Essays in Financial Engineering

Andrew Jooyong Ahn

Submitted in partial fulfillment of the

requirements for the degree

of Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2014

©2014

Andrew Jooyong Ahn

All Rights Reserved

ABSTRACT

Essays in Financial Engineering

Andrew Jooyong Ahn

This thesis consists of three essays in financial engineering. In particular we study problems in option pricing, stochastic control and risk management.

In the first essay, we develop an accurate and efficient pricing approach for options on leveraged ETFs (LETFs). Our approach allows us to price these options quickly and in a manner that is consistent with the underlying ETF price dynamics. The numerical results also demonstrate that LETF option prices have model-dependency particularly in high-volatility environments.

In the second essay, we extend a linear programming (LP) technique for approximately solving high-dimensional control problems in a diffusion setting. The original LP technique applies to finite horizon problems with an exponentially-distributed horizon, T. We extend the approach to fixed horizon problems. We then apply these techniques to dynamic portfolio optimization problems and evaluate their performance using convex duality methods. The numerical results suggest that the LP approach is a very promising one for tackling high-dimensional control problems.

In the final essay, we propose a factor model-based approach for performing scenario analysis in a risk management context. We argue that our approach addresses some important drawbacks to a standard scenario analysis and, in a preliminary numerical investigation with option portfolios, we show that it produces superior results as well.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis addresses three problems in financial engineering. These problems are from the sub-fields of derivatives pricing, dynamic portfolio optimization and risk management, respectively. Our approach to dynamic portfolio optimization actually applies to diffusion-based control problems more generally.

We begin in chapter 2 with the problem of pricing options on a leveraged ETF (LETF) and the underlying security (or ETF) in a consistent manner. We show that if the underlying ETF has Heston dynamics then the LETF also has Heston dynamics so that options on both the ETF and the LETF can be priced analytically using standard transform methods. If the underlying ETF has tractable jump-diffusion dynamics then the dynamics of the corresponding LETF are generally intractable in that we cannot compute a closed-form expression for the characteristic function of the log-LETF price. This is because we need to account for the limited liability of an LETF when we model its dynamics. This is not an issue with diffusion processes but it does become an issue once we introduce jumps. To address this problem we propose tractable approximations to the LETF price dynamics under which the characteristic function of the log-LETF price can be found

in closed form. In a series of numerical experiments including both low and high volatility regimes, we show that the resulting LETF option price approximations are very close to the true prices which we calculate via Monte-Carlo. Because approximate LETF option prices can be computed very quickly our methodology should be useful in practice for pricing and risk-managing portfolios that contain options on both ETFs and related LETFs. Our numerical results also demonstrate the model-dependency of LETF option prices and this is particularly noticeable in high-volatility environments. This model dependency calls into question the market practice of pricing an LETF option using the Black-Scholes formula with the strike and implied volatility scaled by the leverage ratio.

In chapter 3 we study a linear programming (LP) technique to compute good sub-optimal solutions to high-dimensional control problems in a diffusion-based setting. This LP approach was recently introduced by Han and Van Roy (2011). Their problem formulation worked with finite horizon problems where the horizon, T, is an exponentially-distributed random variable. As a result, the time, t, is not a state variable in the associated HJB equation. A direct application of their approach, however, does not work for problems with a fixed and finite horizon because of the dependency of the HJB equation on t in that case. In this chapter we extend their approach to the fixed and finite horizon case and apply it to a series of dynamic portfolio optimization problems. We then simulate the resulting policies to obtain lower bounds on the optimal value functions for these problems. An advantage of considering these portfolio optimization problems is that we can use convex duality methods designed for these problems to construct upper bounds on the optimal value functions. In our numerical experiments we find that the lower and upper bounds are very close. We therefore provide strong evidence (beyond the results of Han and Van Roy 2011) that the LP approach is a very promising approach for high-dimensional diffusion-based control problems.

Chapter 4 discusses our final problem which relates to factor model-based scenario analysis. Scenario analysis is an important and widely used risk-management technique that is used throughout the financial services industry. In the standard version of scenario analysis we shift a small predefined subset of risk factors and compute the resulting profit-and-loss (P&L) on the portfolio. By considering many shifts and many subsets of factors, it is then possible to get a good understanding of the risk profile of the portfolio. Moreover, because this standard form of scenario analysis does not require a probability distribution and produces a P&L for each considered scenario, it is preferred by many practitioners to risk measures such as value-at-risk (VaR) or conditional value-at-risk (CVaR) which are scalar and rely on knowledge of a probability distribution that is often very hard to estimate.

But scenario analysis suffers from at least two important drawbacks: (i) In stressing a small subset of risk factors it implicitly sets the shocks to non-stressed risk-factors to zero. This tends to ignore any conditional dependence structure between the stressed and unstressed risk factors. It also ignores the convexity of the portfolio with respect to the unstressed factors. (ii) In contrast to VaR and CVaR, for example, this standard form of scenario analysis is not testable in that the probability of any given scenario actually occurring is zero. In this sense then it is not possible to quantify the performance of standard scenario analysis. In this chapter we propose a factor-model based scenario analysis which is easy to implement and produces an *expected* P&L for each proposed scenario. This allows us to overcome problem (i). Our factor modeling approach also allows us to estimate *realized* shocks to the risk-factors and therefore compare the *realized* P&L with the P&L we would have *predicted* conditioned on these realizations. We can therefore also address problem (ii). We develop our modeling approach in the context of an options portfolio with a single underlying security but it should be clear that our framework can also be applied in

other contexts. In preliminary numerical tests with S&P 500 options data, our factor model-based scenario analysis performs well and outperforms the standard scenario analysis approach.

Chapter 2

Consistent Pricing of Options on Leveraged ETFs

2.1 Introduction

According to various industry sources, there were more than $4,500$ registered ETFs globally in 2010 with assets under management (AUM) of approximately $1.6 trillion. These ETFs are spread among many asset classes including equity, fixed income, commodity and FX. There were liquid options available on approximately 400 of these ETFs in 2010 and these ETFs accounted for approximately $1 trillion of the $1.6 trillion in AUM. Moreover the total ETF options volume is very large indeed: according to the Chicago Board Options Exchange [16], of the 4 billion exchange traded options contracts in 2010, 1.3 billion were ETF options, with equity options and cash index options accounting for 2.4 billion and 0.3 billion, respectively. In contrast, there were approximately 2 billion contracts traded in 2006, with a split of 1.5 billion equity options, 350 million ETF options and 180 million cash index options. Between 2006 and 2010 the ETF options market therefore grew

by a factor of four and is now a very large market indeed.

An even more recent development has been the introduction of *leveraged* ETFs (LETFs). An LETF is an exchange-traded derivative security based on a single underlying ETF or index. It is intended to achieve a *daily* return of ϕ times the daily return of the underlying ETF and the LETF manager needs to re-balances his portfolio on a daily basis in order to achieve this. The constant ϕ is known as the *leverage ratio* of the LETF. As of 2010, there were approximately 150 LETFs with a total of $30 billion in AUM and approximately 100 of these LETFs have liquid options traded on them. Moreover, a given LETF typically has a very large and liquid ETF or index as its underlying security with options traded on both the LETF and the underlying ETF.

Upon their introduction, there was considerable confusion among investors over the performance of LETFs, particularly during the financial crisis when volatility levels spiked to unprecedented levels. In particular, many investors did not appreciate that LETFs had a negative exposure to the realized variance of the underlying ETF and therefore did not anticipate their potentially poor performance during this period. Cheng and Madhavan [9] and Avellaneda [2] were the first to model and explain this LETF performance. In a continuous-time diffusion framework they obtained an expression (see (2.2) below) that highlighted this negative exposure to realized variance. Based on results in Haugh and Jain [21], Haugh [20] also derived this expression as a simple case of a more general expression for the realized wealth from following a constant proportion trading strategy in a multi-security diffusion setting.

While these papers helped to explain LETF performance, there has been little work on the pricing of LETF options and, in particular, on pricing them in a manner that is *consistent* with the pricing of options on the underlying ETF. One approach for pricing LETF options is based on using the Black-Scholes formula with the implied volatility taken from a related ETF option

and then scaled by the leverage ratio. But this approach is ad-hoc and has not been properly justified. Concurrent with our work is the recent paper of Leung and Sircar [27] who use asymptotic techniques in a diffusion setting to understand the link between implied volatilities of the underlying ETF and related LETFs of a given leverage ratio. They then use the resulting insights to identify possible mispricings in the market-place.

In this chapter we price LETF options quickly and *consistently* with options on the underlying ETF under three different models: (i) Heston's [24] stochastic volatility model (ii) the Bates [5] jump-diffusion model and (iii) an affine jump-diffusion (AJD) model of Duffie, Pan and Singleton [15] which includes jumps in both the volatility and price processes. In the sequel we will often refer to these models as the SV, SVJ and SVCJ models, respectively. It should also be clear that the approximation techniques we develop in this chapter can be applied more generally and that our treatments of the SV, SVJ and SVCJ models may be viewed as applications of a more general approach. For example, other AJD models of Duffie et al. [15] should also be amenable to our approximation techniques.

We show that if the underlying ETF has Heston dynamics then the LETF also Heston dynamics so that options on both the ETF and the LETF can be priced analytically using standard transform methods. If the underlying ETF has tractable jump-diffusion dynamics (as in (ii) and (iii) above) then the dynamics of the corresponding LETF are generally intractable in that we cannot compute a closed-form expression for the log-LETF price. Instead we propose tractable approximations to the LETF dynamics where the characteristic function of the log-LETF price can be found in closed form, thereby implying that we can calculate approximate option prices very quickly. The key to our approach is that under our jump-diffusion models for the underlying ETF, the diffusion component of the LETF dynamics remains "tractable". We therefore only need to focus on approximating the

jump component of the LETF dynamics.

In a series of numerical experiments including both low and high volatility regimes, we show that the resulting LETF option price approximations are very close to the true prices which we calculate via Monte-Carlo. Our approximate LETF option prices can be computed very quickly and therefore should be useful in practice for pricing and risk-managing portfolios that contain options on both ETFs and related LETFs.

Our numerical experiments also show that the ratio of an LETF option implied volatility to the corresponding ETF option implied volatility can be far from the LETF leverage ratio. The difference between the two depends on whether or not the LETF is long or short and is model dependent, thereby emphasizing the path dependence of the LETF price at any given time. In order to illustrate just how model dependent the prices of LETF options can be, we also price these options under the Barndorff-Nielsen and Shephard [3] model in addition to the three models listed above. This model dependency calls into question the market practice of pricing an LETF option using the Black-Scholes formula with the strike and implied volatility scaled by the leverage ratio.

Finally, it is worth emphasizing that our use of the word "consistent" in the title of this cahpter refers to *model* or *internal* consistency. In particular, rather than using separate models for pricing ETF options and LETF options our goal is to show how to consistently price these options at the model level only. We therefore do not claim that any one model can always price these options consistently with market prices. Indeed given the behavior of financial markets, we expect that the only models capable of always fitting to market prices are those models which have too many parameters and therefore tend to over-fit. Moreover, given the need to frequently re-calibrate even parsimonious models throughout the derivatives markets, we suspect that such models may never be found.

The remainder of this chapter is organized as follows. Section 2.2 describes our modeling assumptions for LETF price dynamics. In Sections 2.3, 2.4 and 2.5 we consider the SV, SVJ and SVCJ models, respectively, for the underlying ETF and describe how LETF options can be calculated for each of these models. Section 2.6 describes how we calibrated these models and Section 2.7 provides numerical results confirming the quality of our approximation. We conclude in Section 2.8. The appendices contain further details on our approximation methods as well as some additional numerical results. Our comments in the previous paragraph notwithstanding, in Appendix A.5 we provide a snapshot of how these models perform when they have been calibrated to market data. In particular, we will compare the model prices of LETF options with the corresponding market prices when the models have been calibrated to the market prices of ETF options. We will see that (at least on the day in question) the calibrated models produced very accurate prices for LETF options.

2.2 Modeling Leveraged ETF Dynamics

We let S_t and L_t denote the time t prices of the underlying ETF and LETF, respectively. Rather than working in discrete time we will work instead in continuous-time and assume that the LETF is re-balanced continuously.

Modeling Leveraged ETF Dynamics When the Underlying Has Diffusion Dynamics

If S_t follows a diffusion then the mechanics of the LETF implies that L_t has dynamics

$$\frac{dL_t}{L_t} = \phi \cdot \frac{dS_t}{S_t} + (1-\phi)r\,dt - f\,dt \tag{2.1}$$

where r is the continuously compounded risk-free interest rate and f is the constant expense ratio of the LETF. There is no difficulty incorporating dividends as long as we interpret the dS_t term in (2.1) to include any dividend payments. The $(1-\phi)rdt$ term in (2.1) reflects the cost of funding the leveraged position when $\phi > 1$, or the risk-free income from an inverse ETF when $\phi < 0$.

Assuming general diffusion dynamics of the form $dS_t = \mu_t S_t\ dt + \sigma_t S_t\ dW_t$, Avellaneda and Zhang [2] solved[1] (2.1) to obtain

$$\frac{L_T}{L_0} = \left(\frac{S_T}{S_0}\right)^{\phi} \exp\left((1-\phi)rT - fT + \frac{1}{2}\phi(1-\phi)\int_0^T \sigma_t^2 dt\right). \tag{2.2}$$

They used this expression to explain the empirical performance of LETFs during the financial crisis. Note that for leverage ratios satisfying $|\phi| > 1$ it is clear from (2.2) that a long LETF position is short realized variance for a given value of S_T. Haugh and Jain [21] also derived a more general form of (2.2) in a dynamic portfolio optimization context. It is also easy to show that this negative exposure to variance could be interpreted as a (multiplicative) premium that must be paid for obtaining a payoff of $(S_T/S_0)^{\phi}$ rather than the payoff you would obtain from a buy-and-hold portfolio with initial leverage of ϕ.

Modeling Leveraged ETF Dynamics When the Underlying Can Jump

Note also that if S_t can jump then (2.1) will still be valid as long as we truncate the jumps appropriately to reflect the limited liability of the LETF. But of course the LETF manager must implicitly pay for the truncation of these jumps since otherwise an arbitrage opportunity would exist. When the underlying price process can jump we therefore assume dynamics for L_t of the

[1] Cheng and Madhavan [9] obtained (2.2) under geometric Brownian motion dynamics.

form

$$\frac{dL_t}{L_{t-}} = \phi \cdot \frac{dS_t^*}{S_{t-}} + (1-\phi)r dt - f dt - c_t dt \tag{2.3}$$

where dS_t^* denotes the possibly truncated increment in the underlying price at time t and $c_t dt$ is the insurance premium paid at time t to insure against L_t violating limited liability in the next dt units of time. This premium can be computed directly by calculating the (risk-neutral) expected loss per unit time that the insurer would assume due to a possible jump in L_t to a negative value. We can also calculate c_t implicitly as the drift adjustment that ensures the discounted value of the gains process associated with L_t is a martingale under our risk-neutral probability measure. Note that we can also write (2.3) more explicitly as

$$\frac{dL_t}{L_{t-}} = \phi \cdot \frac{dS_t}{S_{t-}} + (1-\phi)r dt - f dt - c_t dt \quad \text{for } 0 \le t < \tau \tag{2.4}$$

where τ is the first-passage time of the event $\phi\, dS_t/S_{t-} \le -1$. Moreover we assume $L_t \equiv 0$ for all $t \ge \tau$.

2.2.1 Risk-Neutral Dynamics for the Leveraged ETF

The dynamics in (2.1) to (2.4) are all P-dynamics where P is the objective or empirical probability measure. But of course in order to price derivative securities we need to work with an equivalent martingale measure, Q. We will take Q to be the risk-neutral probability measure associated with the cash account as numeraire. We will also assume that the risk-free rate, r, is a constant[2] but note that it would be straightforward to relax this assumption if necessary.

All of our examples in this chapter will assume that the underlying security price has risk-neutral

[2] Given the short expirations that are typical for LETF options the assumption of constant interest rates is easy to justify.

dynamics of the form

$$\frac{dS_t}{S_{t-}} = (r - q - \lambda m)dt + \sqrt{V_t}dW_t^S + dJ_t, \tag{2.5}$$

where q is the dividend yield, λ is the intensity of the jump process, J_t, and V_t is some stochastic volatility process. We will write $J_t := \sum_{i=1}^{N_t}(Y_i - 1)$ so that $Y_i - 1$ represents the relative jump size in the security price at the time of the i^{th} jump. In particular, if the ith jump occurs at time τ_i, then $S_{\tau_i} = S_{\tau_i -} Y_i$. We set $m = \mathbb{E}^Q(Y_i - 1)$ which guarantees that the discounted gains process associated with holding the underlying security is a Q-martingale.

Some simple algebra confirms that jumps, Y_i, in the underlying security that satisfy $\phi(Y_i - 1) < -1$ would cause the LETF to go negative in the absence of limited liability. In the presence of limited liability we must therefore use a jump process for L_t of the form, $J_t^L := \sum_{i=1}^{N_t}(Y_i^L - 1)$ where

$$Y_i^L := \max\left(\phi(Y_i - 1),\ -1\right) + 1.$$

Continuing our insurance analogy, we could imagine the leveraged ETF investor being exposed to all jumps, $\phi(Y_i - 1)$, but that in addition he must insure against any jumps that would cause L_t to go negative. The risk-neutral value of this insurance per unit time[3] is then given by

$$\begin{aligned} c_t &:= \lambda p^* \left(\mathbb{E}^Q[-1 - \phi(Y_i - 1) \,|\, \phi(Y_i - 1) < -1]\right) \\ &= -\lambda p^* \left(\mathbb{E}^Q[\phi(Y_i - 1) \,|\, \phi(Y_i - 1) < -1] + 1\right) \end{aligned} \tag{2.6}$$

where $p^* := P(\phi(Y_i - 1) < -1)$ so that λp^* is the arrival rate for jumps that will drive L_t to zero. The "+1" term on the right-hand-side of (2.6) is required because the insurance only covers that part of the jump *beyond* −1 and indeed the jump event itself will drive the LETF price, L_t, to 0.

[3] Because λ is a constant in our examples c_t is in fact a constant. We could, however, also use our approach for more general point processes such as affine processes which are also tractable.

Substituting (2.6) and (2.5) into (2.4) and also taking the insurance payoff, dJ_t^{ins} say, into account we obtain the following risk-neutral dynamics for L_t

$$\begin{aligned}\frac{dL_t}{L_{t-}} &= \phi\Big((r-q-\lambda m)dt+\sqrt{V_t}dW_t^S+dJ_t\Big)+(1-\phi)rdt-fdt-c_tdt+dJ_t^{ins} \qquad (2.7)\\ &= (r-\phi q-f-\lambda\phi m)dt+\phi\sqrt{V_t}dW_t^S+dJ_t^L-c_tdt\\ &= (r-\phi q-f-\lambda m_L)dt+\phi\sqrt{V_t}dW_t^S+dJ_t^L \qquad (2.8)\end{aligned}$$

where we have used the fact that $\phi dJ_t + dJ_t^{ins} = dJ_t^L$ and used (2.6) and $m = \mathbb{E}^Q(Y_i - 1)$ to obtain

$$\begin{aligned}m_L &:= \phi m + c_t/\lambda\\ &= (1-p^*)\cdot\mathbb{E}^Q[\phi(Y_i-1)|\phi(Y_i-1)>-1]-p^*. \qquad (2.9)\end{aligned}$$

Note that these dynamics are only valid for $0 \le t \le \tau$ and that (2.7) does not contradict (2.4) since the dJ_t^{ins} term (which is absent in (2.4)) is only non-zero at time τ.

In the foregoing analysis we have implicitly assumed that dividends from the underlying ETF will be multiplied by ϕ and then paid out, in the case where ϕ is positive, to investors in the corresponding LETF. If ϕ is negative then the LETF investor will have to pay out these dividends. We make this assumption in order to simplify the exposition but note that in practice the treatment of dividends can vary with each LETF. For example, inverse LETFs with $\phi < 0$ typically have a dividend yield of zero and do not require their investors to make dividend payments while positively leveraged ETFs typically pay a smaller dividend than ϕq. Moreover, because leveraged ETFs often have other sources of income, e.g. interest income from the proceeds of short sales, understanding dividend dynamics needs to be done on a case-by-case basis. We do note that it is also possible to infer an implied LETF dividend yield in the usual manner using put-call parity. For the purpose of this chapter, however, we will assume a dividend yield of $\phi q + f$ as implied by (2.8) and simply note that it would be straightforward to handle other dividend assumptions.

The Path Dependence of Leveraged ETFs

While clear from (2.2) in the case of a diffusion, it is worth emphasizing that the risk-neutral dynamics of (2.8) yield a terminal value of L_T that is path-dependent. In particular L_T cannot be expressed as a function of S_T and so pricing an option on L_T does not amount to simply pricing some derivative of S_T.

2.3 Heston's Stochastic Volatility Model

The first model that we consider is Heston's[24] stochastic volatility (SV) model and we will see that it is particularly easy to price LETF options under this model. We assume the underlying ETF price, S_t, has risk-neutral dynamics given by

$$\frac{dS_t}{S_t} = (r-q)dt + \sqrt{V_t}dW_t^S, \tag{2.10}$$

$$dV_t = \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dW_t^V \tag{2.11}$$

where q is the dividend yield and W_t^S and W_t^V are standard Brownian motions with constant correlation parameter, ρ. Our first result is particularly straightforward and states that if S_t has Heston dynamics then so too does L_t.

Proposition 1. *Suppose the underlying ETF price S_t, has Heston dynamics given by* (2.10) *and* (2.11)*. Then assuming a leverage ratio of ϕ, the LETF price, L_t, has dynamics given by*

$$\frac{dL_t}{L_t} = (r-q_L)dt + sign(\phi)\cdot\sqrt{V_t^L}dW_t^S \tag{2.12}$$

$$dV_t^L = \kappa_L(\theta_L - V_t^L)dt + \gamma_L\sqrt{V_t^L}dW_t^V \tag{2.13}$$

where $V_t^L := \phi^2 V_t$, $q_L := \phi q + f$, $\kappa_L := \kappa$, $\gamma_L := |\phi|\gamma$ and $\theta_L := \phi^2\theta$. In particular the LETF also has Heston dynamics.

Proof : Since S_t follows a diffusion we note that (2.1) and (2.3) are identical. If we therefore substitute (2.10) into (2.3) we obtain $dL_t/L_t = (r - \phi q)dt + \phi\sqrt{V_t}dW_t^S$ which immediately yields (2.12). Similarly, using (2.11) we obtain $dV_t^L = \phi^2\kappa(\theta - V_t)dt + \phi^2\gamma\sqrt{V_t}dW_t^V$ which yields (2.13). ■

Proposition 1 shows that if S_t has Heston dynamics with parameter set $(q, \kappa, \gamma, \theta, V_0, \rho)$ then L_t has Heston dynamics with parameter set

$$(q_L, \kappa_L, \gamma_L, \theta_L, V_0^L, \rho_L) := (\phi q + f, \kappa, |\phi|\gamma, \phi^2\theta, \phi^2 V_0, \text{sign}(\phi) \cdot \rho). \tag{2.14}$$

Since it is easy to price options using transform methods under the Heston model, Proposition 1 implies that we can price options on ETFs and LETFs consistently with each other when the ETF has Heston dynamics. While this result was very easy to derive we have not seen it elsewhere in the literature. In his PhD thesis, for example, Zhang [38] considers the pricing of LETF options when the underlying has Heston dynamics. He does not observe that L_t also has Heston dynamics, however, probably because he worked with (2.2) rather than (2.1). Indeed Zhang proposed a change of measure motivated by (2.2) and observed that L_t had Heston dynamics with time-dependent parameters under this new measure. The time-dependency of the parameters under the new measure does not allow options on the LETF to be calculated via transform methods, however.

One further remark is in order at this point. It should be clear that the tractability that the LETF dynamics inherits from the underlying price dynamics will hold for diffusions in general and not just the Heston model. This should be clear from (2.1).

2.4 The SVJ Model

The Bates [5] stochastic volatility (SVJ) model is an extension of the SV model that allows for the possibility of jumps in the security price process. The risk-neutral dynamics for the SVJ model are

$$\frac{dS_t}{S_{t-}} = (r - q - \lambda m)dt + \sqrt{V_t}dW_t^S + dJ_t, \tag{2.15}$$

$$dV_t = \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dW_t^V \tag{2.16}$$

where W_t^S and W_t^V are standard Brownian motions with correlation coefficient ρ, N_t is a Poisson process with intensity λ, and $J_t := \sum_{i=1}^{N_t}(Y_i - 1)$ so that $Y_i - 1$ represents the relative jump size in the security price at the time of the i^{th} jump. In particular, if the ith jump occurs at time τ_i, then $S_{\tau_i} = S_{\tau_i-}Y_i$. The Y_i's are assumed to be IID log-normally distributed with $\log Y_i \sim \mathrm{N}(a, b^2)$ with $m := \mathbb{E}(Y_i - 1) = \exp\left(a + \frac{b^2}{2}\right) - 1$.

Proposition 2. *If S_t has risk-neutral dynamics given by* (2.15) *and* (2.16) *then the risk-neutral dynamics of L_t satisfy*

$$\frac{dL_t}{L_{t-}} = (r - q_L - \lambda m_L)dt + sign(\phi)\sqrt{V_t^L}dW_t^S + dJ_t^L \tag{2.17}$$

$$dV_t^L = \kappa_L(\theta_L - V_t^L)dt + \gamma_L\sqrt{V_t^L}dW_t^V \tag{2.18}$$

where $V_t^L := \phi^2 V_t$, $q_L := \phi q + f$, $\kappa_L := \kappa$, $\gamma_L := |\phi|\gamma$, $\theta_L := \phi^2\theta$, and $J_t^L := \sum_{i=1}^{N_t}(Y_i^L - 1)$ where

$$\begin{aligned} Y_i^L &:= \max\left(\phi(Y_i - 1),\ -1\right) + 1 \\ m_L &:= (1 - p^*) \cdot \mathbb{E}^Q[\phi(Y_i - 1) | \phi(Y_i - 1) > -1] - p^* \\ \text{and} \quad p^* &:= P(\phi(Y_i - 1) < -1) = \begin{cases} F\left(\log\left(\frac{\phi-1}{\phi}\right); a, b\right), & \text{if } \phi > 0 \\ 1 - F\left(\log\left(\frac{\phi-1}{\phi}\right); a, b\right), & \text{if } \phi < 0 \end{cases} \end{aligned} \tag{2.19}$$

where $F(\,\cdot\,; a, b)$ is the CDF of the $N(a, b)$ distribution.

Proof : (2.17) follows from (2.8). Since $V_t^L := \phi^2 V_t$ it is also clear that (2.18) follows directly from (2.16). ■

We would like to price options on the LETF using standard transform methods based on calculating the characteristic function of the log-LETF price. Since the diffusion component of the LETF dynamics remains Heston, this component is easy to handle. Truncating the jumps to preserve limited liability, however, means that the tractability of the jump component in (2.15) has been lost in (2.17). We will approach this problem by approximating the jump process in (2.17) with a more tractable jump-process.

But first, we will distinguish between two types of jumps. We say that a jump, Y, is of type I if it satisfies $\max(\phi(Y-1)+1, 0) = 0$. Such a jump would drive L_t to zero. Otherwise, it is of type II. A jump is type I with probability p^* and type II with probability $1-p^*$ where p^* is defined in (2.19). Let $N_1(t)$ and $N_2(t)$ denote respectively the number of type I and type II jumps occurring in $[0,t]$. By the thinning property of Poisson processes, $N(t) = N_1(t) + N_2(t)$ where $N_1(t)$ and $N_2(t)$ are independent Poisson processes with rates λp^* and $\lambda(1-p^*)$, respectively. We then have the following proposition.

Proposition 3. *Let $C(L_0, K, T)$ be the time $t = 0$ price of a call option on the LETF with strike K, maturity T and initial LETF price, L_0. Then $C(L_0, K, T) = \exp(-\lambda p^* T)\, \hat{C}(L_0, K, T)$ where*

$$\hat{C}(L_0, K, T) \;:=\; \mathbb{E}_0^Q[e^{-rT}(L_T - K)^+ | N_1(T) = 0] \tag{2.20}$$

is the value of the option given that there are no type I jumps in $[0,T]$.

Proof: The proof is immediate once we note that a type I jump will cause the LETF price to immediately fall to zero so that the call option will expire worthless in that event. ■

We will compute LETF option prices[4] in the SVJ model by approximating $\hat{C}(L_0, K, T)$ and then using (2.20). We will approximate $\hat{C}(L_0, K, T)$ using numerical transform inversion methods[5] applied to the characteristic function of an approximation to the log-LETF price, L_t, *conditional* on $N_1(T) = 0$. As mentioned earlier, the dynamics of the LETF price has a Heston diffusion component which is independent from the jump component. Since we can compute the characteristic function of the log-security price under the Heston model, the only difficulty is in approximating the characteristic function of the jump component of the log-LETF price conditional on $N_1(T) = 0$.

Towards this end, first note that the characteristic function, $\Phi_{J_2^L(T)}$ say, of the jump component of the log-LETF price conditional on $N_1(T) = 0$ is given by

$$\begin{aligned}
\Phi_{J_2^L(T)}(u) &= \mathbb{E}_0^Q\left[\exp\left(iu \cdot \sum_{j=1}^{N(T)} \log(Y_j^L)\right) \middle|\ N_1(T) = 0\right] \\
&= \mathbb{E}_0^Q\left[\exp\left(iu \cdot \sum_{j=1}^{N_2(T)} \log(Y_j^L)\right)\right] \\
&= \exp\left[\lambda(1-p^*)T\left(\Phi_X(u) - 1\right)\right] \qquad (2.21)
\end{aligned}$$

where $\Phi_X(\cdot)$ is the characteristic function of X where

$$X \coloneqq \left(\log(\phi(Y-1)+1) \mid \phi(Y-1)+1 > 0\right). \qquad (2.22)$$

Since we don't have an analytic expression for $\Phi_X(\cdot)$ it follows from (2.21) that we can't compute an analytic expression for $\Phi_{J_2^L(T)}(\cdot)$. Instead we will approximate X with a random variable, $\hat{X}$, whose characteristic function, $\Phi_{\hat{X}}(\cdot)$, is computable analytically. We will therefore approximate

[4] Put prices can then be obtained from put-call parity.

[5] We use the Carr-Madan [8] Fourier inversion approach throughout this chapter.

the characteristic function of the log-LETF price conditional on $N_1(T) = 0$ with

$$\hat{\Phi}_L^{N_1 \equiv 0}(u) \;=\; \exp(-\lambda \hat{m} i u T)\, \Phi_T^{SV}(u\,; r, q_L, \kappa_L, \gamma_L, \theta_L, V_0^L, \rho_L, L_0) \times \exp\left[\lambda(1-p^*)T(\Phi_{\hat{X}}(u) - 1)\right] \tag{2.23}$$

where

$$\hat{m} := -p^* + (1-p^*)\mathbb{E}[\exp(\hat{X}) - 1]. \tag{2.24}$$

We will use $\hat{\Phi}_L^{N_1 \equiv 0}(\cdot)$ and transform methods to approximate $\hat{C}(L_0, K, T)$ as defined in (2.20). Note that our definition of $\hat{m}$ ensures that the (unconditional) gains process from holding the LETF under our approximate dynamics remains a martingale.

2.4.1 The Jump Distribution Approximation

As shown in Figure 2.1 below, the density function of X is skewed to the left irrespective of the sign of ϕ. We will therefore approximate X with

$$\hat{X} = \mathrm{N}(\hat{a}, \hat{b}) - \mathrm{Exp}(\hat{c}) \tag{2.25}$$

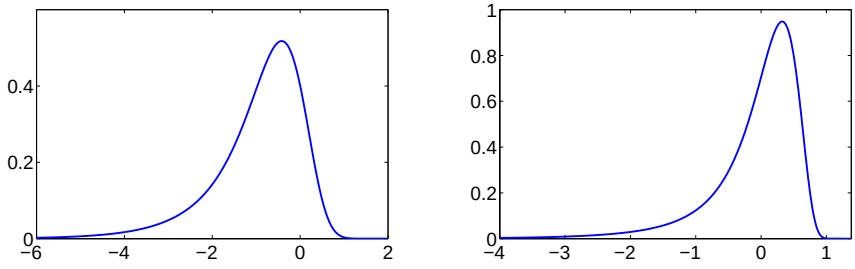
where the normal and exponential random variables are assumed to be independent. We choose $(\hat{a}, \hat{b}, \hat{c})$ by solving the following optimization problem

$$\begin{aligned} \min_{\hat{a},\hat{b},\hat{c}} \quad & \sum_{x \in S} |p(x) - q(x; \hat{a}, \hat{b}, \hat{c})|^2 \\ \text{subject to} \quad & \hat{b}, \hat{c} \geq 0 \end{aligned} \tag{2.26}$$

where S is a pre-specified set of points, and $p(\cdot)$ and $q(\cdot; \hat{a}, \hat{b}, \hat{c})$ are the density functions of X and $\hat{X}$, respectively. The details of this optimization problem can be found in Appendix A.2. Given the optimal solution, (a^*, b^*, c^*), to (2.26), we can then compute

$$\Phi_{\hat{X}}(u) \;=\; \exp\left(a^* i u - \frac{1}{2} {b^*}^2 u^2\right) \cdot \frac{c^*}{c^* + iu} \tag{2.27}$$

Figure 2.1: The Density Function of X



(a) $m = -0.2, b = 0.2, \phi = 3$ (b) $m = 0, b = 0.2, \phi = -3$

and by (2.24)

$$\hat{m} = -p^* + (1 - p^*)\left(\exp\left(a^* + {b^*}^2/2\right) \cdot \frac{c^*}{c^* + 1} - 1\right). \tag{2.28}$$

2.5 The SVCJ Model

The stochastic volatility model (SVCJ) with contemporaneous jumps in price and variance was introduced by Duffie, Pan and Singleton [15]. The risk-neutral dynamics for this model are

$$\frac{dS_t}{S_{t-}} = (r - q - \lambda m)dt + \sqrt{V_t}dW_t^S + dJ_t^S, \tag{2.29}$$

$$dV_t = \kappa(\theta - V_t)dt + \gamma\sqrt{V_t}dW_t^V + dJ_t^V \tag{2.30}$$

where $J_t^S := \sum_{i=1}^{N_t}(Y_i - 1)$, $J_t^V := \sum_{i=1}^{N_t} Z_i$ and N_t is a Poisson process with intensity λ. As before $Y_i - 1$ represents the percentage change in the security price due to the ith jump size and Z_i is the corresponding change in variance. In particular if the ith jumps occur at time τ_i, then $S_{\tau_i} = S_{\tau_i -} Y_i$ and $V_{\tau_i} = V_{\tau_i -} + Z_i$. We also assume the jumps in security price and variance are correlated. More precisely, we assume the Z_i's are exponentially distributed with mean, μ_v, and that conditional on

Z_i, $\log(Y_i)$, is normally distributed with mean, $a + \rho_J Z_i$, and variance, b^2. In other words,

$$Z_i \sim \text{Exp}(\mu_v^{-1}) \tag{2.31}$$

and

$$\log(Y_i) \;\sim\; \text{N}(a, b^2) + \rho_J Z_i \;\sim\; \text{N}(a, b^2) + \text{sign}(\rho_J) \cdot \text{Exp}(c) \tag{2.32}$$

where $c := |\rho_J \mu_v|^{-1}$ and the normal and exponential components in (2.32) are independent. We also have $\text{Corr}(Z_i, \log(Y_i)) = \text{sign}(\rho_J) \cdot ((bc)^2 + 1)^{-1/2}$ which approaches ± 1 as b goes to 0 and see that $m = \mathbb{E}(Y_i - 1) = \exp\left(a + b^2/2\right) c/(c - \text{sign}(\rho_J)) - 1$. Finally note that W_t^S and W_t^V are standard Brownian motions with constant correlation coefficient, ρ. We have the following proposition describing the risk-neutral dynamics of L_t in the SVCJ model.

Proposition 4. *If S_t has risk-neutral dynamics given by* (2.29) *and* (2.30) *then the LETF with leverage ratio ϕ has risk-neutral dynamics*

$$\frac{dL_t}{L_{t-}} = (r - q_L - \lambda m_L)dt + sign(\phi)\sqrt{V_t^L}dW_t^S + dJ_t^L \tag{2.33}$$

$$dV_t^L = \kappa_L(\theta_L - V_t^L)dt + \gamma_L\sqrt{V_t^L}dW_t^V + d(\phi^2 J_t^V) \tag{2.34}$$

where $V_t^L := \phi^2 V_t$, $q_L := \phi q + f$, $\kappa_L := \kappa$, $\gamma_L := |\phi|\gamma$, $\theta_L := \phi^2\theta$, $J_t^L := \sum_{i=1}^{N_t}(Y_i^L - 1)$ and

$$\begin{aligned} Y_i^L &:= \max\left(\phi(Y_i - 1),\, -1\right) + 1 \\ p^* &:= P(\phi(Y_i - 1) < -1) = \begin{cases} P\left(\log(Y_i) < \log\left(\frac{\phi-1}{\phi}\right)\right), & \text{if } \phi > 0 \\ P\left(\log(Y_i) > \log\left(\frac{\phi-1}{\phi}\right)\right), & \text{if } \phi < 0 \end{cases} \\ \text{and} \quad m_L &:= (1 - p^*) \cdot \mathbb{E}^Q[\phi(Y-1) | \phi(Y-1) > -1] - p^*. \end{aligned} \tag{2.35}$$

Proof : (2.33) follows directly from (2.8) and since $V_t^L := \phi^2 V_t$, (2.34) follows immediately from (2.30). ■

The question that now arises is whether or not we can price options on the LETF with dynamics given by (2.33) and (2.34). This appears difficult because truncating the jumps has rendered the model less tractable. As with the SVJ model we will proceed by approximating the dynamics of the LETF with more tractable dynamics. But first note that if we define Type I and Type II jumps as before then Proposition 3 remains valid[6] under the SVCJ model so that (2.20) still holds, i.e. $C(L_0, K, T) = \exp(-\lambda p^* T)\, \hat{C}(L_0, K, T)$ where $\hat{C}(L_0, K, T) := \mathbb{E}_0^Q[e^{-rT}(L_T - K)^+ | N_1(T) = 0]$ is the value of the option given that there are no type I jumps in $[0, T]$. Our goal will be to approximate the option price by approximating $\hat{C}(L_0, K, T)$. To do this let

$$\begin{aligned} X &:= \Big(\big(\log(\phi(Y_i - 1) + 1),\ \phi^2 Z_i\big) \mid \phi(Y_i - 1) + 1 > 0\Big) \\ &= \Big(\big(\log(Y_i^L),\ \phi^2 Z_i\big) \mid \phi(Y_i - 1) + 1 > 0\Big) \end{aligned} \tag{2.36}$$

be the bivariate random vector representing jumps in the log-LETF price and its variance process, respectively. We would like to have a closed-form expression for the characteristic function, $\Phi_X(\cdot)$, of X so that we could then apply the methodology of Duffie et al. [15] to compute the characteristic function of the log-LETF price. We don't have such a closed form expression, however, so we will instead approximate X with another bivariate distribution whose characteristic function is available in closed-form.

2.5.1 The Jump Approximation for X

We approximate X in (2.36) with

$$\hat{X} := (N - E_1, E_2) \tag{2.37}$$

[6] We also note that we could adapt Proposition 3 to handle the situation where the underlying jump process had an affine rather than constant intensity. To do this we would simply need to replace the $\exp(-\lambda p^* T)$ term with the (risk-neutral) probability of 0 type I jumps in $[0, T]$.

where $N \sim \mathrm{N}(\hat{a}, \hat{b})$ and $(E_1, E_2) \sim \mathrm{BVE}(\lambda_1, \lambda_2, \lambda_{12})$ have a bivariate exponential distribution (see [30]) that is independent of N. We are therefore using the same approximation that we used in (2.25) for the log-LETF price jumps in the SVJ model. But using the bivariate exponential distribution also allows us to approximate the variance jumps as well as the correlation between the two components of X in (2.36). In order to determine the parameters $(\hat{a}, \hat{b}, \lambda_1, \lambda_2, \lambda_{12})$ we could solve an optimization problem of the form

$$\begin{aligned} \min_{\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12}} \quad & \sum_{(x,y)\in S} |p(x,y) - q(x,y;\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12})|^2 \\ \text{subject to} \quad & \hat{b},\lambda_1,\lambda_2,\lambda_{12} \geq 0 \end{aligned} \tag{2.38}$$

where S is a pre-determined set of points, $p(\cdot,\cdot)$ is the joint density of X and $q(\cdot,\cdot;\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12})$ is the joint density function of $\hat{X}$ which we compute in Appendix A.3.3. Instead of solving (2.38), however, we prefer instead to use a three-step algorithm which we describe in Appendix A.3.4.

We note that as in the SVJ model, the optimization problems of (2.38) are not convex and so we are only guaranteed to find local mimina. Nonetheless, this was never a problem in our numerical experiments. Moreover, it would be trivial to consider different starting points for each such optimization problem and to only stop when the (squared) errors in (A.12), (A.13) and (A.14) are sufficiently small.

We use $\hat{X}$ rather than X when we model the dynamics of L_t. In order to maintain the martingale property of these dynamics, however, we replace m_L in (2.33) with

$$\begin{aligned} \hat{m} &:= -p^* + (1-p^*)\cdot\mathbb{E}[\exp(N-E_1)-1] \\ &= -p^* + (1-p^*)\left(\exp\left(a^* + \frac{1}{2}b^{*2}\right)\frac{\lambda_1^* + \lambda_{12}^*}{1+\lambda_1^*+\lambda_{12}^*} - 1\right). \end{aligned} \tag{2.39}$$

The characteristic function, $\Phi_{\hat{X}}(\cdot)$, of $\hat{X}$ is easily computed (see Appendix A.3.1) which means

we can employ the approach of Duffie et al. [15] to compute the characteristic function of the log-LETF price conditional on $N_1(T) = 0$. This characteristic function may then be used with the Carr-Madan [8] approach to approximate $\hat{C}(L_0, K, T)$. See Appendix A.3 for further details.

2.6 Model Calibration

We considered three different parameter sets for our numerical experiments. The first set was obtained by calibrating each of the models to 6-month call options on the underlying security in a low volatility environment. The call option strikes ranged from \$60 to \$140. The low volatility regime was characterized by a relatively flat skew and an at-the-money (ATM) volatility of approximately 20%. The second parameter set was obtained by calibrating the three models to 6-month call options on the underlying security in a high volatility environment with a steeper skew and an ATM volatility of approximately 72%. This high volatility environment was typical of the environment that prevailed at the height of the financial crisis of 2008. The third parameter set was obtained by calibrating each of the models to 1-month call option prices in the same high volatility environment that we used for the second parameter set. These environments can be seen in Figure 2.2 where we also assumed the underlying price, S_0, was \$100. To be clear, the three environments were not obtained from any real market data and therefore constitute an artificially created data-set. Nonetheless it is clear from Figure 2.2 that these environments are representative of what might be seen in practice.

In each of our models we assumed $r = 0.01$ and $q = f = 0$. With the exception of ρ, the remaining model parameters in each model were calibrated by minimizing the sum-of-squares between the Black-Scholes implied volatilities in the given environment and the Black-Scholes volatilities implied

by the model. The parameter ρ was fixed in advance as is commonly the case when calibrating Heston-style models. The reason for this is that it is well known that the sum-of-squares objective function tends to have so-called "valleys" or directions along which the objective function changes very little. This tends create a problem for optimization routines and for this reason it is common to fix ρ in advance to some sensible value.

The Heston Model: We set $\rho = -0.7571$. The remaining parameters, $(\kappa, \gamma, V_0 = \theta)$, were obtained by minimizing the sum-of-squares as described above.

The SVJ Model: We set $\rho = -0.7571$ and then solved for the remaining six parameters, $(\kappa, \gamma, V_0 = \theta, \lambda, m, b)$.

The SVCJ Model: We set $\rho = -0.82$ and then solved for the remaining eight parameters, $(\kappa, \gamma, V_0 = \theta, \lambda, m, b, \mu_v, \rho_J)$.

Table 2.1 displays the calibrated parameters for each of the models in the three environments while Figure 2.2 shows that all three models were calibrated successfully to the given implied volatilities in each of the three environments. (The SV model doesn't calibrate quite as well but this is to be expected given that it has fewer parameters than the SVJ and SVCJ models.)

2.7 Numerical Results

In this section we compute LETF options prices by applying the Carr-Madan transform approach to the characteristic function of the approximated log-LETF price. We will compare these approximate prices with exact prices obtained via Monte-Carlo. The advantage of the transform approach is that it is much faster than Monte-Carlo simulation and allows for the consistent calculation and risk-management of ETF and LETF options in real time.

Table 2.1: Model Parameters

Parameter set I (Calibrated to 6-month option prices in the low volatility environment.)

Parameters	SV Model	SVJ Model	SVCJ Model
Risk free rate r	0.01	0.01	0.01
Speed of mean reversion κ	10.95	0.5012	0.6097
Volatility of variance γ	0.2528	0.0895	0.0776
Long run mean variance θ	0.0421	0.0353	0.0393
Initial variance V_0	0.0421	0.0353	0.0393
Correlation ρ	-0.7571	-0.7571	-0.82
Jump arrival rate λ	n/a	1.0808	0.1406
m	n/a	-0.01	-0.0128
b	n/a	0.0745	0.1152
μ_v	n/a	n/a	0.01
ρ_J	n/a	n/a	0.0013

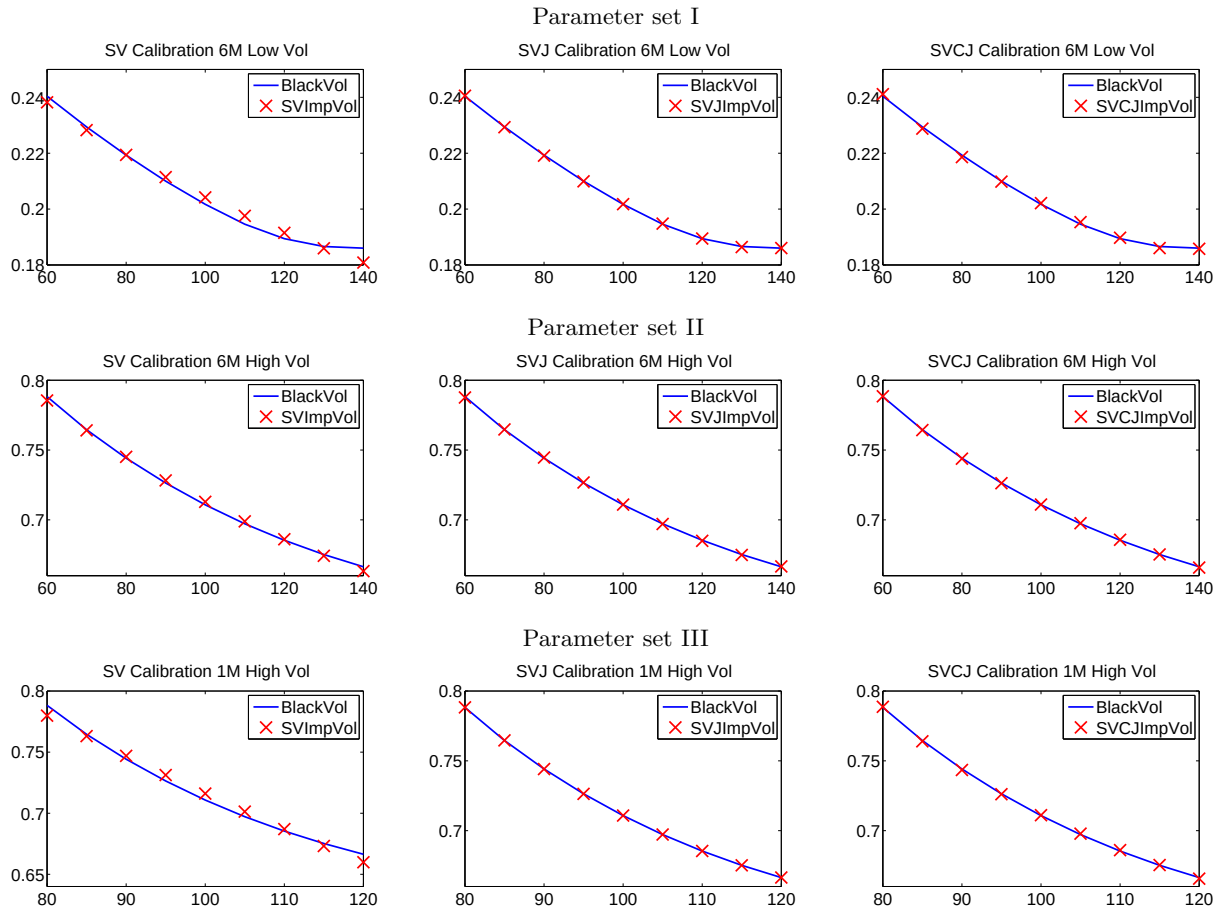
Parameter set II (Calibrated to 6-month option prices in the high volatility environment.)

Parameters	SV Model	SVJ Model	SVCJ Model
Risk free rate r	0.01	0.01	0.01
Speed of mean reversion κ	4.9498	0.6500	0.6500
Volatility of variance γ	1.1478	0.7895	0.3377
Long run mean variance θ	0.5505	0.3969	0.4048
Initial variance V_0	0.5505	0.3969	0.4048
Correlation ρ	-0.7571	-0.7571	-0.82
Jump arrival rate λ	n/a	2.1895	0.4996
m	n/a	-0.0105	-0.2592
b	n/a	0.2719	0.4588
μ_v	n/a	n/a	0.094
ρ_J	n/a	n/a	-0.2713

Parameter set III (Calibrated to 1-month option prices in the high volatility environment.)

Parameters	SV Model	SVJ Model	SVCJ Model
Risk free rate r	0.01	0.01	0.01
Speed of mean reversion κ	10.95	0.3632	0.5474
Volatility of variance γ	1.5086	0.6113	0.5730
Long run mean variance θ	0.5295	0.4156	0.4521
Initial variance V_0	0.5295	0.4156	0.4521
Correlation ρ	-0.7571	-0.7571	-0.82
Jump arrival rate λ	n/a	1.7483	0.5623
m	n/a	-0.1286	-0.2635
b	n/a	0.2384	0.3148
μ_v	n/a	n/a	0.0371
ρ_J	n/a	n/a	0.01

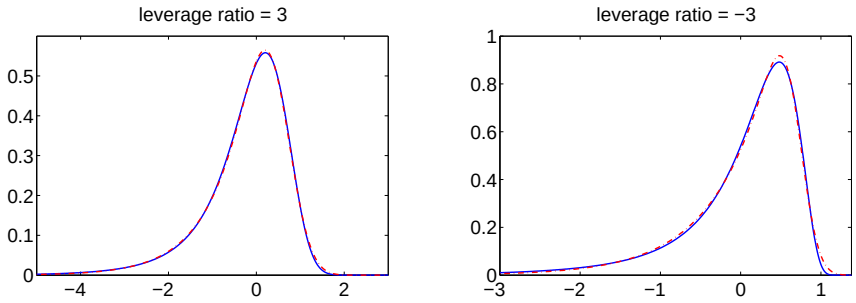
Figure 2.2: Volatility skews for the underlying ETF



2.7.1 Quality of Jump Distribution Approximation

We first evaluate the quality of our jump approximations in the SVJ and SVCJ models since it is the quality of these approximations that largely determines the accuracy of our approximate option prices in the SVJ and SVCJ models. In the case of the SVJ model we can visually compare the density of $\hat{X}$ with the density of X. Figure 2.3 displays these densities for a triple leveraged ETF with $\phi = \pm 3$, and parameter set II which corresponds to the high volatility environment. It is clear that our jump approximation is very accurate and that from a visual inspection there is

Figure 2.3: Jump approximations in the SVJ model for parameter set II. The PDF of $X = (\log(\phi(Y_i - 1) + 1) \mid \phi(Y_i - 1) + 1 > 0)$ is plotted as a continuous curve , and the PDF of $\hat{X}$ is plotted as a dotted curve.



little difference between the true density and the approximated density. These observations were also true for other values of ϕ and the other two parameter sets.

It is more difficult to display the quality of the jump approximation in the SVCJ model as we are dealing with two-dimensional random vectors in this case. Nonetheless our 3-step optimization routine ensures that the marginals are approximated very accurately. Moreover since the log-price marginal was fitted using the same distribution as our SVJ approximation, we obtained a similar good fit in the SVCJ case. We also, however, computed the integrated absolute difference in volumes between the true and approximated density functions using two-dimensional numerical integration. Table A.2 in Appendix A.4 displays these results. The area differences are much smaller than one, the total volume under each of the densities, and comparable to the observed errors for our SVJ approximation.

2.7.2 Computing Approximate LETF Option Prices

We now compare our approximate option prices with the prices obtained via Monte-Carlo. The Monte-Carlo prices were obtained using the scheme of Andersen [1] which was designed to simulate

the Heston model accurately. In the case of the SVJ and SVCJ models, we simply adapted Andersen to account for the independent jump processes. We assumed a time increment of $\Delta = 0.001$ which, assuming 250 trading days per year, corresponds to an interval of a quarter-day. Within the Monte-Carlo we assumed the LETF was re-balanced every 4 periods which is equivalent to the daily re-balancing that is performed in practice. Our first task is to compute option prices on the underlying ETF using the transform approach of Carr-Madan and Andersen's Monte-Carlo scheme. Note that options on the underlying can be priced exactly using the transform approach as the characteristic function of the log-ETF price is available in this case. The reason we compute option prices on the underlying ETF is simply to check that the two sets of prices agree modulo statistical error from the simulation and numerical inversion error from the Carr-Madan scheme.

Tables 2.2 and 2.3 display these ETF option prices for our three models under parameter sets II and III respectively. (All of our results for parameter set I, which corresponds to the low volatility 6-month environment, are deferred to Appendix A.4.) The Monte-Carlo results were based on simulating 10^8 sample paths which took several hours to run. We required this many paths to get sufficiently narrow confidence intervals so as to allow a comparison of the Monte-Carlo prices with the transform prices. It is clear from Tables 2.2 and 2.3 that both methods produce ETF option prices that effectively coincide with one another. Given this agreement we are now in a position to consider how well our approximate pricing of LETFs actually performs.

In Tables 2.4 and 2.5 we display prices of LETF options for parameter sets II and III, respectively, for each of our three models and various leverage ratios. The most computationally demanding task was the the 3-step optimization algorithm that we used to fit the jump approximation for the SVCJ model. Solving this optimization problem for a given leverage ratio, ϕ, and then pricing the options with three different strikes via Carr-Madan took less than a second when coded

Table 2.2: Option prices on underlying ETF for parameter set II computed via Monte-Carlo and transform approaches. Approximate 95% confidence intervals are reported in brackets.

Moneyness	BS vol (%)	BS price	Option price (SV)		Option price (SVJ)		Option price (SVCJ)	
$\frac{K}{S_0}$	Σ_{BS}	C_{BS}	C_{sim}	C_{tran}	C_{sim}	C_{tran}	C_{sim}	C_{tran}
0.75	75.44	33.66	33.66	33.66	33.66	33.66	33.65	33.65
			[33.65, 33.66]	-	[33.65, 33.67]	-	[33.65, 33.66]	
1	71.08	20.04	20.10	20.10	20.05	20.05	20.06	20.05
			[20.09, 20.11]	-	[20.04, 20.06]	-	[20.05, 20.06]	-
1.25	68.03	11.24	11.23	11.23	11.23	11.22	11.25	11.24
			[11.23, 11.24]	-	[11.22, 11.23]	-	[11.24, 11.25]	-

Table 2.3: Option prices on underlying ETF for parameter set III computed via Monte-Carlo and transform approaches. Approximate 95% confidence intervals are reported in brackets.

Moneyness	BS vol(%)	BS price	Option price (SV)		Option price (SVJ)		Option price (SVCJ)	
$\frac{K}{S_0}$	Σ_{BS}	C_{BS}	C_{sim}	C_{tran}	C_{sim}	C_{tran}	C_{sim}	C_{tran}
0.9	74.42	13.98	14.01	14.01	13.98	13.98	13.97	13.98
			[14.01, 14.01]	-	[13.98, 13.99]	-	[13.97, 13.98]	
1	71.08	8.04	8.10	8.10	8.05	8.04	8.05	8.05
			[8.10, 8.10]	-	[8.04, 8.05]	-	[8.04, 8.05]	-
1.1	68.54	4.09	4.10	4.10	4.09	4.09	4.10	4.10
			[4.10, 4.11]	-	[4.09, 4.09]	-	[4.09, 4.10]	-

in Matlab on a standard desktop computer. It was faster for each of the SVJ approximations and obviously faster still for the SV model which did not require any approximation.

Before analyzing the results, we first consider the possible sources of discrepancy between the reported Monte-Carlo prices and the prices obtained via numerical transform inversion. There are four such sources:

(i) Our Monte-Carlo assumed that the leveraged ETFs were re-balanced at a daily frequency as is the case in practice. The transform approach, however, implicitly assumes that the

Table 2.4: Comparing leveraged ETFs option prices with approximate prices in parameter set II. Approximate 95% confidence intervals are reported in brackets.

Leverage ratio	Moneyness		Option price (SV)		Option price (SVJ)		Option price (SVCJ)	
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	C_{sim}	C_{tran}	C_{sim}	C_{tran}	C_{sim}	C_{tran}
2	0.75	0.5	60.74	60.66	61.51	61.41	61.66	61.62
			[60.72, 60.76]	-	[61.49, 61.53]	-	[61.64, 61.68]	-
	1	1	37.87	37.78	38.43	38.41	38.50	38.50
			[37.86, 37.89]	-	[38.41, 38.45]	-	[38.48, 38.51]	-
	1.25	1.5	24.18	24.11	24.45	24.52	24.66	24.72
			[24.16, 24.19]	-	[24.43, 24.46]	-	[24.65, 24.68]	-
3	0.75	0.25	81.98	81.82	83.25	83.08	82.39	82.28
			[81.94, 82.01]	-	[83.21, 83.29]	-	[82.35, 82.43]	-
	1	1	53.09	52.77	54.07	53.98	52.93	52.83
			[53.05, 53.12]	-	[54.03, 54.11]	-	[52.90, 52.97]	-
	1.25	1.75	37.60	37.30	37.84	37.91	37.33	37.32
			[37.57, 37.63]	-	[37.80, 37.87]	-	[37.30, 37.36]	-
-1	0.75	1.25	14.15	14.14	13.93	13.87	12.63	12.59
			[14.14, 14.16]	-	[13.92, 13.94]	-	[12.62, 12.64]	-
	1	1	21.19	21.15	21.16	21.03	20.00	19.91
			[21.18, 21.20]	-	[21.15, 21.17]	-	[19.99, 20.01]	-
	1.25	0.75	32.79	32.73	33.24	33.01	32.25	32.11
			[32.78, 32.80]	-	[33.23, 33.25]	-	[32.24, 32.26]	-
-2	0.75	1.5	32.09	31.88	31.29	30.92	27.94	27.76
			[32.05, 32.14]	-	[31.24, 31.33]	-	[27.91, 27.97]	-
	1	1	41.95	41.68	41.60	41.13	38.81	38.56
			[41.90, 41.99]	-	[41.56, 41.64]	-	[38.78, 38.84]	-
	1.25	0.5	60.25	60.02	61.09	60.56	59.08	58.84
			[60.21, 60.30]	-	[61.04, 61.13]	-	[59.05, 59.11]	-
-3	0.75	1.75	51.48	50.72	49.43	48.47	44.41	43.86
			[51.19, 51.77]	-	[49.18, 49.69]	-	[44.31, 44.51]	-
	1	1	60.88	60.17	59.56	58.57	55.77	55.21
			[60.59, 61.18]	-	[59.31, 59.82]	-	[55.67, 55.87]	-
	1.25	0.25	81.74	81.46	82.59	81.82	80.65	80.39
			[81.45, 82.04]	-	[82.33, 82.85]	-	[80.55, 80.76]	-

Table 2.5: Comparing leveraged ETFs option prices with approximate prices in parameter set III. Approximate 95% confidence intervals are reported in brackets.

Leverage ratio	Moneyness		Option price (SV)		Option price (SVJ)		Option price (SVCJ)	
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	C_{sim}	C_{tran}	C_{sim}	C_{tran}	C_{sim}	C_{tran}
2	0.9	0.8	27.10	27.03	27.32	27.27	27.18	27.14
			[27.10, 27.11]	-	[27.32, 27.33]	-	[27.17, 27.19]	-
	1	1	15.98	15.94	16.06	16.04	15.97	15.96
			[15.98, 15.99]	-	[16.05, 16.06]	-	[15.96, 15.97]	-
	1.1	1.2	8.66	8.66	8.75	8.78	8.71	8.73
			[8.66, 8.67]	-	[8.74, 8.75]	-	[8.70, 8.71]	-
3	0.9	0.7	39.31	39.09	39.70	39.57	39.22	39.08
			[39.30, 39.32]	-	[39.69, 39.71]	-	[39.21, 39.23]	-
	1	1	23.65	23.49	23.78	23.73	23.45	23.38
			[23.65, 23.66]	-	[23.77, 23.79]	-	[23.44, 23.46]	-
	1.1	1.3	13.64	13.58	13.74	13.80	13.56	13.59
			[13.63, 13.65]	-	[13.74, 13.75]	-	[13.56, 13.57]	-
-1	0.9	1.1	4.81	4.84	4.54	4.57	4.57	4.59
			[4.81, 4.81]	-	[4.53, 4.54]	-	[4.57, 4.58]	-
	1	1	8.25	8.24	8.02	8.02	8.05	8.04
			[8.25, 8.25]	-	[8.02, 8.02]	-	[8.05, 8.06]	-
	1.1	0.9	13.58	13.53	13.48	13.45	13.50	13.44
			[13.57, 13.58]	-	[13.48, 13.49]	-	[13.49, 13.50]	-
-2	0.9	1.2	10.39	10.41	9.61	9.63	9.72	9.73
			[10.39, 10.40]	-	[9.60, 9.61]	-	[9.72, 9.73]	-
	1	1	16.58	16.49	15.96	15.87	16.06	15.95
			[16.58, 16.59]	-	[15.95, 15.97]	-	[16.05, 16.06]	-
	1.1	0.8	26.49	26.32	26.24	26.05	26.26	26.07
			[26.48, 26.50]	-	[26.23, 26.25]	-	[26.26, 26.27]	-
-3	0.9	1.3	16.72	16.61	15.19	15.13	15.44	15.34
			[16.71, 16.73]	-	[15.18, 15.20]	-	[15.43, 15.45]	-
	1	1	25.02	24.70	23.81	23.55	24.02	23.72
			[25.00, 25.03]	-	[23.80, 23.82]	-	[24.01, 24.03]	-
	1.1	0.7	38.74	38.31	38.23	37.84	38.31	37.90
			[38.73, 38.76]	-	[38.22, 38.25]	-	[38.30, 38.32]	-

re-balancing takes place continuously. We will see in Table 2.6 below that this is a principal source of the discrepancy between the Monte-Carlo prices and the approximate prices.

(ii) Numerical transform inversion is also a source of error but we believe this error to be very small and on the order of at most 1 or 2 cents. This claim is justified in part by the results in Tables 2.2 and 2.3.

(iii) Statistical error in the reported Monte-Carlo prices. We ensured this error was small by simulating sufficiently many paths so as to ensure that the approximate 95% confidence intervals were just 1 or 2 cents wide.

(iv) The fourth source is of course the errors that arise from our approximations of the jump size distributions in the SVJ and SVCJ models.

The main observation from Tables 2.4 and 2.5 is that the approximate LETF option prices as reported in the C_{tran} columns are very close to the reported Monte-Carlo prices. In particular, any discrepancy between the two should easily fall within the bid-ask spreads found in practice. We also note that the Monte-Carlo prices are generally higher than the transform-based prices. This is presumably due to the fact that prices computed via the transform approach are based on continuous re-balancing of the LETFs whereas the Monte-Carlo prices are computed assuming the LETF is re-balanced daily. We can confirm this observation by examining the option prices in Table 2.6 where we also report Monte-Carlo prices that were estimated assuming the LETF was re-balanced 4 times per day rather than just once per day. In that table we see that the Monte-Carlo prices based on re-balancing four times per day are generally much closer to the transform based prices. Presumably if we were to increase the LETF re-balancing frequency then the Monte-Carlo and transform-based prices would be in even closer agreement. These observations justify our earlier

observation that most of the discrepancy in Tables 2.4 and 2.5 between the Monte-Carlo prices and transform-based prices is due to the differences in re-balancing frequency rather than the quality of our jump approximations. As stated above, however, this discrepancy in prices is sufficiently small as to make little difference in practice. In contrast, we note that for the low-volatility environment of parameter set I, the Monte-Carlo prices were generally lower (by just 1 or 2 cents) than the transform prices. Again, if the LETF was re-balanced more frequently than daily re-balancing in the Monte-Carlo then we would expect this difference to be even smaller.

Another observation from Tables 2.4 and 2.5 is that there is some discrepancy in LETF option prices across the three different models. For example, in Table 2.4 we see that with $\phi = -3$ and $K_L/L_0 = 1.75$ the LETF call option price is approximately \$51, \$49 and \$44 under the SV, SVJ and SVCJ models, respectively. This is despite the fact that all three models were calibrated to the same 6-month implied volatilities. Of course, this observation is not too surprising as the LETF price is path-dependent and so it is not the case that the 6-month LETF option prices will only depend on the risk-neutral distribution of S_t where $t = 6$ months. This difference in LETF option prices across models is less noticeable in the 1-month options of parameter set III in Table 2.5. It is also worth pointing out that the 6-month LETF option prices vary very little by model in the low-volatility environment of parameter set I. These prices are displayed in Appendix A.4. We will return to this issue in Section 2.7.3.

2.7.3 Comparing the LETF Implied Volatilities Across Different Models

We now report the LETF option prices of Section 2.7.2 in terms of their Black-Scholes implied volatilities. We have already seen that there is some variability in these prices across the different models but it would be interesting to see this variability expressed in units of implied volatility.

Table 2.6: Comparison of LETF option prices obtained by Monte-Carlo simulation with different re-balancing frequencies in parameter set II. $C_{sim}^{(1)}$ corresponds to daily re-balancing and $C_{sim}^{(4)}$ corresponds to re-balancing 4 times per day. C_{tran} refers to prices that were obtained via numerical transform inversion.

Leverage ratio	Moneyness		Option price (SV)			Option price (SVJ)			Option price (SVCJ)		
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	$C_{sim}^{(1)}$	$C_{sim}^{(4)}$	C_{tran}	$C_{sim}^{(1)}$	$C_{sim}^{(4)}$	C_{tran}	$C_{sim}^{(1)}$	$C_{sim}^{(4)}$	C_{tran}
2	0.75	0.5	60.74	60.69	60.66	61.51	61.46	61.41	61.66	61.61	61.62
	1	1	37.87	37.81	37.78	38.43	38.39	38.41	38.50	38.46	38.50
	1.25	1.5	24.18	24.13	24.11	24.45	24.42	24.52	24.66	24.65	24.72
3	0.75	0.25	81.98	81.88	81.82	83.25	83.12	83.08	82.39	82.28	82.28
	1	1	53.09	52.87	52.77	54.07	53.92	53.98	52.93	52.80	52.83
	1.25	1.75	37.60	37.39	37.30	37.84	37.73	37.91	37.33	37.24	37.32
-1	0.75	1.25	14.15	14.14	14.14	13.93	13.90	13.87	12.63	12.62	12.59
	1	1	21.19	21.16	21.15	21.16	21.12	21.03	20.00	19.97	19.91
	1.25	0.75	32.79	32.74	32.73	33.24	33.18	33.01	32.25	32.21	32.11
-2	0.75	1.5	32.09	31.93	31.88	31.29	31.06	30.92	27.94	27.81	27.76
	1	1	41.95	41.74	41.68	41.60	41.33	41.13	38.81	38.65	38.56
	1.25	0.5	60.25	60.08	60.02	61.09	60.83	60.56	59.08	58.95	58.84
-3	0.75	1.75	51.48	50.80	50.72	49.43	48.59	48.47	44.41	43.99	43.86
	1	1	60.88	60.24	60.17	59.56	58.72	58.57	55.77	55.36	55.21
	1.25	0.25	81.74	81.43	81.46	82.59	81.96	81.82	80.65	80.48	80.39

This will also allow us to consider the commonly used practice of computing an LETF option price via the Black-Scholes formula with the implied volatility taken (and scaled appropriately) from a corresponding ETF option price. We will also introduce an additional model, namely the Barndorff-Nielsen and Shephard (BNS) model (see [3]), as this model helps to provide an even clearer demonstration of the fact that LETF option prices are strongly path dependent and are not uniquely determined by the implied volatility surface of the underlying ETF. We first describe the BNS model.

The BNS Model: The variance process is modeled by an Ornstein-Uhlenbeck process driven by a Levy process with non-negative increments. In particular we will assume that the variance process is a Gamma-OU process and that the risk-neutral dynamics for the security price and instantaneous variance are

$$\begin{aligned} d\log S_t &= \left(r - q - \frac{a\lambda\rho}{b-\rho} - \frac{V_t}{2}\right)dt + \sqrt{V_t}dW_t + \rho dz_{\lambda\cdot t} \\ dV_t &= -\lambda V_t dt + dz_{\lambda\cdot t} \end{aligned}$$

where W_t is a standard Brownian motion and z_t is a compound Poisson process with $z_t = \sum_{n=1}^{N_t} x_n$ where the Poisson process, N_t, has intensity, a, and the x_i's are IID exponential random variables with mean $1/b$. We also assume $\lambda > 0$ and $V_0 > 0$ so that (since $dz_{\lambda\cdot t}$ is always non-negative) $\inf_{0\le t\le T} V_t \ge \exp(-\lambda T) > 0$. The parameter, ρ, is typically negative to account for the negative correlation between variance and the underlying price process. Note that the variance can only jump upwards and that between jumps it decays exponentially. With a negative value of ρ the security price will jump downwards when a jump in variance occurs and it is worth noting in this case that leveraged ETFs with $\phi < 0$ can then only jump upwards in price. There is therefore no need to truncate the jumps of the LETF price process in this case and indeed the LETF price

process will itself have BNS dynamics. When $\phi > 0$ this will not be true as it will be necessary to truncate the jumps of the underlying ETF. In this case we could try to approximate the jump-process as we did with the SVJ and SVCJ models and then obtain approximate LETF option prices using transform methods. Rather than doing this, however, we will simply price the LETF options using Monte-Carlo because our goal in this section is to simply investigate how model dependent LETF option prices are.

Table 2.7: Comparison for the prices of options on the leveraged ETFs obtained by Monte-Carlo simulation in parameter sets II and III

Leverage ratio	Moneyness		Parameter Set II				Moneyness		Parameter Set III			
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	C_{sim}^{SV}	C_{sim}^{SVJ}	C_{sim}^{SVCJ}	C_{sim}^{BNS}	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	C_{sim}^{SV}	C_{sim}^{SVJ}	C_{sim}^{SVCJ}	C_{sim}^{BNS}
2	0.75	0.5	60.74	61.51	61.66	63.20	0.9	0.8	27.10	27.32	27.18	27.50
	1	1	37.87	38.43	38.50	40.55	1	1	15.98	16.06	15.97	16.26
	1.25	1.5	24.18	24.45	24.66	26.64	1.1	1.2	8.66	8.75	8.71	8.93
3	0.75	0.25	81.98	83.25	82.39	85.18	0.9	0.7	39.31	39.70	39.22	40.50
	1	1	53.09	54.07	52.93	57.81	1	1	23.65	23.78	23.45	24.62
	1.25	1.75	37.60	37.84	37.33	42.29	1.1	1.3	13.64	13.74	13.56	14.47
-1	0.75	1.25	14.15	13.93	12.63	11.53	0.9	1.1	4.81	4.54	4.57	4.39
	1	1	21.19	21.16	20.00	18.92	1	1	8.25	8.02	8.05	7.86
	1.25	0.75	32.79	33.24	33.25	31.48	1.1	0.9	13.58	13.48	13.50	13.35
-2	0.75	1.5	32.09	31.29	27.94	24.88	0.9	1.2	10.39	9.61	9.72	9.18
	1	1	41.95	41.60	38.81	36.15	1	1	16.58	15.96	16.06	15.50
	1.25	0.5	60.25	61.09	59.08	57.72	1.1	0.8	26.49	26.24	26.26	25.88
-3	0.75	1.75	51.48	49.43	44.41	39.00	0.9	1.3	16.72	15.19	15.44	14.36
	1	1	60.88	59.56	55.77	51.48	1	1	25.02	23.81	24.02	22.97
	1.25	0.25	81.74	82.59	80.65	79.48	1.1	0.7	38.74	38.23	38.31	37.64

We calibrated the BNS model to the same implied volatility skews of Figure 2.2 and note here that this calibration was performed successfully so that all four models agreed on the prices of options on the underlying ETF. This agreement can be seen in Tables 2.8 and 2.9 by noting that the columns labeled Σ_S are practically identical across the four models. The prices of the LETF options across the four models and various leverage ratios are displayed in Table 2.7. The same

results, except in terms of implied volatilities, are displayed in Tables 2.8 and 2.9 for parameter sets II and III, respectively. The LETF option implied volatilities are displayed in the columns labeled Σ_L and are calculated using the LETF option prices from Tables 2.4 and 2.5.

In order to compute the implied volatility ratios, $\frac{\Sigma_L}{\Sigma_S}$, we aligned the options on the underlying ETF and the leveraged ETF on a "strike-equivalent-basis" to account for the leverage. For example, we align a 25% OTM option on the underlying ETF with a 50% OTM option on a double-long LETF to account for the higher leverage of $\phi = 2$. We note that the implied volatility ratio tends to be close to the leverage ratio, ϕ, but that there can be a considerable discrepancy between the two. The degree of this discrepancy is model dependent and is very notable for the BNS model (which is why we have included the BNS model here). For a given model, it is also the case that whether or not the ratio $\frac{\Sigma_L}{\Sigma_S}$ is greater than ϕ depends on whether or not the LETF is positively or negatively leveraged. We emphasize again here that these observations are based on parameter set II which models the 6-month, high volatility environment.

2.8 Conclusions

We have shown how to obtain accurate LETF option prices via transform pricing methods for the Heston model as well as two related jump-diffusion models. Our approximations work well in both low and high volatility environments and because they are consistent with the prices of options on the underlying ETF, they permit consistent pricing and risk-management of derivatives portfolios containing both ETF and LETF options. It should also be clear that similar approximation techniques could be applied to other jump-diffusion models and so our examples should be viewed as applications of a more general approximation technique.

In addition to confirming the accuracy of our LETF option prices, our numerical experiments also showed that the ratio of an LETF option implied volatility to the corresponding ETF option implied volatility can be far from the LETF leverage ratio. The difference between the two depends on whether or not the LETF is long or short and is model dependent, thereby emphasizing the path dependence of the LETF price at any given time. This calls into question the market practice of pricing an LETF option using the Black-Scholes formula with the strike and implied volatility from the underlying ETF scaled by the leverage ratio. In particular, it should be clear that using the Black-Scholes formula in this manner amounts to the implicit assumption of (generally unspecified) dynamics for the underlying ETF.

Table 2.8: Comparison of Black-Scholes Implied-Volatilities: Parameter Set II

Leverage ratio	Moneyness		Implied Volatility (SV)			Implied Volatility (SVJ)			Implied Volatility (SVCJ)			Implied Volatility (BNS)		
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$
2	0.75	0.5	75.42	149.10	1.98	75.44	154.61	2.05	75.40	155.73	2.07	75.41	166.68	2.21
	1	1	71.28	139.11	1.95	71.10	141.35	1.99	71.13	141.62	1.99	71.09	149.98	2.11
	1.25	1.5	68.00	133.12	1.96	67.98	134.07	1.97	68.05	134.85	1.98	68.04	141.88	2.09
3	0.75	0.25		224.18	2.97		242.31	3.21		230.05	3.05		270.14	3.58
	1	1		204.20	2.86		208.76	2.94		203.50	2.86		226.66	3.19
	1.25	1.75		195.88	2.88		196.76	2.89		194.89	2.86		213.52	3.14
-1	0.75	1.25		78.48	1.04		77.69	1.03		73.04	0.97		69.07	0.92
	1	1		75.29	1.06		75.20	1.06		70.93	1.00		66.95	0.94
	1.25	0.75		71.11	1.05		73.36	1.08		68.39	1.01		64.44	0.95
-2	0.75	1.5		161.47	2.14		158.55	2.10		146.52	1.94		135.62	1.80
	1	1		155.70	2.18		154.26	2.17		142.90	2.01		132.23	1.86
	1.25	0.5		145.60	2.14		151.60	2.23		137.08	2.01		126.99	1.87
-3	0.75	1.75		250.11	3.32		241.67	3.20		221.67	2.94		201.07	2.67
	1	1		242.04	3.40		235.34	3.31		216.79	3.05		196.86	2.77
	1.25	0.25		220.82	3.25		232.95	3.43		205.07	3.01		187.56	2.76

Table 2.9: Comparison of Black-Scholes Implied-Volatilities: Parameter Set III

Leverage ratio	Moneyness		Implied Volatility (SV)			Implied Volatility (SVJ)			Implied Volatility (SVCJ)			Implied Volatility (BNS)		
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$
2	0.9	0.8	74.70	149.29	2.00	74.43	151.86	2.04	74.33	150.19	2.02	74.47	153.92	2.07
	1	1	71.59	142.28	1.99	71.09	142.98	2.01	71.10	142.18	2.00	71.09	144.82	2.04
	1.1	1.2	68.68	136.37	1.99	68.55	137.13	2.00	68.61	136.78	1.99	68.57	138.82	2.02
3	0.9	0.7		224.42	3.00		229.50	3.08		223.22	3.00		239.89	3.22
	1	1		212.51	2.97		213.65	3.01		210.62	2.96		221.45	3.12
	1.1	1.3		203.50	2.96		204.42	2.98		202.82	2.96		210.92	3.08
-1	0.9	1.1		75.38	1.01		72.78	0.98		73.14	0.98		71.41	0.96
	1	1		72.90	1.02		70.85	1.00		71.16	1.00		69.40	0.98
	1.1	0.9		70.06	1.02		69.04	1.01		69.18	1.01		67.55	0.99
-2	0.9	1.2		152.14	2.04		145.04	1.95		146.08	1.97		141.10	1.89
	1	1		147.73	2.06		142.09	2.00		142.95	2.01		137.94	1.94
	1.1	0.8		142.13	2.07		139.12	2.03		139.43	2.03		134.80	1.97
-3	0.9	1.3		231.00	3.09		217.39	2.92		219.57	2.95		209.93	2.82
	1	1		225.17	3.15		213.95	3.01		215.89	3.04		206.21	2.90
	1.1	0.7		216.96	3.16		210.24	3.07		211.25	3.08		202.31	2.95

Chapter 3

Linear Programming and the Control of Diffusion Processes

3.1 Introduction

Due to the so-called curse-of-dimensionality, solving high dimensional control problems is a notoriously difficult problem. It is not surprising then that sub-optimal control has been an active area of research for many years. Moreover, the advent of ever-increasing computational power has seen many developments[1] in the related area of *approximate dynamic programming* (ADP), particularly for discrete-time control problems. Linear programming (LP) methods have played an important role in the development of several ADP techniques, beginning with Schweitzer and Seidmann [34] and continuing with the important contributions of de Farias and Van Roy [12; 13] among others.

[1] See Bertsekas [6] for a comprehensive introduction to classical sub-optimal control techniques. Bertsekas [7] also contains an excellent treatment of approximate dynamic programming.

Recently, Han and Van Roy [19] proposed an LP-based approach for the approximate solution of the HJB equation that arises from continuous-time control problems. Their approach applies to diffusion problems with an exponentially-distributed horizon, T, and their numerical results were promising, with the LP-based policy outperforming other base-case policies. In this chapter we extend their approach to continuous-time control problems with a fixed horizon, T. We apply these techniques to dynamic portfolio optimization problems and then simulate the resulting policies to obtain primal, i.e. lower, bounds on the optimal value functions. We also use these policies in conjunction with the convex duality methodology[2] of Haugh, Kogan and Wang [22] (hereafter HKW) to construct dual, i.e. upper, bounds on the optimal value functions. By comparing the resulting primal and dual bounds we can easily assess the quality of the sub-optimal policy produced by the LP-approach. In our numerical experiments we find that the primal and dual bounds are very close and so we can conclude that, for these problems at least, the LP approach performs very well indeed.

The remainder of this chapter is organized as follows. In Section 3.2 we formulate the continuous-time portfolio optimization problem and also discuss here the exponentially distributed and fixed horizon versions of the problem. In Section 3.3 we review the approach of Han and Van Roy for approximately solving the HJB equation when the horizon is an exponentially distributed random variable. We extend their methodology to the fixed horizon case in Section 3.4 and our numerical results are presented in Section 3.5. We conclude in Section 3.6. The appendices contain additional details including an overview of the aforementioned dual approach of HKW.

[2]See also Haugh and Jain [21].

3.2 The Portfolio Optimization Problem Formulation

We formulate the dynamic portfolio[3] optimization problems that we will consider throughout the chapter. We will follow the formulation of HKW. There are N risky stocks and an instantaneously risk-free bond in a market. The vector of stock prices is denoted by $P_t = (P_{1t}, \cdots, P_{Nt})^\top$ and the instantaneously risk-free rate of return on the bond is denoted by r_t. Without loss of generality, we assume the stocks pay no dividends. Assets return dynamics depend on the M-dimensional vector of state variables, $Z_t = (Z_{1t}, \cdots, Z_{Mt})^\top$, so that

$$r_t = r(Z_t), \tag{3.1a}$$

$$\frac{dP_t}{P_t} = \mu_P(Z_t)dt + \Sigma_P(Z_t)dB_t, \tag{3.1b}$$

$$dZ_t = \mu_Z(Z_t)dt + \Sigma_Z(Z_t)dB_t \tag{3.1c}$$

where $B_t = (B_{1t}, \cdots, B_{Nt})^\top$ is an N-dimensional standard Brownian motion. $\mu_Z(Z_t)$ and $\mu_P(Z_t)$ are M- and N-dimensional drift vectors, while $\Sigma_Z(Z_t), \Sigma_P(Z_t)$ are $M \times N$ and $N \times N$ diffusion matrices of the state variable and security prices, respectively. We assume the diffusion matrix, $\Sigma_P(Z_t)$, of the asset return process is non-degenerate for each Z_t so that $x^\top \Sigma_P(Z_t)\Sigma_P(Z_t)^\top x \geq \epsilon \|x\|^2$ for all x and some $\epsilon > 0$. We can then define a process, η_t, according to

$$\eta_t(Z_t) := \Sigma_P(Z_t)^{-1}(\mu_P(Z_t) - r(Z_t) \cdot \mathbf{1})$$

where $\mathbf{1} = (1, \cdots, 1)^\top$. In a market without portfolio constraints, η_t corresponds to the market-price-of-risk process (e.g. Duffie 1996 [14], Section 6.G). We make the standard assumption that the

[3] We do note here, however, that the LP-based approach to solving the HJB equation applies to control problems in diffusion settings more generally.

process η_t is square integrable so that[4]

$$\mathbb{E}_0\left[\int_0^T \| \eta_t \|^2 \, dt\right] < \infty.$$

Under this opportunity set, our portfolio consists of positions in the N stocks and the risk-free bond. We also assume that continuous re-balancing of the portfolio is permitted and that $\theta_t(Z_t) \coloneqq (\theta_{1t}(Z_t), \cdots, \theta_{Nt}(Z_t))^\top$ is the vector of risky security weights in the portfolio at time t. To rule out arbitrage, we require the portfolio strategy to satisfy a square integrability condition, namely that $\int_0^T \|\theta\|^2 dt < \infty$. The value of the portfolio, W_t, associated with θ_t then changes according to the SDE:

$$\frac{dW_t}{W_t} = \left[r_t + \theta_t^\top \lambda_t\right]dt + \theta_t^\top \Sigma_{P_t} dB_t \tag{3.2}$$

where $\lambda_t \coloneqq \mu_{P_t} - r_t \cdot \mathbf{1}$.[5] We also assume that the portfolio is constrained so that

$$\theta_t(Z_t) \in \mathbf{K}, \tag{3.3}$$

for all t and where $\mathbf{K}$ is some fixed convex set containing zero.

The portfolio optimization problem is to choose a self-financing trading strategy that maximizes the expected utility of terminal wealth. The horizon, T, is assumed to be finite but it may be either random or deterministic, depending on the specific formulation under consideration. The utility function $u(W)$ is assumed to be strictly increasing, concave and smooth. Moreover, it is assumed to satisfy the Inada conditions at zero and infinity so that $\lim_{W\to 0} u'(W) = \infty$ and $\lim_{W\to\infty} u'(W) = 0$. In this chapter, we will use the constant relative risk aversion (CRRA) utility function so that

$$u(W) \coloneqq \frac{W^{1-\gamma}}{1-\gamma}$$

[4] We use $\mathbb{E}_t[\cdot]$ to denote an expectation conditional on time t information throughout the chapter.

[5] We use $r, \mu_P, \mu_Z, \Sigma_P, \Sigma_Z, \eta, \theta, \lambda$ (or $r_t, \mu_{P_t}, \mu_{Z_t}, \Sigma_{P_t}, \Sigma_{Z_t}, \eta_t, \theta_t, \lambda_t$) instead of $r(Z_t), \mu_P(Z_t), \mu_Z(Z_t), \Sigma_P(Z_t), \Sigma_Z(Z_t), \eta_t(Z_t), \theta_t(Z_t), \lambda(Z_t)$ for simplicity, if their meanings are clear.

with $\gamma \geq 1$.

3.2.1 When the Horizon, T, is Fixed

When the problem has a fixed horizon, T, the investor's portfolio optimization problem at time t is to solve for

$$J^*(w,z,t) = \sup_{\{\theta_s\}} \mathbb{E}_t\left[u(W_T)\right] \qquad \text{subject to} \quad (1),(2) \text{ and } (3) \tag{3.4}$$

where w and z are the wealth and state vector values at time t. A well-known implication of CRRA utility is that J^* is separable in w and (z,t) so that we can write $J^*(w,z,t) = u(w)V^*(z,t)$. The optimal strategy is therefore independent of the wealth process, w_t.

In order to write the HJB equation for this problem we first define the HJB operator

$$\begin{aligned} H_\theta V(z,t) \quad := \quad & (1-\gamma)V(z,t)(\theta^\top \lambda + r) + V_z(z,t)^\top \mu_Z(z) + \frac{1}{2}\gamma(\gamma-1)V(z,t)\theta^\top \Sigma_P \Sigma_P^\top \theta \\ & + (1-\gamma)V_z(z,t)^\top \Sigma_Z \Sigma_P^\top \theta + \frac{1}{2}\mathrm{tr}\left[V_{zz}(z,t)\Sigma_Z \Sigma_Z^\top\right] + V_t(z,t). \end{aligned} \tag{3.5}$$

Note that V_z is the M-dimensional gradient of V with respect to the state variable z. Similarly V_{zz} is the $M \times M$ Hessian matrix of V with respect to z. The HJB equation is then given by

$$0 = HV(z,t) := \inf_{\theta \in \mathbf{K}} H_\theta V(z,t) \tag{3.6}$$

and we note that $V^*(z,t) = J^*(w,z,t)/u(w)$ is a solution to this equation.

3.2.2 When the Horizon, T, is Exponentially Distributed

When the horizon T is an exponentially distributed random variable with mean τ, the investor's problem is identical to (3.4) but now with the understanding that the expectation must also be taken with respect to T. By first taking expectation with respect to T it is easy to see the problem

may also be formulated as

$$J^*(w,z) = \sup_{\{\theta_s\}} \mathbb{E}\Big[\int_{t=0}^{\infty} e^{-t/\tau} u(W_t)dt\Big] \qquad \text{subject to} \quad (1),(2) \text{ and } (3). \tag{3.7}$$

While t is no longer a state variable, J^* is still separable so we can again write $J^*(w,z) = u(w)V^*(z)$. The HJB operator for this problem is then defined as

$$\begin{aligned} H_\theta V(z) \;\coloneqq\; & (1-\gamma)V(z)(\theta^\top \lambda + r) + V_z(z)^\top \mu_Z(z) + \frac{1}{2}\gamma(\gamma-1)V(z)\theta^\top \Sigma_P \Sigma_P^\top \theta \\ & + (1-\gamma)V_z(z)^\top \Sigma_Z \Sigma_P^\top \theta + \frac{1}{2}\mathrm{tr}\big[V_{zz}(z)\Sigma_Z \Sigma_Z^\top\big] - \frac{V(z)}{\tau} + 1. \end{aligned} \tag{3.8}$$

and the HJB equation is given by

$$0 = HV(z) = \inf_{\theta \in \mathbf{K}} H_\theta V(z). \tag{3.9}$$

We note that $V^*(z) = J^*(w,z)/u(w)$ is a solution to this equation.

3.3 Review of Han and Van Roy's LP Approach

In this section we review Han and Van Roy's[19] LP approach for approximately solving (3.7) when the horizon T is exponentially distributed. In a standard argument they show the optimal solution, V^*, to the HJB equation (3.9) is also the unique optimum of the following static optimization problem:

$$\begin{aligned} \max_{V(z)} \quad & \int V(z)\rho(dz) \\ \text{subject to} \quad & H_\theta V(z) \geq 0, \quad \forall\ \theta, z \\ & V \in C^2 \end{aligned} \tag{$\mathcal{P}_1$}$$

where ρ is a pre-specified positive measure for the integral. While the objective and constraints in ($\mathcal{P}_1$) are linear the problem is still very challenging to solve as there are uncountably many decision

variables and constraints; indeed there is one constraint for every (θ, z) pair. We therefore solve an approximation to $(\mathcal{P}_1)$ and this approximation is obtained via the following steps:

1. We first choose a suitable set of basis functions $\{\phi_1(z), \cdots, \phi_k(z)\}$ with the goal of finding a linear combination, $\sum_{j=1}^{k} r_j \phi_j(z)$, that we will use to approximate $V^*(z)$. The original problem then reduces to the problem of solving for k decision variables, $r_1, r_2, \cdots, r_k$. The algorithm is initialized with a predetermined weight vector, $r^{(0)} = (r_1^{(0)}, \cdots, r_k^{(0)})$.

2. We generate a finite sample set $z_1, \ldots, z_Q$ and approximate the integral in $(\mathcal{P}_1)$ by a corresponding finite sum of Q terms. While any positive measure, ρ, can be used in theory, the performance of the algorithm depends on how the samples are generated. Han and Van Roy define

$$\rho(dz) := \frac{1}{\tau}\, \mathbb{E}\left[\int_{t=0}^{\infty} e^{-t/\tau} \mathbf{1}_{\{Z_t \in [z, z+dz]\}} dt \right]$$

and then generate $z_1, \ldots, z_Q$ by simulating (approximately) from this measure. In particular, they first simulate the horizon $T \sim \text{Exp}(1/\tau)$ and then simulate a discrete-time approximation to the dynamics of the state variables (3.1c). The value of the state vector at the simulated time T is taken as one of our Q samples.

3. For each sample, z_j, we choose a single corresponding θ_j as follows: given an approximation, $\Phi r := \phi_1 r_1 + \cdots + \phi_k r_k$, to V^*, we myopically choose a greedy action with respect to Φr. That is, we select $\theta_j \in \underset{\theta \in \mathbf{K}}{\text{argmin}}\ H_\theta\left[(\Phi r)(z_j)\right]$ for each j.

4. Given a weight vector r, we find a new weight vector r' by solving an approximation of $(\mathcal{P}_1)$ (see phase 2 of the algorithm below); this approximation is a linear program that we obtain from steps 1 to 3.

Steps 3 and 4 are repeated to obtain a sequence of weight vectors, $\left\{r^{(0)}, r^{(1)}, r^{(2)}, \cdots\right\}$. We are now ready to define the *adaptive constraint selection algorithm* of Han and Van Roy:

Adaptive Constraint Selection Algorithm I

---

for $i = 1$ *to* ∞ *do*

for $j = 1$ *to* Q *do*

$$\theta_j \in \quad \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \; H_\theta \left[(\Phi r^{(i-1)})(z_j) \right] \tag{phase 1}$$

end for

$$r^{(i)} \in \quad \underset{r \in \mathbb{R}^k}{\operatorname{argmax}} \sum_{j=1}^{Q} (\Phi r)(z_j) \tag{phase 2}$$

$$\text{subject to} \quad H_{\theta_j} \left[(\Phi r)(z_j) \right] \geq 0, \quad \forall j = 1, \cdots, Q$$

end for

This algorithm does not necessarily generate an optimal solution to $(\mathcal{P}_1)$ and there is no[6] theoretical guarantee that the sequence $r^{(i)}$ will converge. However, if it does converge then Han and Van Roy[19] show it must converge to an optimal solution of the following problem which is an approximation of $(\mathcal{P}_1)$:

$$\begin{aligned} \max_r \quad & \sum_{j=1}^{Q} (\Phi r)(z_j) \\ \text{subject to} \quad & H \left[(\Phi r)(z_j) \right] \geq 0, \quad \forall j = 1, \cdots, Q. \end{aligned} \tag{3.10}$$

We now discuss in further detail the steps required to execute phases 1 and 2.

[6] Han and Van Roy did not specify how they handled non convergence. We suspect they simply used a different starting point, $r^{(0)}$, if that occurred.

Phase 1

We can expand the objective of phase 1 using (3.8). If we then remove terms that do not depend on θ and eliminate the common factor, $(\gamma - 1)$, then the problem of phase 1 can be expressed as:

$$\theta_j \in \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \frac{1}{2}\theta^\top \big[\gamma(\Phi r^{(i-1)})(z_j)\Sigma_P \Sigma_P^\top\big]\theta - \Big[(\Phi r^{(i-1)})(z_j)\lambda(z_j) + \Sigma_P \Sigma_Z^\top (\Phi r^{(i-1)})_z(z_j)\Big]^\top \theta. \quad (3.11)$$

If $(\Phi r^{(i-1)})(z_j) > 0$, then (3.11) is a convex quadratic program and therefore easy to solve. Otherwise the objective in (3.11) may be unbounded if $\mathbf{K}$ is not compact and some other heuristic approach for computing θ_j would be required.

Phase 2

If we expand the constraints of the LP in phase 2 using (3.8), then we obtain the following LP:

$$\begin{aligned} \max_{r \in \mathbb{R}^k} \quad & c^\top r \\ \text{subject to} \quad & Ar \geq -\mathbf{1} \end{aligned} \quad (3.12)$$

where

$$\begin{aligned} A_{ij} :=& (1-\gamma)\phi_j(z_i)\big[\theta_i^\top \lambda(z_i) + r(z_i)\big] + \big[(\phi_j)_z(z_i)\big]^\top \mu_Z(z_i) + \frac{1}{2}\gamma(\gamma-1)\phi_j(z_i)\theta_i^\top \Sigma_P \Sigma_P^\top \theta_i \\ & + (1-\gamma)\Big[(\phi_j)_z(z_i)\Big]^\top \Sigma_Z \Sigma_P^\top \theta_i + \frac{1}{2}\mathrm{tr}\Big[(\phi_j)_{zz}(z_i)\Sigma_Z \Sigma_Z^\top\Big] - \frac{\phi_j(z_i)}{\tau} \end{aligned}$$

and $c_i := \phi_i(z_1) + \phi_i(z_2) + \cdots + \phi_i(z_Q)$. This linear program has a k-dimensional decision vector and Q linear constraints.

3.4 Extending the LP Approach to the Case of a Fixed Horizon, T

The LP approach of the previous section applies to problems with an exponentially distributed horizon, T, but we would also like to apply it to the case of a fixed horizon. As we shall see, this extension is not immediate and requires some work due to the fact that time t is also a state variable in this case. We show in Appendix B.1 that under some technical conditions, that the solution to the HJB equation (3.6), V^*, is also the unique optimum to the following optimization problem:

$$\begin{aligned} \max_{V} \quad & \int V(z,t)\rho(dz,dt) \\ \text{subject to} \quad & H_\theta V(z,t) \geq 0, \quad \forall\ \theta, z, t \\ & V(z,T) \leq 1, \quad \forall\ z \qquad \text{(boundary condition)} \\ & V \in C^2 \end{aligned} \tag{$\mathcal{P}_2$}$$

where ρ is again some pre-specified measure. In contrast to problem ($\mathcal{P}_1$), the boundary condition $V(z,T) \leq 1$ is required in this case.

The extension of the adaptive constraint selection algorithm seems straightforward: we choose basis functions $\{\phi_1(z,t), \cdots, \phi_k(z,t)\}$ to approximate $V^*(z,t)$ and generate $\{(z_1,t_1),(z_2,t_2),\cdots,(z_Q,t_Q)\}$ as a representative sample of (z,t). As a simple heuristic for generating this sample, we first generate the t_j's as IID $\sim U[0,T]$ and then, for each t_j, we set $z_j \coloneqq Z_{t_j}$ where Z_{t_j} is obtained by simulating a discrete-time approximation to the market state dynamics (3.1c) and then terminating at time t_j. The adaptive constraint selection algorithm in this case is as follows:

Adaptive Constraint Selection Algorithm II (For a Fixed Horizon T)

---

for $i = 1$ *to* ∞ *do*

for $j = 1$ *to* Q *do*

$$\theta_j \in \quad \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \; H_\theta \Big[(\Phi r^{(i-1)})(z_j, t_j) \Big] \qquad \text{(phase 1)}$$

end for

$$r^{(i)} \in \quad \underset{r \in \mathbb{R}^k}{\operatorname{argmax}} \sum_{j=1}^{Q} (\Phi r)(z_j, t_j) \qquad \text{(phase 2)}$$

$$\text{subject to} \;\; H_{\theta_j} \left[(\Phi r)(z_j, t_j) \right] \geq 0, \qquad \forall \; j = 1, \cdots, Q,$$

$$(\Phi r)(z, T) \leq 1, \qquad \forall \; z$$

end for

We note the boundary condition $(\Phi r)(z, T) \leq 1$ in phase 2 applies to all possible z rather than just the sampled z_j's. We could of course impose this boundary constraint on just a finite subset of z values but we will see later that it is straightforward to impose the general constraint through an appropriate choice of basis functions. For example, we could ensure that each non-constant basis function has a common factor $(T - t)$. As a result, the only contribution to the left-hand side of the constraint $(\Phi r)(z, T) \leq 1$ will come from a constant basis function, say $\phi_1 \equiv 1$. We can then impose the boundary condition by adding the linear constraint $r_1 \leq 1$ to the LP phase 2.

We now discuss the objectives and constraints in phases 1 and 2 in further detail and in particular, why this algorithm is problematic to implement in its current form.

Phase 1

If we substitute (3.5) into the objective function of phase 1, drop all terms that do not depend on θ and eliminate the common factor, $(\gamma - 1)$, then the problem of phase 1 may be written as:

$$\begin{aligned}\theta_j \quad \in \quad \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \; & \frac{1}{2}\theta^\top \big[\gamma(\Phi r^{(i-1)})(z_j,t_j)\Sigma_P\Sigma_P^\top\big]\theta \\ & -\Big[(\Phi r^{(i-1)})(z_j,t_j)\lambda(z_j) + \Sigma_P\Sigma_Z^\top(\Phi r^{(i-1)})_z(z_j,t_j)\Big]^\top\theta.\end{aligned} \tag{3.13}$$

If $(\Phi r^{(i-1)})(z_j,t_j) > 0$, then the phase 1 problem is a convex quadratic program and therefore easy to solve. Otherwise, depending on $\mathbf{K}$, (3.13) may be unbounded. In this case we simply take θ_j to be the myopic portfolio which is described in Appendix B.2. We note here, however, that in our numerical experiments it only rarely[7] occurred that $(\Phi r^{(i-1)})(z_j,t_j) < 0$.

Phase 2

If we substitute (3.5) into the constraints of phase 2, then the problem of phase 2 can be formulated as:

$$\begin{aligned}\max_{r\in\mathbb{R}^k} \quad & c^\top r \\ \text{subject to} \quad & Ar \geq \mathbf{0} \\ & (\Phi r)(z,T) \leq 1, \qquad \forall z\end{aligned} \tag{3.14}$$

where

$$\begin{aligned}A_{ij} =& (1-\gamma)\phi_j(z_i,t_i)\big[\theta_i^\top\lambda(z_i) + r(z_i)\big] + (\phi_j)_z(z_i,t_i)^\top\mu_Z(z_i) + \frac{1}{2}\gamma(\gamma-1)\phi_j(z_i,t_i)\theta_i^\top\Sigma_P\Sigma_P^\top\theta_i \\ & + (1-\gamma)\Big[(\phi_j)_z(z_i,t_i)\Big]^\top\Sigma_Z\Sigma_P^\top\theta_i + \frac{1}{2}\mathrm{tr}\Big[(\phi_j)_{zz}(z_i,t_i)\Sigma_Z\Sigma_Z^\top\Big] + (\phi_j)_t(z_i,t_i),\end{aligned}$$

[7] In the numerical experiments of Section 3.5.1, we considered problems with various combinations of risk-aversion parameter, γ, and horizon T. In the worst-case among all such problems, we observed $(\Phi r^{(i-1)})(z_j,t_j) < 0$ only 1.46% of the time. In most of these problems $(\Phi r^{(i-1)})(z_j,t_j) < 0$ never actually occurred.

$c_i = \phi_i(z_1, t_1) + \phi_i(z_2, t_2) + \cdots + \phi_i(z_Q, t_Q)$. Ignoring the boundary conditions, $(\Phi r)(\cdot, T) \leq 1$, (which we can handle through the choice of basis functions as previously discussed), we see that phase 2 is an LP with a k-dimensional decision vector and Q linear constraints.

It turns out that phase 2 here is very problematic. In particular, the constraint $Ar \geq \mathbf{0}$ of (3.14) is much more difficult to handle than the constraint $Ar \geq -\mathbf{1}$ of (3.12) which occurs in the exponentially distributed horizon case. This latter set of constraints is always satisfied by all points in some ball around the zero vector $\mathbf{0}$. This is not true in the fixed horizon case where the corresponding constraints are $Ar \geq \mathbf{0}$. Since r is k-dimensional, each of the Q constraints in $Ar \geq \mathbf{0}$ defines a k-dimensional closed half-space containing $\mathbf{0}$ on its boundary. Any feasible point must therefore lie in the intersection of these Q half-spaces. Moreover, since Q is typically much larger than k the intersection is generally just a single-point, namely the origin $\{\mathbf{0}\}$. This makes the problem (3.14) trivial to solve but the solution is hardly desirable.

One ad-hoc approach for resolving this issue would be to relax the constraint $Ar \geq \mathbf{0}$ to $Ar \geq -\epsilon \cdot \mathbf{1}$ where ϵ is some small positive number. For example, in our initial numerical experiment of Section 3.5.1, the value $\epsilon = 1$ appeared to yield the best results among several different values of ϵ. We did not, however, have a sensible rule for choosing an appropriate value of ϵ in advance. Moreover, in Section 3.4.1 below we propose an alternative problem formulation that yields superior results.

3.4.1 An Alternative Formulation

We propose here a new problem formulation that is based on the certainty equivalent return, r_{ce}, which is defined as the certain annualized rate of return which makes the investor indifferent between accepting it and following his optimal trading strategy. It is therefore given implicitly via

$$u(w)V^*(z,t) = u\left(we^{r_{ce}(z,t)(T-t)}\right),$$

which implies

$$\ln\left(V^*(z,t)\right) = r_{ce}(z,t)(T-t)(1-\gamma). \tag{3.15}$$

Because $r_{ce}(z,t)$ is generally "less non-linear" than the value function (especially when γ is large), it makes some sense to approximate the log-value function rather than the value function itself. We assume basis functions of the form

$$\{\phi_1(z,t),\cdots,\phi_k(z,t)\} = \left\{(T-t)\tilde{\phi_1}(z,t),\cdots,(T-t)\tilde{\phi_k}(z,t)\right\} \tag{3.16}$$

and will use a linear combination of them to approximate $\ln\left(V^*(z,t)\right)$. We also note that any such linear combination of these functions will automatically satisfy the boundary condition $e^{(\Phi r)(z,T)} \leq 1$ so that this constraint does not need to be explicitly imposed in our new adaptive constraint selection algorithm:

Adaptive Constraint Selection Algorithm III (For a Fixed Horizon T)

---

for $i = 1$ *to* ∞ *do*

for $j = 1$ *to* Q *do*

$$\theta_j \in \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \; H_\theta \left[e^{(\Phi r^{(i-1)})(z_j, t_j)} \right] \qquad \text{(phase 1)}$$

end for

$$r^{(i)} \in \underset{r \in \mathbb{R}^k}{\operatorname{argmax}} \sum_{j=1}^{Q} e^{(\Phi r)(z_j, t_j)} \qquad \text{(phase 2)}$$

$$\text{subject to} \;\; H_{\theta_j} \left[e^{(\Phi r)(z_j, t_j)} \right] \geq 0, \qquad \forall \; j = 1, \cdots, Q$$

end for

We provide further details on the steps required for phases 1 and 2 below.

Phase 1

If we substitute (3.5) into the objective function of phase 1, drop terms that do not depend on θ and then eliminate the common factor, $e^{\Phi r^{(i-1)}}(\gamma - 1)$, then the problem of phase 1 may be reduced to:

$$\theta_j \in \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \; \frac{1}{2} \theta^\top \left[\gamma \Sigma_P \Sigma_P^\top \right] \theta - \left[\lambda(z_j) + \Sigma_P \Sigma_Z^\top \left(\Phi r^{(i-1)} \right)_z (z_j, t_j) \right]^\top \theta \qquad (3.17)$$

where the subscript z in (3.17) denotes a gradient vector. In contrast to the previous algorithm, phase 1 is always a convex quadratic program and therefore easy to solve.

Phase 2

Similarly, if we substitute (3.5) into the constraints of phase 2, eliminate the common factor $e^{\Phi r}$ and rearrange, we obtain:[8]

$$(\Phi r)_z(z,t)^\top \mu_Z(z) + (1-\gamma)(\Phi r)_z(z,t)^\top \Sigma_Z \Sigma_P^\top \theta + \frac{1}{2}\mathrm{tr}\Big[\big((\Phi r)_{zz}(z,t) + (\Phi r)_z(z,t)(\Phi r)_z(z,t)^\top\big)\Sigma_Z\Sigma_Z^\top\Big]$$
$$+ (\Phi r)_t(z,t) \;\geq\; -(1-\gamma)(\theta^\top\lambda + r) - \frac{1}{2}\gamma(\gamma-1)\theta^\top\Sigma_P\Sigma_P^\top\theta. \tag{3.18}$$

for $(z,t) = (z_1,t_1),\ldots,(z_Q,t_Q)$. Note the phase 2 objective function contains the exponential term $e^{(\Phi r)(z_j,t_j)}$ and the constraint (3.18) contains the term $(\Phi r)_z(z,t)(\Phi r)_z(z,t)^\top$ which is quadratic in r. Phase 2 is therefore not an LP.

We resolve this problem by: (i) linearizing the objective function using a first-order Taylor series expansion of e^x around zero and (ii) simply dropping[9] the terms in (3.18) that quadratic in r. This yields the following LP for phase 2

$$\begin{aligned} \max_{r\in\mathbb{R}^k} \quad & c^\top r \\ \text{subject to} \quad & Ar \geq -d \end{aligned} \tag{3.19}$$

where

$$\begin{aligned} A_{ij} &\coloneqq (\phi_j)_z(z_i,t_i)^\top \mu_Z(z_i) + (1-\gamma)\left[(\phi_j)_z(z_i,t_i)\right]^\top \Sigma_Z\Sigma_P^\top\theta_i + \\ &\quad \frac{1}{2}\mathrm{tr}\left[(\phi_j)_{zz}(z_i,t_i)\,\Sigma_Z\Sigma_Z^\top\right] + (\phi_j)_t(z_i,t_i), \\ d_i &\coloneqq (1-\gamma)\left[\theta_i^\top\lambda(z_i) + r(z_i)\right] + \frac{1}{2}\gamma(\gamma-1)\theta_i^\top\Sigma_P\Sigma_P^\top\theta_i, \\ c_i &\coloneqq \phi_i(z_1,t_1) + \phi_i(z_2,t_2) + \cdots + \phi_i(z_Q,t_Q). \end{aligned}$$

[8] Note that if we used the alternative formulation based on approximating the log-value function for the case of the exponentially distributed horizon, we would not be able to eliminate the factor $e^{\Phi r}$ due to the constant term "1" that appears in the HJB operator in (3.8).

[9] We could of course also have used a Taylor expansion to linearize the constraints but in our numerical experiments we obtained very good results by simply dropping the quadratic terms.

In our numerical experiments with this algorithm we will choose $\phi_1(z,t) = \frac{T-t}{T}$ as one of our basis function. It is then easy to see that $A_{1j} = \frac{T-t_j}{T} > 0$ for all j so regardless of d, we can ensure $Ar \geq -d$ holds by taking r_1 sufficiently large. We therefore do not need to relax the constraints, $Ar \geq -d$, as we needed to do with (3.14) in our original problem formulation for the problem with a fixed horizon, T.

3.5 Numerical Experiments

We now illustrate the performance of the LP-based algorithms of Section 3.4. We consider several portfolio optimization problems and assume that in each of them the horizon, T, is fixed. We consider three different trading strategies: the strategies that are greedy with respect to the approximate value functions that are obtained from the adaptive constraint selection algorithms II and III, respectively, as well as the well-known myopic strategy[10] which we will use as a benchmark.

The LB-based strategies at any state (t, Z_t) are found by solving

$$\begin{aligned} \theta_t^{LP1} &= \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \quad \frac{1}{2}\theta^\top \Big[\gamma\left(\Phi r^*\right)(Z_t,t)\Sigma_P\Sigma_P^\top\Big]\theta \\ &\qquad -\Big[\left(\Phi r^*\right)(Z_t,t)\lambda + \Sigma_P\Sigma_Z^\top\left(\Phi r^*\right)_z(Z_t,t)\Big]^\top\theta, \end{aligned} \tag{3.20}$$

$$\theta_t^{LP2} = \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \quad \frac{1}{2}\theta^\top\Big[\gamma\Sigma_P\Sigma_P^\top\Big]\theta - \Big[\lambda + \Sigma_P\Sigma_Z^\top\left(\Phi r^*\right)_z(Z_t,t)\Big]^\top\theta. \tag{3.21}$$

We note that θ_t^{LP1} and θ_t^{LP2} are obtained from (3.13) and (3.17), i.e. phase 1 of Algorithms II and III, respectively, by replacing Φr with Φr^*, where r^* is the solution we obtain from implementing these algorithms. Similarly the myopic strategy in state (t, Z_t) is found by solving the convex

[10] The myopic strategy is known to perform well under the various numerical experiments in HKW [22] and Haugh and Jain [21]. Indeed when $\gamma = 1$ it is known to be the optimal strategy. Further details on the myopic strategy can be found in Appendix B.2.

quadratic program

$$\theta_t^m = \underset{\theta \epsilon \mathbf{K}}{\operatorname{argmin}} \ \frac{1}{2}\theta^\top \left[\gamma \Sigma_P \Sigma_P^\top\right] \theta - \lambda^\top \theta \tag{3.22}$$

where λ is the time t vector of excess returns.

We use a simple Euler scheme to generate sample paths of the security prices and state vector, Z_t. At each time step all three strategies are found by solving (3.20), (3.21) and (3.22), respectively. By simulating many paths and averaging the utility of terminal wealth across all paths for each strategy we can obtain estimates of the value functions associated with each of the strategies. In our numerical results we will report these value functions as certainly equivalent (CE) annualized returns. Since these strategies are all feasible their CE returns are therefore lower bounds on the CE return for the (in general) unknown optimal strategy. Finally we can use the dual approach of HKW[11] to construct upper bounds on the optimal value. These upper bounds are also reported as CE returns.

3.5.1 Example I

Our first numerical example is from HKW who in turn based their model on the discrete-time market model in Lynch [29]. They consider a financial market with three risky assets and a single state variable associated with a four-dimensional Brownian motion. In our framework of Section 3.2 we assumed (without loss of generality) that the volatility matrix, Σ_P, is invertible. We can enforce this here by simply assuming that the state variable is in fact a fourth risky security that we are not allowed to trade. We assume the drift term of risky assets returns is affine in the state variable which itself follows an Ornstein-Uhlenbeck process with a long-term mean of zero. The

[11] See Appendix B.3 for a review of their dual methodology.

asset return dynamics therefore satisfy:

$$\begin{aligned} r_t &\equiv r, \\ \frac{dP_t}{P_t} &= (\mu_0 + Z_t\mu_1)dt + \Sigma_P dB_t, \\ dZ_t &= -kZ_t dt + \Sigma_Z dB_t \end{aligned}$$

where $r = .01, k = 0.366$ and

$$\mu_0 = \begin{bmatrix} 0.142 \\ 0.109 \\ 0.089 \end{bmatrix}, \quad \mu_1 = \begin{bmatrix} 0.065 \\ 0.049 \\ 0.049 \end{bmatrix}, \quad \Sigma_P = \begin{bmatrix} 0.256 & 0 & 0 & 0 \\ 0.217 & 0.054 & 0 & 0 \\ 0.207 & 0.062 & 0.062 & 0 \\ -0.741 & 0.04 & 0.034 & 0.288 \end{bmatrix}, \quad \Sigma_Z = \begin{bmatrix} -0.741 \\ 0.04 \\ 0.034 \\ 0.288 \end{bmatrix}^\top.$$

Note that r, Σ_P and Σ_Z are constant in this model. When performing simulations, we set the initial value, $Z_0 = 0$, of the state variable. We use a discretization time step of $dt = 1/100$ in our simulations as well as in the simulations of the later models of Sections 3.5.2 and 3.5.3. The horizon T is fixed at either 5 or 10 years, and the parameter γ of the CRRA utility function can be either 1.5, 3 or 5. We also consider two sets of trading constraints:

(i) The unconstrained case where the agent does not face any trading constraints (except of course for the fourth asset which is really the state variable and therefore not tradeable). We refer to this as the "Incomplete Markets" case.

(ii) There are no short-sales on all the risk securities as well as a no-borrowing constraint. We refer to this as the "Incomplete Markets + No Short-Sales and No Borrowing" case.

In each of our numerical experiments (here and elsewhere in the chapter), we use $Q = 10,000$ sample points in our two LP-based algorithms. When we use the adaptive constraint selection algorithm

II, we use

$$\mathbb{B} = \{1\} \cup \left\{ P_i(z) \cdot \left(\frac{T-t}{T}\right)^j \mid \ 0 \le i \le 5,\ 1 \le j \le 5 \right\}$$

as our basis functions where $P_i(\cdot)$ is the Chebyshev polynomial of degree i. Note that except for the first one, all of our basis functions contain a factor of $(T-t)$. As stated earlier, this allows us to easily impose the constraint $(\Phi r)(z,T) \le 1$. In phase 2 of algorithm II we set $\epsilon = 1$ to relax the constraints in the linear program (3.14). When using algorithm III associated with our alternative formulation, we use

$$\mathbb{B} = \left\{ P_i(z) \cdot \left(\frac{T-t}{T}\right)^j \mid \ 0 \le i \le 5,\ 1 \le j \le 5 \right\}.$$

as our set of basis functions.

Tables 3.1 and 3.2 present the results. We observe the two trading strategies driven by the LP approach perform better than the myopic strategy even when γ is close to one. In the incomplete markets case it is actually possible to compute the optimal solution by solving a system of ODEs. This optimal solution is reported in the row labeled V^u. If we compare the performances of the LP strategies to the optimal strategy, we see they are generally very close to each other although their performances do deteriorate somewhat with T and γ.

In comparing Tables 3.1 and 3.2 more closely, we also note that algorithm III is clearly superior to algorithm II and this is especially[12] noticeable when $(T,\gamma) = (10,5)$. This seems to suggest that the error due to the linearization in phase 2 of algorithm III is quite small. We also noticed similar behavior in our other numerical experiments and for this reason we will only report results from algorithm III henceforth.

[12] While algorithm III has performed very well, we suspect that we could improve it even further via a more careful linearization of the constraints in phase 2.

Algorithm II with Model I

	$T = 5$			$T = 10$		
	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$
			Incomplete Markets			
LB^{LP}	16.79	10.25	6.98	17.73	11.32	7.80
	(16.77, 16.80)	(10.23, 10.28)	(6.96, 7.00)	(17.71, 17.74)	(11.30, 11.35)	(7.78, 7.83)
UB^{LP}	16.81	10.37	7.25	17.77	11.53	8.29
	(16.78, 16.84)	(10.26, 10.47)	(7.10, 7.40)	(17.74, 17.79)	(11.38 , 11.69)	(7.93, 8.66)
LB^m	16.64	9.87	6.61	17.45	10.58	7.09
	(16.62, 16.66)	(9.86, 9.89)	(6.59, 6.62)	(17.44, 17.47)	(10.56, 10.59)	(7.08, 7.10)
UB^m	16.83	10.43	7.31	17.81	11.65	8.37
	(16.81, 16.86)	(10.32, 10.54)	(7.16, 7.47)	(17.79, 17.83)	(11.48, 11.82)	(7.94, 8.81)
V^u	16.79	10.32	7.06	17.76	11.55	8.12
			Incomplete Markets + No Short-Sales and No Borrowing			
LB^{LP}	10.16	7.82	5.63	10.38	8.48	6.36
	(10.16, 10.17)	(7.81, 7.83)	(5.61, 5.65)	(10.37, 10.38)	(8.47, 8.49)	(6.35, 6.38)
UB^{LP}	10.17	7.87	5.78	10.39	8.59	6.63
	(10.17, 10.18)	(7.84, 7.91)	(5.67, 5.89)	(10.38, 10.39)	(8.56, 8.62)	(6.43, 6.84)
LB^m	10.16	7.63	5.34	10.37	8.17	5.80
	(10.15, 10.16)	(7.63, 7.64)	(5.33, 5.35)	(10.36, 10.37)	(8.16, 8.18)	(5.79, 5.80)
UB^m	10.21	7.98	5.85	10.46	8.85	6.80
	(10.21, 10.22)	(7.94, 8.02)	(5.74, 5.97)	(10.45, 10.46)	(8.80, 8.90)	(6.56, 7.05)

Table 3.1: Algorithm II with Model I: Rows marked LB^{LP} and LB^m report estimates of the CE returns from the strategy determined by algorithm II and the myopic strategy, respectively. Approximate 95% confidence intervals are reported in parentheses. Estimates are based on 1 million simulated paths. The row V^u reports the optimal value function for the problem. Rows marked UB^{LP} and UB^m report estimates of the upper bound on the true value function computed using these strategies.

Algorithm III with Model I

	$T = 5$			$T = 10$		
	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$
			Incomplete Markets			
LB^{LP}	16.79	10.32	7.05	17.77	11.50	8.07
	(16.77, 16.81)	(10.29, 10.35)	(7.00, 7.09)	(17.75, 17.78)	(11.44, 11.55)	(8.01, 8.14)
UB^{LP}	16.79	10.34	7.06	17.77	11.54	8.29
	(16.76, 16.82)	(10.26, 10.43)	(6.87, 7.25)	(17.75, 17.80)	(11.42 , 11.67)	(8.04, 8.55)
LB^m	16.63	9.86	6.59	17.46	10.57	7.09
	(16.61, 16.64)	(9.84, 9.87)	(6.58, 6.61)	(17.45, 17.47)	(10.56, 10.58)	(7.08, 7.10)
UB^m	16.82	10.41	7.12	17.82	11.72	8.50
	(16.79, 16.84)	(10.33, 10.50)	(6.93, 7.32)	(17.80, 17.85)	(11.59, 11.84)	(8.24, 8.77)
V^u	16.79	10.32	7.06	17.76	11.55	8.12
			Incomplete Markets + No Short-Sales and No Borrowing			
LB^{LP}	10.16	7.83	5.68	10.38	8.52	6.55
	(10.15, 10.16)	(7.82, 7.84)	(5.66, 5.71)	(10.38, 10.38)	(8.51, 8.53)	(6.52, 6.58)
UB^{LP}	10.16	7.83	5.67	10.38	8.53	6.63
	(10.15, 10.16)	(7.80, 7.86)	(5.55, 5.80)	(10.38, 10.39)	(8.50, 8.55)	(6.50, 6.77)
LB^m	10.15	7.63	5.33	10.37	8.17	5.80
	(10.15, 10.16)	(7.62, 7.64)	(5.32, 5.34)	(10.36, 10.37)	(8.16, 8.18)	(5.79, 5.80)
UB^m	10.20	7.96	5.77	10.46	8.84	6.98
	(10.20, 10.21)	(7.91, 8.00)	(5.62, 5.91)	(10.45, 10.47)	(8.79, 8.89)	(6.81, 7.15)

Table 3.2: Algorithm III with Model I: Rows marked LB^{LP} and LB^m report estimates of the CE returns from the strategy determined by algorithm III and the myopic strategy, respectively. Approximate 95% confidence intervals are reported in parentheses. Estimates are based on 1 million simulated paths. The row V^u reports the optimal value function for the problem. Rows marked UB^{LP} and UB^m report estimates of the upper bound on the true value function computed using these strategies.

3.5.2 Example II

The second model we consider is taken from from Haugh, Kogan and Wu [23] who in turn based their model and parameters on Wachter and Sangvinatsos [37]. In this model there is only one risky asset which is a long-term bond maturing at time T. The bond has no risk premium and there is a three-dimensional state variable and three-dimensional Brownian motion. In contrast to the previous model (and the duality development in Appendix B.3), we do not explicitly define artificial assets so that the number of risky assets equals the dimension of the Brownian motion (in which case Σ_P will be invertible). Instead we directly set the risk premium of the risky bond as well as the market price of risk process, η_t, to be zero. With these choices, it is clear that $\Sigma_P \eta = \lambda$ will be satisfied. Therefore if necessary we could explicitly define artificial asset price dynamics so that our choice of η, i.e. zero in this example, would be the unique market price of risk process in the unconstrained market. Clearly then we don't need[13] to explicitly define artificial asset price dynamics in order to apply the dual methodology.

Assume that the state variable which follows a three-dimensional Ornstein-Uhlenbeck process reverting to zero vector. More precisely, the asset return dynamics satisfy the following SDEs:

$$\begin{aligned} r_t &= \delta_0 + \delta_1 Z_t, \\ \frac{dP_t}{P_t} &= r_t dt + \Sigma_P dB_t, \\ dZ_t &= -K Z_t dt + \Sigma_Z dB_t \end{aligned}$$

[13] See also the final paragraph of Appendix B.3 where we explain why the choice of artificial asset price dynamics does not impact the dual bound in our numerical examples.

where

$$K = \begin{bmatrix} 0.576 & 0 & 0 \\ 0 & 3.343 & 0 \\ -0.421 & 0 & 0.083 \end{bmatrix}, \quad \Sigma_Z = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \delta_0 = 0.056, \quad \delta_1 = \begin{bmatrix} 0.018 \\ 0.007 \\ 0.010 \end{bmatrix}^\top$$

and $\Sigma_P = -\delta_1 K^{-1}(I - e^{-(T-t)K})\Sigma_Z$, which forces the bond price to equal the face value at maturity. Note that the diffusion vector, Σ_P, of the asset return is time-dependent in this model.

The initial state variable is $Z_0 = 0$, the time to maturity is $T = 5$ years and the risk aversion coefficient is $\gamma = 15$, which reflects a high degree of risk aversion. This is intentional because we can guess in this case that the policy of holding all of the portfolio in the long-term bond should be very close to optimal. In this particular model then can we consider the buy-and-hold policy which invests all in the long-term bond as another benchmark. Note that the myopic strategy in this case will simply invest everything in the cash account, since the risk premium on the long-term bond is zero. In this model (and model III below), we consider two sets of trading constraints. In the first, the investor does not face any trading constraints and simply has an incomplete markets problem. In the second case the investor faces a no-borrowing constraint in addition to an incomplete market.

In applying algorithm III, we choose

$$\mathbb{B} = \left\{ P_i(z_1) \cdot P_j(z_2) \cdot P_k(z_3) \cdot \left(\frac{T-t}{T}\right)^l \mid \ 0 \le i+j+k \le 3, \ 1 \le l \le 3 \right\}$$

as our set the of basis functions where once gain P_i denotes the Chebyshev polynomial of degree i.

Table 3.3 displays the results of this experiment. As expected, due to the high value of γ, we observe that the performance of the buy-and-hold policy on the long-term bond is much better than that of the myopic policy. Surprisingly, however, the LP approach performs even better and produces lower and upper bounds that, to two decimal places at least, are identical.

Algorithm III with Model II

	Incomplete Markets	Incomplete Markets + No Borrowing
LB^{LP}	5.52	5.52
	(5.52, 5.52)	(5.52, 5.52)
UB^{LP}	5.52	5.52
	(5.52, 5.52)	(5.52, 5.52)
LB^{m}	4.41	4.41
	(4.40, 4.41)	(4.41, 4.42)
UB^{m}	5.52	5.52
	(5.52, 5.52)	(5.52, 5.52)
LB^{LT}	5.51	5.51
	(5.51, 5.51)	(5.51, 5.51)
UB^{LT}	5.53	5.58
	(5.52, 5.53)	(5.57, 5.59)

Table 3.3: Algorithm III with Model II: Rows LB^{LP}, LB^{m} and LB^{LT} report estimated CE returns from the strategy determined by algorithm III, the myopic strategy and the buy-and-hold strategy on the long-term bond, respectively. Approximate 95% confidence intervals are reported in parentheses. Estimates are based on 1 million simulated paths. The rows marked UB^{LP},UB^{m} and UB^{LT} report estimates of the upper bound on the true value function computed using these strategies.

3.5.3 Example III

In our final model, which is again taken from Haugh, Kogan and Wu [23], there are three risky assets: two bonds with maturities three years and ten years, respectively, and a stock index. There is a four-dimensional state variable and a five-dimensional Brownian motion. As was the case with Model II, we explicitly define the market price of risk process η_t instead of defining additional artificial risky assets that the investor will not be permitted to trade. We assume the state vector follows a four-dimensional mean-reverting Ornstein-Uhlenbeck process with mean the zero vector. The asset return and the risk premium dynamics satisfy the SDEs:

$$r_t = \delta_0 + \delta_1 Z_t, \tag{3.25a}$$

$$\eta_t = \lambda_0 + \lambda_1 Z_t, \tag{3.25b}$$

$$\frac{dP_t}{P_t} = (r_t \cdot \mathbf{1} + \Sigma_P \eta_t)dt + \Sigma_P dB_t, \tag{3.25c}$$

$$dZ_t = -KZ_t dt + \Sigma_Z dB_t \tag{3.25d}$$

where $Q = K + \Sigma_Z \lambda_1$ and

$$\Sigma_P = \begin{bmatrix} -\delta_1 Q^{-1}(I - e^{-3\cdot Q})\Sigma_Z \\ -\delta_1 Q^{-1}(I - e^{-10\cdot Q})\Sigma_Z \\ -0.0126 \quad 0.0057 \quad -0.0295 \quad 0.143 \quad 0 \end{bmatrix}.$$

The particular forms of the first and second rows of Σ_P imply that we use dynamic rollover strategies for the three-year and ten-year bonds so that the duration of the bonds should be maintained at three and ten years by continuous reinvestment. The other parameter values are $K, \Sigma_Z, \delta_0, \delta_1, \lambda_0$ and λ_1 are reported in Table 3.4. The initial state vector is $Z_0 = 0$, the horizon is $T = 5$ years and the constant relative risk aversion coefficient γ is set to 1.5, 3 and 5.

When we use the adaptive constraint selection algorithm III, we use

$$\mathbb{B} = \left\{ P_i(z_1)P_j(z_2)P_k(z_3)P_l(z_4)\left(\frac{T-t}{T}\right)^m \mid \ 0 \le i+j+k+l \le 2,\ 1 \le m \le 10 \right\}$$

as our set of basis functions.

Parameter	Value				
K	0.576	0	0	0	
	0	3.343	0	0	
	-0.421	0	0.083	0	
	0	0	0	0.080	
Σ_Z	1.0000	0	0	0	0
	0	1.0000	0	0	0
	0	0	1.0000	0	0
	0	0	0	0.1600	0.3664
δ_0	0.056				
δ_1	0.018	0.007	0.010	0	
$\lambda_0^\top$	-0.5630	-0.2450	-0.2190	0.4400	0
$\lambda_1^\top$	0	0	0.5370	0.1110	0
	1.7540	-1.8150	0.3760	0.3050	0
	0	0	-0.0820	-0.0170	0
	0	0	0	0.0700	0

Table 3.4: Parameters for Model III defining the instantaneous risk-free rate, risk premium and state variable processes in (3.25a), (3.25b) and (3.25d), respectively.

Table 3.5 displays the numerical results for this model. The results are consistent with our earlier examples in that, regardless of the market trading constraints, the LP strategy outperforms

Algorithm III with Model III

	Incomplete Markets			Incomplete Markets + No Borrowing		
	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$	$\gamma = 1.5$	$\gamma = 3$	$\gamma = 5$
LB^{LP}	59.47	35.82	24.65	32.26	20.59	15.45
	(59.39, 59.54)	(35.67, 35.98)	(24.20, 25.14)	(32.12, 32.40)	(20.46, 20.72)	(15.31, 15.59)
UB^{LP}	59.49	35.97	25.91	32.31	20.77	15.79
	(59.38, 59.61)	(35.53, 36.41)	(24.97, 26.89)	(32.11, 32.52)	(20.40, 21.15)	(15.15, 16.44)
LB^{m}	59.06	34.59	23.61	32.15	20.23	15.09
	(59.00, 59.13)	(34.52, 34.66)	(23.55, 23.67)	(32.03, 32.28)	(20.13, 20.32)	(15.02, 15.17)
UB^{m}	60.01	37.54	27.40	32.55	21.18	16.34
	(59.90, 60.12)	(37.17, 37.92)	(26.39, 28.44)	(32.34, 32.75)	(20.74, 21.64)	(15.45, 17.27)

Table 3.5: Algorithm III with Model III: Rows LB^{LP} and LB^{m} report estimates of the CE returns from the strategy determined by algorithm III and the myopic strategy, respectively. These estimates are based on 1 million simulated paths for the incomplete market problem and 100 thousand paths for the no-borrowing problem. Approximate 95% confidence intervals are reported in parentheses. Rows UB^{LP} and UB^{m} report the estimates of the corresponding upper bounds on the true value function.

the myopic strategy. The gap between the two trading strategies is more visible here than in Model I, for example. When $\gamma = 5$ in the incomplete markets case, the duality gap of $25.91 - 24.65 = 1.26\%$ suggests that the LP-based strategy is still reasonably far from the optimal strategy. As stated earlier, we suspect that we could improve the LP-based strategy via a more careful linearization of the constraints in phase 2 of algorithm III.

3.6 Conclusions

We have extended the linear programming approach of Han and Van Roy [19] to compute good sub-optimal solutions for high-dimensional control problems in a diffusion-based setting with fixed

time horizons. In considering numerical examples drawn from portfolio optimization, we were able to show that our sub-optimal solutions are indeed very good by using them to construct tight lower and upper bounds on the optimal value functions for these problems. These results suggest that the LP approach is a very promising one for tackling high-dimensional control problems.

There are several possible directions for future research. First, it would be interesting to extend the methodology to jump-diffusions and other more general settings. There is also scope for additional theoretical work in order to better understand the properties of these LP-based algorithms. Given some of the necessary ad-hoc steps of the LP approach in this chapter and the original work of Han and Van Roy [19], this may be particularly challenging.

Chapter 4

A Factor Model-Based Approach to Scenario Analysis

4.1 Introduction

Scenario analysis is an important and widely used risk-management technique that is used throughout the financial services industry. In the standard version of scenario analysis we shift a small pre-defined subset of risk factors and compute the resulting profit-and-loss (P&L) on the portfolio. By considering many shifts and many subsets of factors, it is then possible to get a good understanding of the risk profile of the portfolio. Moreover, because this standard form of scenario analysis does not require a probability distribution and produces a P&L for each considered scenario, it is preferred by many practitioners to risk measures such as value-at-risk (VaR) or conditional value-at-risk (CVaR) which are scalar risk measures and rely on knowledge of a probability distribution that is often very hard to estimate.

But standard scenario analysis suffers from at least two important drawbacks. First, in stressing

a small subset of risk factors it implicitly sets the shocks to non-stressed risk-factors equal to zero. This tends to ignore any conditional dependence structure between the stressed and unstressed risk factors. For example, when the underlying security price falls, the implied volatility surface tends to increase and volatility skews often steepen. A stress test that only shocks the underlying security price will therefore not account properly for likely movements in these other risk factors. It is also important to note that implicitly setting other risk-factors shocks to zero (or their conditional expected values) ignores the convexity of the portfolio with respect to these unstressed factors. This convexity can be significant for derivatives portfolios. A second issue with standard scenario analysis is that it is not testable in that the probability of any given scenario actually occurring is zero. In this sense then it is not possible to quantify the performance of standard scenario analysis. This is in contrast to estimated risk measures such as VaR and CVaR. For example, if we produce a daily estimate of a portfolio's 99% VaR then on average we should see losses exceeding the VaR 1% of the time and it is straightforward to test that this is so.

In this chapter we propose a factor model-based scenario analysis which is easy to implement and produces an *expected* P&L for each proposed scenario. The expected P&L is estimated using a Monte Carlo simulation that conditions on the proposed shocks to the risk factors that we want to stress. This allows us to overcome the first problem above. Our factor modeling approach also allows us to estimate *realized* shocks to the risk-factors and therefore compare the *realized* P&L with the P&L we would have *predicted* conditioned on these realizations. We can therefore also address the second problem.

We develop our modeling approach in the context of an options portfolio with a single underlying security but it should be clear that our framework can also be applied in other contexts. In preliminary numerical tests with S&P 500 options data, our factor model-based scenario analysis

performs well and outperforms the standard scenario analysis approach in that it produces more accurate estimates of realized losses conditioned on realized shocks to the stress factors under consideration. We also mention that our overall approach can be viewed as estimating the joint *stationary* distribution of risk-factor shocks but we would expect better results if we were to estimate the joint distribution of risk-factor shocks *conditional* on current market conditions. This is the main direction we intend to pursue in extending this work.

The literature on factor modeling in finance is vast[1] but it has focused mainly on vanilla equities and fixed income markets. Factor models are also popular among industry practitioners and there are now several companies[2] devoted to constructing and providing factor models to asset managers as well as risk managers. Researchers including Meucci [32] and Da Fonseca and Cont [10] among others have also constructed factor models for derivatives portfolios. Their focus has been more on understanding the nature of these risk factors or on using these models for portfolio construction. Since the financial crisis of 2008 there has been an enormous interest in the area of stress testing and scenario analysis. We can't do justice to that literature here but a notable[3] recent contribution has been the text of Rebonato [33] who adopts a Bayesian approach to stress testing and scenario analysis. It is worth emphasizing that our focus in this chapter, however, is on using a factor model for scenario analysis and then comparing its performance with the performance of a standard scenario analysis. To the best of or knowledge, this aspect of our work is new.

The remainder of this chapter is organized as follows. In Section 4.2 we introduce the implied

[1] This literature begins with the CAPM [35; 28], and has been an active field ever since with other notable contributions including the Fama-French 3-factor model[17].

[2] *BARRA* is a notable player in this space and indeed there are some similarities between their modeling approach (see [4] for example) and the one we propose in this chapter.

[3] The text by McNeil, Frey and Embrechts [31] is also an important reference for quantitative risk management more generally.

volatility surface of a security as well as some associated terminology and notation. In Section 4.3 we describe what we have been referring to as standard scenario analysis while in Section 4.4 we describe our factor-model based approach to scenario analysis. We describe our factor model estimation procedure in Section 4.5 and describe some preliminary numerical results in Section 4.6. We conclude in Section 4.7. The appendix provides some additional details.

4.2 The Implied Volatility Surface

Let $C^{bs}(t, S_t, K, T, \sigma)$ denote the time t Black-Scholes price of a European call option with strike K, maturity T and volatility σ when the underling security price is S_t. The implied volatility of an option whose market price is $C^{mkt}(t, K, T)$ is then defined to be the unique value of $\sigma_t(K, T)$ satisfying

$$C^{bs}(t, S_t, T, K, \sigma_t(K, T)) = C^{mkt}(t, K, T).$$

The mapping $\sigma_t : (K, T) \mapsto \sigma_t(K, T)$ is then called the *implied volatility surface* of the security at time t. Letting $m \coloneqq K/S_t$ denote the so-called moneyness of the option, we note that the implied volatility surface can also be represented in the relative coordinates, (m, τ), and so we define

$$I_t(m, \tau) \coloneqq \sigma_t(mS_t, t + \tau)$$

to be the implied volatility mapping in this coordinate system. For a fixed time-to-maturity, τ, and moneyness, m, we use

$$y_t(m, \tau) = I_t(m, \tau) - I_{t-1}(m, \tau), \tag{4.1}$$

to denote the change in the implied volatility[4] between periods $t-1$ and t. In practice we will observe these changes at a fixed grid of points on $\mathbb{R}^+ \times \mathbb{R}^+$, representing the strike-maturity combinations of liquidly traded options. We use

$$O = \{(m_1, \tau_1), (m_2, \tau_2), \ \dots \ , (m_N, \tau_N)\} \tag{4.2}$$

to denote these strike-maturity combinations and we will use the N-dimensional vector

$$\mathbf{y}_t := (y_t(m_1, \tau_1), y_t(m_2, \tau_2), \ \dots \ , y_t(m_N, \tau_N))^\top \tag{4.3}$$

to represent the volatility surface at time t. In this work we will assume that the investor or risk manager has a portfolio of options on a single underlying stock in addition to a possible position[5] in the stock itself.

4.3 Standard Scenario Analysis

In a standard scenario analysis we fix a number of stresses, M, say which represent the various types of stresses we wish to apply to the implied volatility surface. These stresses might represent parallel "bumps" in the surface, the steepening or flattening of the volatility skew, the steepening or flattening of term structure of volatility etc. Because we are considering portfolios with N possible strike-maturity combinations we can represent all M stresses as N-dimensional vectors $\mathbf{b}_1, \dots, \mathbf{b}_M$.

[4] Cont and Da Fonseca [10] and Da Fonseca and Gottschalk [18] considered the daily variation of the log of the implied volatility and so the defined

$$y_t(m, \tau) := \log I_t(m, \tau) - \log I_{t-1}(m, \tau).$$

One advantage of working in the log space is that positive implied volatilities are always guaranteed when we generate samples of y_t. In our modeling approach, however, we found it more convenient to work with (4.1).

[5] Instead of a position in the stock there may be a position in a futures contract with the stock as the underlying security.

We emphasize here that the b_i's are fixed and pre-defined vectors. If for example the first stress represents a parallel shift in the implied volatility surface then b_1 would be set[6] to a vector of one's.

We can view the standard approach to scenario analysis as *implicitly* assuming the following factor model for $\mathbf{y}_t$:

$$\mathbf{y}_t = \mathbf{B}_m F_{m,t} + \epsilon_{m,t} \tag{4.4}$$

where:

- $\mathbf{B}_m$ is an $N \times M$ matrix with the i^{th} column of $\mathbf{B}_m$ equal to $\mathbf{b}_i$.
- $F_{m,t}$ is an M-dimensional random vector that represents random shocks to the $\mathbf{b}_i$'s. Then the i^{th} component, $F_{m,t}^{(i)}$, of this vector represents the random shock to the i^{th} stress or risk factor.
- $\epsilon_{m,t}$ is an N-dimensional vector representing an idiosyncratic error term so that $\epsilon_{m,t}$ and $F_{m,t}$ are uncorrelated.

It is important to recognize here that risk managers, traders, asset managers etc. may never actually explicitly work with the factor model in (4.4). In considering shocks to the b_i's (which are imposed via the $F_{m,t}^{(i)}$'s) , however, it seems reasonable to assume they have such a model in mind. We do of course also include[7] the return, R_t, on the underlying security as a risk factor. We therefore define

$$S_t = \left\{ R_t, F_{m,t}^{(1)}, \ \ldots, \ F_{m,t}^{(M)} \right\} \tag{4.5}$$

to be the set of all risk factors shocks that apply between periods $t-1$ and t. In our standard form of scenario analysis we will assume that all the shocks we will want to consider are in S_t.

[6] Or equivalently, a vector with each component set equal to the same constant value.

[7] For ease of exposition we will assume throughout this chapter that dividends and risk-free interest rates are deterministic and are therefore not risk factors.

In a standard scenario analysis, it is common to consider many scenarios where each scenario is given by a combination of $L \leq M+1$ stress directions to shock. Suppose, for example, that $L = 2$ and that the two factors we wanted to stress were b_1 and b_3. We would then consider a range of shock values for $F^{(1)}_{m,t}$ and $F^{(3)}_{m,t}$ and then compute the resulting P&L on the options portfolio for each such combination of shocks. An implicit and key assumption of this approach is that $\epsilon_{m,t}$ and all of the other factors would have their shock values, $S_t \setminus \{F^{(1)}_{m,t}, F^{(3)}_{m,t}\}$, set to zero. We could then compute a scenario report of the form

	$F^{(3)}_{m,t} = -10\%$	$\cdots$	$F^{(1)}_{m,t} = 10\%$
$F^{(1)}_{m,t} = -20\%$	$*$	$\cdots$	$*$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$F^{(1)}_{m,t} = 20\%$	$*$	$\cdots$	$*$

where each entry in the main body of the table would be the P&L on the options portfolio in the event that the corresponding stress occurred. For example, the first element in the main body of the table ($F^{(1)}_{m,t} = -20\%$ and $F^{(3)}_{m,t} = -10\%$) would then represent the P&L on the portfolio between periods $t-1$ and t when $F^{(1)}_{m,t} = -20\%$, $F^{(3)}_{m,t} = -10\%$, $\epsilon_{m,t} = 0$ and all random variables in $S_t \setminus \{F^{(1)}_{m,t}, F^{(3)}_{m,t}\}$ are also set to 0.

In practice many of these reports would be computed, using various L-combinations of the $M+1$ possible stresses. We note that these reports can be computed very quickly and can be very informative. This approach to scenario analysis does have some important weaknesses, however. Writing ΔV_t for the portfolio P&L, we note the following:

- Standard scenario analysis implicitly assumes the expected shock values are zero, i.e. $\mathrm{E}[F^{(i)}_{m,t}] = 0$ for $i = 1, \ldots, M$. In general, however, the shocks will be *dependent* so that their conditional

expected values are non-zero. This is ignored in standard scenario analysis when the non-stressed factors simply have their shocks set to zero.

- Since ΔV_t is in general a nonlinear[8] function of S_t, the reported P&L numbers in a given cell in a scenario table may not be at all close to $\mathrm{E}[\Delta V_t \mid F^{(1)}_{m,t} = f_1, F^{(3)}_{m,t} = f_3]$ even if we set the shocks of the non-stressed factors equal to their conditional expected values.

We therefore believe that a scenario analysis that addresses these two weaknesses would be very informative and useful. The goal of the work in this chapter is to do this via a factor-model approach to scenario analysis and to compare empirically our approach with the standard approach. Rather than working with the factor model of (4.4) we prefer to embed it in a factor-model that is better able to capture the dynamics of the implied volatility surface.

4.4 Factor Model-Based Scenario Analysis

As discussed in the previous section we wish to embed the factor model (4.4) in a larger model. An alternative possibility would be to build our own factor model from scratch and to simply ignore (4.4). But that would be problematic because the primary goal of this work is to compare the performance of our factor model-based scenario analysis with the standard scenario analysis of Section 4.3. In order to do that, we need to be able to estimate the realized[9] shocks to each factor on a per-period basis. We can do this by simply regressing the N observations in $\mathbf{y}_t$ on their corresponding components in the b_i's to estimate the realized factor shocks, $F^{(1)}_{m,t}, \ldots, F^{(M)}_{m,t}$. Unfortunately, however, these estimates depend in general on the specific factor model that we're using. In order

[8] This non-linearity of often referred to as simply "option convexity".

[9] The factor shock R_t is of course observable.

to compare the factor model in (4.4) with our proposed factor model, both approaches must be able to agree on their estimates of the realized shocks. To do this we must therefore embed model (4.4) in our proposed factor model.

But this is not enough: in order to agree on estimates of the realized shocks both approaches must use the same model when estimating these realized shocks. For this reason both approaches will use the factor model (4.4) to do this.

4.4.1 Computing Realized Shocks

Given our discussion above, we need to solve the following optimization problem at each time t to estimate the realized factor shocks:

$$\min_{(a_1,\cdots,a_M)} \left\| \mathbf{y}_t - \sum_{i=1}^{M} a_i \mathbf{b}_i \right\|^2 \tag{$\mathcal{P}_1$}$$

That is, we must project $\mathbf{y}_t$ onto the subspace spanned by $\mathbf{b}_1, \cdots, \mathbf{b}_M$. The solution to problem $(\mathcal{P}_1)$ is $\mathbf{a_t} = (\mathbf{B}_m^\top \mathbf{B}_m)^{-1} \mathbf{B}_m^\top \mathbf{y}_t$ with $\mathbf{a_t} = (a_1, \ldots, a_M)$.

4.4.2 New Factor Model

We are now ready to construct our alternative factor model. We define $\mathbf{r}_t$ as that part of $\mathbf{y}_t$ that is not explained by $\mathbf{b}_1, \cdots, \mathbf{b}_M$. That is,

$$\mathbf{r}_t := \mathbf{y}_t - \sum_{i=1}^{M} a_i \mathbf{b}_i. \tag{4.6}$$

where the a_i's are optimal in $(\mathcal{P}_1)$. We then perform a principal component analysis (PCA) on the $\mathbf{r}_t$'s. But first we normalize the $\mathbf{r}_t$'s. If they have sample mean μ and the sample variance of the i^{th} component of $\mathbf{r}_t$ is σ_i^2, then we define

$$\hat{\mathbf{r}}_t := S^{-1}(\mathbf{r}_t - \mu) \tag{4.7}$$

where S is a diagonal matrix with the σ_i's along the diagonal. Each $\hat{\mathbf{r}}_t^{(i)}$) will now have a (sample) mean of zero and a (sample) standard deviation of 1 for all $i = 1, 2, \cdots, N$. Letting Σ denote the covariance matrix of $\hat{\mathbf{r}}_t$ we can perform a spectral decomposition of Σ to obtain

$$\Sigma = \Gamma \Delta \Gamma^\top$$

where

$$\Delta := \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$$

is a diagonal matrix of eigen values of Σ arranged in such a way that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N \geq 0$, and Γ is an $N \times N$ matrix whose i^{th} column contains the unit eigen vector $\hat{\gamma}_i$ with eigen value λ_i so that $\Sigma\hat{\gamma}_i = \lambda_i\hat{\gamma}_i$ and $\hat{\gamma}_i^\top\hat{\gamma}_j = \delta_{i,j}$.

The vector of principal components is defined as

$$P_t = \Gamma^\top \hat{\mathbf{r}}_t, \tag{4.8}$$

and we note that the P_t's have mean zero and variance-covariance matrix equal to Δ. We can easily invert (4.8) to obtain

$$\hat{\mathbf{r}}_t = \Gamma P_t = \sum_{i=1}^{N} \hat{\gamma}_i P_t^{(i)}. \tag{4.9}$$

Combining (4.7) and (4.9), we see that $\mathbf{r}_t$ can be expressed as

$$\mathbf{r}_t = S\ \Gamma P_t + \mu = \sum_{i=1}^{N} \gamma_i P_t^{(i)} + \mu, \tag{4.10}$$

where $\gamma_i := S\hat{\gamma}_i$. A low dimensional model of the residuals $\mathbf{r}_t$ can then be obtained by taking the first p terms ($0 \leq p < N - M$) in the summation in (4.10) and writing

$$\mathbf{r}_t = \sum_{i=1}^{p} \gamma_i P_t^{(i)} + \mu + \epsilon_t, \tag{4.11}$$

where

$$\epsilon_t := \sum_{i=p+1}^{N} \gamma_i P_t^{(i)}, \tag{4.12}$$

is now interpreted as an idiosyncratic noise term. Combining (4.4) and (4.11), we have our new factor model

$$\mathbf{y}_t = \mathbf{B}_m F_{m,t} + \mu + \mathbf{B}_{r,t} F_{r,t} + \epsilon_t, \tag{4.13}$$

where

- $\mathbf{B}_m$ and $F_{m,t}$ are as defined in (4.4).
- $\mathbf{B}_{r,t}$ is an $N \times p$ matrix whose i^{th} column is γ_i.
- $F_{r,t} = \left(P_t^{(1)}, \ \ldots, \ P_t^{(p)}\right)^\top$ is a p-dimensional random vector representing the first p principal components of $\mathbf{r}_t$.
- ϵ_t is an N-dimensional random vector representing an idiosyncratic mean zero noise term that is now assumed to be uncorrelated with the M risk factor shocks and the first p principal components.

4.4.3 Factor model-based Methodology

Our factor model-based scenario analysis is based on the model of (4.13) and we will use this model to generate time $t+1$ options prices and therefore portfolio P&L's. Towards this end let $Y_t := (S_t, I_t(m_i, \tau_i))^\top$ denote the vector of security price and implied volatility that are needed to

price an option with moneyness and maturity m_i and τ_i, respectively. By (4.13) we then have

$$\begin{aligned} Y_{t+1} &= \left(S_t \exp\left(Z_{t+1}^{(1)}\right),\ I_{t+1}(m_i,\tau_i)\right)^\top \\ &= \left(S_t \exp\left(Z_{t+1}^{(1)}\right),\ I_t(m_i,\tau_i) + \mathbf{y}_{t+1}^{(i)}\right)^\top \\ &= \left(S_t \exp\left(Z_{t+1}^{(1)}\right),\ I_t(m_i,\tau_i) + \mu^{(i)} + \sum_{j=1}^{M} \mathbf{b}_j^{(i)} Z_{t+1}^{(j+1)} + \sum_{j=1}^{p} \gamma_j^{(i)} Z_{t+1}^{(j+M+1)} + \epsilon_{t+1}^{(i)}\right)^\top . \end{aligned} \quad (4.14)$$

Returning to scenario analysis, note that the set of risk factors that we might want to shock are the same factors, $b_1, \ldots, b_M$, that we described in Section 4.3. Let us denote by $F_{s,t} \in \mathbb{R}^L$ the vector of L shocks in S_t (defined in (4.5)) that we want to stress. The expected P&L conditional on $F_{s,t} = f$ is then given by

$$\text{P\&L}_{ex} = \mathbb{E}_{t-1}\left[\Delta V_t\left(R_t, F_{m,t}, F_{r,t}, \epsilon_t\right) \mid F_{s,t} = f\right]. \quad (4.15)$$

In general we can't compute (4.15) analytically but we will be able to estimate it by simulating $(R_t, F_{m,t}, F_{r,t}, \epsilon_t)$ (and therefore Y_{t+1} in (4.14)) conditional on $F_{s,t} = f$. Since ϵ_t is an idiosyncratic noise term that is assumed to be independent of R_t and $F_{m,t}$, its distribution conditional on $F_{s,t} = f$ is equal to its unconditional distribution. We therefore need to focus on being able to simulate

$$Z_t = \left(R_t, F_{m,t}^\top, F_{r,t}^\top\right)^\top \quad (4.16)$$

conditional on $F_{s,t} = f$. We will discuss this in Section 4.5 below after we first describe our probabilistic model for Z_t and how we fit this model.

4.5 Modeling the Random Process Z_t

We first introduce the class of *generalised hyperbolic distributions* of which the t and skewed t distributions are examples. A generalised hyperbolic distribution Z has the following convenient

normal mean-variance mixture representation:

$$Z \overset{d}{=} m(W) + \sqrt{W}\mathbf{C}X, \tag{4.17}$$

where

i) $X \sim N_n(0, I_n)$.

ii) W follows a *generalized inverse Gaussian* (GIG) distribution: $W \sim N^-(\lambda, \chi, \theta)$, and its density, $f(w)$, satisfies

$$f(w) \propto w^{\lambda-1} \exp\left(-\frac{1}{2}\left(\chi w^{-1} + \theta w\right)\right) \tag{4.18}$$

iii) $Z \mid W = w \sim N_n(m(W), w\Sigma)$.

iv) $m(W)$, the mean function of Z, can be any $[0, \infty) \to \mathbb{R}^n$ measurable function. We consider the specification

$$m(W) = \mu + W\gamma$$

for $\mu, \gamma \in \mathbb{R}^n$.

v) $\Sigma = \mathbf{C}\mathbf{C}^\top$, with $\mathbf{C}$ an $n \times n$ matrix.

This multivariate generalised hyperbolic family is known to be very flexible. For example,

i) If $\lambda = -\nu/2$, $\chi = \nu$, and $\theta = 0$, then Z has the skewed t distribution. In this case W has an *inverse gamma* distribution, i.e., $W \sim GIG(-\nu/2, \nu, 0) \overset{d}{=} IG(\nu/2, \nu/2)$.

ii) If an addition $\gamma = 0$, then Z has a multivariate t distribution, i.e., $Z \sim t_n(\nu, \mu, \Sigma)$.

In our numerical experiments, we will assume Z_t has a multivariate t or skewed t distributions[10] which are limiting cases of the generalised hyperbolic distribution. A multivariate t distributed

[10] We assume that Z_t are i.i.d. but note that in general this assumption would be violated.

random variable, Z, has three parameters (ν, μ, Σ) with $\mathbb{E}(Z) = \mu$ and $\text{Cov}(Z) = \frac{\nu}{\nu-2}\Sigma$. Note that the covariance matrix is only defined for $\nu > 2$. If Z has a skewed t distribution then there is an additional skewness parameter γ. The mean and the covariance matrix are equal to

$$\begin{aligned} \mathbb{E}(Z) &= \mu + \gamma\frac{\nu}{\nu-2}, \\ \text{Cov}(Z) &= \frac{\nu}{\nu-2}\Sigma + \gamma\gamma'\frac{2\nu^2}{(\nu-2)^2(\nu-4)} \end{aligned}$$

where the mean is only defined for $\nu > 2$ and the covariance matrix is only defined for $\nu > 4$. Since our principal components have mean zero and are uncorrelated, we expect the mean and the covariance matrix of the calibrated model to have approximately the following forms

$$\mathbb{E}(Z) \approx (*, \cdots, *, 0, \cdots, 0)^\top$$

and

$$\mathbb{E}(Z) \approx \begin{pmatrix} * & & \cdots & & * \\ & \ddots & & & \vdots \\ \vdots & & \lambda_1 & \cdots & 0 \\ & & \vdots & \ddots & \vdots \\ * & \cdots & 0 & \cdots & \lambda_n \end{pmatrix}$$

although we did not impose this structure in our estimation procedure. Given historical time-series data $Z_{t_1}, \cdots, Z_{t_n}$, we actually fit the t (or skewed t) distribution to this data using[11] the EM algorithm as described in Hu [25].

[11] Alternatively, we can simply do moment matching for each fixed ν. So we have

$$\begin{aligned} \mu &= \mathbf{E}(Z_t) \\ \Lambda &= \frac{\nu-2}{\nu}\text{Cov}(Z_t) \end{aligned}$$

for each fixed ν. We then compute the profile log-likelibood for ν and find the value, $\hat{\nu}$, that maximizes this profile likelihood.

4.5.1 Distribution of Z_t Conditional on $F_{s,t}$

In this subsection[12] we aim to derive the conditional distribution of Z_t, given $F_{s,t}$ under a few stylized distributional assumptions of Z_t.[13] Even though F_s is a sub-vector of Z in our case, here we will consider the more general situation and assume that $F_s = \mathbf{A}Z$ with an $(M + p + 1) \times L$ loading matrix, $\mathbf{A}$, i.e.

$$F_s^{(i)} = \sum_{j=1}^{M+p+1} \mathbf{A}_{i,j} Z^{(j)}.$$

considering more general cases.

It is well-known that if n-dimensional random vector Z belongs to the *elliptical* family with characteristic generator ψ, then

1) a linear combination of Z, $l(Z)$, remains elliptical with the same characteristic generator ψ. Moreover,

2) the distribution of Z conditional on $l(Z)$ also remains elliptical, but in general with a different characteristic generator $\widetilde{\psi}$.

Although the previous two properties imply that the distribution of $Z_{\mathbf{A},f} := [Z \mid \mathbf{A}Z = f]$ remains in the elliptical family, for practical purposes we will focus our attention on a class of generalised hyperbolic distributions for which we have the following result.

Proposition 1 (Conditional Distribution of Z Given $F_s = f$)**.** *Define*

$$\begin{aligned} Z_f &:= Z \mid F_s = f \\ &= \left(m(W) + \sqrt{W}\mathbf{C}X\right) \;\middle|\; \mathbf{A}Z = f. \end{aligned}$$

[12] The material of this subsection is based on private discussions with Yixi Shi who also provided the explicit calculations that were required for Proposition 1.

[13] In this section, we use Z, F_s instead of $Z_t, F_{s,t}$ if their meanings are clear.

The conditional distribution of Z_f is characterized by the following:

1) The characteristic function of Z_f is given by

$$\phi_{Z_f}(\mathbf{t}) = \mathrm{E}\left[\exp\left(i\mathbf{t}\left(m(W) + \sqrt{W}\mu_{f,W}\right) - \frac{1}{2}W\mathbf{t}^\top\widetilde{\Sigma}\mathbf{t}\right) \;\middle|\; \mathbf{A}Z = f\right], \tag{4.19}$$

where

$$\begin{aligned}\mu_{f,W} &= \Sigma\mathbf{A}^\top\left(\mathbf{A}\Sigma\mathbf{A}^\top\right)^{-1}\left(f - \mathbf{A}m(W)\right)/\sqrt{W} \\ \widetilde{\Sigma} &= \Sigma - \Sigma\mathbf{A}^\top\left(\mathbf{A}\Sigma\mathbf{A}^\top\right)^{-1}\mathbf{A}\Sigma.\end{aligned} \tag{4.20}$$

2) The conditional distribution of W given $F_s = f$ remains a generalised inverse Gaussian (GIG) distribution:

$$\left(W \;\middle|\; \mathbf{A}Z = f\right) \sim N^-\left(\widetilde{\lambda}, \widetilde{\chi}, \widetilde{\theta}\right), \tag{4.21}$$

where

$$\begin{aligned}\widetilde{\lambda} &= \lambda - \frac{1}{2}L \\ \widetilde{\chi} &= \chi + (f - \mathbf{A}\mu)^\top\left(\mathbf{A}\Sigma\mathbf{A}^\top\right)^{-1}(f - \mathbf{A}\mu) \\ \widetilde{\theta} &= \theta + \gamma^\top\mathbf{A}^\top\left(\mathbf{A}\Sigma\mathbf{A}^\top\right)^{-1}\mathbf{A}\gamma\end{aligned} \tag{4.22}$$

Proof. See Appendix. □

Note that (4.19) implies a conditional normal distribution of Z with mean $\mu_{f,W}$ and variance $W\widetilde{\Sigma}$, given W and $\mathbf{A}Z = f$. Therefore Proposition 1 provides us with a very convenient method for simulating the unstressed factors conditional on the stressed scenarios.

Simulating Z Given $F_s = f$

At the beginning of simulation $\#i$, for $i = 1, \ldots, N$ we:

1. Draw W_i from $N^-\left(\widetilde{\lambda}, \widetilde{\chi}, \theta\right)$.

2. Draw Z_i from $N\left(m(W_i) + \sqrt{W_i}\mu_{f,W_i}, W_i\widetilde{\Sigma}\right)$.

This algorithm will allow us to estimate the expected P&L of (4.15).

4.6 Numerical Experiments

As a summary of our work in previous Sections, we describe the overall procedure of our factor model-based scenario analysis as a sequence of the following steps.

1) We get the time series of the daily variation of the implied volatilities, $\mathbf{y}_t$, using daily implied volatility surface data for past T trading days. Given all possible shifts $\mathbf{b}_1, \cdots, \mathbf{b}_M$, we also compute time series data of $\mathbf{r}_t$ via solving the optimization $(\mathcal{P}_1)$.

2) We perform PCA on the sample covariance matrix of $\hat{\mathbf{r}}_t$, normalization of $\mathbf{r}_t$. Time series data of the random vector process $Z_t = \left(R_t, {F_{m,t}}^\top, {F_{r,t}}^\top\right)^\top$ can be obtained from the PCA results as well as the historical underlying asset prices.

3) We fit the multivariate t distribution or skewed t distribution to the time series data of Z_t using the EM algorithm.

4) We estimate a conditional expectation (4.15) of each options portfolio P&L by simulating the idiosyncratic error ϵ_t and

$$[Z_{t+1} \mid F_{s,t+1} = f]$$

via the numerical scheme in Section 4.5.1 for each stressed values f.[14]

[14] Note that using each sample of conditional Z_{t+1}, we can get sample P&L's of options with different strikes and maturities simultaneously.

Now we have an estimate of expectation (4.15) for each stressed values f under our factor model-based methodology. To estimate the conditional expected P&L under the standard methodology as well,

5) We also calculate $\Delta V_{t+1}(R_t, F_{m,t}, \epsilon_{m,t})$ by simply plugging in $F_{s,t+1} = f$ and zeros for the non-stressed factors as described in Section 4.3.

To compare the performance of the standard and our factor model-based methodologies, for each options portfolio,

6) We compute difference between realized P&L and predicted P&L under each methodology. Note that we use the realized stressed values f at time $t+1$ to compute the predicted P&L.

Repeating steps 1-6 each day, we obtain a following 2-dimensional table

	T_1	T_2	$\cdots$	T_{n_t}
K_1	$(a_{1,1}, b_{1,1})$	$(a_{1,2}, b_{1,2})$		(a_{1,n_t}, b_{1,n_t})
K_2	$(a_{2,1}, b_{2,1})$	$(a_{2,2}, b_{2,2})$		(a_{2,n_t}, b_{2,n_t})
$\vdots$				
K_{n_k}	$(a_{n_k,1}, b_{n_k,1})$	$(a_{n_k,2}, b_{n_k,2})$		$(a_{n_k,n_t}, b_{n_k,n_t})$

whose each (i, j) component corresponds to a portfolio of a single *out-of-the-money* option with strike K_i and maturity T_j. In each cell we have two values $(a_{i,j}, b_{i,j})$ for the portfolio. These values are the *averages of the absolute differences* that we have collected in step 6. Each average $a_{i,j}$ is obtained from the standard methodology and each $b_{i,j}$ is obtained from our factor model-based

methodology. We will display tables with actual values as results of our numerical experiments in Section 4.6.3.

4.6.1 Data Set

We took the S&P 500 as the underlying asset and we looked at the daily implied volatilities for options with moneynesses

$$m \in \{0.5, 0.6, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.975, 1, 1.025, 1.05, 1.1, 1.2, 1.25, 1.3, 1.4, 1.5\}$$

and days-to-maturities

$$\tau \in \{30, 60, 90, 180, 365\}.$$

Note that we used the data from *Bloomberg* between October 19$^{\text{th}}$, 2010 and November 26$^{\text{th}}$, 2013. In fact, we used the data after smoothing it by ourselves in order to remove noise. In particular, we smooth the given implied volatility data using Nadaraya-Watson estimator[15] based on an independent bivariate Gaussian kernel. Especially we use bandwidths $h_1 = 0.006, h_2 = 0.14$ as per Da Fonseca and Gottschalk [18]. We simply viewed the smoothed data as true implied volatilities. For the simplicity, we assume constant interest rate $r = 0.02$ and constant dividend yield $q = 0.01$ throughout the period of time. Note that we can easily calculate call, put option prices using implied volatilities and constant r, q.[16]

4.6.2 Numerical Procedure

The following steps are the details of our numerical procedure which was briefly described earlier in this Section.

[15] Please see Appendix C.1.

[16] Put-call parity is automatically satisfied as long as we use the same implied volatility and r, q for call and put options corresponding to the same moneyness and maturity.

1) We use a rolling $T = 250$ days window. So each day t, we obtain time series $\{\mathbf{r}_{t-249}, \cdots, \mathbf{r}_t\}$ of the residual vector using a data set D_t which is a market data of option prices and underlying asset prices between $t - 249$ and t.

2) We then perform PCA on the sample covariance matrix of $\mathbf{r}_u$. Let us denote by $\gamma_1, \gamma_2, \cdots$, the eigenvectors associated with the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$. We include the first 4 principal components in our factor model (4.13), i.e. $p = 4$. Note that we only look at those options which are worth at least 50 cents at time t, i.e. we only use the components of r_u corresponding to those options for PCA.

3) For each out-of-the-money option with given moneyness and days-to-maturity, we determine the position in the option by using 95% historical VaR. Historical VaR is calculated using the daily changes in the price of one unit of option for past 250 days. We determine the position that gives 95% VaR of $ 1 million.

4) Using the data set D_t and PCA results, we get time series $\{Z_{t-249}, \cdots, Z_t\}$ of the random vector process Z_u. We fit the multivariate t distribution and skewed t distribution to this time series via the EM algorithm to find model parameters.

5) Using the fitted parameters, we simulate 10^4 sample Z_{t+1}'s conditional on the realized $F_{s,t+1}$ as per the numerical scheme described in Section 4.5.1. We also simultaneously simulate the idiosyncratic term ϵ_{t+1}.[17] Given each simulated Z_{t+1}, ϵ_{t+1}, we calculate simulated option

[17] In our new factor model (4.13), ϵ_t in (4.12) is assumed to be independent of $F_{m,t}$ and $F_{r,t}$. When simulating ϵ_t numerically, we truncate the summation (4.12) and approximate ϵ_t as

$$\epsilon_t \approx \sum_{i=p+1}^{k} \gamma_i P_t^{(i)}$$

for some $p + 1 \leq k \leq N$ and further assume that each $P_t^{(i)}$ follows an independent normal distribution, $\mathrm{N}(0, \lambda_i)$. Particularly we set $k = 10$. Of course, the other distributions such as a t distribution could be used instead of a

price of each option at $t+1$. Averaging the sample option prices gives an estimation of the conditional expectation (4.15) for each single option portfolio P&L. This is an estimation under our factor model-based methodology.

6) We also estimate a conditional expected P&L under the standard methodology quickly as described in Section 4.3.

7) The above steps all together take about 5 to 10 seconds when coded in Matlab on a standard desktop computer. We repeat these steps each day until the last available trading day. Then we finally create the table described earlier in this Section while considering those options which were worth at least 50 cents for more than half of total number of days.

Note that we consider the following three different shifts in the implied volatility surface after normalization: (i) $\mathbf{b}_1$, one vol point *parallel shift*, (ii) $\mathbf{b}_2$, linear *skew shift* (with positive one vol point shift for options with moneyness 0.5, negative on vol point shift for options with moneyness 1.5 and linearly interpolated in between), (iii) $\mathbf{b}_3$, *term structure shift* proportional to the square root of time-to-maturity.

4.6.3 Numerical Results

We mainly report three tables associated with different stress factors shocks: (i) the underlying return Table 4.1, (ii) the underlying return and the skew shift for Table 4.2, (iii) the skew shift and the term structure shift Table 4.3. Note that we only display our numerical results using t distributions as we could not observe significant improvement when using skewed t distributions.

When we evaluate performance of each methodology in each scenario analysis, we focus on the

normal distribution.

following statistics.

- E_{st}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] for each single out-of-the-money option portfolio under the standard scenario analysis,
- E_{fm}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] for each single out-of-the-money option portfolio under the improved scenario analysis.

Each table consists of four subtables. The first one shows the ratio $E_{fm}^{abs}/E_{st}^{abs}$ for each option. The small numbers below 100% implies that our factor model-based methodology performs better than the standard methodology. The second subtable analyzes the performance of our methodology from a different angle. It shows the proportion of days that the P&L predicted by our factor model-based methodology is closer to the realized P&L than the one predicted by the standard methodology. We would like to highlight that even if one methodology performs better, sometimes the P&L predicted by another methodology could happen to be closer to the realized P&L. However, in the long run, realized P&L should be more likely to be closer to the P&L predicted by the better methodology. The third and last subtables show E_{st}^{abs} and E_{fm}^{abs} for each option.

In all three table, we could observe that our new methodology outperforms the standard methodology. In Table 4.1, we see that the absolute difference between realized and expected P&L has been decreased by 30% on average when using our new methodology. For each option, our new methodology also gives more accurate prediction each day with probability around 60% to 70%. This is a quite high probability considering the fact that the realized P&L can be viewed as a random number drawn from some true distribution. However this improvement is what we have expected since, when the future underlying return is given, the information of the future movement in the implied volatility surface is also incorporated in the given condition. For example,

when the underlying asset price moves down, it is likely that the implied volatilities increase and the volatility curves have steeper skews. Through effectively extracting the information, we could make more accurate predictions than when we simply assume that the implied volatility surface remains unchanged. In the other two tables, we could observe significant improvements for the same reason.

E_{st}^{abs} and E_{fm}^{abs} are smaller in Table 4.2 than in Table 4.1 since we have more information about the future event in the second experiment. E_{st}^{abs} and E_{fm}^{abs} are even smaller in Table 4.3 than in Table 4.2 and it is because the changes in option prices are more sensitive to the implied volatility than the underlying price in general.

4.7 Conclusion

We have developed a factor modeling approach for performing scenario analysis for options portfolios with a single underlying security. We use principal components analysis to estimate the factor model but only after already including a pre-specified set of stress factors. The expected portfolio loss in each possible scenario is estimated using an efficient Monte-Carlo scheme that conditions on the values of the shocks to each of the stressed factors under consideration. Our approach appears to perform very well, and certainly in comparison to the more standard form of scenario analysis. Clearly it can also be extended to other asset classes. We also note that it should also be possible to use other approaches

There are many possible directions for future research. Of immediate concern is to extend the modeling approach to account for current market conditions. This will require a time series

Table 4.1: Numerical results when stress factor is underlying return

Ratio of E_{fm}^{abs} to E_{st}^{abs}

	30D	60D	90D	180D	365D
50%					78%
60%					75%
70%				71%	70%
75%				70%	69%
80%			66%	68%	68%
85%		63%	65%	66%	67%
90%	59%	63%	64%	65%	67%
95%	63%	64%	65%	66%	69%
97.5%	66%	67%	67%	68%	71%
100%	72%	72%	72%	73%	77%
102.5%	67%	68%	69%	70%	73%
105%	65%	67%	68%	69%	71%
110%		69%	70%	71%	69%
115%			67%	76%	70%
120%				81%	73%
125%					75%
130%					77%
140%					76%

Probability that our new method beats the standard method

	30D	60D	90D	180D	365D
50%					63%
60%					64%
70%				63%	65%
75%				65%	66%
80%			66%	66%	67%
85%		67%	66%	65%	68%
90%	72%	68%	68%	66%	67%
95%	68%	68%	68%	67%	66%
97.5%	67%	66%	65%	66%	66%
100%	62%	62%	62%	63%	63%
102.5%	65%	65%	64%	64%	65%
105%	64%	64%	63%	64%	66%
110%		60%	60%	61%	65%
115%			61%	59%	63%
120%				59%	62%
125%					61%
130%					62%
140%					62%

E_{st}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] (unit: $\$10^5$)

	30D	60D	90D	180D	365D
50%					3.0
60%					3.4
70%				2.7	3.8
75%				3.0	4.0
80%			2.9	3.4	4.2
85%		2.9	3.2	3.7	4.3
90%	3.3	3.3	3.6	4.0	4.4
95%	3.5	3.8	4.0	4.1	4.4
97.5%	3.9	4.1	4.1	4.2	4.4
100%	4.1	4.1	4.1	4.1	4.4
102.5%	3.8	3.9	4.0	4.1	4.4
105%	3.2	3.5	3.7	3.9	4.3
110%		2.6	2.9	3.3	4.0
115%			2.3	2.9	3.6
120%				2.5	3.3
125%					3.0
130%					2.8
140%					2.8

E_{fm}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] (unit: $\$10^5$)

	30D	60D	90D	180D	365D
50%					2.4
60%					2.6
70%				1.9	2.7
75%				2.1	2.8
80%			1.9	2.3	2.8
85%		1.8	2.1	2.5	2.9
90%	1.9	2.1	2.3	2.6	3.0
95%	2.2	2.5	2.6	2.7	3.1
97.5%	2.6	2.7	2.8	2.8	3.2
100%	2.9	3.0	3.0	3.0	3.4
102.5%	2.5	2.7	2.7	2.8	3.2
105%	2.1	2.4	2.5	2.7	3.1
110%		1.8	2.0	2.3	2.8
115%			1.5	2.2	2.5
120%				2.1	2.4
125%					2.2
130%					2.1
140%					2.1

Table 4.2: Numerical results when stress factors are underlying return and skew shift

Ratio of E_{fm}^{abs} to E_{st}^{abs}

	30D	60D	90D	180D	365D
50%					84%
60%					76%
70%				62%	68%
75%				60%	65%
80%			54%	58%	62%
85%		53%	55%	57%	60%
90%	48%	55%	55%	56%	58%
95%	56%	56%	57%	56%	56%
97.5%	56%	56%	56%	55%	54%
100%	56%	56%	56%	55%	52%
102.5%	56%	55%	55%	54%	50%
105%	54%	55%	55%	53%	49%
110%		58%	58%	57%	48%
115%			55%	62%	50%
120%				68%	51%
125%					52%
130%					52%
140%					54%

Probability that our new method beats the standard method

	30D	60D	90D	180D	365D
50%					53%
60%					60%
70%				63%	63%
75%				65%	64%
80%			68%	66%	65%
85%		67%	66%	67%	66%
90%	71%	68%	68%	68%	68%
95%	67%	67%	67%	68%	68%
97.5%	67%	68%	68%	68%	70%
100%	67%	68%	67%	69%	71%
102.5%	68%	68%	68%	70%	72%
105%	67%	68%	68%	70%	72%
110%		65%	65%	66%	72%
115%			67%	65%	70%
120%				63%	70%
125%					70%
130%					71%
140%					75%

E_{st}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] (unit: $\$10^5$)

	30D	60D	90D	180D	365D
50%					2.1
60%					2.4
70%				2.1	3.0
75%				2.4	3.4
80%			2.5	2.9	3.7
85%		2.6	2.9	3.4	4.1
90%	3.1	3.1	3.4	3.9	4.6
95%	3.5	3.9	4.1	4.3	5.0
97.5%	4.2	4.4	4.5	4.7	5.4
100%	4.9	4.9	5.0	5.1	6.0
102.5%	4.3	4.6	4.7	5.0	6.0
105%	3.7	4.1	4.4	4.8	6.0
110%		3.1	3.4	4.1	5.5
115%			2.8	3.5	5.1
120%				3.0	4.7
125%					4.4
130%					4.1
140%					3.8

E_{fm}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] (unit: $\$10^5$)

	30D	60D	90D	180D	365D
50%					1.8
60%					1.9
70%				1.3	2.0
75%				1.5	2.2
80%			1.3	1.7	2.3
85%		1.4	1.6	1.9	2.5
90%	1.5	1.7	1.9	2.2	2.6
95%	2.0	2.2	2.3	2.4	2.8
97.5%	2.4	2.5	2.5	2.6	2.9
100%	2.8	2.8	2.8	2.8	3.1
102.5%	2.4	2.5	2.6	2.7	3.0
105%	2.0	2.3	2.4	2.6	2.9
110%		1.8	2.0	2.3	2.7
115%			1.5	2.2	2.5
120%				2.1	2.4
125%					2.3
130%					2.1
140%					2.1

Table 4.3: Numerical results when stress factors are skew and term structure shifts

Ratio of E_{fm}^{abs} to E_{st}^{abs}

	30D	60D	90D	180D	365D
50%					98%
60%					97%
70%				60%	96%
75%				62%	96%
80%			56%	66%	95%
85%		63%	65%	71%	94%
90%	70%	73%	73%	76%	93%
95%	84%	80%	78%	79%	92%
97.5%	86%	80%	78%	78%	89%
100%	87%	81%	78%	77%	86%
102.5%	83%	77%	75%	74%	80%
105%	79%	74%	72%	70%	75%
110%		69%	69%	69%	71%
115%			58%	67%	69%
120%				65%	68%
125%					66%
130%					65%
140%					71%

Probability that our new method beats the standard method

	30D	60D	90D	180D	365D
50%					51%
60%					52%
70%				64%	54%
75%				64%	54%
80%			66%	62%	55%
85%		65%	63%	62%	56%
90%	67%	63%	62%	61%	57%
95%	60%	62%	62%	62%	58%
97.5%	60%	62%	62%	61%	60%
100%	61%	63%	62%	62%	62%
102.5%	62%	64%	63%	63%	62%
105%	62%	63%	62%	63%	64%
110%		61%	61%	62%	66%
115%			65%	62%	65%
120%				62%	65%
125%					65%
130%					65%
140%					66%

E_{st}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] (unit: $\$10^5$)

	30D	60D	90D	180D	365D
50%					1.7
60%					1.9
70%				1.5	2.3
75%				1.8	2.5
80%			1.8	2.2	2.7
85%		1.8	2.1	2.5	3.0
90%	2.0	2.2	2.5	2.9	3.3
95%	2.3	2.8	3.0	3.2	3.5
97.5%	2.7	3.1	3.2	3.4	3.6
100%	2.9	3.2	3.3	3.5	3.8
102.5%	2.8	3.2	3.4	3.6	4.1
105%	2.4	3.0	3.2	3.7	4.4
110%		2.1	2.5	3.2	4.3
115%			2.0	2.7	4.1
120%				2.3	3.9
125%					3.7
130%					3.5
140%					3.3

E_{fm}^{abs} = $\mathbb{E}$ [|realized P&L – expected P&L|] (unit: $\$10^5$)

	30D	60D	90D	180D	365D
50%					1.7
60%					1.8
70%				0.9	2.2
75%				1.1	2.4
80%			1.0	1.4	2.6
85%		1.2	1.4	1.8	2.8
90%	1.4	1.6	1.9	2.2	3.0
95%	2.0	2.2	2.3	2.5	3.2
97.5%	2.4	2.5	2.5	2.6	3.2
100%	2.5	2.6	2.6	2.7	3.3
102.5%	2.3	2.5	2.5	2.7	3.3
105%	1.9	2.2	2.3	2.6	3.3
110%		1.5	1.7	2.2	3.1
115%			1.1	1.8	2.9
120%				1.5	2.6
125%					2.4
130%					2.3
140%					2.4

approach[18] but it should yield superior expected loss forecasts as they should properly account for the stress under consideration as well as current market conditions. We would also like to see how the factor-model approach works for other types of portfolios and asset classes. Finally, we would like to perform a more detailed numerical study than the preliminary study presented here.

[18] See Tsay [36] and McNeil, Frey and Embrechts [31] for time series applications to risk-management.

Bibliography

[1] L. Andersen. Efficient simulation of the Heston stochastic volatility model, January 2007.

[2] M. Avellaneda and S.J.Zhang. Path-dependence of leveraged ETF returns. *SIAM Journal of Financial Mathematics*, 1:586–603, 2010.

[3] O.E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B*, 63(2):167–241, 2001.

[4] BARRA. The Barra US equity model (USE4), 2011.

[5] D. Bates. Exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*, 7:69–107, 1996.

[6] D.P. Bertsekas. *Dynamic Programming and Optimal Control Vol I, 3rd edition.* Athena Scientific, 2005.

[7] D.P. Bertsekas. *Dynamic Programming and Optimal Control Vol II, 4th edition.* Athena Scientific, 2012.

[8] P. Carr and D. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2:61–73, 1998.

[9] M. Cheng and A. Madhavan. The dynamics of leveraged and inverse exchange traded funds. *Journal of Investment Management*, Winter 2009, 2009.

[10] R. Cont and J. Da Fonseca. Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60, 2002.

[11] J. Cox and C.F. Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economics Theory*, 49:33–83, 1989.

[12] D. P. de Farias and B. Van Roy. The linear programming approach to approximate dynamic programming. *Operations Research*, 51(6):850–865, 2003.

[13] D. P. de Farias and B. Van Roy. On constraint sampling in the linear programming approach to approximate dynamic programming. *Mathematics of Operations Research*, 29(3):462–478, 2004.

[14] D. Duffie. *Dynamic Asset Pricing Theory*. Princeton University Press, 1996.

[15] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[16] Chicago Board Options Exchange. CBOE 2010 market statistics, 2010.

[17] E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, 1993.

[18] J. Da Fonseca and K. Gottschalk. A joint analysis of the term structure of credit default swap spreads and the implied volatiltiy surface. *Journal of Futures Markets*, 33(6):494–517, 2013.

[19] J. Han and B. Van Roy. Control of diffusions via linear programming. In Gred Infanger, editor, *Stochastic Programming: The State of the Art, in Honor of George B. Dantzig*, pages 329–354. Springer, 2011.

[20] M.B. Haugh. A note on constant proportion trading strategies. *Operations Research Letters*, 39(3):172–179, 2011.

[21] M.B. Haugh and A. Jain. The dual approach to portfolio evaluation: A comparison of the static, myopic and generalized buy-and-hold strategies. *Quantitative Finance*, 11(1):81–99, 2011.

[22] M.B. Haugh, L. Kogan, and J. Wang. Evaluating portfolio strategies: A duality approach. *Operations Research*, 54(3):405–418, 2006.

[23] M.B. Haugh, L. Kogan, and Z. Wu. Portfolio optimization with position constraints: An approximate dynamic programming approach. *Columbia University and MIT*, Working paper, 2006.

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

[25] W. Hu. *Calibration of Multivariate Generalized Hyperbolic Distributions using the EM Algorithm, with Applications in Risk Management, Portfolio Optimization and Portfolio Credit Risk*. PhD thesis, Florida State University, 2005.

[26] I. Karatzas and S. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, 1998.

[27] T. Leung and R. Sircar. Implied volatility of leveraged ETF options, October 2012.

[28] J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, pages 13–37, 1965.

[29] A. Lynch. Portfolio choice and equity characteristics: Characterizing the hedging demends induced by return predictability. *Journal of Financial Economics*, 62:67–130, 2001.

[30] A.W. Marshall and I. Olkin. A multivariate exponential distribution. *Journal of the American Statistical Association*, 62:30–44, 1967.

[31] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, 2010.

[32] A. Meucci. *Risk and Asset Allocation*. Springer, 2009.

[33] R. Rebonato. A Bayesian approach to stress testing and scenario analysis. *Journal of Investment Management*, 8(3):1–13, 2010.

[34] P. Schweitzer and A. Seidmann. Generalized polynomial approximations in Markovian decision processes. *Journal of Mathematical Analysis and Applications*, 110:568–582, 1985.

[35] W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.

[36] R. S. Tsay. *Analysis of Financial Time Series*, volume 543. John Wiley & Sons, 2005.

[37] J. Wachter and A. Sangvinatsos. Does the failure of the expectations hypothesis matter for long-term investors? *Journal of Finance*, 60:179–230, 2005.

[38] J. Zhang. *Path Dependence Properties of Leveraged Exchange-Traded Funds: Compounding, Volatility and Option Pricing.* PhD thesis, NYU, 2010.

[39] J. Zhu. *Applications of Fourier Transform to Smile Modeling.* Springer, 2010.

Appendix A

Appendix for Chapter 2

A.1 Log-Price Characteristic Functions

The Heston Model: The characteristic function of the log-security price under the Heston model is (see [39], for example) given by

$$\begin{aligned}\Phi_T^{SV}(u;r,q,\kappa,\gamma,\theta,V_0,\rho,S_0) = \exp\left[iu(\log(S_0)+(r-q)T)\right] \\ \times \exp\left[\theta\kappa\gamma^{-2}\left((\kappa-\rho\gamma ui-d)T-2\log((1-g\exp(-dT))/(1-g))\right)\right] \\ \times \exp\left[V_0\gamma^{-2}(\kappa-\rho\gamma ui-d)\left(1-\exp(-dT)\right)/\left(1-g\exp(-dT)\right)\right]\end{aligned} \tag{A.1}$$

where $d := \sqrt{(\rho\gamma ui-\kappa)^2+\gamma^2(iu+u^2)}$ and $g := (\kappa-\rho\gamma ui-d)/(\kappa-\rho\gamma ui+d)$.

The Bates Model: The characteristic function of the log-security price under the SVJ model is

$$\begin{aligned}\Phi_T^{SVJ}(u;r,q,\kappa,\gamma,\theta,V_0,\rho,S_0,\lambda,a,b) = \Phi_T^{SV}(u\ ;r,q,\kappa,\gamma,\theta,V_0,\rho,S_0)\times\exp(-\lambda miuT) \\ \times\exp\left[\lambda T\left(\exp\left(aiu-\frac{b^2u^2}{2}\right)-1\right)\right].\end{aligned} \tag{A.2}$$

where $m \coloneqq \exp\left(a + \frac{b^2}{2}\right) - 1$.

BN-S Model: The characteristic function of the log-security price under the BNS model is

$$\begin{aligned}\Phi_T^{BNS}(u; r, q, a, b, V_0, \lambda, \rho, S_0) = \exp\left[iu\left(\log(S_0) + (r - q - a\lambda\rho(b-\rho)^{-1})T\right)\right] \\ \times \exp\left[-\lambda^{-1}(u^2 + ui)(1 - \exp(-\lambda T))V_0/2\right] \\ \times \exp\left[a(b - f_2)^{-1}\left(b\log\left(\frac{b - f_1}{b - ui\rho}\right) + f_2\lambda T\right)\right]\end{aligned} \tag{A.3}$$

where

$$f_1 = f_1(u) = ui\rho - \lambda^{-1}(u^2 + ui)(1 - \exp(-\lambda T))/2$$

$$f_2 = f_2(u) = ui\rho - \lambda^{-1}(u^2 + ui)/2.$$

A.2 The Jump Approximation for the SVJ Model

First recall that $X \coloneqq [\log(\phi(Y-1)+1) \mid \phi(Y_i - 1) + 1 > 0]$ which we write as

$X = [g(\log(Y_i)) \mid \phi(Y_i - 1) + 1 > 0]$ where $g(x) \coloneqq \log(\phi(e^x - 1) + 1))$. We can now compute the density, $p(\cdot)$, of X. Letting $f(\cdot; a, b)$ denote the density function of $\log(Y) \sim \mathrm{N}(a, b)$ we obtain

$$\begin{aligned} p(x) &= \frac{f(g^{-1}(x); a, b)}{\mathbb{P}(\phi(Y_i - 1) + 1 > 0)} \cdot \left|\frac{d}{dx}(g^{-1}(x))\right| \\ &= \frac{\mathrm{sign}(\phi)}{1 - p^*} \cdot f\left(\log\left(\frac{e^x + \phi - 1}{\phi}\right); a, b\right) \cdot \frac{e^x}{e^x + \phi - 1}. \end{aligned}$$

Similarly we can compute the density function, $q(\cdot)$, of the $\mathrm{N}(\hat{a}, \hat{b}) - \mathrm{Exp}(\hat{c})$ distribution. Indeed it is straightforward to show that

$$q(x; \hat{a}, \hat{b}, \hat{c}) = \hat{c} \cdot \exp\left(\hat{c}(x - \hat{a}) + \frac{\hat{b}^2\hat{c}^2}{2}\right) \cdot F\left(\frac{-x + \hat{a} - \hat{b}^2\hat{c}}{\hat{b}}\right) \tag{A.4}$$

where $F(\cdot)$ is the CDF of the standard normal distribution. We would like to choose $(\hat{a}, \hat{b}, \hat{c})$ by minimizing the

$$\int \left(p(x) - q(x; \hat{a}, \hat{b}, \hat{c})\right)^2 dx \tag{A.5}$$

subject to $\hat{b}, \hat{c} \geq 0$. The interval of integration in (A.5) depends on the sign of ϕ: it is $(-\infty, \infty)$ if $\phi > 0$ and $(-\infty, \log(1-\phi))$ if $\phi < 0$. Instead of minimizing (A.5) we instead solve

$$\begin{aligned} \min_{\hat{a},\hat{b},\hat{c}} \quad & \sum_{i=1}^{n} |p(x_i) - q(x_i; \hat{a}, \hat{b}, \hat{c})|^2 \\ \text{subject to} \quad & \hat{b}, \hat{c} \geq 0 \end{aligned} \tag{A.6}$$

where $x_1 < x_2 < \cdots < x_n$ are equally spaced points with most of the probability mass of p in $[x_1, x_n]$. Since the optimization problem in (A.6) is non-convex we can only obtain a local minimum rather than a global minimum. This was never a problem in our numerical experiments, however. Moreover, it would be easy to choose several different starting points for this problem and terminate once a sufficiently good solution was found. Note that this could all be easily automated.

A.3 The SVCJ Model

Following Duffie et al [15] we can use the SVCJ model to price options on the underlying ETF. Indeed the characteristic function of the log-ETF price under the SVCJ model is given by

$$\Phi_T^{SVCJ}(u; r, q, \kappa, \gamma, \theta, V_0, \rho, S_0, \lambda, a, b, \rho_J, \mu_v) = \exp(A(0,T,u) + iu\log(S_0) + C(0,T,u)V_0) \tag{A.7}$$

where

$$C(t,T,u) = -\frac{a_1(1-e^{-a_4\tau})}{2a_4 - (a_2+a_4)(1-e^{-a_4\tau})}$$

$$\text{and} \quad A(t,T,u) = A_0(t,T,u) - \lambda\tau(1+miu) + \lambda\exp\left(iau - \frac{b^2u^2}{2}\right)A_1(t,T,u)$$

where $a_1 = iu(1-iu)$, $a_2 = i\gamma\rho u - \kappa, a_3 = 1 - i\rho_J\mu_v u$, $a_4 = \sqrt{a_2^2 + a_1\gamma^2}$, $\tau = T - t$ and

$$A_0(t,T,u) = i(r-q)u\tau - \kappa\theta\left(\frac{a_2+a_4}{\gamma^2}\tau + \frac{2}{\gamma^2}\log\left[1 - \frac{a_2+a_4}{2a_4}(1-e^{-a_4\tau})\right]\right),$$

$$A_1(t,T,u) = \frac{a_4-a_2}{(a_4-a_2)a_3+\mu_v a_1}\tau - \frac{2\mu_v a_1}{(a_3a_4)^2-(a_2a_3-\mu_v a_1)^2}\log\left[1 - \frac{(a_2+a_4)a_3-\mu_v a_1}{2a_3a_4}(1-e^{-a_4\tau})\right].$$

A.3.1 The Bivariate Exponential Distribution

The bivariate exponential (BVE) distribution is a bivariate distribution with exponential marginals. It has joint density

$$f(x,y) = \begin{cases} \lambda_2(\lambda_1+\lambda_{12})\bar{F}(x,y), & x > y, \\ \\ \lambda_1(\lambda_2+\lambda_{12})\bar{F}(x,y), & x < y \end{cases}$$

where

$$\bar{F}(s,t) := P(X > s, Y > t) = \exp\left[-\lambda_1 s - \lambda_2 t - \lambda_{12}\max(s,t)\right], \quad s,t > 0. \tag{A.8}$$

The marginal distribution functions then satisfy $\bar{F}_1(x) = e^{-(\lambda_1+\lambda_{12})x}$ and $\bar{F}_2(y) = e^{-(\lambda_2+\lambda_{12})y}$, and we write $(X,Y) \sim \text{BVE}(\lambda_1,\lambda_2,\lambda_{12})$. The characteristic function for the BVE is given by

$$\int_0^\infty \int_0^\infty e^{isx+ity} dF(x,y) = \frac{(\lambda - is - it)(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12}) + st\lambda_{12}}{(\lambda-is-it)(\lambda_1+\lambda_{12}-is)(\lambda_2+\lambda_{12}-it)} \tag{A.9}$$

where $\lambda := \lambda_1+\lambda_2+\lambda_{12}$. Using (A.9) we can calculate the joint characteristic function of $\hat{X}$ in (2.37) to obtain

$$\begin{aligned}
\Phi_{\hat{X}}(u_1,u_2;a,b,\lambda_1,\lambda_2,\lambda_{12}) &= \mathbb{E}[\exp((N-E_1)iu_1 + E_2iu_2)] \\
&= \mathbb{E}[\exp(Niu_1)] \cdot \mathbb{E}[\exp(E_1(-iu_1) + E_2(iu_2))] \\
&= \exp\left(aiu_1 - \frac{1}{2}b^2u_1^2\right) \cdot \frac{(\lambda + iu_1 - iu_2)(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12}) - u_1u_2\lambda_{12}}{(\lambda+iu_1-iu_2)(\lambda_1+\lambda_{12}+iu_1)(\lambda_2+\lambda_{12}-iu_2)}.
\end{aligned} \tag{A.10}$$

A.3.2 The Characteristic Function of the Approximated log-LETF Price

The characteristic function of the approximated log-LETF price conditional on $N_1(T) = 0$ is given by

$$\hat{\Phi}_L^{N_1\equiv 0}(u;\phi,r,q,f,\kappa,\gamma,\theta,V_0,\rho,L_0,\lambda,a,b,\rho_J,\mu_v) = \exp(\hat{A}(0,T,u)+\hat{B}(0,T,u)\log(L_0)+\hat{C}(0,T,u)V_0^L)$$

where $\hat{A},\hat{B},\hat{C}$ satisfy the following ODEs[1]:

$$\begin{aligned}
\frac{d\hat{B}}{dt} &= 0,\\
\frac{d\hat{C}}{dt} &= -\frac{1}{2}\hat{B}^2 - \hat{B}\hat{C}\rho_L\gamma_L - \frac{1}{2}\hat{C}^2\gamma_L^2 + \frac{1}{2}\hat{B} + \kappa_L\hat{C},\\
\frac{d\hat{A}}{dt} &= -(r - q_L - \lambda\hat{m})C - \kappa_L\theta_L\hat{B} + \lambda_L - \lambda_L\cdot\Phi_{\hat{X}}(u,\hat{B};\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12}),
\end{aligned}$$

with boundary conditions $\hat{B}(T,T,u) = iu$, $\hat{C}(T,T,u) = 0$ and $\hat{A}(T,T,u) = 0$, and where

$$(q_L,\kappa_L,\gamma_L,\theta_L,V_0^L,\rho_L,L_0,\lambda_L,\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12}) :=$$

$$(\phi q + f,\kappa,|\phi|\gamma,\phi^2\theta,\phi^2V_0,\mathrm{sign}(\phi)\rho,L_0,\lambda(1-p^*),a^*,b^*,\lambda_1^*,\lambda_2^*,\lambda_{12}^*).$$

Note that $\hat{m}$ and $\Phi_{\hat{X}}$ are specified in equation (2.39) and (A.10). We solved these ODEs and they have the following explicit solution:

$$\begin{aligned}
\hat{B}(t,T,u) &= iu,\\
\hat{C}(t,T,u) &= -\frac{\hat{a}_1(1-e^{-\hat{a}_3\tau})}{2\hat{a}_3-(\hat{a}_2+\hat{a}_3)(1-e^{-\hat{a}_3\tau})},\\
\hat{A}(t,T,u) &= \hat{A}_0(t,T,u) - \lambda\tau(1+\hat{m}iu) + \lambda_L\exp\left(\hat{a}iu - \frac{\hat{b}^2u^2}{2}\right)\hat{A}_1(t,T,u),
\end{aligned}$$

where $\hat{a}_1 = iu(1-iu)$, $\hat{a}_2 = \gamma_L\rho_L iu - \kappa_L$, $\hat{a}_3 = \sqrt{\hat{a}_2^2+\hat{a}_1\gamma_L^2}$, $\tau = T-t$,

$$\alpha(x) = \frac{\hat{a}_3-\hat{a}_2}{x(\hat{a}_3-\hat{a}_2)+\hat{a}_1}\tau - \frac{2\hat{a}_1}{(x\hat{a}_3)^2-(x\hat{a}_3-\hat{a}_1)^2}\log\left[1-\frac{(\hat{a}_2+\hat{a}_3)x-\hat{a}_1}{2\hat{a}_3}(1-e^{-\hat{a}_3\tau})\right]$$

[1] See Duffie et al [15] for the derivation of these ODEs.

and

$$\hat{A}_0(t,T,u) = i(r-q_L)u\tau - \kappa_L\theta_L\left(\frac{\hat{a}_2+\hat{a}_3}{\gamma_L^2}\tau + \frac{2}{\gamma_L^2}\log\left[1-\frac{\hat{a}_2+\hat{a}_3}{2\hat{a}_3}(1-e^{-\hat{a}_3\tau})\right]\right),$$

$$\hat{A}_1(t,T,u) = \alpha(\lambda_2+\lambda_{12})\cdot\frac{\lambda_2+\lambda_{12}}{(\lambda_1+\lambda_{12}+iu)(\lambda_1+iu)}\left[(\lambda_1+\lambda_{12})(\lambda_1+iu)+iu\lambda_{12}\right]$$
$$-\alpha(\lambda_1+\lambda_2+\lambda_{12}+iu)\cdot\frac{iu\lambda_{12}(\lambda+iu)}{(\lambda_1+\lambda_{12}+iu)(\lambda_1+iu)}.$$

A.3.3 The Jump Approximation for the SVCJ Model

Let $h(\cdot)$ denote the density of $\log(Y_i) \sim \mathrm{N}(a,b) + \mathrm{sign}(\rho_J)\cdot\mathrm{Exp}(c)$, where $c = |\rho_J\mu_v|^{-1}$ and where $Y_i - 1$ is the relative jump size in the SVCJ model. It is easy to see that

$$h(x) = \begin{cases} c\exp\left(c(x-a)+\frac{b^2c^2}{2}\right)\cdot F\left(\frac{-x+a-b^2c}{b}\right), & \rho_J < 0, \\ c\exp\left(-c(x-a)+\frac{b^2c^2}{2}\right)\cdot F\left(\frac{x-a-b^2c}{b}\right), & \rho_J > 0 \end{cases} \tag{A.11}$$

where as before $F(\cdot)$ denotes the standard normal CDF. Recall that $p_1(\cdot)$ is the marginal density of the first component of X as defined in (2.36). It satisfies

$$p_1(x) = \frac{\mathrm{sign}(\phi)}{1-p^*}\cdot h\left(\log\left(\frac{e^x+\phi-1}{\phi}\right)\right)\cdot\frac{e^x}{e^x+\phi-1}$$

where $h(\cdot)$ is given in (A.11). Note that $h(x)$ is well-defined only for $x < \log(1-\phi)$ when $\phi < 0$. We can also see that $q_1(\cdot)$, the marginal density of the first component of $\hat{X}$ in (2.37), is given by

$$q_1(x;\hat{a},\hat{b},\hat{c}) = \hat{c}\exp\left(\hat{c}(x-\hat{a})+\frac{\hat{b}^2\hat{c}^2}{2}\right)\cdot F\left(\frac{-x+\hat{a}-\hat{b}^2\hat{c}}{\hat{b}}\right)$$

which of course has the same form as (A.4). Recall also that $p_2(\cdot)$ is the marginal density of the second component of X. It may be calculated as

$$p_2(y) = \begin{cases} \frac{1}{1-p^*}\cdot\left(1-F\left(\log\left(\frac{\phi-1}{\phi}\right)-\frac{\rho_J}{\phi^2}y;\ a,b\right)\right)\cdot\frac{1}{\phi^2\mu_v}\exp\left(-\frac{1}{\phi^2\mu_v}y\right), & \phi > 0, \\ \frac{1}{1-p^*}\cdot F\left(\log\left(\frac{\phi-1}{\phi}\right)-\frac{\rho_J}{\phi^2}y;\ a,b\right)\cdot\frac{1}{\phi^2\mu_v}\exp\left(-\frac{1}{\phi^2\mu_v}y\right), & \phi < 0 \end{cases}$$

for all $y \geq 0$. The marginal density of the second component of $\hat{X}$ is simply $q_2(y;\hat{d}) = \hat{d}\exp(-\hat{d}y)$. Finally, we need to compute the joint densities. The joint density of X is given by

$$p(x,y) = \frac{\text{sign}(\phi)}{1-p^*} \cdot f\left(\log\left(\frac{e^x+\phi-1}{\phi}\right) - \frac{\rho_J}{\phi^2}y; a, b\right) \cdot \frac{1}{\phi^2\mu_v}\exp\left(-\frac{1}{\phi^2\mu_v}y\right) \cdot \frac{e^x}{e^x+\phi-1}.$$

The joint density of $\hat{X}$ can be calculated as

$$\begin{aligned} q(x,y;\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12}) = &\int_0^y \frac{1}{\sqrt{2\pi\hat{b}}}\exp\left(-\frac{(x+z-\hat{a})^2}{2\hat{b}^2}\right)\cdot\lambda_1(\lambda_2+\lambda_{12})\exp\left(-\lambda_1 z-(\lambda_2+\lambda_{12})y\right)dz \\ &+\int_y^\infty \frac{1}{\sqrt{2\pi\hat{b}}}\exp\left(-\frac{(x+z-\hat{a})^2}{2\hat{b}^2}\right)\cdot\lambda_2(\lambda_1+\lambda_{12})\exp\left(-(\lambda_1+\lambda_{12})z-\lambda_2 y\right)dz \\ &+\frac{1}{\sqrt{2\pi\hat{b}}}\exp\left(-\frac{(x+y-\hat{a})^2}{2\hat{b}^2}\right)\cdot\lambda_{12}\exp\left(-(\lambda_1+\lambda_2+\lambda_{12})y\right) \end{aligned}$$

which after some algebra yields

$$\begin{aligned} q(x,y;\hat{a},\hat{b},\lambda_1,\lambda_2,\lambda_{12}) = &\lambda_1(\lambda_2+\lambda_{12})\exp\left((x-\hat{a})\lambda_1+\frac{\hat{b}^2\lambda_1^2}{2}-(\lambda_2+\lambda_{12})y\right) \\ &\cdot\left(F\left(\frac{-x+\hat{a}-\hat{b}^2\lambda_1}{\hat{b}}\right)-F\left(\frac{-x-y+\hat{a}-\hat{b}^2\lambda_1}{\hat{b}}\right)\right) \\ &+\lambda_2(\lambda_1+\lambda_{12})\exp\left((x-\hat{a})(\lambda_1+\lambda_{12})+\frac{\hat{b}^2(\lambda_1+\lambda_{12})^2}{2}-\lambda_2 y\right) \\ &\cdot F\left(\frac{-x-y+\hat{a}-\hat{b}^2(\lambda_1+\lambda_{12})}{\hat{b}}\right) \\ &\lambda_{12}\exp\left(-(\lambda_1+\lambda_2+\lambda_{12})y\right)\cdot f\left(\frac{x+y-\hat{a}}{\hat{b}}\right). \end{aligned}$$

A.3.4 Determining the Optimal Parameters for the SVCJ Approximation

We use the following three-step algorithm to solve the optimization problem (2.38) of Section 2.5.1.

Step 1: First fit the marginal distribution of the price jump, $N - E_1$. More specifically, we solve

$$\begin{aligned} \min_{\hat{a},\hat{b},\hat{c}} \quad & \sum_{x \in S_1} |p_1(x) - q_1(x; \hat{a}, \hat{b}, \hat{c})|^2 \\ \text{subject to} \quad & \hat{b}, \hat{c} \geq 0 \end{aligned} \tag{A.12}$$

where S_1 is a set of equally spaced points $x_1, \ldots, x_n$ with almost all of the probability mass falling in $[x_1, x_n]$, $p_1(\cdot)$ is the marginal density of the first component of X in (2.36) and $q_1(\cdot; \hat{a}, \hat{b}, \hat{c})$ is the marginal density of the first component of $\hat{X}$ in (2.37).

Step 2: We then fit the marginal distribution of the variance jump, E_2. We do this by solving

$$\begin{aligned} \min_{\hat{d}} \quad & \sum_{y \in S_2} |p_2(y) - q_2(y; \hat{d})|^2 \\ \text{subject to} \quad & \hat{d} \geq 0 \end{aligned} \tag{A.13}$$

where S_2 is a set of equally spaced points $0 = y_1, \ldots, y_n$ with almost all of the probability mass falling in $[y_1, y_n]$, $p_2(\cdot)$ is the marginal density of the second component of X in (2.36) and $q_2(\cdot; \hat{d})$ is the marginal density of the second component of $\hat{X}$ in (2.37).

Step 3: The last step is to fit the joint distribution but keeping the solutions to steps 1 and 2 fixed. In particular, if the optimal solutions for (A.12) and (A.13) are (a^*, b^*, c^*) and d^*, respec-

tively, then the third step solves the following optimization problem over $(\hat{a}, \hat{b}, \lambda_1, \lambda_2, \lambda_{12})$:

$$\begin{aligned} \min_{\lambda_1,\lambda_2,\lambda_{12}} \quad & \sum_{(x,y)\in S} |p(x,y) - q(x,y;a^*,b^*,\lambda_1,\lambda_2,\lambda_{12})|^2 \\ \text{subject to} \quad & \lambda_1 + \lambda_{12} = c^*, \\ & \lambda_2 + \lambda_{12} = d^*, \\ & \lambda_1, \lambda_2, \lambda_{12} \geq 0. \end{aligned} \tag{A.14}$$

where $S = \{(x_i, y_j) \,|\, i = 1, 2, \cdots, m \text{ and } j = 1, 2, \cdots, n\}$ is a set of equally spaced sample points representing almost all of the probability mass of X. If the optimal solution to (A.14) is $(\lambda_1^*, \lambda_2^*, \lambda_{12}^*)$, then $(a^*, b^*, \lambda_1^*, \lambda_2^*, \lambda_{12}^*)$ constitutes our fitted parameter set in (2.37).

A.4 Additional Numerical Results

A.4.1 Jump Approximation Parameters for the SVJ and SVCJ Models

We report here the optimized jump approximation parameters for the three parameter sets of Section 2.7.

A.4.2 Results for Parameter Set I

We report here our numerical results for the low-volatility environment of parameter set I.

Table A.1: Optimized Jump Approximation Parameters for the SVJ and SVCJ Models

Parameter Set	Leverage Ratio	SVJ model			SVCJ model				
	ϕ	a^*	b^*	c^*	a^*	b^*	λ_1^*	λ_2^*	λ_{12}^*
I	2	0.0506	0.1321	12.1236	0.101	0.1936	6.4174	24.9997	0.0003
	3	0.111	0.1801	5.8972	0.2149	0.2492	3.0934	11.111	0.0002
	-1	0.0589	0.0577	19.2356	0.1013	0.0785	10.3518	99.9389	0.0611
	-2	0.1302	0.1029	8.1678	0.2196	0.131	4.5105	24.9996	0.0004
	-3	0.2076	0.1375	4.8989	0.34	0.1655	2.8205	11.1111	0
II	2	0.3611	0.3712	1.849	0.269	0.5753	0.8521	2.6362	0.0902
	3	0.6249	0.4201	1.1857	0.601	0.6237	0.8409	1.2026	0.0323
	-1	0.2721	0.1085	3.4312	0.5015	0.071	3.5228	10.625	0
	-2	0.52	0.1542	1.8643	0.8448	0.0945	2.3258	2.647	0
	-3	0.7233	0.1802	1.4727	1.1011	0.1064	1.9714	1.1733	0
III	2	0.1157	0.3574	1.7521	0.0084	0.4603	1.0777	6.7364	0.0001
	3	0.3202	0.4239	1.0788	0.2361	0.5379	0.8916	2.9924	0
	-1	0.2968	0.0914	5.1514	0.4163	0.0775	5.0428	26.9542	0
	-2	0.5503	0.1315	2.7564	0.7243	0.1069	3.022	6.7387	0
	-3	0.7583	0.1534	2.0794	0.9625	0.1221	2.4156	2.9951	0

Table A.2: The absolute volume between the density functions of the true and approximated conditional joint jump distribution in the SVCJ model.

Parameter set	Leveraged ratio ϕ	Volume
I	2	0.0177
	3	0.0228
	-1	0.0373
	-2	0.0363
	-3	0.0364
II	2	0.0254
	3	0.0358
	-1	0.0661
	-2	0.0692
	-3	0.0674
III	2	0.0187
	3	0.0353
	-1	0.0451
	-2	0.0479
	-3	0.0499

Table A.3: Option prices on underlying ETF for parameter set I computed via Monte-Carlo and transform approaches. Approximate 95% confidence intervals are reported in brackets.

Moneyness	BS vol(%)	BS price	Option price (SV)		Option price (SVJ)		Option price (SVCJ)	
$\frac{K}{S_0}$	Σ_{BS}	C_{BS}	C_{sim}	C_{tran}	C_{sim}	C_{tran}	C_{sim}	C_{tran}
0.85	21.47	16.37	16.38	16.38	16.37	16.37	16.37	16.37
			[16.38, 16.38]	-	[16.37, 16.37]	-	[16.37, 16.37]	-
1	20.17	5.92	5.99	5.99	5.92	5.92	5.94	5.94
			[5.99, 5.99]	-	[5.92, 5.92]	-	[5.94, 5.94]	-
1.15	19.20	1.23	1.27	1.27	1.22	1.22	1.23	1.23
			[1.27, 1.27]	-	[1.22, 1.22]	-	[1.23, 1.23]	-

Table A.4: Comparison of Black-Scholes implied-volatilities: parameter set I

Leverage ratio	Moneyness		Implied Volatility (SV)			Implied Volatility (SVJ)			Implied Volatility (SVCJ)		
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_S	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$
2	0.85	0.7	21.51	43.11	2.00	21.43	43.14	2.01	21.43	43.21	2.02
	1	1	20.41	40.66	1.99	20.17	40.22	1.99	20.22	40.33	1.99
	1.15	1.3	19.44	38.82	2.00	19.18	38.26	1.99	19.24	38.41	2.00
3	0.85	0.55		64.92	3.02		65.48	3.06		65.22	3.04
	1	1		60.79	2.98		60.26	2.99		60.27	2.98
	1.15	1.45		58.17	2.99		57.30	2.99		57.48	2.99
-1	0.85	1.15		21.43	1.00		21.20	0.99		21.25	0.99
	1	1		20.49	1.00		20.21	1.00		20.28	1.00
	1.15	0.85		19.39	1.00		19.27	1.00		19.25	1.00
-2	0.85	1.3		42.91	1.99		42.38	1.98		42.51	1.98
	1	1		41.16	2.02		40.60	2.01		40.75	2.02
	1.15	0.7		38.77	1.99		38.81	2.02		38.64	2.01
-3	0.85	1.45		64.45	3.00		63.59	2.97		63.82	2.98
	1	1		62.00	3.04		61.18	3.03		61.40	3.04
	1.15	0.55		58.04	2.99		58.82	3.07		58.24	3.03

Table A.5: Comparison for the prices of options on the leveraged ETFs obtained by Monte-Carlo simulation and transform approach in parameter set I. Approximate 95% confidence intervals are reported in brackets.

Leverage ratio	Moneyness		Option price (SV)		Option price (SVJ)		Option price (SVCJ)	
ϕ	$\frac{K_S}{S_0}$	$\frac{K_L}{L_0}$	C_{sim}	C_{tran}	C_{sim}	C_{tran}	C_{sim}	C_{tran}
2	0.85	0.7	31.81	31.80	31.81	31.82	31.82	31.80
			[31.80, 31.81]	-	[31.80, 31.81]	-	[31.81, 31.82]	-
	1	1	11.65	11.65	11.53	11.56	11.56	11.55
			[11.65, 11.66]	-	[11.53, 11.54]	-	[11.56, 11.57]	-
	1.15	1.3	2.91	2.91	2.80	2.82	2.83	2.83
			[2.91, 2.91]	-	[2.79, 2.80]	-	[2.82, 2.83]	-
3	0.85	0.55	46.75	46.74	46.80	46.82	46.78	46.76
			[46.74, 46.76]	-	[46.79, 46.81]	-	[46.77, 46.78]	-
	1	1	17.23	17.21	17.08	17.13	17.08	17.08
			[17.22, 17.23]	-	[17.07, 17.09]	-	[17.08, 17.09]	-
	1.15	1.45	4.98	4.98	4.79	4.84	4.83	4.84
			[4.97, 4.98]	-	[4.78, 4.79]	-	[4.82, 4.83]	-
-1	0.85	1.15	1.66	1.66	1.61	1.63	1.62	1.63
			[1.66, 1.66]	-	[1.61, 1.61]	-	[1.62, 1.62]	-
	1	1	6.01	6.01	5.93	5.96	5.95	5.96
			[6.01, 6.02]	-	[5.93, 5.94]	-	[5.95, 5.96]	-
	1.15	0.85	16.10	16.10	16.09	16.10	16.08	16.09
			[16.10, 16.10]	-	[16.08, 16.09]	-	[16.08, 16.08]	-
-2	0.85	1.5	3.78	3.80	3.67	3.72	3.69	3.71
			[3.78, 3.78]	-	[3.66, 3.67]	-	[3.69, 3.70]	-
	1	1	11.79	11.79	11.64	11.69	11.68	11.69
			[11.79, 11.80]	-	[11.63, 11.64]	-	[11.67, 11.68]	-
	1.15	0.5	31.34	31.33	31.35	31.37	31.33	31.33
			[31.34, 31.35]	-	[31.34, 31.35]	-	[31.32, 31.34]	-
-3	0.85	1.75	6.43	6.46	6.23	6.30	6.28	6.31
			[6.43, 6.44]	-	[6.22, 6.23]	-	[6.28, 6.29]	-
	1	1	17.56	17.55	17.33	17.38	17.39	17.38
			[17.55, 17.56]	-	[17.33, 17.34]	-	[17.39, 17.40]	-
	1.15	0.25	46.22	46.20	46.27	46.28	46.23	46.21
			[46.21, 46.22]	-	[46.26, 46.28]	-	[46.22, 46.24]	-

A.5 Calibration to Market Data

The focus of this chapter has been on model-consistent pricing of ETF and LETF options. While not our main focus, it is also of interest to see how the LETF option prices generated by these (calibrated) models compare with the corresponding LETF option prices in the market-place. In this appendix we perform such a study using *Bloomberg* price data as of the market close on June 14^{th}, 2013. We emphasize that the observations we make only apply to the market data as of that date and that a more thorough[2] empirical study would be required to investigate how these models perform across time and different market regimes.

We took the SPDR S&P 500 ETF (ticker SPY) as our underlying security and considered 3-month options on the corresponding[3] double long, double short, triple long and triple short LETFs. The underlying price at the close was 163.18, the 3-month call option strikes ranged from 150 to 174 and the ATM volatility was approximately 15.2%. The 3-month risk-free rate was 0.29%. According to *Bloomberg* the dividend yield for the underlying ETF was 1.63%. Rather than using this value and the LETF dynamics in (2.4) to determine the dividend yields of the LETFs we again used the dividend yields provided by *Bloomberg*. These yields were 0.06% and .03% for SSO and UPRO, respectively. The inverse LETFs, SDS and SPXU, do not pay dividends and therefore have dividend yields of zero.

In calibrating the SV, SVJ and SVCJ models we again fixed the parameter ρ and determined the remaining parameters by minimizing the sum-of-squares between the market implied volatilities and the model's implied volatilities. Table A.6 displays the calibrated parameters for the three models

[2]We also investigated how the model performed on June 30^{th}, 2013 and obtained similar results.

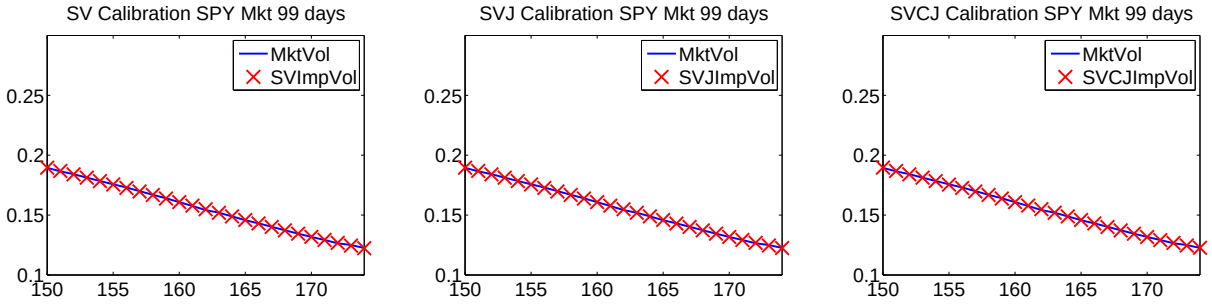
[3] With tickers SSO, SDS, UPRO and SPXU, respectively. Note also that all options had 99 days to-maturity rather than exactly 3 months.

while Figure A.1 displays the market and model implied volatility skews on the calibration date. Table A.7 also reports the calibration performance for the three models in terms of pricing error and implied volatility error. We see that the calibration was successful for each model with any pricing error well inside the bid-offer spread we would expect to see in the market.

Table A.6: Calibrated Model Parameters

Parameters	SV Model	SVJ Model	SVCJ Model
Risk free rate r	0.0029	0.0029	0.0029
Speed of mean reversion κ	4.5858	2.7509	6.3404
Volatility of variance γ	0.6284	0.5498	0.7197
Long run mean variance θ	0.0279	0.0270	0.0230
Initial variance V_0	0.0279	0.0270	0.0230
Correlation ρ	-0.7571	-0.7571	-0.82
Jump arrival rate λ	n/a	1.0926	1.7665
m	n/a	-0.0252	-0.0121
b	n/a	0.0100	0.0411
μ_v	n/a	n/a	0.0117
ρ_J	n/a	n/a	-0.6917

Figure A.1: Volatility skews for SPY options



In the case of the calibrated SVJ and SVCJ models, we first had to solve for the optimal jump distribution approximations. The optimal parameters are reported in Table A.8. We then

Table A.7: Market prices and implied volatilities for SPY options versus corresponding calibrated model prices and model implied volatilities. Root-mean-squared errors (RMSE) are reported in the final row.

Moneyness	Implied Volatility(%)				Option Price			
	Mkt ImpVol	SV ImpVol	SVJ ImpVol	SVCJ ImpVol	Mkt Price	SV Price	SVJ Price	SVCJ Price
$\frac{K}{S_0}$	Σ_{mkt}	Σ_{SV}	Σ_{SVJ}	Σ_{SVCJ}	C_{mkt}	C_{SV}	C_{SVJ}	C_{SVCJ}
0.919	18.95	18.96	18.96	18.97	21.00	20.97	20.93	21.01
0.925	18.65	18.68	18.68	18.69	20.09	20.05	20.02	20.08
0.932	18.42	18.41	18.41	18.41	19.16	19.14	19.12	19.17
0.938	18.12	18.12	18.13	18.12	18.25	18.24	18.22	18.26
0.944	17.84	17.84	17.84	17.84	17.37	17.35	17.34	17.36
0.950	17.57	17.55	17.56	17.55	16.46	16.47	16.46	16.46
0.956	17.30	17.26	17.27	17.26	15.57	15.59	15.59	15.58
0.962	17.00	16.97	16.97	16.97	14.72	14.73	14.73	14.71
0.968	16.69	16.68	16.68	16.68	13.85	13.87	13.89	13.85
0.974	16.41	16.38	16.38	16.38	13.00	13.03	13.05	13.00
0.981	16.08	16.09	16.09	16.09	12.17	12.20	12.23	12.17
0.987	15.76	15.79	15.79	15.79	11.36	11.38	11.41	11.35
0.993	15.42	15.49	15.49	15.49	10.56	10.58	10.62	10.54
0.999	15.20	15.19	15.19	15.19	9.76	9.79	9.83	9.76
1.005	14.89	14.89	14.89	14.89	3.25	3.23	3.19	3.26
1.011	14.57	14.60	14.59	14.60	2.77	2.75	2.70	2.78
1.017	14.31	14.30	14.30	14.30	2.34	2.32	2.26	2.35
1.023	14.03	14.01	14.01	14.01	1.95	1.93	1.87	1.95
1.030	13.73	13.73	13.72	13.72	1.61	1.58	1.54	1.60
1.036	13.45	13.45	13.45	13.44	1.30	1.29	1.25	1.29
1.042	13.18	13.18	13.18	13.17	1.04	1.03	1.01	1.03
1.048	12.92	12.92	12.92	12.91	0.82	0.82	0.82	0.81
1.054	12.65	12.68	12.68	12.68	0.64	0.64	0.66	0.63
1.060	12.49	12.45	12.46	12.46	0.49	0.50	0.53	0.49
1.066	12.25	12.24	12.25	12.27	0.38	0.39	0.43	0.39
RMSE		**0.02**	**0.02**	**0.03**		**0.01**	**0.01**	**0.01**

Table A.8: Optimized jump approximation parameters in the SVJ and SVCJ models

Ticker	Leverage Ratio	SVJ model			SVCJ model				
	ϕ	a^*	b^*	c^*	a^*	b^*	λ_1^*	λ_2^*	λ_{12}^*
SSO	2	-0.0501	0.0205	597.1613	0.01	0.0773	17.0741	12.3376	9.1189
SDS	-2	0.0559	0.0175	145.5139	0.0729	0.0666	19.0482	21.4565	0
UPRO	3	-0.0768	0.0316	524.4489	0.0301	0.1105	10.1514	6.485	3.0512
SPXU	-3	0.0836	0.0252	89.927	0.1159	0.0935	11.3475	9.5362	0

computed the calibrated model prices and implied volatilities of European call options on the LETFs and compared them with their corresponding market prices and implied volatilities. These prices are reported in Tables A.9 to A.12 for SSO (double long), SDS (double short), UPRO (triple long) and SPXU (triple short), respectively, and for different values of money-ness.

We note that for each model-LETF combination, the root-mean-squared pricing error is relatively small and typically within the (non-reported) bid-ask price ranges that we see in the market. On the date in question at least, we can conclude that all three calibrated models are capable of reproducing market prices of LETF options. We also note, however, that the current market regime could be characterized as a low-volatility regime (compare Figure A.1 with the low-volatility graphs in Figure 2.2) and so it is not surprising given the results in Table A.5 that the LETF option prices broadly agree across all three models. In contrast, the results in Tables 2.4 and 2.5 suggest that we would expect considerably less agreement among the three models in a high volatility regime.

We might also conclude that in the current low-volatility environment, market participants could use any of the three models to hedge an ETF-LETF option portfolio by trading in the underlying ETF only rather than in the leveraged ETF. This might be preferable as the underlying ETF

typically has a lower bid-offer spread than the corresponding LETFs. This of course also allows the possibility of combining an ETF options trading book with LETF options trading books and risk-managing them together rather than separately. That said, we do acknowledge that a more rigorous study would be required before we would recommend risk-managing such portfolios using these models. Moreover, in a high volatility environment we expect that only one of our three models might be suitable for hedging.

More specific observations can also be made. We note that the pricing error is generally smaller for the negative LETFs than for the positive LETFs on the date in question. The root-mean-squared pricing error for the negative LETFs is less than 7 cents in all three models whereas the corresponding number for the positive LETFs is 18 cents. We also observe that the implied volatility ratios (the ratios of implied volatilities of the LETF to the corresponding implied volatilities of the underlying ETF) are generally close to the leverage multiple. In our particular data-set we see that the ratios for the negative LETFs are almost always higher than the ratios for the positive leverage ETFs. Whether or not these are persistent features would require a more thorough time-series study which would also include a broader cross-section of ETFs.

Table A.9: SSO (Double Long): Market Prices and Implied Volatilities Versus Calibrated Model Prices and Implied Volatilities.

Leverage ratio	Moneyness	Implied Volatility (%)								Option Price			
		Mkt ImpVol		SV Imp Vol		SVJ ImpVol		SVCJ ImpVol		Mkt Price	SV Price	SVJ Price	SVCJ Price
ϕ	$\frac{K_L}{L_0}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	C_{mkt}	C_{SV}	C_{SVJ}	C_{SVCJ}
2	0.850	35.67	1.88	37.67	1.99	37.68	1.99	37.48	1.98	13.30	13.51	13.51	13.49
	0.862	34.95	1.87	37.06	1.98	37.07	1.98	36.90	1.97	12.45	12.68	12.68	12.66
	0.875	34.37	1.87	36.44	1.98	36.46	1.98	36.31	1.97	11.63	11.87	11.87	11.86
	0.888	34.23	1.92	35.82	2.01	35.84	2.01	35.72	2.00	10.88	11.08	11.08	11.07
	0.900	33.65	1.92	35.21	2.01	35.22	2.01	35.14	2.00	10.10	10.30	10.30	10.29
	0.913	33.15	1.92	34.59	2.00	34.60	2.00	34.55	2.00	9.35	9.55	9.55	9.54
	0.926	32.49	1.91	33.97	2.00	33.98	2.00	33.97	2.00	8.60	8.81	8.81	8.81
	0.938	32.16	1.93	33.34	2.00	33.36	2.00	33.39	2.00	7.92	8.10	8.10	8.10
	0.951	31.71	1.93	32.72	2.00	32.74	2.00	32.81	2.00	7.25	7.40	7.41	7.42
	0.964	31.09	1.93	32.10	2.00	32.11	2.00	32.24	2.00	6.58	6.74	6.74	6.76
	0.976	30.54	1.94	31.48	1.99	31.49	1.99	31.67	2.01	5.95	6.10	6.10	6.13
	0.989	30.01	1.95	30.86	1.99	30.87	1.99	31.10	2.01	5.35	5.49	5.49	5.53
	1.002	29.47	1.94	30.25	1.99	30.25	1.99	30.54	2.01	4.78	4.91	4.91	4.95
	1.014	28.98	1.95	29.64	1.99	29.64	1.99	29.98	2.01	4.25	4.36	4.36	4.41
	1.027	28.35	1.95	29.04	1.99	29.04	1.99	29.43	2.02	3.73	3.84	3.84	3.91
	1.040	27.75	1.94	28.44	1.99	28.44	1.99	28.90	2.02	3.25	3.36	3.36	3.43
	1.052	27.26	1.94	27.86	1.99	27.86	1.99	28.38	2.03	2.82	2.91	2.91	3.00
	1.065	26.73	1.95	27.28	1.99	27.29	1.99	27.88	2.03	2.42	2.50	2.51	2.60
	1.078	26.23	1.95	26.73	1.99	26.73	1.99	27.40	2.04	2.06	2.13	2.13	2.23
	1.091	25.77	1.96	26.20	1.99	26.20	1.99	26.94	2.05	1.74	1.80	1.80	1.90
	1.103	25.21	1.95	25.69	1.99	25.69	1.99	26.52	2.05	1.44	1.50	1.50	1.61
	1.116	24.77	1.96	25.20	1.99	25.21	1.99	26.14	2.06	1.19	1.24	1.24	1.36
	1.129	24.23	1.94	24.75	1.99	24.77	1.99	25.79	2.07	0.96	1.02	1.02	1.14
	1.141	23.88	1.95	24.33	1.99	24.35	1.99	25.49	2.08	0.78	0.83	0.83	0.95
RMSE				**1.14**		**1.15**		**1.36**			**0.15**	**0.15**	**0.18**

Table A.10: SDS (Double Short): Market Prices and Implied Volatilities Versus Calibrated Model Prices and Implied Volatilities.

Leverage ratio	Moneyness	Implied Volatility (%)								Option Price			
		Mkt ImpVol		SV Imp Vol		SVJ ImpVol		SVCJ ImpVol		Mkt Price	SV Price	SVJ Price	SVCJ Price
ϕ	$\frac{K_L}{L_0}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	C_{mkt}	C_{SV}	C_{SVJ}	C_{SVCJ}
-2	0.874	25.14	2.01	25.41	2.04	25.41	2.04	26.45	2.12	5.45	5.46	5.46	5.51
	0.899	26.95	2.09	26.52	2.05	26.49	2.05	27.24	2.11	4.75	4.72	4.72	4.77
	0.924	28.02	2.08	27.75	2.06	27.70	2.06	28.22	2.10	4.10	4.08	4.08	4.11
	0.949	29.58	2.11	29.02	2.07	28.97	2.07	29.30	2.09	3.57	3.53	3.52	3.55
	0.974	30.84	2.12	30.30	2.08	30.25	2.07	30.43	2.08	3.10	3.06	3.05	3.07
	0.999	32.33	2.13	31.55	2.08	31.50	2.07	31.55	2.08	2.72	2.66	2.65	2.66
	1.024	33.55	2.13	32.76	2.07	32.71	2.07	32.65	2.07	2.38	2.31	2.31	2.31
	1.049	34.61	2.11	33.92	2.07	33.87	2.07	33.72	2.06	2.08	2.02	2.02	2.01
	1.074	35.84	2.11	35.02	2.06	34.98	2.06	34.75	2.05	1.84	1.77	1.77	1.75
	1.099	36.97	2.10	36.07	2.06	36.03	2.05	35.75	2.04	1.63	1.56	1.56	1.54
	1.124	37.93	2.09	37.08	2.05	37.03	2.04	36.70	2.03	1.44	1.38	1.37	1.35
	1.149	38.92	2.09	38.03	2.04	37.99	2.03	37.62	2.01	1.28	1.22	1.21	1.19
RMSE				**0.69**		**0.73**		**0.92**			**0.05**	**0.06**	**0.07**

Table A.11: UPRO (Triple Long): Market Prices and Implied Volatilities Versus Calibrated Model Prices and Implied Volatilities.

Leverage ratio	Moneyness	Implied Volatility (%)								Option Price			
		Mkt ImpVol		SV Imp Vol		SVJ ImpVol		SVCJ ImpVol		Mkt Price	SV Price	SVJ Price	SVCJ Price
ϕ	$\frac{K_L}{L_0}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	C_{mkt}	C_{SV}	C_{SVJ}	C_{SVCJ}
3	0.772	57.67	3.04	56.56	2.98	56.59	2.98	56.49	2.98	16.80	16.71	16.71	16.70
	0.779	55.60	2.98	56.16	3.01	56.19	3.01	56.10	3.00	16.25	16.30	16.30	16.29
	0.787	54.76	2.94	55.76	2.98	55.79	2.99	55.72	2.98	15.80	15.88	15.89	15.88
	0.795	56.20	3.05	55.36	3.01	55.40	3.01	55.34	3.01	15.55	15.48	15.48	15.47
	0.802	53.02	2.88	54.96	2.99	55.00	2.99	54.96	2.99	14.90	15.07	15.08	15.07
	0.810	57.03	3.10	54.56	2.96	54.61	2.97	54.58	2.97	14.90	14.67	14.67	14.67
	0.818	56.03	3.09	54.17	2.99	54.21	2.99	54.20	2.99	14.45	14.27	14.28	14.28
	0.825	54.50	3.01	53.77	2.97	53.82	2.97	53.83	2.97	13.95	13.88	13.88	13.88
	0.833	53.99	3.03	53.38	2.99	53.42	2.99	53.45	3.00	13.55	13.49	13.49	13.50
	0.841	51.96	2.91	52.99	2.97	53.03	2.97	53.08	2.98	13.00	13.10	13.11	13.11
	0.848	51.90	2.91	52.59	2.95	52.64	2.95	52.71	2.95	12.65	12.72	12.73	12.73
	0.856	50.84	2.89	52.20	2.97	52.25	2.98	52.34	2.98	12.20	12.34	12.35	12.36
	0.864	51.16	2.91	51.81	2.95	51.85	2.95	51.97	2.96	11.90	11.97	11.97	11.99
	0.879	49.82	2.88	51.03	2.96	51.07	2.96	51.24	2.97	11.10	11.24	11.24	11.26
	0.917	48.85	2.93	49.10	2.94	49.14	2.95	49.42	2.96	9.45	9.48	9.48	9.52
	0.955	46.80	2.91	47.19	2.93	47.23	2.94	47.65	2.96	7.80	7.85	7.86	7.91
	0.993	44.10	2.86	45.32	2.93	45.35	2.93	45.92	2.96	6.20	6.37	6.37	6.45
	1.032	42.51	2.86	43.50	2.92	43.52	2.92	44.25	2.97	4.90	5.03	5.04	5.14
	1.070	41.20	2.88	41.75	2.92	41.76	2.92	42.67	2.98	3.80	3.87	3.87	4.00
	1.108	40.18	2.99	40.09	2.98	40.10	2.98	41.19	3.07	2.90	2.89	2.89	3.03
	1.146	37.99	2.94	38.54	2.98	38.56	2.98	39.87	3.09	2.02	2.08	2.09	2.24
	1.184	37.12	2.97	37.15	2.98	37.18	2.98	38.75	3.11	1.45	1.45	1.46	1.62
	1.223	35.92	2.93	35.94	2.94	35.99	2.94	37.87	3.09	0.98	0.98	0.99	1.15
RMSE				**1.06**		**1.07**		**1.39**			**0.11**	**0.11**	**0.15**

Table A.12: SPXU (Triple Short): Market Prices and Implied Volatilities Versus Calibrated Model Prices and Implied Volatilities.

Leverage ratio	Moneyness	Implied Volatility (%)								Option Price			
		Mkt ImpVol		SV Imp Vol		SVJ ImpVol		SVCJ ImpVol		Mkt Price	SV Price	SVJ Price	SVCJ Price
ϕ	$\frac{K_L}{L_0}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	Σ_L	$\frac{\Sigma_L}{\Sigma_S}$	C_{mkt}	C_{SV}	C_{SVJ}	C_{SVCJ}
-3	0.797	35.40	2.89	37.75	3.09	37.79	3.08	39.64	3.23	5.05	5.10	5.10	5.14
	0.839	38.44	2.97	39.62	3.07	39.58	3.06	40.82	3.16	4.30	4.34	4.34	4.37
	0.881	42.16	3.14	41.76	3.10	41.68	3.10	42.47	3.16	3.70	3.68	3.68	3.71
	0.923	44.21	3.15	43.98	3.14	43.89	3.13	44.33	3.17	3.15	3.14	3.14	3.16
	0.965	46.40	3.18	46.18	3.16	46.08	3.16	46.26	3.17	2.70	2.69	2.68	2.69
	1.007	48.95	3.17	48.29	3.12	48.19	3.11	48.16	3.11	2.35	2.32	2.31	2.31
	1.049	50.52	3.14	50.28	3.13	50.19	3.12	49.98	3.11	2.02	2.01	2.00	1.99
	1.091	52.16	3.13	52.15	3.13	52.07	3.12	51.72	3.10	1.75	1.75	1.75	1.73
	1.133	54.29	3.09	53.91	3.07	53.83	3.07	53.38	3.04	1.55	1.53	1.53	1.51
	1.175	55.59	3.07	55.56	3.07	55.47	3.06	54.94	3.03	1.35	1.35	1.34	1.32
	1.217	57.29	3.07	57.10	3.06	57.01	3.05	56.43	3.02	1.20	1.19	1.19	1.16
RMSE				**0.85**		**0.88**		**1.57**			**0.02**	**0.03**	**0.04**

Appendix B

Appendix for Chapter 3

B.1 Outline Proof of the Unique Optimality of V^* in $(\mathcal{P}_2)$

Applying Itô's lemma to $J(W_t, Z_t, t) = u(W_t)V(Z_t, t)$ for an arbitrary twice-differentiable V satisfying the constraints of $(\mathcal{P}_2)$, we have

$$\begin{aligned} J(W_T, Z_T, T) &= J(W_t, Z_t, t) + \int_t^T u(W_u) H_{\theta^*}\left[V(Z_u, u)\right] du \\ &\quad + \int_t^T \left(J_w(W_u, Z_u, u) W_u \theta_u^{*\top} \Sigma_{P_u} + J_z^\top(W_u, Z_u, u)\Sigma_{Z_u}\right) dB_u, \end{aligned} \tag{B.1-1}$$

if W_t is the wealth process associated with the optimal policy, θ_t^*. Since $H_{\theta^*}V \geq 0$ and $V(Z_T, T) \leq 1$, (B.1-1) implies

$$u(W_T) \leq J(W_t, Z_t, t) + \int_t^T \left(J_w(W_u, Z_u, u) W_u \theta_u^{*\top} \Sigma_{P_u} + J_z^\top(W_u, Z_u, u)\Sigma_{Z_u}\right) dB_u. \tag{B.1-2}$$

By the martingale property of the stochastic integral, the second term on the right-hand side of (B.1-2) has mean zero. Taking expectations conditional on $\mathcal{F}_t$ on both sides of (B.1-2), we obtain

$$E_{t,\theta^*}[u(W_T)] \leq J(W_t, Z_t, t).$$

Since θ^* is the optimal policy, we have

$$J^*(W_t, Z_t, t) = E_{t,\theta^*}[u(W_T)] \leq J(W_t, Z_t, t),$$

or equivalently since $u(W_T) < 0$ (since $\gamma > 1$)

$$V^*(Z_t, t) \geq V(Z_t, t).$$

Noting that $V^*(z, t)$ is also feasible for $(\mathcal{P}_2)$, the result follows.

B.2 The Myopic Trading Strategy

The myopic strategy assumes the instantaneous moments of asset returns are fixed at their current values for the remainder of the investment horizon. It therefore ignores the hedging component of the optimal trading strategy. At each time t, the myopic strategy, θ_t^m, is obtained as the solution to the following quadratic optimization problem:

$$\theta_t^m = \underset{\theta \in \mathbf{K}}{\operatorname{argmin}} \quad \frac{1}{2}\theta^\top \left[\gamma \Sigma_P \Sigma_P^\top\right] \theta - \lambda^\top \theta. \tag{B.2-3}$$

where $\lambda = \mu_{P_t} - r_t \cdot \mathbf{1}$. The optimization problem in (B.2-3) is the HJB equation that the myopic investor formulates at time t if she observes the instantaneous moments of asset returns, μ_{P_t} and Σ_{P_t}, at that time and then assumes these moments are fixed thereafter. Because we do not have a closed-form expression for the terminal wealth resulting from the myopic strategy, we estimate its expected utility by simulating the stochastic differential equations for Z_t, P_t and W_t, and solving (B.2-3) at each point on each simulated path.

The myopic strategy is popular[1] in the financial literature for several reasons. For example, it is expected that the the myopic policy will be very close to optimal when the hedging component

[1] See, for example, Kroner and Sultan (1993), Lioui and Poncet (2000), Brooks, Henry and Persand (2002) and Basak and Chabakauri (2008)

of the optimal trading strategy is not significant. The myopic strategy can also be used to estimate the magnitude of this hedging component. Moreover it's clear from (B.2-3) that solving for the myopic policy is very straightforward.

B.3 Review of Duality Theory and Construction of Upper Bounds

Here we briefly review the duality approach of HKW [22] for analyzing the quality of a suboptimal strategy. This is done by using the suboptimal strategy to construct a lower and upper bound on the true value function. If the difference between the two bounds is large, i.e. the duality gap is wide, then it suggests that the suboptimal policy is not close to the optimal solution. If the duality gap is narrow, then (i) we know that the suboptimal strategy is close to optimal and (ii) we know approximately the optimal value function. In this chapter we will use the myopic policy and the policy driven by LP approach to construct upper bounds on the optimal dynamic trading strategy and compare performances of these two trading strategies.

Starting with the portfolio optimization problem of Section 3.2, we can define a fictitious problem $(\mathcal{P}^{(\nu)})$, based on a different financial market and without the portfolio constraints. First we define the *support function* of the trading constraint $\mathbf{K}$, $\delta(\cdot) : \mathbb{R}^N \to \mathbb{R} \cup \infty$, by setting

$$\delta(\nu) = \sup_{x \in \mathbf{K}} (-\nu^\top x). \tag{B.3-4}$$

The effective domain of the support function is given by

$$\widetilde{\mathbf{K}} = \{\nu \in \mathbf{K} : \ \delta(\nu) < \infty\}. \tag{B.3-5}$$

Because the constraint set $\mathbf{K}$ is convex and contains zero, the support function is continuous and bounded from below on its effective domain $\widetilde{\mathbf{K}}$. We then define the set $\mathbf{D}$ of $\mathcal{F}_t$-adapted $\mathbb{R}^N$ valued

processes to be

$$\mathbf{D} = \left\{ \nu_t, 0 \le t \le T : \ \nu_t \in \widetilde{\mathbf{K}}, \ \mathrm{E}_0\left[\int_0^T \delta(\nu_t)\, dt \right] + \mathrm{E}_0\left[\int_0^T \|\nu_t\|^2\, dt \right] < \infty \right\}. \tag{B.3-6}$$

For each process ν in $\mathbf{D}$, we define a fictitious market $M^{(\nu)}$, in which the N stocks and the risk-free bond are traded. The diffusion matrix of stock returns in $M^{(\nu)}$ is the same as in the original market. However, the risk-free rate and the vector of expected stock returns are different. In particular, the risk-free rate process and the market price of risk in the fictitious market are defined respectively by

$$r_t^{(\nu)} = r(Z_t) + \delta(\nu_t) \tag{B.3-7a}$$

$$\eta_t^{(\nu)} = \eta_t + \Sigma_P^{-1} \nu_t \tag{B.3-7b}$$

where $\delta(\nu)$ is the support function defined in (B.3-4). We assume that $\eta_t^{(\nu)}$ is square-integrable. Following Cox and Huang [11], the state-price density process $\pi_t^{(\nu)}$ in the fictitious market is given by

$$\pi_t^{(\nu)} = \exp\left(- \int_0^t r_s^{(\nu)}\, ds - \frac{1}{2} \int_0^t \eta_s^{(\nu)\top} \eta_s^{(\nu)}\, ds - \int_0^t \eta_s^{(\nu)\top}\, dB_s \right) \tag{B.3-8}$$

and the vector of expected returns is given by

$$\mu_{Pt}^{(\nu)} = r_t^{(\nu)} + \Sigma_P\, \eta_t^{(\nu)}.$$

The dynamic portfolio choice problem in the fictitious market without position constraints can be equivalently formulated in a static form[2]:

$$V^{(\nu)} \ \equiv \ \sup_{\{W_T\}} \mathrm{E}_0\left[u(W_T)\right] \quad \text{subject to} \quad \mathrm{E}_0\left[\pi_T^{(\nu)} W_T\right] \le W_0. \tag{$\mathcal{P}^{(\nu)}$}$$

[2] See Cox and Huang [11], or Section 3 of Karatzas and Shreve [26].

Due to its static nature, the problem $(\mathcal{P}^{(\nu)})$ is easy to solve. Since we assume that the utility function is of the CRRA type with relative risk aversion γ so that $u(W) = W^{1-\gamma}/(1-\gamma)$, the corresponding value function in the fictitious market is given explicitly by

$$V_0^{(\nu)} = \frac{W_0^{1-\gamma}}{1-\gamma} \; \mathrm{E}_0\left[\pi_T^{(\nu)\,\frac{\gamma-1}{\gamma}}\right]^{\gamma}. \tag{B.3-9}$$

It is easy to see that for any admissible choice of $\nu \in \mathbf{D}$, the value function in (B.3-9) gives an upper bound for the optimal value function of the original problem. In the fictitious market, the wealth dynamics of the portfolio are given by

$$dW_t^{(\nu)} = W_t^{(\nu)}\left[\left(r_t^{(\nu)} + \theta_t^{\top}\Sigma_P\eta_t^{(\nu)}\right)dt + \theta_t^{\top}\Sigma_P\, dB_t\right]$$

so that

$$\frac{dW_t^{(\nu)}}{W_t^{(\nu)}} - \frac{dW_t}{W_t} = \left[\left(r_t^{(\nu)} - r_t\right) + \theta_t^{\top}\Sigma_P\left(\eta_t^{(\nu)} - \eta_t\right)\right]dt = \left(\delta(\nu_t) + \theta_t^{\top}\nu_t\right)dt.$$

The last expression is non-negative according to (B.3-4) since $\theta_t \in \mathbf{K}$. Therefore, $W_t^{(\nu)} \geq W_t$ $\forall\ t \in [0,T]$ and so

$$V_0^{(\nu)} \geq V_0. \tag{B.3-10}$$

Under fairly general assumptions, it can be shown that there exists a process, ν^*, such that (B.3-10) holds with equality. While one can pick any fictitious market from the admissible set $\mathbf{D}$ to compute an upper bound, HKW showed how a given suboptimal strategy, $\widetilde{\theta}_t$, may be used to select a particular $\widehat{\nu}_t \in \mathbf{D}$. If the suboptimal strategy is in fact optimal, then the lower bound associated with the suboptimal strategy will equal the associated upper bound, thereby demonstrating its optimality.

Given an approximation to the optimal portfolio policy $\widetilde{\theta}_t$, one can compute the corresponding approximation to the value function, $\widetilde{J}(w,z,t)$, defined as the conditional expectation of the utility

of terminal wealth, under the portfolio policy $\widetilde{\theta}_t$. Assuming that the approximate value function $\widetilde{J}$ is sufficiently smooth, we can define[3] $\widetilde{\eta}_t$ as

$$\widetilde{\eta}_t := -W_t \left(\frac{\partial_{ww} \widetilde{J}}{\partial_w \widetilde{J}} \right) \Sigma_P^\top \widetilde{\theta}_t - \left(\partial_w \widetilde{J} \right)^{-1} \Sigma_Z^\top \left(\partial_{wz} \widetilde{J} \right)$$

where ∂_w denotes the partial derivatives with respect to W, and ∂_{wz} and ∂_{ww} are corresponding second partial derivatives. We then define $\widetilde{\nu}_t$ as a solution to (B.3-7b) where $\eta_t^{(\nu)}$ replaced by $\widetilde{\eta}_t$.

Since we consider a CRRA utility function, the expression for $\widetilde{\eta}_t$ simplifies. In the case of a CRRA utility function, for a given trading strategy, $\widetilde{\theta}_t$, the corresponding value function is of the following form

$$\widetilde{J}(W_t, Z_t, t) = u(W_t) \widetilde{V}(Z_t, t).$$

Hence, the market price of risk in the dual problem simplifies to

$$\widetilde{\eta}_t = \gamma \Sigma_P^\top \widetilde{\theta}_t - \frac{\Sigma_Z^\top}{\widetilde{J}} \left(\frac{\partial \widetilde{J}}{\partial Z_t} \right) = \gamma \Sigma_P^\top \widetilde{\theta}_t - \frac{\Sigma_Z^\top}{\widetilde{V}} \left(\frac{\partial \widetilde{V}}{\partial Z_t} \right) \tag{B.3-11}$$

and one only needs to compute the first derivative of the value function with respect to the state variables, Z_t, to evaluate the second term in (B.3-11). This simplifies numerical implementation, since it is easier to estimate first-order than second-order partial derivatives of the value function. In the case of LP approach, we can approximate the value function through the adaptive constraint selection algorithm. But for more general strategies such as the myopic strategy, we do not have an approximation of an analytical solution for the value function and its derivatives.

Obviously, $\widetilde{\eta}_t$ is a candidate for the market price of risk in the fictitious market. However, there is no guarantee that $\widetilde{\eta}_t$ and the corresponding process $\widetilde{\nu}_t$ belong to the feasible set $\mathbf{D}$ defined by (B.3-6). In fact, for many important classes of problems the support function $\delta(\nu_t)$ may be infinite

[3] See HKW [22] who motivate this definition of $\widetilde{\eta}_t$.

for some values of its argument. We therefore look for a price-of-risk process $\widehat{\eta}_t \in \mathbf{D}$ that is "close" to $\widetilde{\eta}_t$ by formulating a simple quadratic optimization problem. Depending on the portfolio constraints, this problem may be solved analytically. Otherwise, we solve it numerically at each discretization point on each simulated path of the underlying SDE's. The lower bound is then computed by simulating the given portfolio strategy. The same simulated paths of the SDE's are then used to estimate the upper bound given by (B.3-9). At each discretization point on each simulated path we solve a quadratic optimization problem to find the appropriate $\widehat{\eta}_t \in \mathbf{D}$. See HKW for further details.

It is also worth noting here that the definition of $\widetilde{\eta}_t$ in (B.3-11) will not depend on any artificial assets that we might add to the market in order to make[4] it complete. To see this, suppose the last n rows of Σ_P correspond to some artificial assets that were introduced to complete the market. Since these artificial assets cannot be traded the corresponding components of $\widetilde{\theta}_t$ will be zero and so the final n rows of Σ_P will not contribute on the right-hand-side of (B.3-11).

B.3.1 Trading Constraints

We consider three different sets of trading constraints, $\mathbf{K}$, in the numerical experiments of Section 3.5. For each $\mathbf{K}$, we describe here how to compute a price-of-risk process, $\hat{\eta}_t$, associated with $\hat{\nu}_t \in \mathbf{D}$ that is close to $\widetilde{\eta}_t$. The objective is to minimize the Euclidean distance between the process $\hat{\eta}$, used to compute the upper bound, and the candidate process for the market price of risk, $\widetilde{\eta}$. To find $\hat{\eta}$ and $\hat{\nu}$, we must solve:

[4] In numerical examples with incomplete markets it would be necessary to add artificial assets (that cannot be traded) in order to be consistent with the assumption of complete markets made in Section 3.2 and also assumed here. This assumption of complete markets is without loss of generality since incompleteness can then be modeled via our choice of $\mathbf{K}$.

$$\min_{\hat{\eta}} \quad \frac{1}{2}(\hat{\eta}-\widetilde{\eta})^{\top}(\hat{\eta}-\widetilde{\eta}) \tag{B.3-12}$$

subject to

$$\hat{\eta} = \eta + \Sigma_P^{-1}\hat{\nu}$$

$$\hat{\nu} \in \widetilde{\mathbf{K}}$$

where $\widetilde{\mathbf{K}}$ is the effective domain defined in (B.3-5).

Incomplete Markets

Assume that only the first L stocks are traded. Then the set of feasible portfolio policies is given by

$$\mathbf{K} = \{\theta \mid \theta_i = 0 \text{ for } L < i \leq N\}$$

and hence the support function $\delta(\nu)$ in (B.3-4) is equal to zero if $\nu_i = 0,\ 1 \leq i \leq L$ and is infinite otherwise. So the effective domain is:

$$\widetilde{\mathbf{K}} = \{\nu \mid \nu_1 = \nu_2 = \cdots = \nu_L = 0\}.$$

In this case, the constraints $\hat{\eta} = \eta + \Sigma_P^{-1}\hat{\nu}$ in the optimization problem (B.3-12) can be simplified as $A\hat{\eta} = A\eta$ where A is an $L \times N$ matrix consisting of the first L rows of Σ_P. Relaxing this constraint with a Lagrange multiplier λ it's easy to see that the optimal $\hat{\eta}$ and λ must satisfy the following system of linear equations:

$$\begin{aligned} \hat{\eta} - \widetilde{\eta} &= A^{\top}\lambda \\ A\hat{\eta} &= A\eta. \end{aligned}$$

Therefore the optimal solution to (B.3-12) is $\hat{\eta} = \widetilde{\eta} + A^{\top}\left(AA^{\top}\right)^{-1}A\left(\eta - \widetilde{\eta}\right)$.

Incomplete Markets and No Borrowing Constraints

Consider the market in which only the first L stocks can be traded and borrowing is not allowed. The set of admissible portfolios is given by:

$$\mathbf{K} = \{\theta \mid \theta_1 + \theta_2 + \cdots + \theta_L \leq 1,\ \theta_i = 0 \text{ for } L < i \leq N\}.$$

The support function is equal to $-\nu_1$ when $\nu_1 = \nu_2 = \cdots = \nu_L \leq 0$, and is infinite otherwise. So the effective domain is:

$$\widetilde{\mathbf{K}} = \{\nu \mid \nu_1 = \nu_2 = \cdots = \nu_L \leq 0\}.$$

In this case, the constraints of (B.3-12) can be simplified as:

$$A\hat{\eta} = A\eta + c \cdot \mathbf{1}, \quad c \leq 0$$

where $\mathbf{1} = (1, \cdots, 1)^\top$. Then using the Lagrangian relaxation of (B.3-12) with Lagrangian multiplier λ and its first-order conditions, one can show that, in the optimal case, the following system of linear equations:

$$\hat{\eta} - \widetilde{\eta} - A^\top \lambda = 0,$$

$$A(\hat{\eta} - \eta) - c \cdot \mathbf{1} = 0,$$

$$\mathbf{1}^\top \lambda = 0$$

should be satisfied with $c \leq 0$. Otherwise we must set $c = 0$. After some calculation, we obtain the following optimal solution:

$$\hat{\eta} = \widetilde{\eta} + A^\top \left(AA^\top\right)^{-1} A(\eta - \widetilde{\eta}) + c \cdot A^\top \left(AA^\top\right)^{-1} \mathbf{1},$$

where

$$c := \min\left(-\frac{\mathbf{1}^\top (AA^\top)^{-1} A(\eta - \widetilde{\eta})}{\mathbf{1}^\top (AA^\top)^{-1} \mathbf{1}},\ 0\right).$$

Incomplete Markets with No Short-Sales and No Borrowing Constraints

Consider the same market as in the previous case, but in addition neither short-sales nor borrowing are now allowed. The set of admissible portfolios is then given by:

$$\mathbf{K} = \{\theta \mid \theta \geq 0,\ \theta_1 + \theta_2 + \cdots + \theta_L \leq 1,\ \theta_i = 0 \text{ for } L < i \leq N\}.$$

The support function is given by $\epsilon(\nu) = \max(0, -\nu_1, \cdots, -\nu_L)$, which is finite for any vector ν. So the effective domain is $\widetilde{\mathbf{K}} = \mathbb{R}^N$. In this case $\hat{\eta} = \widetilde{\eta}$ is feasible for (B.3-12) and therefore solves that problem.

Appendix C

Appendix for Chapter 4

C.1 Smoothing Volatility Surfaces

We smoothed the implied volatilities in a non-parametric way using *Nadaraya-Watson estimator* based on an independent bivariate Gaussian kernel, i.e.

$$\hat{I}_t(m,\tau) = \frac{\sum_{i=1}^{N} I_t(m_i,\tau_i) g(m-m_i,\tau-\tau_i)}{\sum_{i=1}^{N} g(m-m_i,\tau-\tau_i)}$$

where $g(x,y) = \frac{1}{2\pi}\exp\left(-\frac{x^2}{2h_1}-\frac{y^2}{2h_2}\right)$ is the Gaussian kernel and $\{(m_i,\tau_i)\,;\ i=1,\cdots,N\}$ is the set of moneyness, time-to-maturity pairs of options available at time t. The important parameters are the bandwidth parameters h_1, h_2, which determine the degree of smoothing. For the optimal choice of these parameters we may refer to the standard literature mentioned in Cont and Da Fonseca [10]. However, Da Fonseca and Gottschalk [18] found it more convenient to set the values directly. Especially they took $h_1 = 0.006$ and $h_2 = 0.14$ for the S&P 500.

C.2 Proof of Proposition 1

Proof. We first prove (4.19). The characteristic function for Z_f is

$$\begin{aligned}
\phi_{Z_f}(\mathbf{t}) &= \mathrm{E}\left[\exp\left(i\mathbf{t}\left(m(W)+\sqrt{W}\mathbf{C}X\right)\right) \;\middle|\; \mathbf{A}Z = f\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\exp\left(i\mathbf{t}\left(m(W)+\sqrt{W}\mathbf{C}X\right)\right) \;\middle|\; W, \mathbf{A}Z = f\right] \;\middle|\; \mathbf{A}Z = f\right] \\
&= \mathrm{E}\left[\mathrm{E}\left[\exp\left(i\mathbf{t}\left(m(W)+\sqrt{W}\mathbf{C}X\right)\right) \;\middle|\; W, \mathbf{A}\sqrt{W}\mathbf{C}X = f_m\right] \;\middle|\; \mathbf{A}Z = f\right] \qquad \text{(C.2-1)}
\end{aligned}$$

Let

$$Y \coloneqq \mathbf{C}X \mid W, \mathbf{A}\mathbf{C}X = \frac{f_m}{\sqrt{W}}.$$

where $f_m \coloneqq f - \mathbf{A}m(W)$. Since Y is a multivariate normal given another multivariate normal, it is not difficult to show that

$$Y \sim N\left(\mu_{f,W}, \widetilde{\Sigma}\right), \qquad \text{(C.2-2)}$$

where $\mu_{f,W}$ and $\widetilde{\Sigma}$ are defined in (4.20).

Therefore, continuing from (C.2-1), we have

$$\begin{aligned}
\phi_{Z_f}(\mathbf{t}) &= \mathrm{E}\left[\mathrm{E}\left[\exp\left(i\mathbf{t}^\top\left(m(W)+\sqrt{W}\mathbf{C}X\right)\right) \;\middle|\; W, \mathbf{A}\sqrt{W}\mathbf{C}X = f_m\right] \;\middle|\; \mathbf{A}Z = f\right] \\
&= \mathrm{E}\left[e^{i\mathbf{t}^\top m(W)}\mathrm{E}\left[\exp\left(i\mathbf{t}^\top\sqrt{W}Y\right) \;\middle|\; W, \mathbf{A}\sqrt{W}\mathbf{C}X = f_m\right] \;\middle|\; \mathbf{A}Z = f\right] \\
&= \mathrm{E}\left[\exp\left(i\mathbf{t}^\top\left(m(W)+\sqrt{W}\mu_{f,W}\right) - \frac{1}{2}W\mathbf{t}^\top\widetilde{\Sigma}\mathbf{t}\right) \;\middle|\; \mathbf{A}Z = f\right] \qquad \text{(C.2-3)}
\end{aligned}$$

It remains to prove (4.21). Note that

$$
\begin{aligned}
& \mathbb{P}\left(W \in w+dw \;\middle|\; \mathbf{AC}\sqrt{W}X + \mathbf{A}m(W) = f\right) \\
= \quad & \frac{\mathbb{P}\left(W \in w+dw, \mathbf{AC}\sqrt{W}X + \mathbf{A}m(W) \in f+df\right)}{\mathbb{P}\left(\mathbf{AC}\sqrt{W}X + \mathbf{A}m(W) \in f+df\right)} \\
\propto \quad & \mathbb{P}\left(\mathbf{AC}\sqrt{W}X + \mathbf{A}m(W) \in f+df \;\middle|\; W \in w+dw\right)\mathbb{P}\left(W = w+dw\right) \\
= \quad & \mathbb{P}\left(\mathbf{AC}X \in w^{-1/2}f - w^{-1/2}\mathbf{A}m(w) + w^{-1/2}df\right)\mathbb{P}\left(W = w+dw\right) \\
\propto \quad & w^{-L/2}\left[w^{\lambda-1}\exp\left(-\frac{1}{2}\left(\frac{\chi}{w}+\theta w\right)\right)\right] \\
& \cdot\left[\exp\left(-\frac{1}{2}\frac{1}{w}(f-\mathbf{A}m(w))^{\top}\left(\mathbf{A}\Sigma\mathbf{A}^{\top}\right)^{-1}(f-\mathbf{A}m(w))\right)\right] \\
\propto \quad & w^{-L/2}\left[w^{\lambda-1}\exp\left(-\frac{1}{2}\left(\frac{\chi}{w}+\theta w\right)\right)\right] \\
& \cdot\exp\left(-\frac{1}{2}\frac{(f-\mathbf{A}\mu)^{\top}\left(\mathbf{A}\Sigma\mathbf{A}^{\top}\right)^{-1}(f-\mathbf{A}\mu)}{w}\right)\cdot\exp\left(-\frac{1}{2}\gamma^{\top}\mathbf{A}^{\top}\left(\mathbf{A}\Sigma\mathbf{A}^{\top}\right)^{-1}\mathbf{A}\gamma w\right) \\
= \quad & w^{\widetilde{\lambda}-1}\exp\left(-\frac{1}{2}\left(\frac{\widetilde{\chi}}{w}+\widetilde{\theta}w\right)\right), \qquad \text{(C.2-4)}
\end{aligned}
$$

for $\widetilde{\lambda}$, $\widetilde{\chi}$ and $\widetilde{\theta}$ defined in (4.22). □