QUANTITATIVE FINANCE
RESEARCH CENTRE

UNIVERSITY OF
TECHNOLOGY SYDNEY

QUANTITATIVE FINANCE RESEARCH CENTRE

Research Paper 276 May 2010

Optimal Investment Strategies under Stochastic Volatility – Estimation and Applications

Carl Chiarella and Chih-Ying Hsiao

ISSN 1441-8010 www.qfrc.uts.edu.au

Optimal Investment Strategies under Stochastic Volatility – Estimation and Applications

Carl Chiarella* Chih-Ying Hsiao †

Version of April 2007
Updated in February 2010

Abstract

This paper studies the impact of stochastic volatility (SV) on optimal investment decisions. We consider three different SV models: an extended Stein/Stein model, the Heston Model and an extended Heston Model with a constant elasticity variance (CEV) process and derive the the long-term optimal investment strategies under each of these processes. Since volatility is not a directly observable quantity, extended Kalman filter techniques are adopted to deal with this partial information problem. Optimal investment strategies based on the CEV volatility model are obtained by adopting the Backward Markov Chain approximation method since analytical solutions are no longer available. We find in the empirical investigation that the Heston model is favored as a more parsimonious model compared with the other two models. All three investment strategies based on the three SV models contain a positive intertemporal hedging term in addition to the static mean-variance portfolio. However, in their details the three investment strategies differ from each other. We also find that the investment strategies are sensitive to the CEV parameter.

Key words: Asset allocation, stochastic volatility, partial information problem, extended Kalman filter, the Heston model, CEV process.

1 Introduction

Stochastic volatility has been recognized recently as an important feature for asset price modelling because it is seen as an explanation of a number of well-known empirical findings such as volatility clustering, the thick-tailed nature of return distributions and the "volatility smile". This paper studies how stochastic volatility will affect long-term optimal investment strategies.

In the continuous-time framework, the mainstream approach to modelling stochastic volatility is to assume the volatility itself follows a stochastic process. Representative papers of this modelling approach include Stein and Stein (1991) with a Gaussian volatility process, Heston (1993) with a square-root variance process and Jones (2003) with a constant elasticity variance (CEV) process. Such models explain well some of the empirical features of the

*Email: carl.chiarella@uts.edu.au. School of Finance and Economics, University of Technology, Sydney, Australia

†Corresponding authro. Email: chih-ying.hsiao@uts.edu.au. School of Finance and Economics, University of Technology, Sydney, Australia

joint time series behavior of stock and option prices.

Volatility of asset returns is not observed directly so we need either to adopt filtering techniques to extract volatility from market data, or use extra data such as option prices or market volatility indices to infer volatility. This paper implements maximum likelihood estimation based on the extended Kalman filter to extract volatility from stock returns. Peng et al (2005) adopted a similar approach to estimate volatility in foreign exchange rates. Application of the Kalman filtering techniques in finance can be found in Harvey (1990). Other approaches to estimation of stochastic volatility includes use the volatility index VIX as a proxy in Duan and Yeh (2008), Ait-Sahalia and Kimmel (2007) and Bakshi et al. (2008)**?**. Pan (2002) adopted "implied-state" generalized method of moments (ISGMM) estimation and Chacko and Viceira (2003) adopted spectral GMM estimation. Ait-Sahalia and Kimmel (2007) and Bakshi et al (2008) **?** employed maximum likelihood estimation based on close-form approximations.

Our construction of an optimal long-term portfolio is based on the intertemporal framework of Merton (1971,1973), where long term effects on optimal investment strategies can be obtained quantitatively. Based on the intertemporal framework Liu (2007) provided some advances in obtaining analytical solutions to intertemporal optimal investment problems under a fairly broad class of model specifications including stochastic volatility. However, analytical solutions for the extended model with a CEV process are not available. This paper adopt a *backward Markov chain approximation method* proposed in Chiarella and Hsiao (2006) to provide a computational solution for optimal strategies.

This paper considers three models for modelling stochastic volatility (SV) in stock prices: an extended Stein-Stein model, the Heston model and an extended Heston model based on a CEV process. Based on data for the Australian Stock Index S&P/ASX200 from 1996 to 2006, the Heston is strongly supported empirically as a more parsimonious model against the extended Stein-Stein model as well as the extended Heston model. Optimal long-term investment strategies are constructed based on the all three SV models.

The structure of the paper is organized as follows. Section 2 sets up the framework for the paper including the intertemporal asset allocation problem and the three SV models. In Section 3 we investigate the three SV models based on Australian stock market data and provide a comparison between them. Section 4 provides concrete optimal investment strategies based on the estimation results using both analytical solutions and computational solutions. We will compare the different strategies based on the different models and draw some conclusions in Section 5.

2 The Framework

The investment model is based on within Merton's (1971,1973) intertemporal asset allocation framework. There is one security, the stock S_t, in the investment opportunity set following the dynamics

$$\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dB_t^S \,, \tag{1}$$

where B_t^S is a one-dimensional Wiener process.

2.1 Three models for stochastic volatility

We adopt three models to model stock prices with stochastic volatility (SV).

Model 1: An extended Stein-Stein model
The stock price follows the dynamics

$$\frac{dS_t}{S_t} = \big(R_t + \lambda_t \sigma_t\big)dt + \sigma_t dB_t^S \ , \tag{2}$$

where R_t is the instantaneous risk-less rate, σ_t is the volatility process and λ_t is the process of market price of the risk B_t^S. The quantities σ_t and λ_t are assumed to follow the processes

$$d\sigma_t = \kappa_\sigma(\overline{\sigma} - \sigma_t)dt + \beta_{\sigma S} dB_t^S + g_\sigma dB_t^\sigma \ , \tag{3}$$

$$d\lambda_t = \kappa_\lambda(\overline{\lambda} - \lambda_t)dt + \beta_{\lambda S} dB_t^S + \beta_{\lambda\sigma} dB_t^\sigma + g_\lambda dB_t^\lambda \ . \tag{4}$$

where $(B_t^S, B_t^\sigma, B_t^\lambda)$ are orthogonal multidimensional Wiener processes. Not only is the volatility process σ_t time-varying, this model extends the Stein and Stein (1991) model by allowing the market price of risk λ_t to follow also a stochastic process.

Model 2: The Heston model
The Heston model consider a variance process V_t in the stock price, the dynamics of which are given by

$$\frac{dS_t}{S_t} = \big(R_t + lV_t\big)dt + \sqrt{V_t} dB_t^S \ , \tag{5}$$

where l is a constant. The variance process V_t follows a square-root process

$$dV_t = \kappa_V(\overline{V} - V_t)dt + \sqrt{V_t}\big(b_{SV} dB_t^S + h_V dB_t^V\big) \ , \tag{6}$$

with B_t^V orthogonal of the price risk B_t^S.

The market price of the risk B^S is $l\sqrt{V_t}$ so that the excess instantaneous return in (5) is $l\sqrt{V_t}\sqrt{V_t} = lV_t$. The Heston model can be considered as a special case of the extended Stein-Stein model with the market price of risk (4) perfectly proportional to the volatility $\lambda_t = c\sigma_t$. We will come back to this point later in the empirical section.

Model 3: An extended Heston Model with a CEV process
We extend the Heston model by considering a constant elasticity variance (CEV) process

$$dV_t = \kappa_V(\overline{V} - V_t)dt + \overline{V}^{\frac{1}{2}-\eta} V_t^\eta \big(b_{SV} dB_t^S + h_V dB_t^V\big) \ , \tag{7}$$

with a CEV parameter $\eta \geq 0$. The re-scaling term $\overline{V}^{\frac{1}{2}-\eta}$ sets the diffusion coefficient $\overline{V}^{\frac{1}{2}-\eta} V_t^\eta$ to the value $\overline{V}^{\frac{1}{2}}$ at $\overline{V}$. Setting $\eta = \frac{1}{2}$ the model goes back to the Heston model. The CEV parameter η characterizes the sensitivity of the diffusion coefficient with respect to its level. In Figure 1 we see that for a higher η the diffusion coefficient changes more sensitively with respect to its level V_t. We can also express the η-effect in this way. When V_t is high ($V_t > \overline{V}$), the variance process with a higher η will have higher volatility. While when V_t is low ($V_t < \overline{V}$), the process with a higher η now has lower volatility.

2.2 Investment decision

There are identical rational agents who make investment decisions dynamically over the investment horizon $t \in [0, T]$. The investment weight at time t relative to agents' wealth is

htp

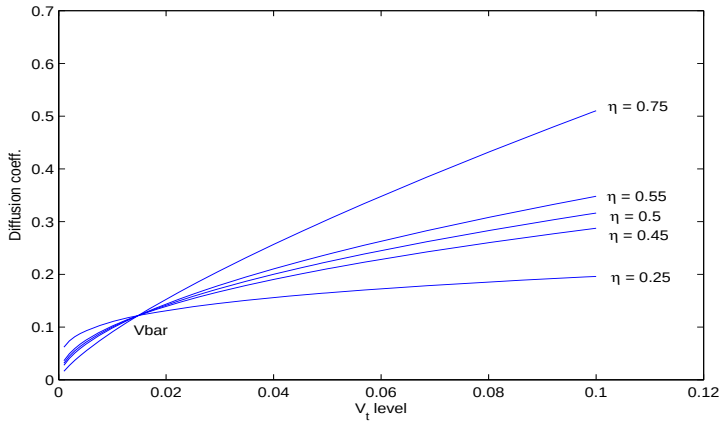


Figure 1: Sensitivity of the CEV diffusion coefficient
The figures describes how the diffusion term $\overline{V}^{\frac{1}{2}-\eta}V_t^{\eta}$ depends on its current level V_t for various η's. The higher the η the more sensitively changes the diffusion term with respect to the variance level V_t.

denoted by α_t which can be re-balanced at every point of time without transaction costs. Under the investment decision α_t the wealth, denoted by W_t, develops according to

$$\frac{dW_t}{W_t} = \alpha_t\Big(\mu_t dt + \sigma_t dB_t^S\Big) + (1-\alpha_t)R_t \ . \tag{8}$$

The optimal investment weight α_t, $t \in [0,T]$ is constructed in order to achieve the maximal expected utility of the final wealth W_T, that is

$$\max_{\alpha_t, t\in[0,T]} \mathbf{E}_0\Big[e^{-\delta T}U(W_T)\Big] \ , \tag{9}$$

where $\delta(>0)$ is the agents' subjective discount rate, The utility function U is assumed to be of CRRA (Constant Relative Risk Aversion) form given by

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma} \ . \tag{10}$$

with the risk aversion parameter $\gamma(>0)$. The larger is γ the more risk averse are the agents.

At each time t, the optimal intertemporal investment weight can be decomposed into an *optimal static investment weight*, the optimal investment weight if the stock volatility were constant, and an *intertemporal hedging term*, the adjustment to investment required to hedge stochastic volatility. This decomposition goes back to Merton (1973) and for more recent results see Liu (2007). We will present this decomposition in the optimal investment solutions in Section 4.

3 Estimating the SV models

Estimation is carried out for the three SV models described in Section 2 using the Australian data. We will see that the empirical results favour the Heston model.

3.1 Data

The data adopted for the empirical investigation are the *S&P ASX 200 Index*. The Australian Stock Exchange (ASX) was formed in 1987 and trading on the ASX has been fully automated since 1996. The S&P ASX 200 Index was introduced in April 2000 and has been maintained by the Standard and Poor's Australian Index Committee since then. It covers approximately 78% of the Australian equities market and is considered as an ideal proxy for the total market. Besides its role as a benchmark index, the S&P ASX 200 also serves as an index for investment purposes because of its high liquidity.

The data are from the weekly S&P ASX 200 Price Index, from 23 Sep 1996 until 25 Sep 2006 consisting of 523 observations, which are provided by Thomas Financial – DATASTREAM. Although the S&P ASX 200 Index was introduced in 2000, the data for the index for the period before the introduction can still be replicated based on the stock prices of the same 200 companies.

The log return of the index is shown in Figure 2. The weekly stock returns have a mean 0.1613% and standard deviation of 1.7336%. The first order autocorrelation is very weak, equal to -0.0624 which can be accepted as indicative of no autocorrelation.

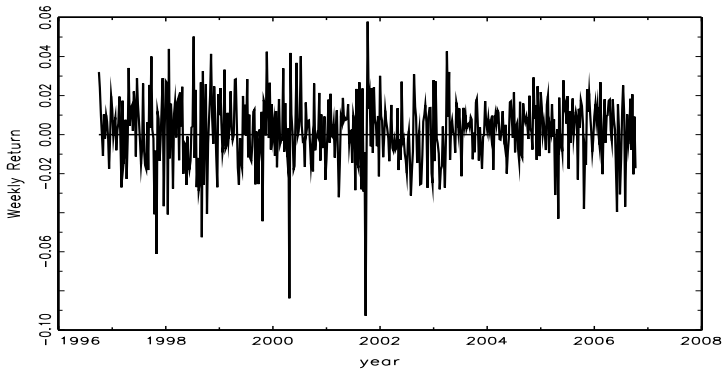


Figure 2: Weekly Log Return of the S&P/ASX 200

3.2 Estimation with filtering techniques

Based on the data we estimate the hidden volatility/variance processes for the three models using maximum likelihood estimation based on the *extended Kalman filter*. We outline the algorithm in detail for the extended Stein-Stein model.

3.2.1 Estimation of Model 1 (the extended Stein-Stein Model)

We consider the log price $s_t := \ln S_t$, the dynamics of which, by application of Ito's Formula to (1) are given by

$$ds_t = \big(R + \lambda_t\sigma_t - \frac{\sigma_t^2}{2}\big)dt + \sigma_t dB_t^S \,. \tag{11}$$

We discretize the equations (11), (3) and (4) using the Euler-Maruyama approximation scheme to obtain the system dynamics for the *state equations*,

$$\begin{pmatrix} \Delta s_{t+\Delta} \\ \Delta\sigma_{t+\Delta} \\ \Delta\lambda_{t+\Delta} \end{pmatrix} = A_t \begin{pmatrix} s_t \\ \sigma_t \\ \lambda_t \end{pmatrix} \Delta + \begin{pmatrix} R_t \\ \kappa_\sigma \overline{\sigma} \\ \kappa_\lambda \overline{\lambda} \end{pmatrix} \Delta + G_t \Delta B_{t+\Delta} \,, \tag{12}$$

where

$$A_t = \begin{pmatrix} 0 & -\frac{1}{2}\sigma_t & \sigma_t \\ 0 & -\kappa_\sigma & 0 \\ 0 & 0 & -\kappa_\lambda \end{pmatrix} \,, \; G_t := \begin{pmatrix} \sigma_t & 0 & 0 \\ \beta_{\sigma S} & g_\sigma & 0 \\ \beta_{\lambda S} & \beta_{\lambda\sigma} & g_\lambda \end{pmatrix} \quad \text{and} \quad \Delta B_{t+\Delta} := \begin{pmatrix} \Delta B^S_{t+\Delta} \\ \Delta B^\sigma_{t+\Delta} \\ \Delta B^\lambda_{t+\Delta} \end{pmatrix}. \tag{13}$$

The *observation equation* is given by

$$p_t = s_t + \epsilon_t \,, \tag{14}$$

where p_t is the logarithm of the observed market stock price and ϵ_t is the measurement error which is assumed to be i.i.d. $\mathcal{N}(0, \sigma_\epsilon)$ distributed and is distributed independently of s_t and of all other risk sources $(B^S_t, B^\sigma_t, B^\lambda_t)$.

Since the coefficients A_t and G_t are now time-dependent but not constant, we cannot use the standard *Kalman filter* technique. We employ here the *extended Kalman filter* as freezing the coefficients A_s and G_s as constant for $s \in [t, t+\Delta)$. The extended Kalman filter method[1] filters out the unobservable state variables $(s_t, \sigma_t, \lambda_t)$ from the observed time series p_t.

The estimation results for the parameters are displayed in Table 1.[2] The filtered instantaneous volatility σ_t and the filtered market price of risk λ_t are plotted in Figure 3 where we see rapid changes in the trajectories of σ_t and λ_t. This coincides with the estimation results in Table 1 that the processes have fast mean-reversion speed (that is large κ_σ and κ_λ) and large volatility (that is large g_σ and $\beta_{\lambda\sigma}$). The estimated process σ_t remains positive over the whole period. The standard deviation of the measurement error σ_ϵ is extremely small ($9.8{*}10^{-8}$). We note there are two risk sources ΔB^S_t and ϵ_t in the observation equation (14),

$$p_t = s_t + \epsilon_t = s_{t-\Delta} + \big(R + \lambda_{t-\Delta}\sigma_{t-\Delta} - \frac{1}{2}\sigma^2_{t-\Delta}\big)\Delta + \sigma_{t-\Delta}\Delta B^S_t + \epsilon_t \,.$$

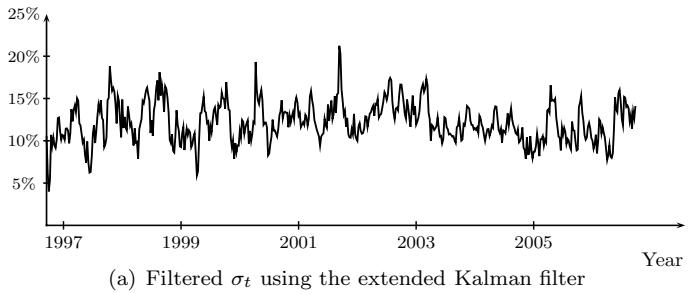
Given the small value estimated for σ_ϵ the risk is then mainly absorbed by the ΔB^S_t.

[1] For reference see Harvey (1990).

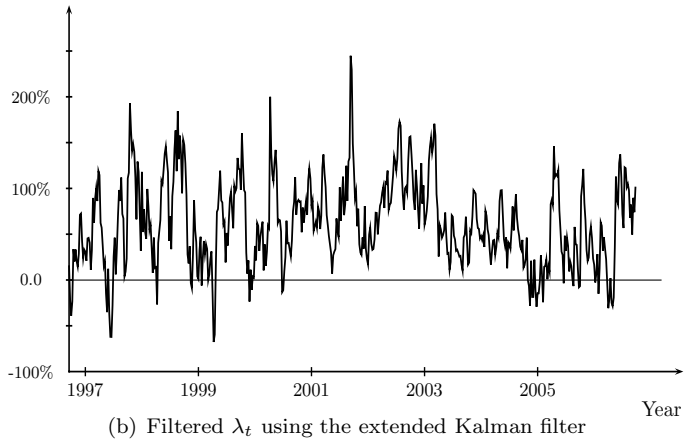
[2] For the estimation we choose R_t in (12) is equal to zero. Subsequent research we should use the returns for the 90 day bank bill rate.

Parameter	Estimate	t-stat.
κ_σ	11.088004	2.845593
κ_λ	9.790118	2.749497
$\overline{\sigma}$	0.122406	17.464097
$\overline{\lambda}$	0.629087	2.369271
g_σ	18.049795	11.044842
g_λ	0.059410	0.522316
$\beta_{\sigma S}$	0.023847	1.106185
$\beta_{\lambda S}$	-1.972936	-1.402033
$\beta_{\lambda\sigma}$	8.406405	5.200823
σ_ϵ	$9.8*10^{-8}$	0.001965
Log-Likelihood	1400.2525	

Table 1: Estimates for Model 1, the Extended Stein/Stein Model



(a) Filtered σ_t using the extended Kalman filter



(b) Filtered λ_t using the extended Kalman filter

Figure 3: The Filtered Factors in the Extended Stein-Stein Model

The estimation has been carried out by use of the GAUSS package. The estimated likelihood of 1400.2525 has been checked against the global maximum obtained by implementing the Genetic Algorithm[3].

[3] The Genetic Algorithm (GA) is used in Chiarella, Hung and Tô (2005) to estimate a nonlinear model using filtering methods. For the estimation here the algorithm provided a log-likelihood value of 1400.1908 based on an implementation using 1,000 generations and population size of 1,000.

Testing time-variation of λ_t and σ_t

We test whether the processes σ_t and λ_t can be held as constant. Table 2 gives the estimation results which are obtained directly by standard maximum likelihood estimation. The estimated reverting means $\overline{\sigma}$ and $\overline{\lambda}$ are very similar to the estimation results of the benchmark model given in Table 2. The result of the Likelihood Ratio (LR) test[4] rejects significantly that the processes σ_t and λ_t are held as constants. The p-value is equal to zero (with accuracy $1.0 * 10^{-8}$).

Model	
$\overline{\sigma}$	0.1250
$\overline{\lambda}$	0.6084 .
Log-Likelihood	1376.5187
p-value(LR-test)	0.0

Table 2: Estimated Results for the Benchmark Model

3.2.2 Towards the Heston Model

Visual inspection indicates that the trajectories of the filtered σ_t and λ_t fluctuate in a very similar way. Their correlation is in fact $\mathrm{Cor}[\sigma_{t|t}, \lambda_{t|t}] = 0.99363091$, indicating that the two processes are extremely highly correlated. This suggests that there holds a linear relation

$$\lambda_t = c\sigma_t \ , \tag{16}$$

and then the two factors collapse into one factor. Define the new factor $V_t := \sigma_t^2$. The last two terms in the drift coefficient in equation (11) become

$$\lambda_t\sigma_t - \frac{1}{2}\sigma_t^2 = (c - \frac{1}{2})\sigma_t^2 = l\sigma_t^2 = lV_t \ , \ \ withl := c - \frac{1}{2} \ ,$$

so we rewrite (11) to

$$ds_t = (R_t + lV_t)dt + \sqrt{V_t}dB_t^S \ . \tag{17}$$

The dynamics of V_t is given by (6). This dynamics can be also obtained by calculating $d\sigma_t^2$ with minor modification[5]. In this sense, the Heston model can be considered as a restricted version of the extended Stein-Stein model with the restrictions

[4] The statistic $\mathcal{LR}_T$ of the Likelihood Ratio Test is defined as

$$\mathcal{LR}_T := -2N\Big(\mathrm{LogLik}(\hat{\theta}_r) - \mathrm{LogLik}(\hat{\theta})\Big) \ , \tag{15}$$

where N is the number of data points, $\hat{\theta}_r$ denotes the estimate in the restricted model and $\hat{\theta}$ denotes the estimate in the unrestricted model. According to the theory of the likelihood ratio test, see for example, Hamilton (1994), the distribution of $\mathcal{LR}_T$ is $\chi(q)$ where q denotes the number of restrictions on the parameters. In our case, the parameter restrictions for a constant σ_t and λ_t are

$$\kappa_\sigma = \kappa_\lambda = g_\sigma = g_\lambda = \beta_{\sigma S} = \beta_{\lambda S} = \beta_{\lambda\sigma} = 0 \ .$$

Also, we let $\sigma_\epsilon = 0$ so $p_t = s_t$. Therefore, we have eight parameter restrictions. The likelihood value of the unrestricted model is given in Table 3 whereas that of the restricted model is given in Table 1.

[5] Recall $d\sigma_t = \kappa(\overline{\sigma} - \sigma_t)dt + \beta dB_t$. Apply Itô's Lemma to σ^2 we have

$$\begin{aligned} dV_t &= d\sigma_t^2 = 2\sigma_t\big(\kappa(\overline{\sigma} - \sigma_t)dt + \beta dB_t\big) + \beta^2 dt \\ &= 2\kappa\big(\sigma_t\overline{\sigma} + \frac{\beta^2}{4} - \sigma_t^2\big)dt + 2\beta\sigma_t dB_t \ . \end{aligned} \tag{18}$$

If we modify (18) by setting $\sigma_t\overline{\sigma} + \frac{\beta^2}{4} := \overline{V}$ as a constant, we obtain the dynamics (6).

$$\kappa_\sigma = \kappa_\lambda, \quad , g_\lambda = 0 \quad , \overline{\lambda} = c\overline{\sigma}, \quad , \beta_{\lambda S} = c\beta_{\sigma S}, \quad \beta_{\lambda\sigma} = cg_\sigma \quad . \tag{19}$$

Estimation of the Heston Model (Model 2)

The estimation is also carried out by maximum likelihood estimation based on the extended Kalman filter technique. The Euler-Maruyama discretization is applied to the dynamics (17) and (6). The estimation results are given in Table 3. The V_t process still has a large mean-reversion speed indicated by a large estimated κ_V and has a reversion mean around 0.0147. The parameter h_V is estimated to be zero which means the two risk sources are perfectly correlated. The LR test is carried out for the the restriction (19) and yields p-value equal to 49.564686%. So we accept the Heston model as a more parsimonious model for modelling stochastic volatility.

Parameters	estimates	t-statistic
κ_V	9.540729	4.386914
$\overline{V}$	0.014735	6.694009
l	5.597853	2.048038
b_{SV}	-0.211266	-3.896279
h_V	0.000000	0.000000
σ_ϵ	0.002714	1.146935
Log Likelihood	1398.56017	
p-value (against Model 1)	49.56%	
p-value (against Model 3)	90.16%	

Table 3: Estimates for Model 2, the Heston Model
The first p-value is from the LR-test against the Model 1 and the second is from the LR-test against the Model 3.

3.2.3 Estimation of the Extended Heston Model (Model 3)

The estimation of Model 3 follows similarly to that Model 2. Table 4 gives the estimation results for Model 3. All of them are similar with those of Model 2 given in Table 3, especially the CEV parameter estimate of $\eta = 0.4546$, which is close to 0.5.[6] We test the restriction $\eta = \frac{1}{2}$ against a free η using the LR ratio. The test has p-value equal to 90.16% (given in Table 3), which indicates an acceptance of the Heston model in its restricted version.

[6] We should point out that the estimate of the CEV parameter is insignificant with a t-statistic equal to 1.1 in Table 4.

Parameters	estimates	t-statistic
κ_V	9.638299	3.781803
$\overline{V}$	0.014753	6.721543
λ	5.573966	2.016290
$b_{S\sigma}$	-0.209753	-3.838356
h_V	0.000000	0.000000
σ_ϵ	0.002645	0.991274
η	0.454551	1.100192
Log-Likelihood	1398.56782	

Table 4: Estimates for Model 3, the Extended Heston Model

Summary
The Heston model passes the LR-test against the extended Heston model with a CEV process as well as against the extended Stein-Stein model for the data. The empirical investigation favors the Heston model for the data we used.

4 The Optimal Investment Strategies

This section provides some concrete investment strategies having exposure to stochastic volatility based on the estimation results in Section 3. Optimal portfolios based on Models 1 and 2 are available in analytical forms while optimal portfolios for Model 3 are obtained through adopting *Backward Markov chain approximation* methods as in Chiarella and Hsiao (2006).

4.1 Computational Solutions

4.1.1 Backward algorithm based on discretization

The optimal investment strategy which maximized the expected utility (9) can be obtained through the *backward Markov chain approximation* scheme, which is proposed by Kushner and Dupius (2000)**?**. Chiarella and Hsiao (2006) applied the scheme to portfolio asset allocation problem. Here we briefly introduce the scheme.

We approximate the optimization problem given (9) by a discretization scheme. First we discretize $[0, T]$ as
$\{0, \Delta, 2\Delta, \cdots, N\Delta(= T)\}$ and agents re-balances their portfolio only at these points. Let $J^T(k\Delta, V_{k\Delta}, W_{k\Delta})$ be the *value function*

$$J^T(k\Delta, V_{k\Delta}, W_{k\Delta}) := \max_{\alpha_{k'\Delta}, k'=k,\cdots,N} \mathbf{E}_{k\Delta}\left[e^{-\delta T} U(W_T) \right], \tag{20}$$

which is the achieved optimized value over the investment horizon $[k\Delta, T]$ given the initial values $(V_{k\Delta}, W_{k\Delta})$. In the discretized system the value function satisfies *Bellman's principle of optimality*

$$J^T(t, W_t, V_t) = \max_{\alpha_t} \mathbf{E}_t\left[J^T\big(t + \Delta, \hat{W}_\Delta(W_t, V_t, \alpha_t), \hat{V}_\Delta(V_t)\big) \right], \tag{21}$$

where $\hat{W}_\Delta$ and $\hat{V}_\Delta$ are discrete time evolutions of the wealth and the volatility obtained by use of the *Eular-Maruyama* approximation as given by

$$\hat{W}_\Delta(W_t, V_t, \alpha_t) := W_{t+\Delta} := W_t + R\Delta + \alpha_t\big(lV_t\Delta + \sqrt{V_t}\Delta B_t^S\big)\ , \tag{22}$$

and

$$\hat{V}_\Delta(V_t) := V_{t+\Delta} := V_t + \kappa_V(\overline{V} - V_t)\Delta + \overline{V}^{\frac{1}{2}-\eta}V_t^\eta\big(b_{SV}\Delta B_t^S + h_V\Delta B_t^V\big) \tag{23}$$

for $t = k\Delta, k = 0, \cdots, N-1$. The Brownian motions ΔB^S and ΔB^V are approximated by binomial trees. The Eular-Maruyama method works well for small discretization interval Δ against the other higher order approximations such as the Milstein method or the Runge-Kutta method.

The terminal condition is given by

$$J^T(T, W_T, V_T) = e^{-\delta T}U(W_T)\ .$$

For the utility function of CRRA type (10), we can decompose the value function into the form as shown in Proposition 5 in Chiarella and Hsiao (2006)

$$J^T(t, W_t, V_t) = e^{-\delta t}U(W_t)\Phi^T(t, V_t)\ , \tag{24}$$

so that $\Phi^T(V_t)$ represents the value function when given initial wealth $W_t = 1$. The value function iteration (21) can then be rewritten as [7]

$$\Phi^T(t, V_t) = (1-\gamma)e^{-\delta\Delta}\max_{\alpha_t}\mathbf{E}_t\Big[U\big(\,1 + \mu_\Delta(V_t, \alpha_t)\,\big)\Phi^T\big(\,t+\Delta, \hat{V}(V_t)\,\big)\Big]\ . \tag{25}$$

The computational codes to implement (25) have been written in the programming language "GAUSS"[8]. The main code implementation of the dynamic programming algorithm has been designed by the authors while the optimization routines have been adopted from the GAUSS application package "Optimization".

4.1.2 Performance of the algorithm

This section will choose discretization steps used in the backward Markov chain approximation method and check its numerical performance. As an example we consider a three-year investment plan where investors can decide to invest in two assets: one being a money account with a riskless rate of $R = 2\%$[9] and the other one being a stock. The stock price is generated according to the dynamics given in Eq (5) and (6) with the parameter values given in Table 3.

Three types of discretization errors will be involved in the approximation method: *time* distretization errors, state space discretization errors and state space truncation errors. The

[7] Using the decomposition (24), we rewrite the iteration scheme (21) as

$$e^{-\delta t}U(W_t)\Phi^T(t, V_t) = \max_{\alpha_t}\Big[e^{-\delta(t+\Delta)}W_t^{1-\gamma}U\big(\,1 + \mu_\Delta(V_t, \alpha_t)\,\big)\Phi^T\big(\,t+\Delta, \hat{V}(V_t)\,\big)\Big]\ .$$

[8] Provided by Aptech, see www.aptech.com.

[9] This level is close to the average interest payment on cash management accounts for amounts of AU$ 10000. We also point out that the riskless rate R does not appear in the analytical solutions for the benchmark model, Models 1 and 2.

second and last errors arise because we can only calculate Φ^T (25) on finite points of V_t[10]. Truncation errors occur as we force the process V_t to remain between an upper and lower reflecting boundary. They can actually be managed relatively well because the square-root process V_t given in (6) has a stationary distribution[11]. In Figure 4 we plot the stationary distribution of V_t based on the estimated results given in Table 3 where we can see that most values lie within the range $[0.0, 0.04]$. In our computational exercise we take a range $[0.0, 0.08]$ for V_t.

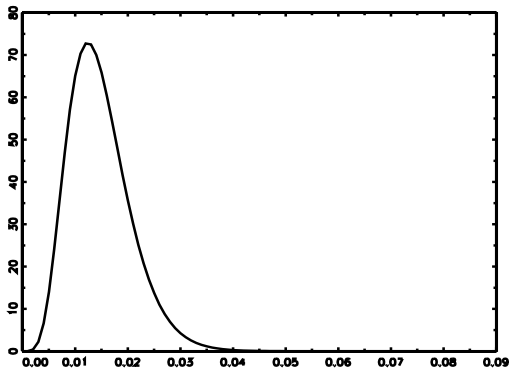


Figure 4: Stationary distribution of the V_t process (given in Eq. (6) with parameters given in Table 3.

To investigate the discretization errors for both state space and time space we consider the grid sizes $\Delta V = 0.0025$ and 0.00125 and time discretizations $\Delta t = 0.01$ and 0.001. In order to investigate the computational performance, we compare the solutions for both the optimal investment proportions and the value function. The solution for the optimal investment proportion and the value function are given by Properties 3 and 4 respectively. In Table 5 we summarize the average relative absolute errors which is the average of the absolute relative error

$$\frac{|\alpha_{computational} - \alpha_{analytical}|}{\alpha_{analytical}} .$$

Investment α_t %	$\Delta V = 0.0025$	$\Delta V = 0.00125$
$\Delta t = 0.01$	0.655%	0.655%
$\Delta t = 0.001$	0.061%	0.067%

Table 5: Average relative absolute errors due to Discretizations.

We can see from Table 5 that a reasonable discretixation would be $\Delta t = 0.001$ and $\Delta V = 0.0025$ for the numerical computation to be carried out later since refining ΔV does not improve the performance very much and $\Delta = 0.001$ provides already relatively small average absolute errors.

4.2 Optimal Investment Weights

We will provide the optimal investment decisions and investigate the impact of stochastic volatility via our three models. In order to gauge the impact of stochastic volatility we start

[10] We linearly interpolate for the other values which are not calculated on grid points. For other sophisticated interpolation methods and the implementation details see Rust (1996).

[11] The formula of the stationary distribution is given in equation (49) in the Appendix.

by providing the optimal investment weight for the constant volatility case. Throughout this section the risk aversion parameter γ is chosen to be 7. Munk et al. (2004) estimate γ to be around 4.8 for moderate investors and 7.6 for aggressive investors.

4.2.1 Benchmark Model

As we hold $\lambda_t = \overline{\lambda}, \sigma_t = \overline{\sigma}$ constant over time the optimal investment weight is also constant.

Property 1. *The optimal portfolio for the reference model is given by*

$$\alpha_t^* = \frac{1}{\gamma}\frac{\lambda}{\sigma}\,, \tag{26}$$

See Merton (1973).

Since there is no stochastic volatility in this model the optimal intertemporal investment weight is equal to the optimal static investment weight. For the parameters given in Table 2 and risk aversion parameter $\gamma = 7$ the optimal investment proportion in the stock is 72.12%, which means 72.12% of the entire wealth.

4.2.2 Model 1 : The extended Stein/Stein Model

Here we calculate the optimal investment weight based on the extended Stein-Stein model.

Property 2. *The optimal portfolio for Model 1 is given by*

$$\alpha^*(t) = \underbrace{\frac{1}{\gamma}\frac{\lambda_t}{\sigma_t}}_{Static} + \underbrace{\frac{\beta_{\lambda S}}{\sigma_t}\big(d_\lambda(T-t) + Q_{\lambda\lambda}(T-t)\lambda_t\big)}_{Intertemp.\ Hedging}\,, \tag{27}$$

where[12]

$$Q_{\lambda\lambda}(\tau) = -\frac{2\big(\exp(\xi\tau)-1\big)}{(K+\xi)\big(\exp(\xi\tau)-1\big)+2\xi}\,\delta\,, \tag{28}$$

$$d_\lambda(\tau) \simeq \frac{\kappa_\lambda\overline{\lambda}\,\overline{q}}{K_q}\big(1-e^{-K_q\tau}\big)\,, \tag{29}$$

with

$$\begin{aligned} K &= \kappa_\lambda - \beta_{\lambda S}\frac{1-\gamma}{\gamma}\,,\\ \xi &= \sqrt{K^2 + 2\big(\beta_{\lambda S}^2 + \gamma\beta_{\lambda\sigma}^2 + \gamma g_\lambda^2\big)}\\ \delta &= \frac{\gamma-1}{2\gamma^2}\,,\\ K_q &= K - \overline{q}\big(\beta_{\lambda S}^2 + \gamma\beta_{\lambda\sigma}^2 + \gamma g_\lambda^2\big)\,,\\ \overline{q} &= \frac{-2\delta}{K+\xi}\,. \end{aligned}$$

□

Proof see Appendix.

Figure 5 illustrates the optimal investment strategy implied by Eq. (27) based on the

[12]The symbol $\simeq$ represents an approximate relation, which here is almost an equality for $\tau > 1$.

estimation results for the whole investment horizon (which is equal to our observation horizon). The optimal investment weight evolves over time because the intertemporal hedging term depends on the λ_t. We can see that the optimal portfolio strategy is very volatile. We decompose the optimal strategy into an optimal static portfolio and an intertemporal hedging term according to Eq. 27) and draw them respectively in Figure 6. The optimal static investment weight is volatile and dominates (for the given risk aversion parameter $\gamma = 7$). The volatile behavior of the optimal static investment weight ($= \lambda_t/\sigma_t$) is due to the volatile trajectories of λ_t and σ_t (although these two processes are highly correlated). Note that the market price of risk attains negative values at the beginning of the observation period because of the bad performance of the index. This leads to a negative suggested investment proportion during this period. The optimal static portfolio proportions fluctuate around 73.42% which is calculated at their means $\sigma_t = \overline{\sigma}, \lambda_t = \overline{\lambda}$.

The optimal investment weight relating to intertemporal hedging term is positive. The positive intertemporal hedging term mathematically comes from a negative correlation between the asset return shock B_t^S and the shock to the market price of risk B_t^λ (that shows up as the negative value of $\beta_{\lambda S}$ in Table 1) and a negative $d_\lambda(\tau)$ term. This investment weight starts around 6.16%, remains at the same level and then decreases shortly before the end of the investment. As the final time $t = T$ the intertemporal hedging term is zero according to Eq. (27). We see in Figure 7 the processes $d_\lambda(T-t)$ and $Q_{\lambda\lambda}(T-t)$ have similar movements. They stay at the levels equal to $d_\lambda(\infty)$ and $Q_{\lambda\lambda}(\infty)$ until time t comes very close to the final time T. This kind of the movement can be explained by a large mean reversion parameter κ_λ. A large κ_λ leads to a large K_q in Eq. (29) so that $d_\lambda(\tau)$ goes to $d_\lambda(\infty)$ very quickly.

4.2.3 Model 2: The Heston Model

Now we provide the optimal investment weight for the Heston model.

Property 3. *The optimal portfolio for Model 2 is given by*

$$\alpha^*(t) = \underbrace{\frac{l+\frac{1}{2}}{\gamma}}_{Static} + \underbrace{b_{SV} d_V(T-t)}_{Intertemp.\ Hedging} \quad , \tag{30}$$

where

$$d_V(\tau) = -\frac{2\big(\exp(\xi_V\tau)-1\big)}{(\tilde{K}_V+\xi_V)\big(\exp(\xi_V\tau)-1\big)+2\xi_V}\delta_V \ , \tag{31}$$

where

$$\tilde{K}_V = \kappa_V - \frac{1-\gamma}{\gamma}\sigma_V(l+\frac{1}{2}) \ , \tag{32}$$

$$\sigma_V = \sqrt{b_{SV}^2 + h_V^2} \ , \tag{33}$$

$$\delta_V = -\frac{1-\gamma}{2\gamma^2}(l+\frac{1}{2})^2 \ , \tag{34}$$

$$\xi_V = \sqrt{\tilde{K}_V^2 + 2\delta_V\sigma_V^2} \ . \tag{35}$$

Property 4. *The value function for Model 2 is given by*

$$\Phi^T(V_t) = \exp\Big(c_V(T-t) + d_V(T-t)V_t\Big) \ , \tag{36}$$

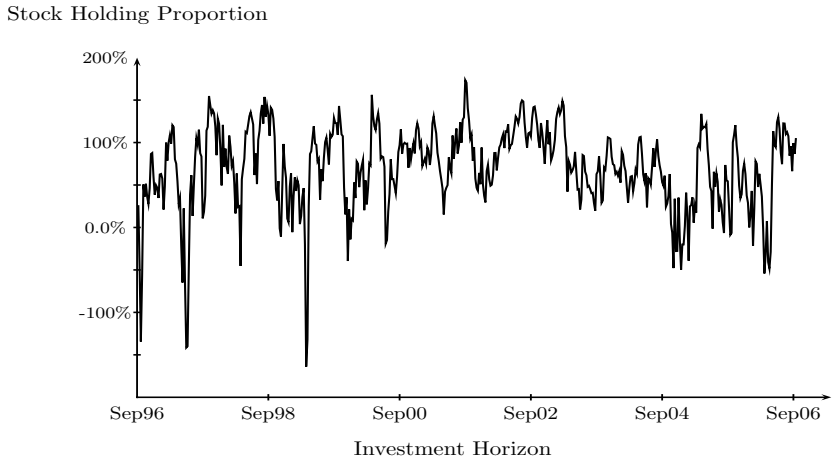


Figure 5: The Optimal Investment Strategy for Model 1

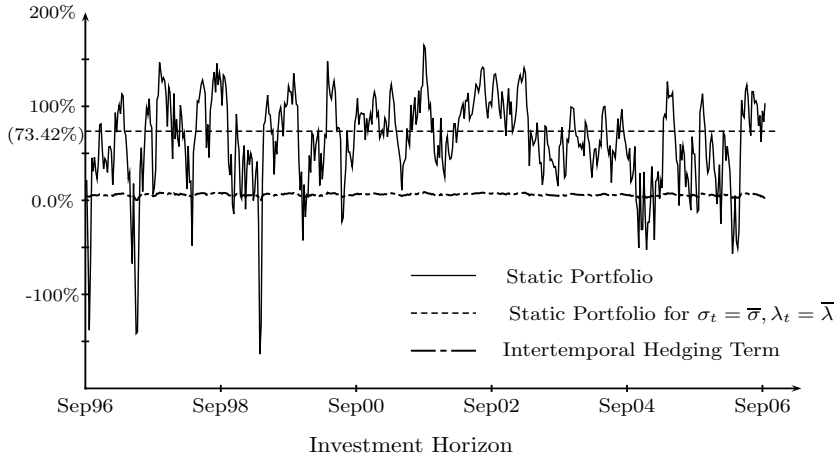
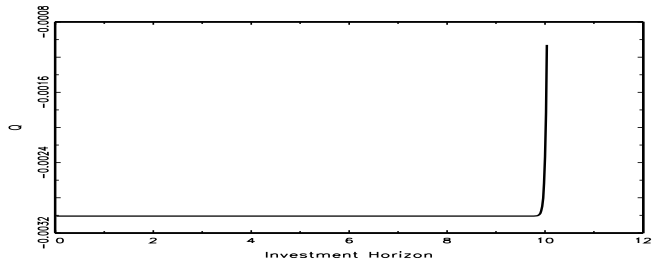
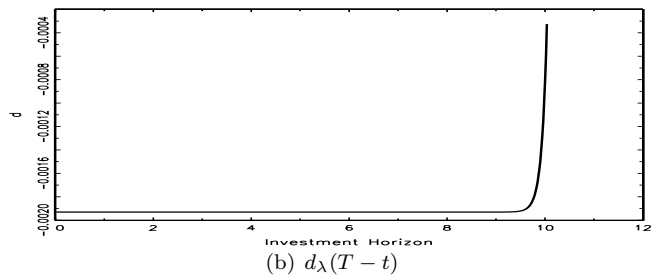


Figure 6: Decomposition for the Optimal Investment Strategy of Model 1



(a) $Q_{\lambda\lambda}(T-t)$



(b) $d_\lambda(T-t)$

Figure 7: Coefficients in the value function for Model 1

where

$$\begin{aligned} & c_V(\tau) \\ = \quad & -\frac{2\kappa_V \overline{V}}{\sigma_V^2} \ln\left(\frac{\tilde{K}_V + \xi_V}{2\xi_V}\exp(\xi_V \tau) + \frac{\xi_V - \tilde{K}_V}{2\xi_V}\right) + \frac{2\tau\delta_V \kappa_V \overline{V}}{\xi_V - \tilde{K}_V} + \frac{R - \gamma R - \delta}{\gamma}\tau, \end{aligned} \tag{37}$$

with $\tilde{K}_V, \sigma_V, \delta_V, \xi_V$ *defined in equations (32) to (35).*

See Liu (2007) Corollary 3 on P.29 with $\lambda_s = l + \frac{1}{2}$.

In contrast to Model 1, the optimal investment proportion (30) depends only on the time to maturity as shown in Figure 8. The optimal static investment $(l+\frac{1}{2})/\gamma = 79.97\%$ is constant over time based on the estimation results and the assumption $\gamma = 7$. The intertemporal hedging term starts at its long-term level $b_{SV} d_V(\infty) = 4.75\%$ and then declines to zero as the investment period ends. This positive hedging position, similar to the result for Model 1, is also due to the negative correlation between the asset return shocks and the shock to the market price of risk, and the negative d_V.

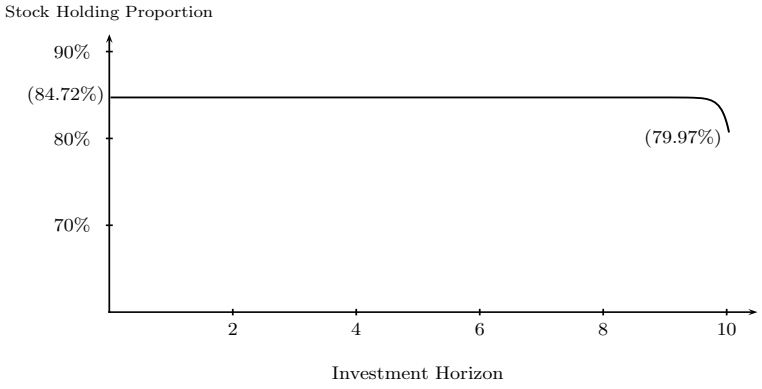


Figure 8: The Optimal Investment Strategy based on Model 2

Comparing the optimal investment weights of Models 1 and 2, the optimal static investment weights of Model 2 are at a similar level to the mean optimal static investment weights of Model 1. The optimal static investment weights of Models 1 and 2 also have a similar scale with the benchmark model with constant σ and λ in Section 4.2.1. We summarize the in Table 6. The intertemporal term is also of a similar scale and has similar movement.

	Optimal Static Invest. Weight	Intertemp Hedging far from T
The Benchmark Model	72.12%	–
Model 1	73.42%	6.16%
Model 2	79.97%	4.57%

Table 6: Comparison of Investment Strategies based on Models 1 and 2

4.2.4 Model 3, the Extended Heston Model

Here we provide computational results for the optimal investment strategies for different CEV parameters. Since the optimal investment proportion does not change much until the

investment approaches maturity as we saw in Fig. 8 we reduce computational burden by reducing the investment horizon from ten years to three years . The discretization steps are still given by $\Delta t = 0.001, \Delta V = 0.0025$.

Figures 9 and 10 summarize the computational results where each graph depicts how the optimal investment weights change from time $t = 1.6$ until $t = 2.9999$ against various levels of the variance V_t. The risk aversion parameter γ is still set to be 7. The first graph in Figure 9 shows the optimal investment proportions based for $\eta = \frac{1}{2}$ (the Heston model). They are not dependent on V_t therefore they are horizontal lines. The optimal investment proportion in the stock is 84.7% at $t = 1.6$ and declines to 79.97% at $t = 2.9999$. These optimal investment weights coincide exactly with those in Figure 8, which indicates that the backward Markov chain approximation method performs well. The sharp drop in Figure 8 can be also seen in the first graph in Figure 9. For the case of $\eta < 0.5$ shown in the second and third graphs in Figure 9 we see that the investment weights now depend on the variance level V_t. The weights decrease with the level of V_t. The intertemporal hedging term also decreases with the level of V_t. The term can be graphically represented by the difference to the "static portfolio" at $T = 2.9999$. This effect becomes more significant as η moves more further away from $\frac{1}{2}$. Recall that in Figure 1 that for an η smaller than 0.5, the CEV process (7) has a higher diffusion coefficient than the square-root process (6) at low V_t levels. This gives rise to a stronger intertemporal hedging term than that of the Heston model at low V_t levels. When we now consider the opposite case $\eta > 0.5$ given in the second and third graphics in Figure 10 the situation reverses. The intertemporal term now increases with the variance level V_t. This effect can again be explained through Figure 1 where the CEV process has higher volatility at high levels V_t.

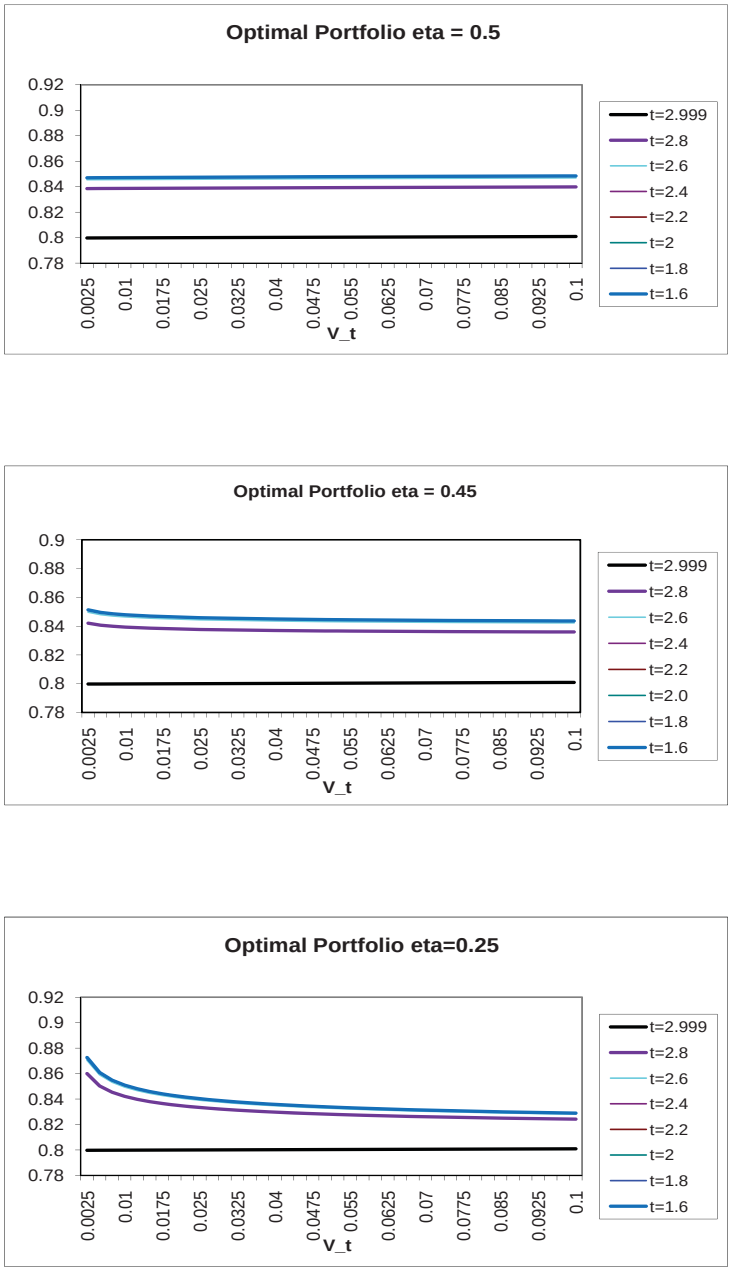


Figure 9: Optimal Portfolio for $\eta \leq \frac{1}{2}$.
Each figure plots the optimal investment weights against different variance levels V_t for different investment times from $t = 1.6$ until $t = 2.9999$ for a given η. The investment matures at $T = 3.0$. The risk aversion parameter is set to be $\gamma = 7$. For the case $\eta = 0.25$ in the third graph the optimal investment proportion at investment time $t = 1.6$ is 87.26% (of the whole wealth) for a low variance level $V_t = 0.0025$ and declines to 82.98% for a high volatility level $V_t = 0.1$.

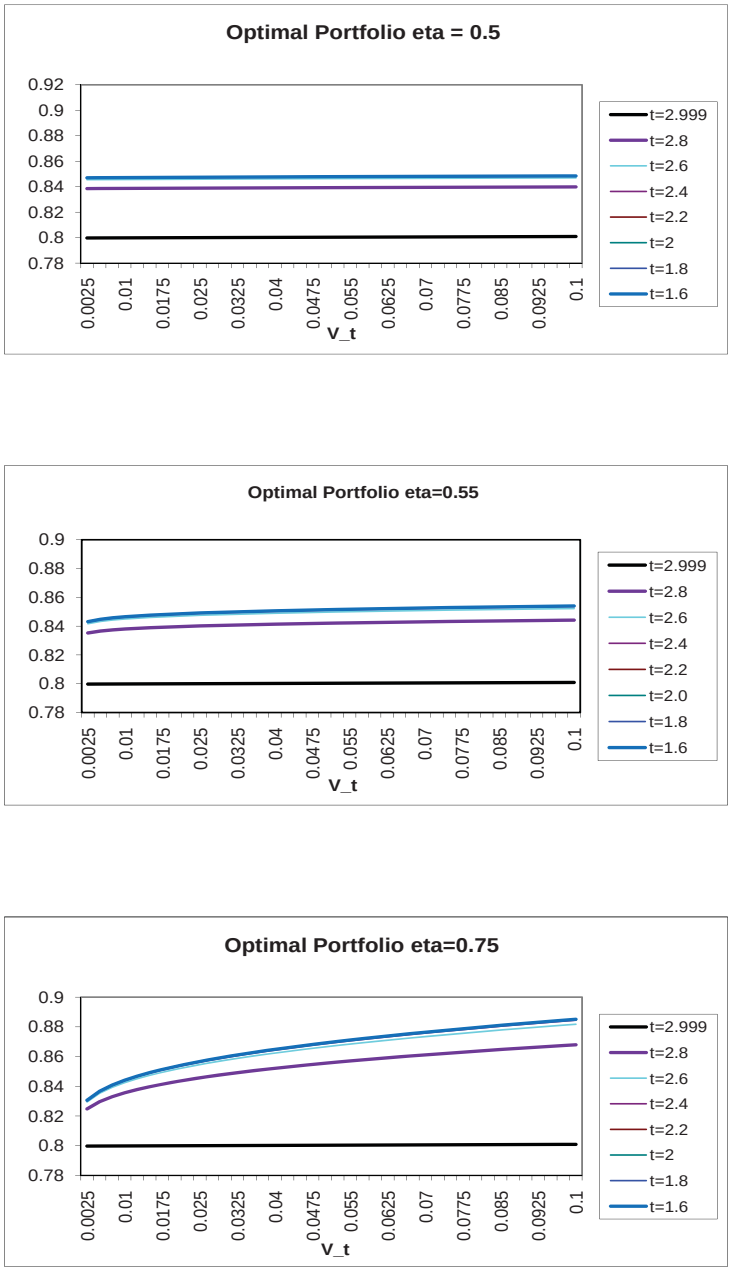


Figure 10: Optimal Portfolio for $\eta \geq \frac{1}{2}$.
Each figure plots the optimal investment weights against different variance levels V_t for different investment times from $t = 1.6$ until $t = 2.9999$ for a given η. The investment matures at $T = 3.0$. The risk aversion parameter is set to be $\gamma = 7$. For the case $\eta = 0.75$ in the third graph the optimal investment proportion at investment time $t = 1.6$ is 83.06% (of the whole wealth) for a low variance level $V_t = 0.0025$ and increases to 88.51% for a high volatility level $V_t = 0.1$.

5 Conclusions

The paper has investigated the effect of stochastic volatility via three models, an extended Stein-Stein, the Heston model and an extended Heston model with a CEV process. The backward Markov chain approximation method was adopted in order to obtain optimal investment strategies based on the extended Heston model. Regarding modelling of stochastic volatility, our empirical results based on the ASX200 data favors the Heston model as a more parsimonious model. Results for optimal investment strategies are summarized in the following four points:

- For all three SV models, stochastic volatility induces a positive intertemporal hedging term for the all three SV models. The positive intertemporal hedging term is explained by a negative correlation between the volatility process/variance process and the asset return process. The intertemporal hedging terms of all three SV models are of a similar scale.
- Regarding optimal static investment weights, the extended Stein-Stein model suggests a very volatile process of weights depending on the level of σ_t and λ_t. The Heston model and extended Heston model suggest a constant weights independent of V_t.
- For all three SV models, the longer the time to maturity the stronger the intertemporal hedging effect. The intertemporal hedging terms remain at the same level for the most of investment time, then decline only shortly before the end of the investment period, and then go down to zero at the end of the investment.
- The CEV parameter η has an impact on the intertemporal hedging term. For the case $\eta = \frac{1}{2}$ the intertemporal hedging term does not depend on the level of the variance V_t. While for $\eta < \frac{1}{2}$ the term decreases with the level of the variance V_t and increases with V_t for $\eta > \frac{1}{2}$.

6 Appendix

Proof of Property 2

Following Liu (2007), when considering Model 1, see (3) and (4), the value function $J(t,T,W_t,\sigma_t,\lambda_t)$ for the optimization problem (9) is defined as

$$J(t,T,W_t,\sigma_t,\lambda_t) := \max_{\alpha_s, s\in[t,T]} \mathbf{E}_t\big[e^{-\delta T}U(W_T)\big] \,.$$

The solution turns out to be of the form

$$J(t,T,W_t,\sigma_t,\lambda_t) = e^{-\delta T}U(W_t)\Phi(t,T,\sigma_t,\lambda_t)$$

where $\Phi(t,T,\sigma_t,\lambda_t)$ has the form

$$\begin{aligned}\Phi(t,T,\sigma_t,\lambda_t) \;&=\; \exp\Big(\; c(t) + d_\sigma(t)\,\sigma_t + d_\lambda(t)\,\lambda_t \\ &\qquad + \frac{1}{2}Q_{\sigma\sigma}(t)\,\sigma_t^2 + Q_{\sigma\lambda}(t)\,\sigma_t\lambda_t + \frac{1}{2}Q_{\lambda\lambda}(t)\,\lambda_t^2 \;\Big)\,.\end{aligned} \tag{38}$$

with the terminal conditions

$$c(T) = d_\sigma(T) = d_\lambda(T) = Q_{\sigma\sigma}(T) = Q_{\sigma\lambda}(T) = Q_{\lambda\lambda}(T) = 0\,. \tag{39}$$

The task is to determine the coefficients $c(t), d_\sigma(t), d_\lambda(t), Q_{\sigma\sigma}(t), Q_{\sigma\lambda}(t)$ and $Q_{\lambda\lambda}(t)$ such that J satisfies the *HJB equation*[13].

Inserting the solution form (38) into the HJB equation and comparing coefficients, it turns out that the coefficients $Q_{\sigma\sigma}(t)$ and $Q_{\sigma\lambda}(t)$ are zero[14]

$$Q_{\sigma\sigma}(t) \equiv 0\,,\;\; Q_{\sigma\lambda}(t) \equiv 0\,, \qquad \forall t = [0,T] \tag{40}$$

The coefficient $Q_{\lambda\lambda}(t)$ satisfies the non-linear ODE of Riccati-type

$$\frac{d}{dt}Q_{\lambda\lambda}(t) + Q_{\lambda\lambda}(t)\big(\beta_{\lambda S}\frac{1-\gamma}{\gamma} - \kappa_\lambda\big) + \frac{Q_{\lambda\lambda}^2(t)}{2}\big(\beta_{\lambda S}^2 + \gamma\beta_{\lambda\sigma}^2 + \gamma g_\lambda^2\big) + \frac{1-\gamma}{\gamma} = 0\,, \tag{41}$$

which has the solution given in (28). For the linear coefficients $d_\sigma(t)$ and $d_\lambda(t)$, one of the coefficients is solved by

$$d_\sigma(t) = 0 \qquad \forall t = [0,T], \tag{42}$$

with the same reasoning as above. The coefficient $d_\lambda(t)$ satisfies the linear ODE

$$d_\lambda'(t) + \Big(\beta_{\lambda S}\frac{1-\gamma}{\gamma} - \kappa_\lambda + \big(\beta_{\lambda S}^2 + \gamma\beta_{\lambda\sigma}^2 + \gamma g_\lambda^2\big)Q_{\lambda\lambda}(t)\Big)d_\lambda(t) + \kappa_\lambda\overline{\lambda}Q_{\lambda\lambda}(t) = 0\,. \tag{43}$$

For our case the function $Q_{\lambda\lambda}(\tau)$ approaches its limit very quickly, as shown in Figure 7, the value of which is given by

$$\lim_{\tau\to\infty} Q_{\lambda\lambda}(\tau) = \frac{-2\delta}{K+\xi} =: \overline{q}\,. \tag{44}$$

[13] See, for example Liu (2007) or Hsiao (2006)**?**.

[14] From the comparison of coefficients, the equation which $Q_{\sigma\sigma}(t)$ and $Q_{\sigma\lambda}(t)$ have to satisfy is given by

$$-\kappa_\sigma Q_{\sigma\sigma}(t) + \frac{\beta_{\sigma S}^2 + \gamma g_\sigma^2}{2}Q_{\sigma\sigma}(t)^2 + \big(\beta_{\sigma S}\beta_{\lambda S} + g_\sigma\beta_{\lambda\sigma}\big)Q_{\sigma\sigma}(t)Q_{\sigma\lambda}(t) + \frac{\beta_{\lambda S}^2 + \gamma g_\lambda^2}{2}Q_{\sigma\lambda}(t)^2 = 0\,.$$

We remark that can check easily that the trivial solutions (40) satisfy this PDE.

This allows us to consider an approximation where we replace $Q_{\lambda\lambda}(t)$ in eqn (43) with $\overline{q}$ and then obtain the approximate solution

$$d_\lambda(\tau) = \frac{\kappa_\lambda \overline{\lambda q}}{K_q}\big(1 - e^{-K_q \tau}\big) \,, \tag{45}$$

where

$$\begin{aligned} \tau &= T - t \,, \\ K_q &= K - \overline{q}\big(\beta_{\lambda S}^2 + \gamma \beta_{\lambda\sigma}^2 + \gamma g_\lambda^2\big) \,. \end{aligned}$$

□

Proof of Property 4
The calculation is based on Liu's (2005) **?** example Sec.3.4 on pp. 22-23.
In order to calculate c_V, we use equation (15) in Liu (2007) and obtain

$$\frac{d}{dt} c_V(t) + \kappa_V \overline{V} d(t) + \frac{R - \gamma R - \delta}{\gamma} = 0 \,. \tag{46}$$

□ **Properties of the CIR process**
Consider the CIR process

$$dV_t = \kappa_V(\overline{V} - V_t)dt + \sigma_V dW_t \,, \tag{47}$$

where $\kappa_V > 0, \overline{V} > 0$ and $\sigma_V > 0$.

The following results are adopted from Cox et al. (1985) p.391. The transition density of the CIR process is given by

$$p(V_t, t|V_0, 0) = \phi(t) \exp\big(-c(t)(\hat{v}_0 + V_t)\big)\big(\frac{V_t}{\hat{v}_0}\big)^{\frac{\nu}{2}} I_\nu\Big(2\phi(t)\sqrt{\hat{v}_0 V_t}\Big) \,, \tag{48}$$

where

$$\hat{v}_0 = e^{-\kappa_V t} V_0 \,, \quad \phi(t) = \frac{2\kappa_V}{\sigma_V^2(1 - e^{-\kappa_V t})} \,, \quad \nu = \frac{2\overline{V}\kappa_V}{\sigma_V^2} - 1 \,,$$

and I_ν is the *modified Bessel function of the first type* and given by

$$I_\nu(z) = (\frac{z}{2})^\nu \sum_{j=0}^{\infty} \frac{(\frac{z}{2})^{2j}}{j!\Gamma(j + \nu + 1)} \,.$$

We require here $\overline{V}\kappa_V > \frac{\sigma_V^2}{2}$.

The CIR process in this case has a stationary distribution, which is given by

$$p(x) = \frac{1}{\Gamma(\eta)} \, \overline{\phi}^\eta x^{\eta - 1} e^{-\overline{\phi} x} \,, \tag{49}$$

where $\overline{\phi} := 2\kappa_V/\sigma_V^2 = \lim_{t\to\infty} \phi(t)$ and the *gamma* function $\Gamma(\eta) = \int_0^\infty x^{\eta-1} e^{-x} dx$.

References

Ait-Sahalia, Y. and Kimmel, R. (2007), 'Maximum likelihood estimation of stochastic volatility models', *Journal of Financial Economics* **83**(2), 413–452.

Chacko, G and Viceira, L. M. (2003), 'Spectral gmm estimation of continuous-time processes', *Journal of Econometrics* **116**, 259–292.

Chiarella, C. and Hsiao, C.-Y. (2006), 'The impact of short-sale constraints on asset allocation strategies via the backward markov chain approimxation method', *Computational Economics* **28**(2), 113–137.

Chiarella, C, Hung, H. and Tô, T. (2009), 'The volatility structure of the fixed income market under the HJM framework: A non-linear filtering approach', *Computational Statistics and Data Analysis* **53**, 2075–2088.

Cox, J.C. , Ingersoll, J. and Ross, S. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–407.

Duan, J. and Yeh, C. (2008), 'Jump and volatility risk premiums implied by vix', *SSRN online source* `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=966682` .

Hamilton, J. (1994), *Time Series Analysis*, Princeton.

Harvey, A.C. (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge.

Heston, S.L. (1993), 'A closed-form solution for options with stochastic volatility with application to bond and currency options', *The Review of Financial Studies* **6**(2), 327–343.

Jones,C. (2003), 'The dynamics of stochastic volatility: evidence from underlying and options markets', *Journal of Econometrics* **116**, 181–224.

Liu, J. (2007), 'Portfolio selection in stochastic environments', *Review of Financial Studies* **20**(1), 1–39.

Merton, R.C. (1971), 'Optimum consumption and portfoliio rules in a continuous-time model', *Journal of Economic Theory* **3**, 373–413.

Merton, R.C. (1973), 'An intertemporal capital asset pricing model', *Econometrica* **41**(5), 867–887.

Munk, C., Sörensen, C. and Vinther, T. (2004), 'Dynamic asset allocation under mean-reverting returns, stochastic interest rates and inflation uncertainty: Are populer recommdendations consistent with rational behavior?', *International Review of Economics and Finance* **13**(141-166).

Pan, J (2002), 'The jump-risk premia implicit in options: evidence form tan intergrated time-series study', *Journal of Financial Economics* **63**, 3–50.

Peng, H, Tamura, Y., Gui, W. and Ozaki, T. (2005), 'Modelling and asset allocation for financial markets based on a stochastic volatility microstructure model', *International Journal of System Science* **36**(6), 315–327.

Rust, J. (1996), *Handbook of Computational Economics*, Elsevier, North Holland, chapter Numerical Dynamic Programming in Economics.

Stein, E.M. and Stein, J. (1991), 'Stock price distributions with stochastic volatility: An analyic approach', *The Review of Financial Studies* **4**(4), 727–752.