3 - 3b - (4b^2 + 2b) \ln b.$$

Taking derivatives,

$$f'(b) = 9b^2 - 4b - 5 - (8b + 2) \ln b$$

$$f''(b) = 18b - 8 \ln b - 12 - \frac{2}{b}$$

$$f'''(b) = 18 - \frac{8}{b} + \frac{2}{b^2}.$$

Clearly, $f'''(b) > 0$ for $b > 1$. Since $f''(1) = 4$, we have $f''(b) > 0$ for $b \geq 1$. Next, since $f'(1) = 0$, $f'(b) > 0$ for $b > 1$. Now since $f(1) = 0$, we have $f(b) > 0$ for $b > 1$. Therefore, $F(b) > 0$ for $b > 1$, which implies that $J_1(\kappa_2)$ is increasing in κ_2 .

References

- Anderson, E., L. Hansen, and T. Sargent. 2003. A Quarter of Semi-Groups for Model Specification, Detection, Robustness, and the Price of Risk. *Journal of the European Economics Association* 1:68–123.
- Bank, P., and N. El. Karoui. 2004. A Stochastic Representation Theorem with Applications to Optimization and Obstacle Problems. *Annals of Probability* 32:1030–67.
- Bernardo, A., and O. Ledit. 2000. Gain, Loss, and Asset Pricing. *Journal of Political Economy* 108:144–72.
- Bizid, A., E. Jouini, and P.-F. Koehl. 1999. Pricing of Non-Redundant Derivatives in a Complete Market. Working Paper, CREST-ENSAE, 15 Bd Gabriel Péri, 92241 Malakoff Cedex, France.
- Bondarenko, O. 2003. Statistical Arbitrage and Securities Prices. *Review of Financial Studies* 16:875–919.
- Cagetti, M., L. Hansen, T. Sargent, and N. Williams. 2002. Robustness and Pricing with Uncertain Growth. *Review of Financial Studies* 15:363–404.
- Cerney, A. 2003. Generalized Sharpe Ratios and Asset Pricing in Incomplete Markets. *European Finance Review* 7:191–233.
- Chacko, G., and L. Viceira. 2005. Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Market. *Review of Financial Studies* 18:1369–1402.
- Cochrane, J., and J. Saá-Requejo. 2000. Beyond Arbitrage: Good-Deal Asset Price Bounds in Incomplete Markets. *Journal of Political Economy* 108:79–119.

- Cox, J., J. Ingersoll, and S. Ross. 1985. An Intertemporal General Equilibrium Model of Asset Prices. *Econometrica* 53:363–84.
- Dawson, D., and S. Feng. 1998. Large Deviation for the Fleming-Viot Process with Neutral Mutation and Selection. (I). *Stochastic Processes and Applications* 77:207–32.
- Dawson, D., and S. Feng. 2001. Large Deviation for the Fleming-Viot Process with Neutral Mutation and Selection. (II). *Stochastic Processes and Applications* 92:131–82.
- Dembo, A., and O. Zeitouni. 2003. *Large Deviations Techniques and Applications*. 2nd ed. Applications of Mathematics, vol. 38. New York: Springer.
- Feng, S., and J. Xiong. 2002. Large Deviations and Quasi-Potential of a Fleming-Viot Process. *Electronic Communication in Probability* 7:13–25.
- Frey, R., and C. Sin. 1999. Bounds on European Option Prices Under Stochastic Volatility. *Mathematical Finance* 9:97–116.
- Garlappi, L., R. Uppal, and T. Wang. 2007. Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach. *Review of Financial Studies* 20:41–81.
- Hansen, L., and R. Jagannathan. 1981. Implications for Security Market Data for Models of Dynamic Economies. *Journal of Political Economy* 99:225–62.
- Hansen, L., and R. Jagannathan. 1997. Assessing Specification Errors in Stochastic Discount Factor Models. *Journal of Finance* 52:557–90.
- Hansen, L., and T. Sargent. 1995. Discounted Linear Exponential Quadratic Gaussian Control. *IEEE Transactions and Automatic Control* 40:141–53.
- Hansen, L., and T. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91:60–66.
- Heston, S. 1993. A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies* 6:327–43.
- Hodrick, R., and X. Zhang. 2001. Evaluating the Specification Errors of Asset Pricing Model. *Journal of Financial Economics* 62:327–76.
- Hull, J., and A. White. 1987. The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance* 42:281–300.
- Johnson, M., and K. Sircar. 2002. Partial Hedging in a Stochastic Volatility Environment. *Mathematical Finance* 12:375–409.
- Jouini, E., and C. Napp. 1999. Continuous Time Equilibrium Pricing of Nonredundant Assets. Working Paper, CREST-ENSAE, 15 Bd Gabriel Péri, 92241 Malakoff Cedex, France.
- Levy, H. 1985. Upper and Lower Bounds of Put and Call Option Value: Stochastic Dominance Approach. *Journal of Finance* 40:1197–1217.
- Liu, J. 2007. Portfolio Selection in Stochastic Environments. *Review of Financial Studies* 20:1–39.
- Lucas, R. 1978. Asset Prices in an Exchange Economy. *Econometrica* 46:1429–45.
- MacKay, A., and E. Prisman. 1997. Estimating Valuation Operators in Incomplete Markets: Can Noise Complete the Market? Working Paper, Faculty of Management, University of Toronto.
- Maenhout, P. 2004. Robust Portfolio Rules and Asset Pricing. *Review of Financial Studies* 17:951–83.
- Merville, L., and D. Piepeta. 1989. Stock Price Volatility, Mean-Reverting Diffusion, and Noise. *Journal of Financial Economics* 24:193–214.
- Rogers, L. 1994. Equivalent Martingale Measures and No-Arbitrage. *Stochastics and Stochastics & Reports* 51:41–49.
- Ross, S. 2005. *Neoclassical Finance*. Princeton, NJ: Princeton University Press.

Stein, E. 1989. Overactions in the Options Market. *Journal of Finance* 44:1011–23.

Stein, E., and J. Stein. 1991. Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies* 4:727–52.

Stutzer, M. 1995. A Bayesian Approach to Diagnosis of Asset Pricing Models. *Journal of Econometrics* 68:367–97.

Stutzer, M. 1996. A Simple Nonparametric Approach to Derivative Security Valuation. *Journal of Finance* 51:1633–52.

Uppal, R., and T. Wang. 2003. Model Misspecification and Underdiversification. *Journal of Finance* 58:2465–86.

Copyright of Review of Financial Studies is the property of Society for Financial Studies and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.