A Multifactor Stochastic Volatility Model with Time-Varying Conditional Skewness

Bruno Feunou∗ Roméo Tédongap†
Duke University Stockholm School of Economics

Previous Version: November 2008
This Version: March 2009

Abstract

We develop a conditional arbitrage pricing theory (APT) model where factors and idiosyncratic noises are both heteroscedastic and asymmetric. The model features both stochastic volatility and conditional skewness (SVS model), as well as conditional leverage effects. We explicitly allow asset prices to be asymmetric conditional on current factors and past information, termed contemporaneous asymmetry. Conditional skewness is driven by conditional leverage effects (through factor loadings) and contemporaneous asymmetry (through idiosyncratic skewness). We estimate and test three versions of the SVS model using several equity and index daily returns, as well as daily index option data. Results suggest that contemporaneous asymmetry is particularly important in several dimensions. It helps to match sample return skewness, negative and significant cross-correlations between returns and squared returns, as well as positive and significant cross-correlations between returns are cubed returns. Further diagnostics suggest that SVS models with contemporaneous asymmetry show a better option pricing performance compared to contemporaneous normality and existing affine GARCH models, especially, but not only, for in-the-money call options and short-maturity contracts.

Keywords: APT, Discrete-Time Model, Continuous-Time Limit, GARCH, GMM

JEL Classification: G12, C01, C22, C51

∗Duke University, Economics Department, 213 Social Sciences Building, Durham, NC 27708, USA. Email: Bruno.Feunou@duke.edu.
†Corresponding Author: Stockholm School of Economics, Finance Department, Sveavägen 65, 6th floor, Box 6501, SE-113 83 Stockholm, Sweden. Email: Romeo.Tedongap@hhs.se.
1 Introduction

Three relevant stylized facts have emerged from the analysis of financial time series, namely, time-varying conditional variance (or heteroscedasticity), time-varying conditional leverage effect, and time-varying conditional skewness. Since these time series characteristics are common to many financial assets, and given that these assets are likely to be affected by same economic risk factors, time series properties of the factors combined with asset’s systematic risk and idiosyncratic characteristics, will have important implications for the time series of asset returns. This article develops a conditional arbitrage pricing theory (APT) model where factors and idiosyncratic noises are both heteroscedastic and asymmetric. Heteroscedasticity in the factors implies heteroscedasticity in asset returns, as well as time-varying conditional skewness and leverage effect. Our approach does not tackle independently time series and cross-sectional characteristics of asset returns. In fact, leverage effect arises from asset systematic risk (asset’s factor loading or beta), heteroscedasticity results from asset’s beta and idiosyncratic volatility, and conditional skewness relates to both asset’s beta, idiosyncratic volatility and idiosyncratic skewness.

The autoregressive conditional heteroscedasticity (ARCH, Engle (1982)) and its generalization (GARCH, Bollerslev (1986)) have been widely used in modeling time-series variation in conditional variance. While return volatility is completely determined as a function of past observed returns in ARCH and GARCH models, an alternative approach, which has become more popular recently, is the stochastic volatility (SV) model, where return volatility is an unobserved component which undergoes shocks from a different source other than return shocks. Most empirical applications of SV and GARCH models assume that the conditional distribution of returns is symmetric. Even if these models help matching the observed unconditional kurtosis in actual data, they fail to match unconditional asymmetries (skewness and leverage effects). Allowing for conditional leverage effect in GARCH models (Nelson (1991) and Engle and Ng (1993)) helps to match these unconditional asymmetries. Heston and Nandi (2000, hereinafter HN), Christoffersen et al. (2008) and Christoffersen, Heston and Jacobs (2006, hereinafter CHJ) are examples of GARCH models which belong to the discrete-time affine class, and feature conditional leverage effect (the three of them) and conditional skewness (only CHJ). Conditionally nonsymmetric return innovations are critically important as in option pricing for example, heteroscedasticity and leverage effect alone do not suffice to explain the option smirk. However, skewness in the IG GARCH model of CHJ is still deterministically related to volatility and both undergo return shocks.

Existing GARCH and SV models are univariate and do not have a straightforward generalization to multiple returns and multiple volatility components without loosing their
main advantage. They also focus on explaining time-series characteristics of returns and loose interest on the cross-sectional dimension. As argued at the beginning, financial assets are likely to be affected by same economic risk factors. Then, time series properties of the factors combined with asset’s systematic risk and idiosyncratic characteristics, will have important implications for the time series of asset returns. Our model belongs to the discrete-time affine class, features both stochastic volatility and skewed return innovations (SVS model), and appropriately takes part into multiple assets and multiple factors. The affine property of the model allows for a closed-form derivation of asset’s risk premium and option prices under no arbitrage. We derive the risk-neutral version of our conditional APT model and show that asset’s risk premium and option prices are also function of asset’s beta, idiosyncratic volatility and idiosyncratic skewness. The model then allows for a direct analysis of the sensitivity of an individual asset’s option prices to asset’s beta, idiosyncratic volatility and idiosyncratic skewness. The affine property of the model also leads to a GMM estimation based on exact moment conditions (see Jiang and Knight (2002) for the case of continuous-time processes, and Feunou and Tédongap (2008) for the discrete-time setting). We distinguish agent and econometrician information sets in our SV setting and provide explicit GARCH counterparts of volatility, conditional skewness and leverage effects.

Harvey and Siddique (1999, hereinafter HS) also consider a nonsymmetric conditional distribution of return with volatility and skewness as two separate factors which follow GARCH-type processes. Their autoregressive conditional skewness is a simple way to model conditional asymmetry and provides an easy methodology to estimate time-varying conditional skewness because of the availability of the likelihood function. However, the non-affinity of their model is a practical drawback, for example for solving option pricing models. The price of a European call option does not exist in closed-form, as opposed to affine GARCH models previously cited. Then, solving such a price would involve numerical methods or simulation techniques which are time-consuming. Our model can also be viewed as a convenient affine alternative to autoregressive conditional skewness, where skewness and volatility are affine combinations of the same factors. We assume that factors follow a multivariate autoregressive gamma process and that idiosyncratic noises are combinations of inverse gaussian shocks which variance and skewness are also functions of the factors. In consequence, all conditional moments of returns are affine combinations of the factors, with coefficients given by cross-sectional characteristics of the asset. Interestingly, our discrete-time conditional APT model has several continuous-time limits, including affine diffusion models with jumps with stochastic intensities.

We apply the GMM procedure suggested by Feunou and Tédongap (2008) to estimate a single factor univariate SVS model using several equity and index daily returns. Because
we only use asset returns at this stage, this corresponds to the historical dynamics. This estimation technique permits a direct evaluation of the model performance in replicating well-known stylized facts as the persistence of volatility through the autocorrelation of squared returns as shown in Figure 2, the negative correlation between returns and future squared returns as shown in Figure 3, the positive correlation between returns and cubed returns, especially for small stocks, as shown in Figure 4. We apply the unscented Kalman filter to estimate cumulants of the factors conditional on observable returns, as they are necessary to evaluate GARCH counterparts of volatility and conditional skewness. We further estimate the single factor and the two-factor SVS models using index daily option data. This corresponds to the risk-neutral dynamics. We test for a specification that allows for contemporaneous asymmetry, and also for a specification with contemporaneous normality. We compare the SVS model performance to the GARCH(1,1) model of Heston and Nandi (2000) and the IG GARCH model of Christoffersen, Heston and Jacobs (2006).

Fitting the historical dynamics, model parameters are significantly estimated and model implications are striking. We find that contemporaneous asymmetry is positive, and this result is robust across all assets under consideration. Contemporaneous asymmetry is particularly important to match sample return skewness, as well as negative and significant cross-correlations between returns and squared returns. When contemporaneous normality is allowed, unconditional skewness is not matched. We also find that the HN GARCH and the IG GARCH models have the same performance as the SVS with contemporaneous asymmetry in matching significant return moments, but only when cross-correlations between return and cubed returns are not important. The SVS model with contemporaneous asymmetry performs better in matching significant cross-correlations between return and cubed returns in addition to other relevant moments of returns. The positive contemporaneous asymmetry is the SVS model dominates negative components of the conditional skewness, and leads to a positive historical conditional skewness, although unconditional skewness is negative and well matched. However, when contemporaneous normality is allowed, conditional skewness becomes negative, consistent with the IG GARCH model of CHJ. However, the model does not match unconditional skewness and short-term leverage effects, and tends to be rejected at conventional level of significance.

Fitting the risk-neutral dynamics using option data, we find that, explicitly allowing for contemporaneous asymmetry leads to substantial gains in option pricing, compared to existing GARCH models with equal or superior number of parameters. The single factor SVS model with contemporaneous asymmetry performs well in-sample, compared to the HN GARCH and the IG GARCH models. The two-factor SVS model has the best in-sample performance, which is not surprising since it nests the single factor SVS model and provides
more flexibility in conditional skewness modeling. Contemporaneous asymmetry is negative and this also is not surprising since a more negative risk-neutral conditional skewness is needed to capture strong biases in short-term options. Empirical evidence show that in-the-money call prices are relatively high compared to the Black-Scholes price, a stylized fact often represented by the well-known “volatility smirk”. Our results suggest that all SVS models outperform the HN GARCH and the IG GARCH in fitting the actual Black-Scholes implied volatility for in-the-money and deep-in-the-money calls, when looking into short-maturity contracts (less than three months).

The rest of the chapter is organized as follows. Section 2 presents the general affine multivariate latent factor model of asset returns. Section 3 introduces our discrete-time SVS model, discusses continuous-time limits, derives GARCH counterparts of volatility and skewness, and discuss the filtering method. Section 4 presents assets risk-neutral valuation and derive the closed-form option pricing formula consistent with SVS model. Section 5 estimates univariate SVS, SV and GARCH models using several equity and index daily returns and provides comparisons and diagnostics. Section 6 estimates univariate SVS, SV and GARCH models using index daily option data and provides comparisons and diagnostics. Section 7 concludes. The appendix contains technical material and proofs.

2 Discrete-Time Affine Models

2.1 Definition and Overview

We consider a discrete-time affine multivariate latent factor model of returns with time-varying conditional moments, characterized by its conditional cumulant generating function:

\[ \Psi_t(x, y; \theta) = \ln E_t \left[ \exp \left( x^\top r_{t+1} + y^\top l_{t+1} \right) \right] = A(x, y; \theta) + B(x, y; \theta)^\top l_t, \]  

where \( E_t[\cdot] \equiv E[\cdot | I_t] \) denotes the expectation conditional to a well-specified information set \( I_t, r_t = (r_{1t}, \ldots, r_{Nt})^\top \) is the vector of observable returns, \( l_t = (l_{1t}, \ldots, l_{Kt})^\top \) is the vector of latent factors and \( \theta \) is the vector of parameters.\(^1\) Notice that the conditional moment generating is exponentially linear in the latent variables \( l_t \) but not necessarily in the observed returns \( r_t \). The vector process \( (r_t^\top, l_t^\top)^\top \) is then semi-affine in the sense of Bates (2006). The conditional cumulant generating function of a fully affine process would be

\(^1\)Darolles, Gourieroux and Jasiak (2006) study in details conditions for the stationarity in distribution of vector affine processes. The vector process \( (r_t^\top, l_t^\top)^\top \) is stationary in distribution if the conditional moment-generating function \( E_t \left[ \exp \left( x^\top r_{t+\tau} + y^\top l_{t+\tau} \right) \right] \) converges to the unconditional moment-generating function \( E \left[ \exp \left( x^\top r_{t+\tau} + y^\top l_t \right) \right] \) as \( \tau \) approaches infinity.
also linear in $r_t$. In all what follows, the parameter $\theta$ is withdrawn from functions $A$ and $B$ for expository purposes.

In practice, such processes are specified through the joint dynamics of observable returns $r$ and latent factors $l$, from which the cumulant generating function (2.1) obtains. In general, all conditional moments of returns are affine functions of the latent factors. In particular, a latent factor $l_i$ itself can be a specific conditional return moment, equivalent to the fact that derivatives of the functions $A(x, y)$ and $B_i(x, y)$ also satisfy specific conditions. Proposition 2.1 below gives necessary and sufficient conditions under which the latent factor $l_i$ is the conditional variance or the conditional asymmetry of the return $r_j$.

**Proposition 2.1** The factor $l_i$ is the conditional variance of returns $r_j$ if and only if

$$\left. \frac{\partial^2 A(x, y)}{\partial x_j^2} \right|_{x=0, y=0} = 0$$

and

$$\left. \frac{\partial^2 B_k(x, y)}{\partial x_j^2} \right|_{x=0, y=0} = 1_{\{k=i\}}. \quad (2.2)$$

The factor $l_i$ is the central conditional third moment of returns $r_j$ if and only if

$$\left. \frac{\partial^3 A(x, y)}{\partial x_j^3} \right|_{x=0, y=0} = 0$$

and

$$\left. \frac{\partial^3 B_k(x, y)}{\partial x_j^3} \right|_{x=0, y=0} = 1_{\{k=i\}}. \quad (2.3)$$

Especially, affine models of the form (2.1) with a single return and a single latent factor corresponding to the conditional variance have been widely studied in the literature as GARCH and stochastic volatility models. An extensive review of this literature is given in Shephard (2005). Example 2.1 below lists most common affine GARCH and SV models with great success in the literature.

**Example 2.1** Stochastic Volatility.

Discrete-time semi-affine univariate latent factor models of returns considered in several empirical studies, are the following stochastic volatility models. Return dynamics is given by:

$$r_{t+1} = \mu_r - \lambda_h \mu_h + \lambda_h h_t + \sqrt{h_t} u_{t+1} \quad (2.4)$$

where the volatility process satisfies one of the followings:

$$h_{t+1} = (1 - \phi_h) \mu_h - \alpha_h + (\phi_h - \alpha_h \beta_h^2) h_t + \alpha_h \left( \varepsilon_{t+1} - \beta_h \sqrt{h_t} \right)^2, \quad (2.5)$$

$$h_{t+1} = (1 - \phi_h) \mu_h + \phi_h h_t + \sigma_h \varepsilon_{t+1}, \quad (2.6)$$

$$h_{t+1} = (1 - \phi_h) \mu_h + \phi_h h_t + \sigma_h \sqrt{h_t} \varepsilon_{t+1}, \quad (2.7)$$

and where $u_{t+1}$ and $\varepsilon_{t+1}$ are two i.i.d standard normal shocks. The parameter vector $\theta$ is $(\mu_r, \lambda_h, \mu_h, \phi_h, \alpha_h, \beta_h, \rho_{rh})^\top$ with the volatility dynamics (2.5) whereas it is $(\mu_r, \lambda_h, \mu_h, \phi_h, \sigma_h)^\top$ with the autoregressive gaussian volatility (2.6) and $(\mu_r, \lambda_h, \mu_h, \phi_h, \sigma_h, \rho_{rh})^\top$ with the square-root volatility (2.7), where $\rho_{rh}$ denotes the conditional correlation between the shocks $u_{t+1}$.
and \( \varepsilon_{t+1} \). The particular case \( \rho_{rh} = 1 \) in the volatility dynamics (2.5) corresponds to the Heston and Nandi (2000)’s GARCH. For this reason we refer to the dynamics (2.5) as HN-S volatility. Christoffersen, Heston and Jacobs (2006) also study an affine GARCH model specified by:

\[
\begin{align*}
    r_{t+1} &= \alpha_h + \lambda_h h_t + \eta_h y_{t+1} \\
    h_{t+1} &= w_h + b_h h_t + c_h y_{t+1} + a_h \frac{h_t^2}{y_{t+1}}
\end{align*}
\]

(2.8) (2.9)

where, given the available information at time \( t \), \( y_{t+1} \) has an inverse gaussian conditional distribution with degrees of freedom parameter \( h_t / \eta_h^2 \). As in the original paper, we refer to this specification as IG GARCH.

The \( A \) and \( B \) functions characterizing the cumulant generating functions for these GARCH and SV models are explicitly given in Appendix A. One should notice that the volatility processes (2.6) and (2.7) are not well defined since \( h_t \) can take negative values. In simulations, one should be careful when using a reflecting barrier at a small positive number to ensure positivity of simulated volatility samples.\(^2\) This can also arise with the process (2.5) unless parameters satisfy a couple of constraints. Note also that if the volatility shock \( \varepsilon_{t+1} \) in (2.6) is allowed to be correlated to the return shock \( u_{t+1} \) in (2.4), then the model becomes non-affine. The HN-S and the IG GARCH specifications will be examined in more details in the empirical part.

A known case of a well-defined affine stochastic volatility model assumes that \( h_t \) follows an autoregressive gamma process (see Gourieroux and Jasiak (2001) for more details). However, when combined with the return process (2.4), the model presumes that within a period, return and volatility shocks are mutually independent, what appears to be a counterfactual assumption against the well-documented conditional leverage effect (Black(1976) and Christie (1982)). As discussed above, the autoregressive gaussian dynamics (2.6), coupled with the return equation (2.4) cannot allow for leverage effect without the model losing its affine property. This counterfactual assumption is not required for classical SV models (Taylor (1986), Andersen (1994)) and GARCH models (Bollerslev (1986), Nelson (1991), Engle and Ng (1993)). However, these latter models are less tractable in empirical studies because of their non affine property. Then, there has always been a trade-off between tractable affine models with counterfactual assumptions and non-tractable non-affine models that do not require these assumptions. In this paper, we aim at combining both the affine model and the ability of a SV model to take into account important features of the data (fat-tailedness, asymmetry and leverage effect) in a coherent way.

\(^2\)Because of this limitation, autoregressive gaussian and squared-root stochastic volatility models have been mainly explored in continuous time. To avoid negative values of \( h_t \) in simulations for examples, one relies on the true dynamics of \( \ln h_t \) using the Itô lemma and works through the logarithmic model.
2.2 Modeling Conditional Skewness and Leverage Effect in Affine SV Models

While return models of Example 2.1 are such that the vector \((r_{t+1}, h_{t+1})\) of returns and volatility is affine, the conditional skewness of returns in these models is zero (only the IG GARCH is an exception and we will come back into this in subsequent sections). The literature on asset return models has evolved so far and empirical evidence upon path dependence of conditional skewness as well as its importance and contribution to risk management and asset pricing rose in recent studies. Higher moments, and especially skewness, are implicitly priced in nonlinear asset pricing models (Bansal and Viswanathan (1993), Bansal, Hsieh and Viswanathan (1993), Harvey and Siddique (2000)). HS show that conditional skewness is time-varying and significant in asset prices, and that it impacts the persistence in conditional variance. In their original paper, CHJ highlight the fact that, while specification (2.4) creates negative conditional skewness in multi-period returns when combined with volatility dynamics (2.7) and (2.5) for example, single-period innovations remain gaussian in these models, and the models cannot explain the strong biases in short-term options. The necessity to model return skewness has become of first order importance. HS model conditional skewness as a GARCH process and the IG GARCH model in Example 2.1 restricts conditional skewness to be deterministically related to volatility \((s_t = 3\eta_h/\sqrt{h_t})\). Liesenfeld and Jung (2000) introduce SV models with conditional heavy tails. However, SV models with conditional asymmetry have received less attention so far. We depart from previous literature by allowing skewness, as well as other higher order moments, to undergo unobservable shocks, which in general can be uncorrelated or linearly independent to returns and volatility shocks. Most importantly, we keep the affine property of the overall system, with a straightforward generalization to a cross-section of returns. In this section, we explain our approach for accounting for both conditional skewness and leverage effect in a general affine univariate SV model. It is the same modeling technique we use in next section for our model.

Existing affine SV models basically lead to a couple of equations of the form:

\[
\begin{align*}
    r_{t+1} &= e(h_t) + \sqrt{h_t}u_{t+1} \\
    h_{t+1} &= m(h_t) + \sqrt{v(h_t)}\varepsilon_{t+1}
\end{align*}
\]

where \(u_{t+1}\) and \(\varepsilon_{t+1}\) are two errors with mean zero and unit variance. Written in this form, the conditional skewness of returns is zero unless \(u_{t+1}\) is conditionally asymmetric. These models do not allow for the leverage effect unless the shocks \(u_{t+1}\) and \(\varepsilon_{t+1}\) are correlated. However, it is generally assumed that \(u_{t+1}\) is gaussian and therefore “unfeasible” to assume a conditional correlation when at least one of the shocks is non-gaussian. This is a potential limitation that typically arises when \(u_{t+1}\) is gaussian and equation (2.11) is such that \(h_t\)
is an autoregressive gamma process. Since the leverage effect is the nonzero conditional covariance between returns and volatility, projecting $r_{t+1}$ onto $h_{t+1}$ should lead to a nonzero slope coefficient. Therefore, we suggest to account for skewness and leverage effect in asset returns by projecting returns $r_{t+1}$ onto volatility $h_{t+1}$ and characterizing the projection error. This will basically lead to a return equation of the form:

$$ r_{t+1} = g(h_t) + \beta h_{t+1} + \sqrt{h_t - \beta^2 v(h_t)} u_{t+1} $$

(2.12)

where $u_{t+1}$ is an error with mean zero and unit variance. One could still endow $u_{t+1}$ with a suitable distribution conditional on $\langle h_{t+1}, I_t \rangle$ such that combining (2.11) with (2.12) leads to an affine stochastic volatility model of asset returns. The model will now account for the leverage effect through $\beta$. The conditional skewness will also depend on $\beta$ as well as on the asymmetry of the shock $u_{t+1}$ conditional on $\langle h_{t+1}, I_t \rangle$, if any. We refer to the asymmetry of observable returns conditional on current factors and past information as the contemporaneous asymmetry.

It is more easier to think to a semi-affine one-factor SV model as in Example 2.1, with a directly specified equation for volatility dynamics, precisely because of tractable properties of the standard normal distribution appearing in both return and volatility shocks. However, it is more challenging to think to a semi-affine two-factor model with stochastic skewness as additional factor, such that both equations for volatility and skewness dynamics are directly specified. The reason is that, while conditional asymmetry of returns appears to be a necessary and sufficient condition to generate time-variation in conditional skewness, asymmetric distributions are not as tractable as the normal distribution. A strategy to get equations which explicitly characterize the joint dynamics of returns, volatility and skewness would be to first specify a semi-affine two-factor model with arbitrary linearly independent latent factors, more easier to think at, and:

- find volatility and conditional skewness in terms of the two arbitrary factors,
- then, invert the previous relationship to determine the two arbitrary factors in terms of volatility and skewness,
- and finally, replace the arbitrary factors in the initial return model to get the joint dynamics of returns, volatility and skewness.

In the next section, we develop a semi-affine multivariate latent factor model of returns such that both conditional variance $h_t$ and conditional skewness $s_t$ are stochastic. Moreover, the vector $\left( r_{t+1}, h_{t+1}, s_{t+1}h_{t+1}^{3/2} \right)^\top$ is affine in the case of a single return and two linearly independent latent factors.
3  An Affine Multivariate Latent Factor Model with Stochastic Skewness

3.1  General Setup

The dynamics of returns in our model is built upon shocks drawn from a standardized inverse gaussian distribution. The cumulant generating function of a discrete random variable which follows a standardized inverse gaussian distribution of parameter \(s\), denoted \(SIG(s)\), is given by:

\[
\psi(u; s) = \ln E[\exp(uX)] = -3s^{-1}u + 9s^{-2}\left(1 - \sqrt{1 - \frac{2}{3}su}\right). \quad (3.1)
\]

For such a random variable, one has \(E[X] = 0, E[X^2] = 1\) and \(E[X^3] = s\), meaning that \(s\) is the skewness of \(X\). In addition to the fact that the \(SIG\) distribution is directly parameterized by its skewness, the limiting distribution when the skewness \(s\) tends to zero is the standard normal distribution, that is \(SIG(0) \equiv \mathcal{N}(0,1)\). This particularity makes the \(SIG\) an ideal building block for studying departures from normality.

For each variable in all what follows, the time subscript denotes the date from which the value of the variable is observed by the economic agent. We assume that components of the vector \(r_t\) of \(N\) returns on financial assets follow the dynamics:

\[
r_{j,t+1} = \ln \frac{S_{j,t+1}}{S_{j,t}} = \mu_{j0} + \sum_{i=1}^{K} \lambda_{ji} (\sigma_{i,t}^2 - \mu_i) + \sum_{i=1}^{K} \beta_{ji} (\sigma_{i,t+1}^2 - \mu_i) + \sum_{i=1}^{K} \gamma_{ji}\sigma_{i,t+1}u_{ji,t+1}
\]

(3.2)

where \(S_{jt}\) is the price of the \(j^{\text{th}}\) asset and \(u_{ji,t+1} = \langle \sigma_{i,t+1}^2, I_t \rangle \sim SIG(\eta_{ji}(\gamma_{ji}\sigma_{i,t+1})^{-1})\). The components of the latent vector \(\sigma_t^2\) are \(K\) linearly independent positive factors driving all returns’ dynamics. For identification, we impose \(\gamma_{ii} = 1, \forall i\). The \(NK\) return shocks \(u_{ji,t+1}\) are mutually independent conditionally on \(\langle \sigma_{i,t+1}^2, I_t \rangle\). If \(\eta_{ji} = 0\), then \(u_{ji,t+1}\) is a standard normal shock. The time \(t\) information set \(I_t\) contains past realizations of returns \(r_t = \{r_t, r_{t-1}, \ldots\}\) and latent factors \(\sigma_t = \{\sigma_t^2, \sigma_{t-1}^2, \ldots\}\). The return dynamics (3.2) can also be written in vector forms:

\[
r_{j,t+1} = \delta_j + \beta_j^\top\sigma_{t+1}^2 + \sigma_{t+1}^\top(\gamma_j u_{j,t+1}) \quad \text{or} \quad r_{t+1} = \delta_t + \beta_t^\top\sigma_{t+1}^2 + (\gamma u_{t+1})^\top\sigma_{t+1}
\]

(3.3)

where \(\delta_j = \mu_{j0} - (\lambda_j + \beta_j)^\top\mu + \lambda_j^\top\sigma_t^2\) and \(\delta_t = \mu_0 - (\lambda + \beta)^\top\mu + \lambda^\top\sigma_t^2\). The vector \(\mu\) is the unconditional mean of the stationary process \(\sigma_t^2\). In consequence \(\mu_0\) is the vector of unconditional expected returns. \(\lambda, \beta\) and \(\eta\) are \(K \times N\) matrices such that \(\lambda^\top = [\lambda_{ji}], \beta^\top = [\beta_{ji}]\) and \(\eta^\top = [\eta_{ji}]\), and \(\lambda_j, \beta_j\) and \(\eta_j\) are the \(j^{\text{th}}\) column of the matrices \(\lambda, \beta\) and \(\eta\) respectively. Similarly, \(\gamma u_{t+1}\) is the \(K \times N\) matrix process such that \((\gamma u_{t+1})^\top = [\gamma_{ji}u_{ji,t+1}]\) and \(\gamma_{ji}u_{ji,t+1}\) represents the \(j^{\text{th}}\) column of \(\gamma u_{t+1}\).
Under previous assumptions on $u_{t+1}$, the cumulant generating function of returns conditional to $\langle \sigma^2_{t+1}, I_t \rangle$ is given by:

$$\ln E \left[ \exp \left( x^T r_{t+1} \right) \mid \sigma^2_{t+1}, I_t \right] = x^T \delta_t + \sum_{i=1}^{K} \sum_{j=1}^{N} \left( \beta_{ji} x_j + \psi \left( x_j; \eta_{ji} \right) \gamma^2_{ji} \right) \sigma^2_{i,t+1}.$$  \hspace{1cm} (3.4)

The process $\sigma^2_t$ is assumed to be affine with the conditional cumulant generating function

$$\Psi^\sigma_t (y) = \ln E \left[ \exp \left( y^T \sigma^2_{t+1} \right) \mid I_t \right] = a \left( y \right) + b \left( y \right)^T \sigma^2_t.$$  \hspace{1cm} (3.5)

In this case, the vector $\left( r^T_{t+1}, \left( \sigma^2_{t+1} \right)^T \right)^T$ is semi-affine in the sense of Bates (2006). Its conditional cumulant generating function is given by:

$$\Psi_t (x, y) = \ln E \left[ \exp \left( x^T r_{t+1} + y^T \sigma^2_{t+1} \right) \mid I_t \right] = A \left( x, y \right) + B \left( x, y \right)^T \sigma^2_t,$$

with

$$A \left( x, y \right) = \left( \mu_0 - (\lambda + \beta)^\top \mu \right)^\top x + a \left( f \left( x, y \right) \right)$$  \hspace{1cm} (3.6)

$$B \left( x, y \right) = \lambda x + b \left( f \left( x, y \right) \right)$$  \hspace{1cm} (3.7)

where $f \left( x, y \right) = \left( f_1 \left( x, y_1 \right), \ldots, f_K \left( x, y_K \right) \right)^\top$ with $f_i \left( x, y_i \right) = y_i + \sum_{j=1}^{N} \left( \beta_{ji} x_j + \psi \left( x_j; \eta_{ji} \right) \gamma^2_{ji} \right).$

Since the factors are positive, we assume that the vector $\sigma^2_t$ follows a multivariate autoregressive gamma process. This process also represents the discrete-time counterpart to continuous-time multivariate square root processes that have previously been examined in the literature.\(^3\) Its log conditional Laplace transform has the form (3.5) with:

$$a \left( y \right) = - \sum_{i=1}^{K} \nu_i \ln \left( 1 - \alpha_i y_i \right) \quad \text{and} \quad b_i \left( y \right) = \sum_{j=1}^{K} \phi_{ij} y_j \left( 1 - \alpha_j y_j \right).$$  \hspace{1cm} (3.8)

The $K \times K$ matrix $\Phi = [\phi_{ij}]$ represents the persistence matrix of the vector $\sigma^2_t$ and the autoregressive gamma processes $\sigma^2_{it}$ are mutually correlated if the off-diagonal elements of $\Phi$ are nonzero. The factors are mutually independent conditional on $I_t$ if the off-diagonal elements are zero. In this latter case we note $\phi_i = \phi_{ii}$. In the single factor case, the factor $\sigma^2_{it}$ has the conditional cumulant generating function $\psi^\sigma_{it} (y_1) = a \left( y_1 \right) + b_1 \left( y_1 \right) \sigma^2_{it}$, where $a \left( y_1 \right) = - \nu_1 \ln \left( 1 - \alpha_1 y_1 \right)$ and $b_1 \left( y_1 \right) = \phi_1 y_1 \left( 1 - \alpha_1 y_1 \right).$ The parameter $\phi_1$ is the persistence of the factor and the parameters $\nu_1$ and $\alpha_1$ are related to persistence and unconditional mean $\mu_1$ and variance $\omega_1$ as $\nu_1 = \mu_1^2 / \omega_1$ and $\alpha_1 = (1 - \phi_1) \omega_1 / \mu_1$.

\(^3\)See for example Singleton (2001).
Although our empirical focus in this article will be on the time series dynamics of a single return, it is important to notice that equation (3.3) is a multifactor conditional arbitrage-pricing model. In fact, we assume that a true conditional multifactor representation of expected returns in the cross-section is such that log returns are linear in the factors and the idiosyncratic noise. The vector $\beta_j$ represents the loadings of asset $j$ on the factors, and this asset’s conditional beta is time-invariant. The factors are heteroscedastic and the idiosyncratic noise is a combination of independent heteroscedastic and asymmetric shocks. This constitutes a substantial depart from previous literature, as the true data generating process in existing APT models is, in general, specified such that factors as well as idiosyncratic shocks are implicitly or explicitly homoscedastic and normally distributed. Considering latent factors is also appealing as, in the original APT model of Ross (1976), factors are unknown. Also, focusing on positive factors is not restrictive as any arbitrary economic factor, say $F_t$, can be written as a difference of two nonnegative factors, say $\sigma^2_{1t} - \sigma^2_{2t}$, where $\sigma^2_{1t} = \max(F_t, 0)$ and $\sigma^2_{2t} = \max(-F_t, 0)$.

### 3.2 Expected Returns, Volatility, Conditional Skewness and Leverage Effects

In the previous section, we do not model directly volatility and conditional skewness as well as other higher moments of returns. Instead, we relate returns to a finite number of stochastic linearly independent positive factors. In this section, we relate expected returns, volatility, conditional skewness and leverage effects to these factors and discuss important features of the model.

**Proposition 3.1** Conditional on $I_t$, the mean $\mu^r_j$, the variance $h_j$ and the skewness $s_j$ of returns $r_j$ are expressed as follows:

\[
\mu^r_{jt} = \mu_{j0} - (\lambda_j + \beta_j) \mu + \lambda_j \sigma^2_t + \beta_j m^\sigma_t = c_{j0,\mu} + \sum_{i=1}^{K} c_{ji,\mu} \sigma^2_{it} \tag{3.9}
\]

\[
h_{jt} = \beta_j^\top V^\sigma_t \beta_j + (\gamma^2_{j}) \top m^\sigma_t = c_{j0,h} + \sum_{i=1}^{K} c_{ji,h} \sigma^2_{it} \tag{3.10}
\]

\[
s_{jt} h^{3/2}_{jt} = (\beta_j \otimes \beta_j) \top S^\sigma_t \beta_j + 3 (\gamma^2_{j}) \top V^\sigma_t \beta_j + (\gamma^2_{j}) \top m^\sigma_t = c_{j0,s} + \sum_{i=1}^{K} c_{ji,s} \sigma^2_{it} \tag{3.11}
\]

where the coefficients $c_{jn,t}$ depend on model parameters,

\[
m^\sigma_t = E \left[ \sigma^2_{t+1} \mid I_t \right], \quad V^\sigma_t = E \left[ \left( \sigma^2_{t+1} - m^\sigma_t \right) \left( \sigma^2_{t+1} - m^\sigma_t \right)^\top \mid I_t \right]
\]

and

\[
S^\sigma_t = E \left[ \left( \left( \sigma^2_{t+1} - m^\sigma_t \right) \otimes \left( \sigma^2_{t+1} - m^\sigma_t \right) \right) \left( \sigma^2_{t+1} - m^\sigma_t \right)^\top \mid I_t \right].
\]
The linearity of expected returns, volatility and conditional asymmetry of returns in terms of the factors results from the fact that components of the vector $m_t^\sigma$, and of the matrices $V_t^\sigma$ and $S_t^\sigma$ are also linear in terms of the $\sigma_i^2$’s. This is a consequence of the affine structure of the process $\sigma_t^2$. Also, note that the bivariate vector $(h_{jt}, s_j h_{jt}^{3/2})^\top$ is not deterministically related to contemporaneous and past returns as for GARCH-type processes as in Harvey and Siddique (1999) and Feunou and Tédongap (2009), as well as many other authors.\footnote{Hansen (1994), Jondeau and Rockinger (2003), and Leon, Rubio and Serna (2004), do not explicitly model conditional skewness, but related shape parameters of the conditional return distribution using GARCH-type dynamics.}

For this the reason, we label the present model, stochastic volatility and skewness (SVS).

**Proposition 3.2** Conditional on $I_t$, the covariance between returns $r_j$ and volatility $h_j$ (leverage effect) and the covariance between returns $r_j$ and skewness $s_jh_{jt}^{3/2}$ are given by:

\[
\operatorname{Cov}(r_{j,t+1}, h_{j,t+1} \mid I_t) = c_{j,h}^\top V_t^\sigma \beta_j = c_{j0,rh} + \sum_{i=1}^K c_{ji,rh} \sigma_i^2 = c_{j0,rh} + c_{j,rh}^\top \sigma_i^2 
\] (3.12)

\[
\operatorname{Cov}(r_{j,t+1}, s_{j,t+1} h_{jt+1}^{3/2} \mid I_t) = c_{j,s}^\top V_t^\sigma \beta_j = c_{j0,rs} + \sum_{i=1}^K c_{ji,rs} \sigma_i^2 = c_{j0,rs} + c_{j,rs}^\top \sigma_i^2 
\] (3.13)

where the coefficients $c_{jn,rl}$ depend on model parameters.

It is not surprising that the parameter $\beta_j$ governs conditional leverage effect as it represents the slope of linear projection of returns on current factors. For a negative correlation between spot returns and variance, and consistently with the postulate of Black (1976) and the leverage effect documented by Christie (1982) and others, the parameter $\beta_j$ may be expected to be negative, in particular for the single-factor case.

It should be noted that, in our SVS model, although the parameter $\eta_j$ dictates contemporaneous asymmetry of returns (that is, the asymmetry of returns conditional on current factors and past information), it is not the only parameter determining conditional skewness as shown in equation (3.11). The parameter $\beta_j$, which alone characterizes leverage effect, also plays a central role in generating conditional asymmetry in returns, even if returns are contemporaneously normally distributed, that is when $\eta_j = 0$. In contrast to existing SV models with leverage effect as discussed in Example 2.1, where leverage effect generates skewness only in multiple-period returns, in our setting, leverage effect invokes skewness in single-period returns as well. If $\beta_j = 0$, there is no leverage effect. In addition, there is also no skewness unless $\eta_j \neq 0$. Then, contemporaneous asymmetry in this model reinforces the effect of the leverage parameter $\beta_j$ in generating conditional skewness. In other words, time-varying conditional skewness in this model is a combination of conditional leverage effect (through $\beta_j$) and contemporaneous asymmetry (through $\eta_j$).
To better understand the flexibility of the SVS model in generating conditional skewness, we refer to the single-factor SVS. Equation (3.11) shows that conditional skewness is the sum of three terms. The first two terms have the same sign, which is the sign of $\beta_j$ as components of the matrices $V_\sigma^t$ and $S_\sigma^t$ are positive. The last term has the sign of $\eta_j$ as $m_t^\sigma$ is positive. As discussed previously, a negative value of $\beta_j$ is necessary to generate the documented negative leverage effect. If so, the first two terms in (3.11) are negative. The sign of conditional skewness will then depend on contemporaneous asymmetry $\eta_j$. If $\eta_j$ is zero or negative, then conditional skewness is negative over time as in the IG GARCH model. This also arises if $\eta_j$ is positive, but not enough that the third term dominates the first two. If it does, then conditional skewness is positive over time. Also remark that skewness of the $j^{th}$ financial asset may change sign over time if $\eta_j$ is positive and such that $c_{j0,s}c_{j1,s} < 0$. There are lower and upper positive bounds on $\eta_j$ such that this latter condition holds. This will then be consistent with the empirical evidence in Harvey and Siddique (1999) that conditional skewness changes sign over time. Feunou and Tédongap (2009) findings also suggest that, although conditional skewness is centered around a negative value, return innovations are conditionally normal or weakly positively skewed most of the time, but undergo unfrequent and large drops in conditional skewness. However, it is recognizes in the literature that a negative conditional skewness is particularly important for explaining strong biases in option prices.

While $\sigma_1^2 t, ..., \sigma_K^2 t$ are the primitive predictive variables in our SVS model, predictability when $K \geq 2$ can also be directly related to conditional variance and skewness which are economically interpretable. For example, empirical facts tend to support that an increase in volatility drives up expected returns, as people require more premium when it becomes more riskier to invest in stocks. As well as people dislike high return volatility, they prefer positive skewness (extreme positive returns are more likely to realize than extreme negative returns). Therefore, people would pay a premium in exchange of positive skewness, and require a premium to compensate for negative skewness. In the two-factor case, $K = 2$, and if $c_{j1,h}c_{j2,s} \neq c_{j1,s}c_{j2,h}$ without loss of generality, one can invert relations (3.10) and (3.11) to obtain $\sigma_1^2 t$ and $\sigma_2^2 t$ in terms of $h_{jt}$ and $s_{jt}h_{jt}^{3/2}$. Using inverted relations in (3.9) expresses expected returns in terms of volatility and skewness, economically meaningful, instead of arbitrary factors. The IG GARCH does not separate skewness from volatility whereas the two-factor SVS disentangles these two measures while maintaining a semi-affine structure of the model. This separation results from the decomposition of return shocks into two linearly independent IG components with individual conditional variances having specific affine dynamics.
3.3 Continuous-Time Limits

Although the present SVS model is written in discrete time and easily applicable to discrete data, we are interested in its continuous-time versions. Following several papers which derive continuous-time limits of discrete-time processes (Nelson (1991), Foster and Nelson (1994), and others), we write the model for a small time interval, and let the time interval shrink to zero. For a small time interval \( \Delta \), the return equation (3.2) becomes:

\[
\ln \frac{S_{j,t+\Delta}}{S_{j,t}} = \mu_{j0} + \sum_{i=1}^{K} \lambda_{ji} (\sigma_{i,t}^2 - \mu_i) + \sum_{i=1}^{K} \beta_{ji} (\sigma_{i,t+\Delta}^2 - \mu_i) + \sum_{i=1}^{K} \gamma_{ji} \sigma_{i,t+\Delta} u_{ji,t+\Delta}.
\]

For simplicity we assume that the factors are independent. Let

\[
\mu_{j0}(\Delta) = \mu_{j0} \Delta, \quad \beta_{ji}(\Delta) = \frac{\beta_{ji}}{\Delta}, \quad \lambda_{ji}(\Delta) = \lambda_{ji} - \beta_{ji} \exp\left(-\frac{\kappa_{ji} \Delta}{\Delta}\right)
\]

\[
\phi_i(\Delta) = \exp\left(-\kappa_{i} \Delta\right), \quad \alpha_i(\Delta) = \frac{\omega_i \Delta^2}{2 \exp(-\kappa_{i} \Delta)}, \quad \nu_i(\Delta) = \theta_i (1 - \exp(-\kappa_{i} \Delta)) \frac{2 \exp(-\kappa_{i} \Delta)}{\omega_i \Delta}.
\]

Letting \( v_{it} = \sigma_{i,t}^2 / \Delta \) represent factors per unit time, it follows that:

\[
dv_{it} = \kappa_i (\theta_i - v_{it}) \, dt + \sqrt{\omega_i} \sqrt{v_{it}} dB_{it},
\]

where \( dB_{it} \) is a Wiener process. It can also be established that the term \( \sigma_{i,t+\Delta} \sqrt{\Delta} u_{ji,t+\Delta} \) has two different continuous time limits depending on the value of the parameter \( \eta_{ji} \). If \( \eta_{ji} = 0 \) then \( \sigma_{i,t+\Delta} \sqrt{\Delta} u_{ji,t+\Delta} \) converges to \( \sqrt{v_{it}} dW_{ji,t} \) as \( \Delta \) shrinks to zero, where \( W_{ji,t} \) is a Wiener process. To the contrary if \( \eta_{ji} \neq 0 \), then \( \sigma_{i,t+\Delta} \sqrt{\Delta} u_{ji,t+\Delta} \) converges to \( -(3 \gamma_{ji} v_{it} / \eta_{ji}) \, dt + (\eta_{ji} / 3 \gamma_{ji}) dJ_{ji,t} \) as \( \Delta \) shrinks to zero, where \( J_{ji,t} \) is a pure jump inverse gaussian process with degree of freedom \( 9 \gamma_{ji}^2 v_{it} / \eta_{ji}^2 \) on interval \([t, t + dt]\). We then show that the limiting distribution of the SVS model in continuous time is a stochastic volatility process where the return is a sum of diffusion and pure jump inverse gaussian processes:

\[
d \ln S_{jt} = \left[ \mu_{j0} + \sum_{i=1}^{K} \lambda_{ji} (v_{it} - \theta_i) - \sum_{i: \eta_{ji} \neq 0} \frac{3 \gamma_{ji}^2 v_{it}}{\eta_{ji}} \right] \, dt + \sum_{i=1}^{K} \beta_{ji} \sqrt{\omega_i} \sqrt{v_{it}} dB_{it}
\]

\[
+ \sum_{i: \eta_{ji} = 0} \gamma_{ji} \sqrt{v_{it}} dW_{ji,t} + \sum_{i: \eta_{ji} \neq 0} \frac{\eta_{ji}}{3} dJ_{ji,t}.
\]

3.4 GARCH versus SVS: Filtering the Unobservable Factors

In GARCH models, the information set \( I_t \) is exactly \( r_t \) so that both the economic agent and the econometrician view the same information. This is a strong assumption that is implicit in GARCH models. In SV models in general, and the present SVS model in particular, the
econometrician does not observe $\sigma_t^2$, only known by the economic agent. While the moments in Proposition 3.1 are conditional on information $I_t = r_t \cup \sigma_t^2$, one can also derive their GARCH counterparts, meaning same return moments now conditional on econometrician’s information, $r_t$ only. Without loss of generality, we derive these conditional moments for the case of a single return ($N = 1$). However, the formulas can be generalized to multiple returns as well. Let $\mu_t^G$, $h_t^G$ and $s_t^G$ respectively denote the mean, the variance and the skewness of $r_{t+1}$ conditional on $r_t$. One has:

$$\mu_t^G = c_{0\mu} + c_{\mu}^T G_{\mu}$$

and

$$h_t^G = c_{0h} + c_h^T G_{\mu} + c_{\mu}^T G_{ht} c_{\mu},$$

$$s_t^G \left( h_t^G \right)^{3/2} = c_{0s} + c_s^T G_{\mu} + c_{\mu}^T G_{ht} c_{h} + \left( c_{\mu} \otimes c_{\mu} \right)^T G_{st} c_{\mu}$$

where

$$G_{\mu} = E \left[ \sigma_t^2 \mid r_t \right] \quad \text{and} \quad G_{ht} = E \left[ \sigma_t^2 \left( \sigma_t^2 \right)^T \mid r_t \right] - E \left[ \sigma_t^2 \mid r_t \right] E \left[ \sigma_t^2 \mid r_t \right]^T,$$

$$G_{st} = E \left[ \left( \sigma_t^2 \otimes \sigma_t^2 \right) \left( \sigma_t^2 \right)^T \mid r_t \right] - 3 E \left[ \left( \sigma_t^2 \otimes \sigma_t^2 \right) \mid r_t \right] E \left[ \sigma_t^2 \mid r_t \right]^T + 2 \left( E \left[ \sigma_t^2 \mid r_t \right] \otimes E \left[ \sigma_t^2 \mid r_t \right] \right) E \left[ \sigma_t^2 \mid r_t \right]^T.$$

are mean, variance and third central moment of the latent vector $\sigma_t^2$ conditional upon observed returns $r_t$.

Disentangling agent and econometrician information sets in return modeling can be crucial. In the single-factor SVS model, return conditional variance and third central moment are perfectly correlated to the agent, whereas it is the contrary to the econometrician, unless returns are unpredictable by the factor ($c_{\mu} = 0$). Otherwise ($c_{\mu} \neq 0$), the SVS model generates, conditional to observed returns, an asymmetry that is not perfectly correlated to the variance, although this correlation remains high for a persistent factor. In contrast, conditional variance and third central moment are perfectly correlated in the IG GARCH, given past observed returns.

GARCH counterparts of leverage effect and of conditional covariance between returns and skewness are defined by:

$$Cov \left( r_{t+1}, h_{t+1}^G \mid r_t \right) \quad \text{and} \quad Cov \left( r_{t+1}, s_{t+1}^G \left( h_{t+1}^G \right)^{3/2} \mid r_t \right).$$

These two quantities are difficult to express in terms of the moments of the latent vector $\sigma_t^2$ conditional on observed returns $r_t$, and instead, we consider the following two quantities:

$$Cov \left( r_{t+1}, h_{t+1} \mid r_t \right) = c_{0,rh} + c_{rh}^T G_{\mu} + c_{\mu}^T G_{ht} \Phi^T c_h$$

$$Cov \left( r_{t+1}, s_{t+1} h_{t+1}^{3/2} \mid r_t \right) = c_{0,rs} + c_{rs}^T G_{\mu} + c_{\mu}^T G_{ht} \Phi^T c_s.$$
where $\Phi$ represents the persistence matrix of the latent vector.

We now describe how to compute expectations in (3.19) and (3.20). Various strategies to deal with non-linear state-space systems have been proposed in the filtering literature: the Extended Kalman Filter, the Particle Filter and more recently the Unscented Kalman Filter that we apply in this paper.\footnote{See Leippold and Wû (2003) and Bakshi, Carr and Wû (2005) for application in finance, Julier et al. (1995) and Jullier and Uhlmann (1996) for details and Wan and van der Merwe (2001) for textbook treatment.} Since our SVS model has the standard state space representation, one can use Kalman Filter-based techniques to compute $G_t$, $G_{ht}$ and $G_{st}$.

As these methods will not guarantee that $E\left[\sigma^2_{it}\mid r_t\right]$ is positive, it would be more convenient to filter $\omega_t = (\omega_{1t}, ..., \omega_{kt})^\top$. Let $\omega_t = (\omega_{1t}^*, ..., \omega_{kt}^*)^\top$.

The basic framework of Kalman filter techniques involves estimation of the state of a discrete-time nonlinear dynamic system of the form:

$$r_{t+1} = H(\omega_{t+1}, u_{t+1}^*)$$
$$\omega_{t+1} = F(\omega_t, \varepsilon_{t+1}^*)$$

(3.23)

(3.24)

where $u_{t+1}^*$ and $\varepsilon_{t+1}^*$ are not necessarily but conventionally two gaussian noises. For this reason, we log-normally approximate our model, which in the one-factor case leads to:

$$H(\omega_{1,t+1}, u_{1,t+1}^*) = \mu_0 + \beta_1 \exp(\omega_{1,t+1}) + \exp\left(\frac{\omega_{1,t+1}}{2}\right)\left[\exp\left(\ln\left(\frac{9}{s(\omega_{1,t+1})^2 + 9}\right)\right) + \sqrt{\ln\left(\frac{9}{s(\omega_{1,t+1})^2 + 9}\right)\frac{u_{1,t+1}^*}{s(\omega_{1,t+1})} - \frac{3}{s(\omega_{1,t+1})}}\right]$$

(3.25)

and

$$F(\omega_{1t}, \varepsilon_{1,t+1}^*) = \ln\left(\frac{m(\omega_{1t})}{\sqrt{m(\omega_{1t})^2 + v(\omega_{1t})}}\right) + \ln\left(\frac{m(\omega_{1t})^2 + v(\omega_{1t})}{m(\omega_{1t})^2} \varepsilon_{1,t+1}^*\right)$$

(3.26)

where

$$s(\omega_{1,t+1}) = \eta_1 \exp\left(-\frac{\omega_{1,t+1}}{2}\right)$$

$$m(\omega_{1t}) = (1 - \phi_1) \mu_1 + \phi_1 \exp(\omega_{1t})$$

$$v(\omega_{1t}) = (1 - \phi_1)^2 \sigma_1^2 + \frac{2(1 - \phi_1) \phi_1 \sigma_1^2}{\mu_1} \exp(\omega_{1t})$$

Details on this log-normal approximation of the one-factor SVS model are provided in appendix C.

Let $\omega_{1\tau}$ be the estimate of $\omega_t$ using returns up to and including time $\tau$, $r_{\tau\tau}$, and let $P_{\tau\tau}^{\omega\omega}$ be its covariance. Given the joint distribution of $(\omega_{t}^\top, u_{t+1}^\top, \varepsilon_{t+1}^\top)^\top$ conditionally to $r_{\tau\tau}$,
the filter predicts what future state and returns will be using process models. Optimal predictions and associated mean squared errors are given by:

\[
\omega_{t+1|t} = E \left[ \omega_{t+1} \mid r_t \right] = E \left[ F (\omega_t, \varepsilon_{t+1}^*) \mid r_t \right] \tag{3.27}
\]

\[
r_{t+1|t} = E \left[ r_{t+1} \mid r_t \right] = E \left[ H (\omega_{t+1}, \varepsilon_{t+1}^*) \mid r_t \right] \tag{3.28}
\]

\[
P_{t+1|t} = E \left[ (\omega_{t+1} - \omega_{t+1|t}) (\omega_{t+1} - \omega_{t+1|t})^\top \mid r_t \right] \tag{3.29}
\]

\[
P_{t+1|t} = E \left[ (r_{t+1} - r_{t+1|t}) (r_{t+1} - r_{t+1|t})^\top \mid r_t \right] \tag{3.30}
\]

\[
P_{t+1|t} = E \left[ (\omega_{t+1} - \omega_{t+1|t}) (r_{t+1} - r_{t+1|t})^\top \mid r_t \right]. \tag{3.31}
\]

The join distribution of \((\omega_t^\top, u_{t+1}^\top, \varepsilon_{t+1}^\top)\) conditionally to \(r_t\) is conventionally assumed gaussian. To the contrary of the standard Kalman filter where the functions \(H\) and \(F\) are linear, the precise values of the conditional moments (3.27) to (3.31) cannot be determined analytically in our model because the functions \(H\) and \(F\) are strongly nonlinear. Alternative methods produce approximations of these conditional moments.

The Extended Kalman Filter linearizes the functionals \(H\) and \(F\) in the state-space system to determine the conditional moments analytically. While this simple linearization maintains a first-order accuracy, it can introduce large errors in the true posterior mean and covariance of the transformed random variable which may lead to sub-optimal performance and sometimes to divergence of the filter. The Particle Filter uses Monte-Carlo simulations of the relevant distributions to get estimates of moments. In contrast, the Unscented Kalman Filter addresses the approximation issues of the Extended Kalman filter and the computational issues of the Particle Filter. It represents the distribution of \((\omega_t^\top, u_{t+1}^\top, \varepsilon_{t+1}^\top)\) conditional on \(r_t\) by a minimal set of carefully chosen points. This reduces the computational burden but maintain second-order accuracy. Details on the Unscented Kalman Filter are provided in appendix E.

The next step is to use current returns to update estimate (3.27) of the state. In the Kalman filter, a linear update rule is specified, where the weights are chosen to minimize the mean squared error of the estimate. This rule is given by:

\[
\omega_{t+1|t+1} = \omega_{t+1|t} + K_{t+1} (r_{t+1} - r_{t+1|t}) \tag{3.32}
\]

\[
P_{t+1|t+1} = P_{t+1|t} - K_{t+1} P_{t+1|t} K_{t+1}^\top \tag{3.33}
\]

\[
K_{t+1} = P_{t+1|t} (P_{t+1|t}^{-1} \tag{3.34}
\]

Once the Kalman recursion outlined above delivers the estimates \(\omega_{tt}\) and \(P_{tt}\) for the whole sample, the statistics \(G_{tt}\), \(G_{ht}\) and \(G_{st}\) can be computed using approximations of moments of a nonlinear function of a gaussian random variable. Without loss of generality, appendix D derives corresponding formulas in the univariate case.

4 Asset Pricing with Stochastic Skewness

In the context of asset and derivative pricing, one would like to find a probability measure under which the expected gross return on any risky security equals the gross return on a safe
security. It is sufficient to define a change of measure \( Z_{t,t+1} \) from historical to risk-neutral, or equivalently to defining a stochastic discount factor \( M_{t,t+1} \) from which investors value financial payoffs (see Gourieroux and Monfort (2006) and Christoffersen et al. (2006)). The change of measure \( Z_{t,t+1} \) satisfies the following conditions:

\[
E[Z_{t,t+1} \mid I_t] = 1 \quad \text{and} \quad E^* [\exp (r_{j,t+1}) \mid I_t] \equiv E[Z_{t,t+1} \exp (r_{j,t+1}) \mid I_t] = \exp (r_{f,t+1}).
\]  

(4.1)

where \( r_{j,t+1} \) and \( r_{f,t+1} \) refer to the \( j^{th} \) risky return and the risk-free rate from date \( t \) to date \( t + 1 \), respectively, and where \( E^* [\cdot \mid I_t] \equiv E[Z_{t,t+1} (\cdot) \mid I_t] \) denotes the risk-neutral expectation associated with the density \( Z_{t,t+1} \).

Given the historical return dynamics (3.2), we would like to find a change of measure such that risk-neutral return dynamics is also an affine SVS model similar to (3.2). Exploiting the affine property, we assume that the change of measure \( Z_{t,t+1} \) is given by:

\[
Z_{t,t+1} = \exp \left(-A (\kappa, \pi) - B (\kappa, \pi) \mathbf{\hat{\sigma}}_t^2 + \kappa^\top r_{t+1} + \pi^\top \mathbf{\hat{\sigma}}_{t+1}^2 \right) \quad \text{and} \quad E[Z_{t,t+1} \mid \langle \mathbf{\hat{\sigma}}_{t+1}^2, I_t \rangle] = 1.
\]  

(4.2)

In particular, an implication of the second equation is that the moment generating function of \( \mathbf{\hat{\sigma}}_{t+1}^2 \), conditional to \( I_t \), does not change from the physical to the risk-neutral measure. Thus, the factors still follow the same multivariate autoregressive gamma under the risk-neutral dynamics. Appendix B finally shows that the risk-neutral dynamics of returns is given by:

\[
r_{j,t+1} = r_f - a \left( q^2 \theta^*_j \right) - \sum_{i=1}^{K} b_i \left( q^2 \theta^*_j \right) \mathbf{\hat{\sigma}}_t^2 + \sum_{i=1}^{K} (\beta_{j,i} q^2 \theta^*_j) \mathbf{\hat{\sigma}}_{t+1}^2 + \sum_{i=1}^{K} \gamma_{j,i} \mathbf{\hat{\sigma}}_{t+1}^2 u^*_{j,i,t+1}
\]  

(4.4)

with \( u^*_{j,i,t+1} \mid \langle \mathbf{\hat{\sigma}}_{t+1}^2, I_t \rangle \sim SIG \left( \eta^*_j, \gamma_{j,i}^2 \sigma_{i,t+1} \right) \), and where \( q^2 \theta^*_j \) denotes the \( K \times 1 \) vector with components \( q^2 \theta^*_j \). Risk-neutral parameters are defined by:

\[
q^2_{ji} = (1 - (2/3) \eta_{ji} \kappa_j)^{-3/2} \quad \text{and} \quad \theta^*_{ji} = \beta^*_{ji} = \psi (1; \eta^*_j) \gamma^2_{ji}
\]

\[
\beta^*_{ji} = (\beta_{ji} + \psi (k_j; \eta_j) \gamma^2_{ji}) / q^2_{ji}, \quad \eta^*_j = \eta_j / (1 - (2/3) \eta_{ji} \kappa_j) \quad \text{and} \quad \gamma^*_{ji} = q_{ji} \gamma_{ji}.
\]

The return dynamics (3.3) and the no-arbitrage restrictions (4.1) lead to the characterization of the asset’s risk premium, which in our model is given by:

\[
\mu_{0,j} - r_f = -a \left( q^2 \theta^*_j \right) + \left( \beta_{j} - b \left( q^2 \theta^*_j \right) \right) ^\top \mu.
\]  

(4.5)

In most of empirical studies, ingredients of the return dynamics that are important for explaining actual time series properties of returns, and which turn to be relevant also for
explaining characteristics of observed option prices (for example leverage effects and conditional skewness), are studied separately from features that relate to actual cross-sectional properties of asset returns. We argue that time series and cross-sectional properties of returns result from the same features, and that these features should not be model independently. We see for example that, if factors are heteroscedastic and idiosyncratic shocks are heteroscedastic and asymmetric, as in our model, leverage effects are determined by asset’s factor loadings ($\beta_j$), and conditional skewness is determined by both factor loadings and idiosyncratic skewness ($\eta_j$). In addition, no-arbitrage equilibrium restrictions imply that asset’s risk premium depends both on factor loadings, idiosyncratic volatility ($\gamma_j$) and idiosyncratic skewness. Our model offers a tractable framework to address simultaneously time series and cross-sectional properties of asset returns as well as of asset’s option prices.

Assuming that factors are independent ($\phi_{ij} = 0$ for $i \neq j$), it is convenient to write the risk-neutral return dynamics as:

$$r_{1,t+1} = r_f - a^*(\theta^*_1) - \sum_{i=1}^{K} b^*_i (\theta^*_i) \sigma^*_it^2 + \sum_{i=1}^{K} \beta^*_i \sigma^*_it_{i,t+1} + \sum_{i=1}^{K} \sigma^*_it_{i,t+1} u^*_{i,t+1} \quad (4.6)$$

$$r_{j,t+1} = r_f - a^*(\theta^*_j) - \sum_{i=1}^{K} b^*_i (\theta^*_j) \sigma^*_it^2 + \sum_{i=1}^{K} \beta^*_j \sigma^*_it_{i,t+1} + \sum_{i=1}^{K} \sigma^*_it_{i,t+1} u^*_{j,i,t+1}, \quad 2 \leq j \leq N$$

where $\sigma^*_it = q_{ii} \sigma_{it}$ with $q_{ii}^2 = (1 - (2/3) \eta_{ii} \kappa_i)^{-3/2}$, $u^*_{i,t+1} \sim SIG(\eta_{ii}^*, \sigma_{i,t+1}^{-1})$ and $u^*_{j,i,t+1} \sim SIG(\gamma_{ji}^*, \sigma_{i,t+1}^* \gamma_{ji}^{-1})$. The vector process $\sigma^*_it^2$ is a multivariate autoregressive gamma under the risk-neutral measure, with parameters of the $i^{th}$ factor, $\sigma^*_it^2$, given by:

$$\alpha^*_i = q_{ii}^2 \alpha_i, \quad \nu^*_i = \nu_i \quad \text{and} \quad \phi^*_i = \phi_i.$$  

Parameters in the first risk-neutral return equation ($j = 1$) are given by:

$$\beta^*_ii = (\beta_{ii} + \psi^*(1; \eta_{ii}^*)) / q_{ii}^2, \quad \eta^*_i = \eta_i / (1 - (2/3) \eta_{ii} \kappa_i) \quad \text{and} \quad \theta^*_ii = \beta^*_ii + \psi (1; \eta_{ii}^*) .$$

The functions $a^*(\cdot)$ and $b^*(\cdot)$ are analogue to the functions $a(\cdot)$ and $b(\cdot)$ in (3.5), and similarly characterize the cumulant generating function of the multivariate autoregressive gamma process $\sigma^*_it^2$ under the risk-neutral dynamics. Parameters in the second risk-neutral return equation ($2 \leq j \leq N$) are given by:

$$\beta^*_jj = \beta^*_j q_{jj}^2 / q_{ii}^2, \quad \eta^*_j = \eta_j / (1 - (2/3) \eta_{jj} \kappa_j) \quad \text{and} \quad \theta^*_jj = \beta^*_jj + \psi (1; \eta_{jj}^*) \gamma_{jj}^2.$$  

Because $\theta^*_i$ is related to $\beta^*_i$ and $\eta^*_i$, and $\theta^*_j$ is related to $\beta^*_j$, $\gamma^*_j$ and $\eta^*_j$ for $2 \leq j \leq N$, the risk-neutral dynamics of every asset has $K$ parameters less compared to its historical dynamics. This is analogous to the IG GARCH risk-neutral model of CHJ (which is a single-factor model) where the parameter governing conditional skewness is a function of

---

A similarly argument can be found in Santos and Veronezi (2008). The authors argue that the equity premium puzzle and the value premium puzzle cannot be tackled independently, as any economic mechanism proposed to address one of them immediately has general equilibrium implications for the other.
other parameters. This result is in fact more general, and, as we show in this article, for a $K$ factor model, the risk-neutral dynamics of an individual asset return has $K$ independent parameters less compared to the physical model.

In the single return case ($N = 1$), the joint dynamics of returns and factors under the risk-neutral distribution is characterized by the following cumulant generating function:

$$
\Psi^*_t (x, y) = \ln E^* \left[ \exp \left( x r_{t+1} + y \sigma_t^2 \right) \ | \ I_t \right] = A^*_r (x, y) + B^*_r (x, y) \sigma_t^2
$$

where the functions $A^*(\cdot, \cdot)$ and $B^*(\cdot, \cdot)$ are analogue to the functions $A(\cdot, \cdot)$ and $B(\cdot, \cdot)$ in (3.6) and (3.7) respectively. Let $\Psi^*_t (x)$ denotes the conditional log moment generating function of aggregate returns $\sum_{i=1}^{\tau} r_{t+i}$, under the risk-neutral measure. One has

$$
E^* \left[ \exp \left( x \sum_{i=1}^{\tau} r_{t+i} \right) \ | \ I_t \right] = \exp \left( \Psi^*_t (x) \right) = \exp \left( A^*_r (x; \tau) + B^*_r (x; \tau) \sigma_t^2 \right),
$$

where the sequence of functions $A^*_r (x; \tau)$ and $B^*_r (x; \tau)$ satisfy the following recursion:

$$
A^*_r (x; \tau) = A^*_r (x; \tau - 1) + A^* (x, B^*_r (x; \tau - 1)) \quad \text{and} \quad B^*_r (x; \tau) = B^* (x, B^*_r (x; \tau - 1))
$$

with $A^*_r (x; 1) = A^* (x, 0)$ and $B^*_r (x; 1) = B^* (x, 0)$.

The price at date $t$ of a European call option with strike price $X$ and maturity $\tau$, is given by

$$
C \left( \tau, \sigma_t^2, \frac{X}{S_t} \right) = \exp (-r \tau) \left[ \frac{1}{2} \exp (r \tau) + C_1 \left( \tau, \sigma_t^2, \frac{X}{S_t} \right) - X \left( \frac{1}{2} + C_2 \left( \tau, \sigma_t^2, \frac{X}{S_t} \right) \right) \right]
$$

where

$$
C_1 \left( \tau, \sigma_t^2, \frac{X}{S_t} \right) = \int_0^{+\infty} \frac{1}{\pi u} \text{Im} \left[ \exp \left( A^* (1 + iu; \tau) + B^* (1 + iu; \tau) \sigma_t^2 - iu \ln \left( \frac{X}{S_t} \right) \right) \right] \ du
$$

$$
C_2 \left( \tau, \sigma_t^2, \frac{X}{S_t} \right) = \int_0^{+\infty} \frac{1}{\pi u} \text{Im} \left[ \exp \left( A^* (iu; \tau) + B^* (iu; \tau) \sigma_t^2 - iu \ln \left( \frac{X}{S_t} \right) \right) \right] \ du.
$$

5 Estimation and Comparison of Affine SVS, SV and GARCH Models Using Equity and Index Daily Returns

5.1 Estimation Methodology and Data

Return unconditional moments can be computed in closed-form in a discrete-time affine multivariate latent factor model, as shown in Feunou and Tédongap (2008). All these moments are functions of the parameter vector $\theta$ that governs both returns and factors dynamics. We can then choose $M$ pertinent moments to perform GMM estimation of the model. Assuming a single return, we choose $M$ moments of the form $\mu_{r,j} (n, m) = E \left[ r_t^n r_{t+j}^m \right]$ such that $1 \leq j \leq J$, $0 \leq n \leq Q$ and $0 < m \leq Q - n$, that means $M$ among $Q + JQ (Q - 1) / 2$ moments of order less than or equal to $Q$. Since moments of observed returns implied by a given model can directly be compared to their sample equivalent, our
estimation setup is more likely to evaluate the performance of a given model in replicating well-known stylized facts like autocorrelation of squared returns, absence of autocorrelation of returns, leverage effect which can be captured via autokoskewness, unconditional fat-tailness and asymmetry of returns. Model performance in replicating these empirical facts is assessed by including corresponding moments.

Let \( g_t(\theta) = \left[ r_t^{n_i} r_t^{m_i} - \mu_{r,j_i}(n_i,m_i) \right]_{1 \leq i \leq M} \) denotes the \( M \times 1 \) vector of retained moments. We have \( E[g_t(\theta)] = 0 \) and we define the sample counterpart of this moment condition as follows:

\[
\hat{g}(\theta) = \left( \begin{array}{c}
\hat{E}[r_t^{n_1} r_t^{m_1} - \mu_{r,j_1}(n_1,m_1)] \\
\vdots \\
\hat{E}[r_t^{n_M} r_t^{m_M} - \mu_{r,j_M}(n_M,m_M)]
\end{array} \right),
\]

(5.1)

Given the \( M \times M \) weighting matrix \( \hat{W} \), the GMM estimator \( \hat{\theta} \) of the parameter vector is given by:

\[
\hat{\theta} = \arg \min_{\theta} \hat{g}(\theta)^\top \hat{W} \hat{g}(\theta).
\]

(5.2)

Interestingly, the variance-covariance matrix of \( g_t(\theta) \) does not depend on the vector of parameter \( \theta \). This is a huge advantage since with a nonparametric empirical variance-covariance matrix of moment conditions, the optimal GMM procedure is readily implemented in one step. In addition, for two different models estimated via same moment conditions and weighting matrix, the minimum value of the GMM objective function itself is a criterion for comparison of alternative models.

In some cases, this GMM procedure also has a huge numerical advantage compared to the maximum likelihood estimation even when the likelihood function can be derived. Maximum likelihood estimation becomes difficult to perform numerically especially when the support of the likelihood function is parameter-dependent. This is the case in the IG-GARCH model of Christoffersen, Heston and Jacobs (2006) which can also be estimated through this GMM method. On the other hand, the maximum likelihood estimation of semi-affine latent variable models of Bates (2006) and the quasi-maximum likelihood estimation based on the Kalman recursion have the downside that critical unconditional higher moments (skewness and kurtosis) of returns can be poorly estimated due to the second order approximation of the distribution of the latent variable conditional on observable returns. Moreover, in single-stage estimation and filtering methods like the Unscented Kalman Filter and the Bates (2006)’s algorithm, one can argue that approximations affect both parameter and state estimations. Instead, our GMM procedure matches critical higher moments exactly and requires no approximation for parameter estimation. Provided with the GMM estimates of model parameters, Bates (2006)’s procedure or any other filtering procedure like the Unscented Kalman Filter can be followed for the state estimation. In this sense, approximations required by these techniques will only affect state estimation.

We estimate the single-factor SVS, the HN-S volatility and the IG GARCH models using daily returns on S&P500 and CRSP indexes, as well as daily returns on the six Fama and French size and book-to-market sorted portfolios. As explained in Fama and French
(1993), the six portfolios are the outcome of the intersection of two independent sorts. Stocks are sorted into two size groups—\(S\) (small; that is, market capitalization below the NYSE median) and \(B\) (big; that is, market capitalization above the NYSE median)—and into three book-to-market groups—\(G\) (growth; that is, in the bottom 30 percent of the NYSE book-to-market), \(N\) (neutral; that is, in the middle 40 percent of the NYSE book-to-market) and \(V\) (value; that is, in the top 30 percent of the NYSE book-to-market). The six portfolios are commonly labelled \(SG\), \(SN\), \(SV\), \(BG\), \(BN\) and \(BV\). Table 1 summarizes basic descriptive statistics of these returns. It shows the well-documented facts that asset returns are negatively skewed and fat-tailed. Small stocks are generally more negatively skewed than big stocks and a growth portfolio has lower average returns and higher negative skewness compared to a value portfolio of the same size.

### 5.2 Parameter Estimation

To perform the GMM procedure for each series, we need to decide which moments to choose. To achieve this task, we refer to the relative importance of return moments. We consider the moments

\[
\left\{ E\left[r^j_t\right]\right\}_{j\geq 1}
\]

in order to match the critical first moments of asset returns. Indeed, we do not estimate the unconditional mean of returns \(\mu_0\), which we set to its sample counterpart. Figure 2 displays autocorrelations of square returns which as shown are significant up to the twentieth lag. As the positive and significant autocorrelation of square returns appears to be a critical empirical fact, we consider the moments

\[
\left\{ E\left[r^2_tr^2_{t+j}\right]\right\}_{j\geq 1}
\]

in order to match these autocorrelations. For this set, we choose up to five leads for all stocks under consideration. To assess the ability of our SVS models to match significant autocoskewness and autocokurtosis, we add the set of moments

\[
\left\{ E\left[r_tr^2_{t+j}\right], E\left[r_t^3\right]\right\}_{j\geq 1}.
\]

The negative and significant cross-correlation between returns and square returns for various leads as shown in Panel A of Figure 3 is an empirical fact characterizing the well-known leverage effect. We choose up to five leads for small stocks and up to three leads for big stocks and market indexes. As shown in Panel B of Figure 3, similar cross-correlations for various lags are not significant. The cross-correlation between returns and cube returns is positive and significant, at least for the first three leads as shown in Panel A of Figure 4, especially for small stocks. Panel B of Figure 4 shows that similar cross-correlations for various lags are not significant.

We weigh moments using the diagonal of the inverse of their covariance matrix:

\[
\hat{W} = Diag \left\{ \left( \text{Var} \left[ g_t \right] \right)^{-1} \right\}.
\]

This matrix is nonparametric and puts more weight on moments with low magnitude. Estimation results for one-factor SVS models are shown in Table 2 for small stocks, in Table
3 for big stocks and in Table 4 for market indexes, both for contemporaneous asymmetry, contemporaneous normality, as well as alternative SV and GARCH models. For single factor SVS models, the parameter $\lambda_1$ is not estimated. The reason is that, due to the high expected persistence of the factor, it would be difficult in the return equation (3.2) to identify $\lambda_1$ and $\beta_1$ separately. To avoid this identification problem, we set $\lambda_1 = 0$.

We first focus on the first panel of Tables 2, 3 and 4, for estimation results in the context of contemporaneous asymmetry of returns, that is, $\eta_1 \neq 0$ is estimated. Starting with the measure equation (3.2), estimation output confirms that projecting returns onto the latent factor results in a significant and negative coefficient $\beta_1$, then corroborating the story that an increase in contemporaneous volatility lowers asset payoffs. The coefficient $\eta_1$ in the return equation is significant and positive and this result is robust across all stocks under consideration. This suggests that the distribution of daily returns conditional upon contemporaneous volatility is asymmetric. This result contrasts with the findings of Bollerslev and Forsberg (2002) that daily returns are normal conditional upon current realized volatility. Coming to the state dynamics, estimation results show that the factor governing daily return dynamics is highly persistent, with significant estimates of the coefficient of persistence, 0.963 and 0.948 for S&P500 and CRSP indexes respectively. This also means that daily return volatility and conditional asymmetry as perceived by agents are highly persistent as well, since they are linear in the factor. All estimates for the single factor SVS with contemporaneous asymmetry are significant. In addition, the $J$-test of over-identifying restrictions does not reject the model, but on small value stocks.

We now assess how important is contemporaneous asymmetry for asset return modeling. The second panel of Tables 2, 3 and 4 shows estimation results in the context of contemporaneous normality of returns, that is, with the constraint $\eta_1 = 0$. As for contemporaneous asymmetry, all parameters are significantly estimated and, in comparison, there is a decrease in the magnitude of the leverage parameter and an increase in the persistence of the factors—estimate of the persistence even becomes unrealistically greater or equal to 1 for some assets under consideration as we do not explicitly impose a restriction on this parameter in our estimation. For estimation results with a realistic persistence of factor, models are or tend to be rejected in the data. The maximum $p$-value for the $J$-test of over-identifying restrictions is 0.05. The sharp decrease in GMM criterion from contemporaneous normality to contemporaneous asymmetry also suggests that tests favor the latter compared to the former. The GMM criterion falls from 94.72 to 34.12 for CRSP index, and from 101.87 to 24.06 for S&P500 index.

The third panel of Tables 2, 3 and 4 shows estimation results for the HN-S volatility specification discussed in Section 2.1, and the fourth panel of Table 4 shows estimation results for the IG GARCH specification also discussed in Section 2.1. For big stocks and market indexes, HN-S volatility and IG GARCH specifications are comparable to the single factor SVS with contemporaneous asymmetry. Notice that, cokurtosis moments are not significant for these stocks and therefore not included for the GMM estimation. Instead, cokurtosis moments are significant for small stocks and, when included for the GMM estimation, results show that the single factor SVS model is preferred to the HN-S volatility specification.

For small stocks, Table 6 compares model unconditional moments to their sample counterparts across different models. Tables 7 and 8 show similar comparisons for big stocks.
and market indexes. Tables show ratios of model unconditional moments to their sample counterparts. The closer to one is the ratio, the better the model matches the moment. Mean, variance and kurtosis are perfectly matched by all models and this is robust across all stocks. It is also the case for autocorrelations of squared returns. A straightforward remark is how accurate the model with $\eta_1 \neq 0$ matches selected moments better than the model with $\eta_1 = 0$. In particular, Table 6 shows that skewness (moment 3) is not well matched by the model with contemporaneous normality, and, as shown in Tables 7 and 8, this matching is the worst when autocokurtosis is not significant. Contemporaneous normality matches autocoskewness better at long horizons ($j > 2$) than at short horizons ($j \leq 2$), while it is the contrary for contemporaneous asymmetry.

Finally, as we mentioned previously, the choice of the moments to be used in the GMM procedure is crucial when intended to reproduce important empirical facts. While the cross-correlation between returns and cubed returns is in general not significant for big stocks and market indexes, these moments are not matched by the GMM estimates when selected for estimation, except for the first lead where it appears weakly significant for some of these stocks. However, for small stocks, this moment is significant empirically as shown in Panel A of Figure 4 for the three first leads, and Table 6 shows that the GMM estimates reproduce the moments as well. Next, we filter the latent factors using the GMM estimates of parameters.

5.3 State Estimation

We use the Unscented Kalman Filter algorithm with our GMM estimates to filter the latent factor $\sigma_t^2$ that we use to compute GARCH counterparts of volatility and conditional skewness, i.e $h_t^G$ and $s_t^G$. Figure 5 displays the time series of GARCH counterparts of volatility and conditional skewness for the CRSP and the S&P500 indexes, for contemporaneous asymmetry ($\eta_1 \neq 0$) as well as for contemporaneous normality ($\eta_1 = 0$). Asset returns in our sample as plotted in Figure 1, are characterized by moderately high volatility at the beginning of the sample (1990-1992), followed by low volatility (1993-1996), then high volatility (1997-2003) and low or moderately high volatility at the end of the sample (2004-2005). This volatility pattern is well-matched by the volatility time series plotted in the first and the second rows of Figure 5. Also notice the slightly difference between volatility time series in different columns of the figure, due to the effect of contemporaneous asymmetry. Volatility is more persistent for contemporaneous normality.

The third and the fourth rows of the figure show the pattern of the GARCH counterpart of conditional skewness. Overall results are striking. As shown in the figure, conditional skewness is negative for contemporaneous normality, and this is consistent with the IG-GARCH model of Christoffersen, Heston and Jacobs (2006). We also recall that critical unconditional third order moments of returns, skewness and leverage effects, are not well-matched by GMM estimates under contemporaneous normality. In contrast, if contemporaneous asymmetry is allowed, we find that GMM estimates match unconditional skewness and leverage effects very well and, in this case, Figure 5 shows that conditional skewness is positive, and its mean has a larger magnitude compared to the contemporaneous normality case. Figure 6 confirms that these results hold for individual portfolios as well.
6 Estimation and Comparison of Affine SVS, SV and GARCH Models Using Index Option Prices

We conduct our empirical analysis using 9 years of data on S&P500 index call options. We use option data for Wednesday only in the the period from 1996 to 2004. If Wednesday is a holiday, we use the next trading day. Using only Wednesday data allows us to study a long time series, which is useful considering the highly persistent factors. Besides, using Wednesday is common practice in the literature to limit the impact of holidays and day-of-the-week effects (see Heston and Nandi (2000), Christoffersen and Jacobs (2004, 2006)).

Table 9 presents the number of contracts used by moneyness (Panel (a)) and maturity (Panel (b)), and provides a cross-tabulation across moneyness and maturity (Panel (c)). From panel (a), we can observed a skewed Black-Scholes IV curve which is materialized by a high in-the-money call option implied volatility, compared to out-of-the-money. This suggests a necessity to model risk-neutral skewness in a flexible way. This pattern differs substantially across maturity groups. Short maturity options display an asymmetric smile pattern with high deep out-of-the-money implied volatility. This pattern reverses gradually when maturity increases to finally yield a smirked curve with high deep in-the-money implied volatilities.

Implied volatility show little variation across maturitie s, but a variable pattern across moneyness classes. For deep out-of-the-money call, implied volatility decreases with maturity, while it shows a smile-shape for out-of-the-money. For in-the-money, implied volatility increases with maturity while it shows an asymmetric smile pattern for deep in-the-money. Later, we evaluate empirically the ability of different models to replicate these observed patterns. Smirked implied volatility patterns for short maturities suggest a skewed one-step ahead conditional return distribution, while the reversion of this pattern and its persistence for longer maturities suggest more than one factor in risk-neutral conditional return distribution. Conditional skewness controls short-term properties while multiple factors control the long-term.

In this section, we estimate risk-neutral versions of the following models: single factor SVS both with contemporaneous asymmetry (SVS1f) and contemporaneous normality (SVS1f, $\eta = 0$), two-factor SVS with contemporaneous asymmetry (SVS2f), HN GARCH and IG GARCH. One challenge facing with unobservable factors is the joint estimation of risk-neutral parameters and latent factors. Several methods have been used in the literature and can be divided into two categories. The first approach considers latent factors as parameters (Bakshi, Cao and Chen (1997), Bates (2000) Huang and Wu (2004) and Christoffersen Heston and Jacobs (2007)), and the second approach filters latent factors using time series of underlying returns in a Bayesian framework (see Jones(2003) and Eraker (2004)).

6.1 Estimation Methodology

We follow the first approach here described. Without loss of generality we describe the method for an SVS model, as the same approach is applied to others. Consider a sample of $T$ Wednesdays of option data ($T = 463$ corresponds to the number of Wednesdays in our sample). Given starting values for the structural parameter vector $\theta^*$ and the vector $\sigma_t^*$ of
latent factors under the risk-neutral model, the iterative procedure proceeds as follows:

**Step 1:** For a given set of structural parameters, \( \theta \), solve \( T \) sums of squared pricing errors optimization problems of the form:

\[
\hat{\sigma}^*_t = \arg\min_{n=1}^{N_t} \left(C_{nt} - C_n(\theta^*, \sigma^*_t) \right)^2, \quad t = 1, 2, ..., T. \tag{6.1}
\]

where \( C_{nt} \) is the observed price of contract \( n \) on day \( t \) and \( C_n(\theta^*, \sigma^*_t) \) is the corresponding model price. \( N_t \) is the number of contracts available on day \( t \).

**Step 2:** For a given estimated factor \( \hat{\sigma}^*_t \) obtained from Step 1, solve one aggregate sum of squared pricing errors optimization problem of the form:

\[
\hat{\theta}^* = \arg\min_{t=1}^{T} \sum_{n=1}^{N_t} \left(C_{nt} - C_n(\theta^*, \hat{\sigma}^*_t) \right)^2. \tag{6.2}
\]

The procedure iterates between Step 1 and Step 2 until no further significant decreases in the overall objective function in Step 2 are obtained.

### 6.2 Risk-Neutral Volatility and Conditional Skewness

Table 10 shows parameter estimates of risk-neutral models. As shown in the table, all models deliver persistent factors. Even for the two-factor SVS model, the second factor is still very persistent. For the two-factor model, estimates of the two parameters driving contemporaneous asymmetry of returns (\( \eta^*_1 \) and \( \eta^*_2 \)) are negative, but lower in absolute value, compared to the corresponding parameter in the single-factor SVS model. This negative contemporaneous asymmetry contrasts with the positive contemporaneous asymmetry found when estimating the single-factor SVS using return data. Because the conditional risk neutral distribution is highly negatively skewed, a negative contemporaneous asymmetry is needed. Also in the two-factor risk-neutral SVS, the factor associated with the lowest (in absolute value) negative contemporaneous skewness is the most persistent.

Figure 7 shows annualized time series of risk-neutral volatility and conditional skewness for all models. The figure shows high co-movements of volatility and skewness across different risk-neutral models. Meanwhile, the level of volatility increases with the flexibility in conditional skewness modeling. The two-factor SVS model generates the highest level of volatility, then follows the SVS1F, the IG GARCH, the HN GARCH and finally the SVS1F with \( \eta = 0 \). The risk-neutral one-day ahead conditional skewness is the highest with the IG GARCH, while still comparable to the SVS1F, and the lowest in the SVS1F with \( \eta = 0 \). By construction, it is zero in the HN GARCH model. Conditional skewness in the IG GARCH increases (in absolute value) as volatility lowers. Relaxing this link as in other specifications reduces the level of conditional skewness.

### 6.3 Model Diagnostics

For all models, Table 11 shows the relative root mean squared error (RRMSE), that is the root mean squared error (RMSE) divided by the sample mean of call price, across different
moneyness and maturity classes. As expected, The SVS2F model has the best in-sample fit in every moneyness class as well as every maturity class. It provides more flexibility than other models as it has a superior number of parameters/factors, twice the number of parameters/factors in other specifications. Using all available option data, the RRMSE for the SVS2F is 5.8%, followed by 7.8% for the SVS1F, 8.5% for the IG GARCH, 9.2% for the HN GARCH and finally 10.3% for the SVS1F with $\eta = 0$. In this order, we argue that flexibility in conditional asymmetries modeling also reduces option pricing errors. For short It should be noticed that, even if the one-day ahead conditional skewness is zero in the HN GARCH model, it is not the case for the multi-day ahead conditional skewness, due to the leverage effect. A negative one-day ahead conditional skewness reduces pricing errors for short-term option contracts. This intuition is confirmed in CHJ (2006), where the authors find an improvement of the IG GARCH over the HN GARCH on short-maturity options. Our results shed light once more on this, and we argue that flexibility in one-day ahead conditional skewness modeling decreases the RRMSE for short maturities. As shown in Panel (b) of Table 11, for maturities less than one month, the best performance measured by the RRMSE is attributable to the SVS2F (6.2%), then SVS1F models (10.6% and 10.5%), the IG GARCH (11.6%) and the HN model (11.8%). The two-factor model has the best performance along all dimensions as also shown in Panel (a) of Table 11.

We summarize the model relative bias (the bias divided by the average price) in Table 12. Although the bias is generally low for all models under consideration, for maturities less than one month, it is the highest for the IG GARCH, which is comparable to that of the SVS1F, $\eta = 0$, and more than twice the bias for the SVS2F. The HN GARCH does better among single factor models for these maturities, and it is comparable to the SVS2F, although both models bias observed prices in opposite directions. For deep in-the-money call options, the SVS1F bias the least, followed by the IG GARCH, then the SVS2F, the HN GARCH and the SVS1F, $\eta = 0$.

We finally represent, the observed and model’s Black-Scholes implied volatilities along different dimensions. We retain our analysis to in-the-money call options and short-maturity contracts. In Figure 8 we fix the maturity class and represent implied volatility as function of moneyness. Using all available option data, the first panel of Figure 8 shows that all SVS models outperform HN and IG GARCH models in fitting deep in-the-money implied volatility, and have comparable performances for in-the-money implied volatility. For same options but maturities less than one month, the SVS1F and the HN GARCH fits perfectly in-the-money are outperform the IG GARCH deep in-the-money. For maturities between one and two months, the SVS1F, $\eta = 0$ perfectly fits deep in-the-money, while all SVS models outperform GARCH models and have comparable fits in-the-money. The SVS1F maintains its lead over GARCH models for maturities between two and three months, both in-the-money and deep in-the-money.

In Figure 9 we fix moneyness class and represent implied volatility as function of maturity. For short-maturity contracts as shown in the first panel of the figure, the SVS2F outperforms all models at every maturity. The SVS1F and the HN GARCH on one hand, and the SVS1F, $\eta = 0$ and the IG GARCH on the other hand, have comparable fits for maturities less than one month. Coming to in-the-money call options in the fifth panel, the SVS1F and the HN GARCH have comparable fits of the IV curve for maturities less than three months and outperform the IG GARCH. For deep in-the-money call options of
less than three months of maturity, the SVS1F and the SVS2F have the lead over GARCH models in fitting observed Black-Scholes implied volatilities. Finally, these implied volatility curves confirm main analyses and model rankings resulting from RRMSE. The SVS1F, $\eta = 0$ even seems to do better in terms of implied volatility fit.

7 Conclusion and Future Work

This paper presents a new approach for modeling conditional skewness in a discrete-time affine multivariate latent factor model with volatility. The model explicitly allows returns to be asymmetric conditional on current factors and past information. This contemporaneous asymmetry is shown to be particularly important for the model to fit both return and option data. An empirical investigation suggests that the flexibility that the model offers for conditional skewness, increases its option pricing performance relative to existing affine GARCH and SV models. In particular, SVS models with contemporaneous asymmetry outperform existing affine GARCH and SV models especially, but not only, for in-the-money calls and short-maturity contracts.

Although the model is flexible enough to accommodate both multiple returns and multiple factors, our analysis in this paper focuses on the single return case. In a future research, it would be interesting to study the implications of the model for a parsimonious multiple returns setting as well.
A Cumulant Generating Functions of Affine SV and GARCH Models

The $A$ and $B$ functions characterizing the cumulant-generating functions for GARCH and SV models in Example 2.1 are given by:

$$A(x, y) = (\mu_r - \lambda_h \mu_h) x + ((1 - \phi_h) \mu_h - \alpha_h) y - \frac{1}{2} \ln (1 - 2\alpha_h y) \quad (A.1)$$

$$B(x, y) = \lambda_h x + (\phi_h - \alpha_h \beta_h^2) y + \frac{1}{2} x^2 + \frac{\alpha_h y}{1 - 2\alpha_h y} (\beta_h - \rho r_h x)^2 \quad (A.2)$$

for the HN-S specification,

$$A(x, y) = (\mu_r - \lambda_h \mu_h) x + (1 - \phi_h) \mu_h y + \frac{1}{2} \sigma_h^2 y^2 \quad (A.3)$$

$$B(x, y) = \lambda_h x + \phi_h y + \frac{1}{2} x^2 \quad (A.4)$$

for the autoregressive gaussian specification ,

$$A(x, y) = (\mu_r - \lambda_h \mu_h) x + (1 - \phi_h) \mu_h y \quad (A.5)$$

$$B(x, y) = \lambda_h x + \phi_h y + \frac{1}{2} (x^2 + 2\rho r_h \sigma_h x y + \sigma_h^2 y^2) \quad (A.6)$$

for the square-root specification and finally,

$$A(x, y) = \alpha_h x + w_h y - \frac{1}{2} \ln (1 - 2\alpha_h \eta_h^4 y) \quad (A.7)$$

$$B(x, y) = \lambda_h x + b_h y + \frac{1}{2 \eta_h^2} \left[ \sqrt{(1 - 2\alpha_h \eta_h^4 y)(1 - 2\eta_h x - 2c_h y)} \right] \quad (A.8)$$

for the IG GARCH specification.

B Change of Measure, Risk-Neutral Dynamics of Returns and Option Pricing

For the multifactor SVS model, we assume a change of measure $Z_{t,t+1}$ given by:

$$Z_{t,t+1} = \exp \left( -A(\kappa, \pi) - B(\kappa, \pi)^\top \sigma_i^2 + \kappa^\top r_{t+1} + \pi^\top \sigma_i^2 r_{t+1} \right), \quad (B.1)$$

and which, by definition and specification, satisfies $E[Z_{t,t+1} | I_t] = 1$. We are interested in deriving return dynamics under the risk-neutral measure. We first look at expected returns conditional to $\langle \sigma_{t+1}^2, I_t \rangle$. For the return $r_j$ we show that:

$$E^* \left[ r_{j,t+1} | \langle \sigma_{t+1}^2, I_t \rangle \right] = E^* \left[ Z_{t,t+1} r_{j,t+1} | \langle \sigma_{t+1}^2, I_t \rangle \right]$$

$$= \delta_{j,t} + \sum_{i=1}^{\kappa} \left( \beta_{ji} + \psi' (\kappa_j; \eta_{ji}) \gamma_{ji} \right) \sigma_i^2 r_{t+1} \exp \left( -\Psi_i^* (f(\kappa, \pi)) + f(\kappa, \pi)^\top \sigma_i^2 r_{t+1} \right). \quad (B.2)$$
where $\psi' (\cdot; \cdot)$ is the derivative of $\psi (\cdot; \cdot)$ with respect to its first argument.

Expected returns are linear on factors under the physical measure. They will also be linear on factors under the risk-neutral measure if and only if $f (\kappa, \pi) = 0$. We assume that the parameters $\kappa$ and $\pi$ in our change of measure specification satisfies this condition. Also notice that this condition implies the followings:

$$Z_{t,t+1} = \exp \left( -\kappa^\top \delta_t + \kappa^\top r_{t+1} + \pi^\top \sigma_{t+1}^2 \right) \quad \text{and} \quad E \left[ Z_{t,t+1} \mid \langle \sigma_{t+1}^2, I_t \rangle \right] = 1. \quad (B.3)$$

In particular, an implication of the second equation is that the moment generating function of $\sigma_{t+1}^2$, conditional to $I_t$, does not change from the physical to the risk-neutral measure. Thus, the factors still follow a multivariate autoregressive gamma under the risk-neutral measure. Return innovations under the risk-neutral measure, and conditional to $\langle \sigma_{t+1}^2, I_t \rangle$, are given by:

$$r_{j,t+1} - E^* \left[ r_{j,t+1} \mid \langle \sigma_{t+1}^2, I_t \rangle \right] = \sum_{i=1}^{K} \gamma_{ji} \sigma_{i,t+1} \left( u_{ji,t+1} - \psi' (\kappa_j; \eta_{ji}) \gamma_{ji} \sigma_{i,t+1} \right). \quad (B.4)$$

Finding the distribution of the terms $\gamma_{ji} \sigma_{i,t+1} \left( u_{ji,t+1} - \psi' (\kappa_j; \eta_{ji}) \gamma_{ji} \sigma_{i,t+1} \right)$ will achieve the return dynamics under the risk-neutral measure. This distribution can be detected through their moment generating function. We show that:

$$E^* \left[ \exp \left( x_j \gamma_{ji} \sigma_{i,t+1} \left( u_{ji,t+1} - \psi' (\kappa_j; \eta_{ji}) \gamma_{ji} \sigma_{i,t+1} \right) \right) \mid \langle \sigma_{t+1}^2, I_t \rangle \right] = \exp \left( \left( \psi (x_j + \kappa_j; \eta_{ji}) - \psi (\kappa_j; \eta_{ji}) - x_j \psi' (\kappa_j; \eta_{ji}) \gamma_{ji}^2 \sigma_{i,t+1}^2 \right) \right) = \exp \left( \psi (x_j; \eta_{ji}^*) q_{ji}^2 \gamma_{ji}^2 \sigma_{i,t+1}^2 \right) \quad (B.5)$$

where we have also shown that

$$\psi (x_j + \kappa_j; \eta_{ji}) - \psi (\kappa_j; \eta_{ji}) - x_j \psi' (\kappa_j; \eta_{ji}) = q_{ji}^2 \psi (x_j; \eta_{ji}^*) \quad (B.6)$$

with

$$q_{ji}^2 = \left( 1 - \frac{2}{3} \eta_{ji} \kappa_j \right)^{-3/2} \quad \text{and} \quad \eta_{ji}^* = \eta_{ji} \left( 1 - \frac{2}{3} \eta_{ji} \kappa_j \right)^{-1}. \quad (B.7)$$

Equation (B.5) means that the term $\gamma_{ji} \sigma_{i,t+1} \left( u_{ji,t+1} - \psi' (\kappa_j; \eta_{ji}) \gamma_{ji} \sigma_{i,t+1} \right)$ can also be written

$$\gamma_{ji} \sigma_{i,t+1} \left( u_{ji,t+1} - \psi' (\kappa_j; \eta_{ji}) \gamma_{ji} \sigma_{i,t+1} \right) = \gamma_{ji}^* \sigma_{i,t+1} u_{ji,t+1}^* \quad (B.8)$$

where

$$\gamma_{ji}^* = q_{ji} \gamma_{ji} \quad \text{and} \quad u_{ji,t+1}^* \mid \langle \sigma_{t+1}^2, I_t \rangle \sim \text{SIG} \left( \eta_{ji}^* \left( \gamma_{ji}^* \sigma_{i,t+1} \right)^{-1} \right). \quad (B.9)$$

The return $r_j$ should also satisfies the no-arbitrage condition:

$$E^* \left[ \exp (r_{j,t+1}) \mid I_t \right] = \exp (r_{f,t+1}) \quad (B.10)$$
where $r_{f,t+1}$ is the risk-free rate. We show that:

$$E^* [\exp (r_{j,t+1} \mid I_t) = \exp (\delta_{jt} + \Psi_t^* (f (e_j + \kappa, \pi))) = \exp (\delta_{jt} + \Psi_t^* (\theta_j))$$

(B.11)

where $e_j$ is the $N \times 1$ vector with all components equal to zero but the $j^{th}$ component equals one, and $\theta_j = f (e_j + \kappa, \pi)$ is the $K \times 1$ vector which components are given by:

$$\theta_{ji} = \beta_{ji} + (\psi (1 + \kappa_j; \eta_{ji}) - \psi (\kappa_j; \eta_{ji})) \gamma_{ji}^2.$$

(B.12)

We further show that

$$\theta_{ji} = q_{ji}^2 \theta_{ji}^*$$

where $\theta_{ji}^* = \beta_{ji}^* + \psi (1; \eta_{ji}) \gamma_{ji}^2$ and $\beta_{ji}^* = (\beta_{ji} + \psi' (\kappa_j; \eta_{ji}) \gamma_{ji}^2) / q_{ji}^2.$

(B.13)

Assuming that the risk-free rate is constant and equal to $r_f$, equation (B.11) implies that

$$\lambda_j = -b (\theta_j) \quad \text{and} \quad \mu_{0,j} = r_f + (\lambda_j + \beta_j) \mu - a (\theta_j).$$

(B.14)

Finally, the dynamics of returns under the risk-neutral measure is given by:

$$r_{j,t+1} = r_f - a (q_{ji}^2 \theta_{ji}^*) - \sum_{i=1}^K b_i (q_{ji}^2 \theta_{ji}^*) \sigma_{it}^2 + \sum_{i=1}^K \beta_{ji}^* q_{ji}^2 \sigma_{i,t+1}^2 + \sum_{i=1}^K \gamma_{ji} \sigma_{i,t+1} u_{ji,t+1}^*.$$

(B.15)

where $q_{ji}^2 \theta_{ji}^*$ denotes the $K \times 1$ vector with components $q_{ji}^2 \theta_{ji}^*$.

C Second Order Log-normal Approximation of Positive Random Variables

The second order lognormal approximation of a positive random variable $X$ with mean $\mu_x$ and variance $\sigma_x^2$ is given by:

$$X \approx \exp \left( \ln \left( \frac{\mu_x^2}{\sqrt{\mu_x^2 + \sigma_x^2}} \right) + \sqrt{\ln \left( \frac{\mu_x^2 + \sigma_x^2}{\mu_x^2} \right)} \varepsilon_X \right)$$

(C.1)

where $\varepsilon_X$ is a standard normal random variable.

Given (C.1), the second order lognormal approximation of a standardized inverse gaussian random variable $u$ with positive skewness $s$ is given by:

$$u \approx \exp \left( \ln \left( \frac{9}{s \sqrt{s^2 + 9}} \right) + \sqrt{\ln \left( \frac{s^2 + 9}{9} \varepsilon \right)} - \frac{3}{s} \right)$$

(C.2)

where $\varepsilon$ is a standard normal random variable.

Given (C.1), the second order lognormal approximation for the dynamics of a stationary univariate autoregressive gamma process $X_{t+1}$ with mean $\mu_x$, variance $\sigma_x^2$ and persistence $\phi_x$ is given by:

$$X_{t+1} \approx \exp \left( \ln \left( \frac{m (X_t)^2}{\sqrt{m (X_t)^2 + v (X_t)}} \right) + \sqrt{\ln \left( \frac{m (X_t)^2 + v (X_t)}{m (X_t)^2} \varepsilon_{X,t+1} \right)} \right)$$

(C.3)
where

\[ m(X_t) = (1 - \phi_x) \mu_x + \phi_x X_t \] (C.4)

\[ v(X_t) = (1 - \phi_x)^2 \sigma_x^2 + \frac{2 (1 - \phi_x) \phi_x \sigma_x^2}{\mu_x} X_t \] (C.5)

and \( \varepsilon_{X,t+1} \) is a i.i.d. standard normal shock.

D The Unscented Kalman Filter

The Unscented Kalman Filter is essentially an approximation of a nonlinear transformation of probability distribution coupled with the Kalman Filter. It has been introduced in the engineering literature by Julier et al. (1995) and Julier and Uhlmann (1996). (See also Wan and van der Merwe (2001) for general introduction) and, to our knowledge, was first imported in Finance by Leippold and Wu (2003).

The Unscented Filter selects a set of sigma points in the distribution of \((\omega_t^T, u_{t+1}^T, \varepsilon_{t+1}^T)^T\) conditional on \(r_t\). This distribution is assumed gaussian with mean

\[ \tilde{\chi} = (\omega_t^T, \bar{u}_t^T, \bar{\varepsilon}_t^T)^T \]

and variance

\[ P_{\chi\chi} = \begin{pmatrix} P_{\omega\omega} & P_{\omega u} & P_{\omega \varepsilon} \\ P_{u\omega} & P_{uu} & P_{u \varepsilon} \\ P_{e\omega} & P_{e u} & P_{ee} \end{pmatrix} \].

Following Julier et al. (1995) we consider the \(2n+1\) sigma points \( \chi_i = (\omega_{i,t}^T, u_{i,t+1}^T, \varepsilon_{i,t+1}^T)^T \)

with associated weights \( W_i \) defined by:

\[
\begin{align*}
\chi_0 &= \tilde{\chi}, \\
W_0 &= \kappa / (n + \kappa) \\
\chi_i &= \tilde{\chi} + \left( \sqrt{(n + \kappa) P_{\chi\chi}} \right)_i, \quad W_i = 1/2 (n + \kappa) \\
\chi_{i+n} &= \tilde{\chi} - \left( \sqrt{(n + \kappa) P_{\chi\chi}} \right)_i, \quad W_i = 1/2 (n + \kappa),
\end{align*}
\] (D.1)

where \( n \) is the dimension of the vector \((\omega_t^T, u_{t+1}^T, \varepsilon_{t+1}^T)^T\), \( \kappa \) is an appropriately chosen real number and \( \left( \sqrt{(n + \kappa) P_{\chi\chi}} \right)_i \) is the \( i \)th column of the matrix \((n + \kappa) P_{\chi\chi}\).

These sigma points are transformed through state and observation functions to obtain:

\[ \omega_{i,t+1|t} = F \left( \omega_{i,t|t}, u_t \right) \quad \text{and} \quad r_{i,t+1|t} = H \left( \omega_{i,t+1|t}, \varepsilon_i \right) \]
from which approximations of predicted means and covariances are computed as:

$$\hat{\omega}_{t+1|t} = \sum_{i=0}^{2n} W_i \omega_{i,t+1|t} \text{ and } \hat{r}_{t+1|t} = \sum_{i=0}^{2n} W_i r_{i,t+1|t}$$  \hspace{1cm} (D.2)

$$\hat{p}_{t+1|t} = \sum_{i=0}^{2n} W_i (\omega_{i,t+1|t} - \hat{\omega}_{t+1|t}) (\omega_{i,t+1|t} - \hat{\omega}_{t+1|t})^\top$$  \hspace{1cm} (D.3)

$$\hat{p}_{r_{t+1|t}} = \sum_{i=0}^{2n} W_i (r_{i,t+1|t} - \hat{r}_{t+1|t}) (r_{i,t+1|t} - \hat{r}_{t+1|t})^\top$$  \hspace{1cm} (D.4)

$$\hat{p}_{\omega_{t+1|t}} = \sum_{i=0}^{2n} W_i (\omega_{i,t+1|t} - \hat{\omega}_{t+1|t}) (r_{i,t+1|t} - \hat{r}_{t+1|t})^\top.$$  \hspace{1cm} (D.5)

### E Approximated Moments of a Function of a Normal Random Variable

Consider a normal random variable $X$ with mean $\mu_x$ and variance $\sigma_x^2$. Let $Y = f(X)$, where $f$ is a twice differentiable real function. The variable $Y$ admits the second order Taylor approximation

$$Y = f(\mu_x) + f'(\mu_x)(X - \mu_x) + \frac{1}{2} f''(\mu_x)(X - \mu_x)^2$$  \hspace{1cm} (E.1)

which implies that the mean of $Y$ can be approximated by:

$$\mu_y = E[Y] = f(\mu_x) + \frac{1}{2} f''(\mu_x) \sigma_x^2$$  \hspace{1cm} (E.2)

It follows that:

$$(Y - \mu_y)^2 = f'(\mu_x)^2 (X - \mu_x)^2 + f''(\mu_x) (X - \mu_x)^3 - \sigma_x^3$$  \hspace{1cm} (E.3)

$$(Y - \mu_y)^3 = f'(\mu_x)^3 (X - \mu_x)^3 + \frac{3}{2} f'(\mu_x)^2 f''(\mu_x) (X - \mu_x)^4 - \sigma_x^4$$  \hspace{1cm} (E.4)

$$(Y - \mu_y)^4 = f'(\mu_x)^4 (X - \mu_x)^4 + \frac{3}{2} f'(\mu_x)^3 f''(\mu_x) (X - \mu_x)^5 - 3\sigma_x^5$$  \hspace{1cm} (E.5)

The third and fifth central moments of $X$ are zero whereas the fourth and sixth central moments of $X$ are respectively $3\sigma_x^4$ and $15\sigma_x^6$. Based on that, taking expectations of (E.4) and (E.5) gives the following approximations for the variance and the third moment of $Y$:

$$\sigma_y^2 = Var[Y] = f'(\mu_x)^2 \sigma_x^2 + \frac{1}{2} f''(\mu_x)^2 \sigma_x^4$$  \hspace{1cm} (E.6)

$$E[(Y - \mu_y)^3] = 3 f'(\mu_x)^2 f''(\mu_x) \sigma_x^4 + f''(\mu_x)^3 \sigma_x^6.$$  \hspace{1cm} (E.7)
References


35


Table 1: **Summary Statistics of Daily Stocks Returns for the Period 1990-2005.** The entries of the table are summary statistics of daily stock returns. Means and standard deviations are annualized values using 252 trading days per year, whereas skewness and kurtosis are daily values. Return data span the period January 2, 1990 to December 30, 2005.

<table>
<thead>
<tr>
<th>r</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>SG</td>
<td>6.11</td>
<td>19.12</td>
<td>-0.46</td>
<td>6.61</td>
</tr>
<tr>
<td>SN</td>
<td>13.84</td>
<td>13.45</td>
<td>-0.46</td>
<td>6.36</td>
</tr>
<tr>
<td>SV</td>
<td>15.45</td>
<td>12.50</td>
<td>-0.61</td>
<td>6.99</td>
</tr>
<tr>
<td>BG</td>
<td>10.05</td>
<td>17.24</td>
<td>-0.06</td>
<td>6.78</td>
</tr>
<tr>
<td>BN</td>
<td>11.30</td>
<td>14.34</td>
<td>-0.19</td>
<td>7.04</td>
</tr>
<tr>
<td>BV</td>
<td>10.88</td>
<td>14.09</td>
<td>-0.30</td>
<td>6.88</td>
</tr>
<tr>
<td>CRSP</td>
<td>10.10</td>
<td>15.54</td>
<td>-0.21</td>
<td>7.21</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>9.17</td>
<td>16.09</td>
<td>-0.01</td>
<td>6.69</td>
</tr>
</tbody>
</table>
Table 2: Estimation Results on Small Stocks.
The entries of the table are GMM parameter estimates of the single factor SVS, the HN-S volatility and the HN GARCH(1,1) models on small stocks. To perform the GMM estimation for all the three models, we use the same sixteen moment conditions that are all significant for small stocks as depicted in Figures 2, 3 and 4. The corresponding moments are \( \{ E \left[ r_t^2 \right] \}_{j=2}^4 \), \( \{ E \left[ r_t^2 r_{t+j}^2 \right], E \left[ r_t r_{t+j}^2 \right] \}_{j=1}^5 \) and \( \{ E \left[ r_t r_{t+j}^3 \right] \}_{j=1}^3 \). Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations. Standard errors are given below the estimates.

<table>
<thead>
<tr>
<th></th>
<th>SVS1f, ( \eta \neq 0 )</th>
<th>SVS1f, ( \eta = 0 )</th>
<th>HN-S Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>SN</td>
<td>SV</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-24.25</td>
<td>-36.51</td>
<td>-52.77</td>
</tr>
<tr>
<td>( \eta_1 )</td>
<td>5.25E-03</td>
<td>3.34E-03</td>
<td>3.79E-03</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>1.35E-04</td>
<td>6.82E-05</td>
<td>5.76E-05</td>
</tr>
<tr>
<td>( \phi_1 )</td>
<td>0.979</td>
<td>0.979</td>
<td>0.930</td>
</tr>
<tr>
<td>( \sqrt{\omega_1} )</td>
<td>1.42E-04</td>
<td>6.75E-05</td>
<td>5.54E-05</td>
</tr>
<tr>
<td>( \beta_h )</td>
<td>476.36</td>
<td>767.80</td>
<td>333.12</td>
</tr>
<tr>
<td>Criterion</td>
<td>125.04</td>
<td>128.66</td>
<td>126.01</td>
</tr>
<tr>
<td>p-value</td>
<td>0.06</td>
<td>0.11</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Table 3: Estimation Results on Big Stocks.

The entries of the table are GMM parameter estimates of the single factor SVS and the HN-S volatility models on big stocks. To perform the GMM estimation for all the three models, we use the same eleven moment conditions that are all significant for big stocks and market indexes as depicted in Figures 2, 3 and 4. The corresponding moments are \( \{E[r_t^j]\}_{j=2}^4 \), \( \{E[r_t^2r_{t+j}^2]\}_{j=1}^5 \) and \( \{E[r_t^2r_{t+j}]\}_{j=1}^3 \). Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations. Standard errors are given below the estimates.

<table>
<thead>
<tr>
<th></th>
<th>SVS1f, ( \eta \neq 0 )</th>
<th>SVS1f, ( \eta = 0 )</th>
<th>HN-S Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BG</td>
<td>BN</td>
<td>BV</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>-26.45</td>
<td>-21.95</td>
<td>-19.74</td>
</tr>
<tr>
<td>( \eta_1 )</td>
<td>8.83E-03</td>
<td>4.77E-03</td>
<td>2.97E-03</td>
</tr>
<tr>
<td>( \mu_1 )</td>
<td>1.09E-04</td>
<td>7.81E-05</td>
<td>7.61E-05</td>
</tr>
<tr>
<td>( \phi_1 )</td>
<td>0.946</td>
<td>0.973</td>
<td>1.005</td>
</tr>
<tr>
<td>( \sqrt{\omega_1} )</td>
<td>1.12E-04</td>
<td>8.69E-05</td>
<td>8.48E-05</td>
</tr>
<tr>
<td>( \rho_h )</td>
<td>1.45E-05</td>
<td>1.43E-05</td>
<td>1.58E-05</td>
</tr>
<tr>
<td>Criterion</td>
<td>33.12</td>
<td>16.43</td>
<td>10.45</td>
</tr>
<tr>
<td>p-value</td>
<td>0.11</td>
<td>0.26</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Table 4: Estimation Results on Market Indexes.
The entries of the table are GMM parameter estimates of the single factor SVS, the HN-S volatility and the IG GARCH models on market indexes. To perform the GMM estimation for all the three models, we use the same eleven moment conditions that are all significant for big stocks and market indexes as depicted in Figures 2, 3 and 4. The corresponding moments are \( \{E[r_j^4]\}_{j=2}, \{E[r_t^2 r_{t+j}^2]\}_{j=1}^5 \) and \( \{E[r_t^2 r_{t+j}^2]\}_{j=1}^3 \). Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations. Standard errors are given below the estimates.

<table>
<thead>
<tr>
<th></th>
<th>SVS1f, ( \eta \neq 0 )</th>
<th>SVS1f, ( \eta = 0 )</th>
<th>HN-S Volatility</th>
<th>IG GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_1 )</td>
<td>-30.09</td>
<td>-25.29</td>
<td>-17.19</td>
<td>-10.46</td>
</tr>
<tr>
<td>( \lambda_h )</td>
<td>6.44</td>
<td>5.84</td>
<td>3.76</td>
<td>3.11</td>
</tr>
<tr>
<td>( \eta_1 )</td>
<td>7.65E-03</td>
<td>7.76E-03</td>
<td>3.76</td>
<td>3.11</td>
</tr>
<tr>
<td>( \phi_1 )</td>
<td>0.948</td>
<td>0.963</td>
<td>0.990</td>
<td>0.968</td>
</tr>
<tr>
<td>( \sqrt{\omega_1} )</td>
<td>9.43E-05</td>
<td>9.89E-05</td>
<td>1.02E-04</td>
<td>1.10E-04</td>
</tr>
<tr>
<td>Criterion</td>
<td>34.12</td>
<td>24.06</td>
<td>94.72</td>
<td>101.87</td>
</tr>
<tr>
<td>J-Stat</td>
<td>8.24</td>
<td>7.19</td>
<td>13.49</td>
<td>17.50</td>
</tr>
<tr>
<td>p-value</td>
<td>0.14</td>
<td>0.21</td>
<td>0.04</td>
<td>0.01</td>
</tr>
</tbody>
</table>
Table 5: Estimation Results. $c$’s coefficients.

The entries of the table are loadings of expected returns, volatility, asymmetry and leverage effects on factors, using GMM parameter estimates of the single factor SVS, the HN-S volatility and the IG GARCH models on small stocks and market indexes. To perform the GMM estimation for all three models on small stocks, we use the same sixteen moment conditions that are all significant for small stocks as depicted in Figures 2, 3 and 4. The corresponding moments are $\{E[r_i]\}_{j=2}^4, \{E[r_i^2]\}_{j=1}^5$ and $\{E[r_i^3]\}_{j=1}^3$. For all four models on market indexes, we use the same eleven moment conditions that are all significant for big stocks and market indexes as depicted in Figures 2, 3 and 4. The corresponding moments are $\{E[r_i]\}_{j=2}^4, \{E[r_i^2]\}_{j=1}^5$ and $\{E[r_i^3]\}_{j=1}^3$. Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations. Standard errors are given below the estimates.

<table>
<thead>
<tr>
<th>SVS1f, $\eta \neq 0$</th>
<th>SVS1f, $\eta = 0$</th>
<th>HN-S Volatility</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{0\mu}$</td>
<td>3.44E-03</td>
<td>2.99E-03</td>
</tr>
<tr>
<td>$c_{1\mu}$</td>
<td>-23.75</td>
<td>-35.76</td>
</tr>
<tr>
<td>$c_{0h}$</td>
<td>2.80E-06</td>
<td>1.41E-06</td>
</tr>
<tr>
<td>$c_{1h}$</td>
<td>0.983</td>
<td>0.983</td>
</tr>
<tr>
<td>$c_{0s}$</td>
<td>1.41E-08</td>
<td>4.49E-09</td>
</tr>
<tr>
<td>$c_{1s}$</td>
<td>4.70E-03</td>
<td>2.98E-03</td>
</tr>
<tr>
<td>$c_{0r}$</td>
<td>-2.10E-04</td>
<td>-7.09E-05</td>
</tr>
<tr>
<td>$c_{1r}$</td>
<td>-1.47E-04</td>
<td>-9.87E-04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SVS1f, $\eta \neq 0$</th>
<th>SVS1f, $\eta = 0$</th>
<th>HN-S Volatility</th>
<th>IG GARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_{0\mu}$</td>
<td>2.91E-03</td>
<td>2.72E-03</td>
<td>$c_{0\mu}$</td>
</tr>
<tr>
<td>$c_{1\mu}$</td>
<td>-28.52</td>
<td>-24.34</td>
<td>$c_{1\mu}$</td>
</tr>
<tr>
<td>$c_{0h}$</td>
<td>4.61E-06</td>
<td>3.63E-06</td>
<td>$c_{0h}$</td>
</tr>
<tr>
<td>$c_{1h}$</td>
<td>0.957</td>
<td>0.967</td>
<td>$c_{1h}$</td>
</tr>
<tr>
<td>$c_{0s}$</td>
<td>3.29E-08</td>
<td>2.70E-08</td>
<td>$c_{0s}$</td>
</tr>
<tr>
<td>$c_{1s}$</td>
<td>6.35E-03</td>
<td>6.91E-03</td>
<td>$c_{1s}$</td>
</tr>
<tr>
<td>$c_{0r}$</td>
<td>-7.30E-10</td>
<td>-3.48E-10</td>
<td>$c_{0r}$</td>
</tr>
<tr>
<td>$c_{1r}$</td>
<td>-3.01E-04</td>
<td>-1.85E-04</td>
<td>$c_{1r}$</td>
</tr>
</tbody>
</table>
Table 6: Moment Matching for Small Stocks.
The entries of the table are ratios of model unconditional moments to their empirical counterparts, based on parameter estimates of the single factor SVS and the HN-S volatility models on small stocks. To perform the GMM estimation for all the three models, we use the same sixteen moment conditions that are all significant for small stocks as depicted in Figures 2, 3 and 4. The corresponding moments are identified by a 1 in the third column. Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations.

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>SN</th>
<th>SV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\eta \neq 0$</td>
<td>$\eta = 0$</td>
<td>HN</td>
</tr>
<tr>
<td>$E[r_{t+1}]$</td>
<td>1 0</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>$E[r_{t+2}^2]$</td>
<td>2 1</td>
<td>1.01</td>
<td>1.01</td>
</tr>
<tr>
<td>$E[r_{t+3}^2]$</td>
<td>3 1</td>
<td>1.04</td>
<td>1.67</td>
</tr>
<tr>
<td>$E[r_{t+4}^2]$</td>
<td>4 1</td>
<td>1.03</td>
<td>1.00</td>
</tr>
<tr>
<td>$E[r_t r_{t+5}]$</td>
<td>5 0</td>
<td>-7.90</td>
<td>-6.75</td>
</tr>
<tr>
<td>$E[r_t^2 r_{t+5}]$</td>
<td>6 0</td>
<td>-10.96</td>
<td>-10.55</td>
</tr>
<tr>
<td>$E[r_{t+1}^2]$</td>
<td>7 1</td>
<td>1.29</td>
<td>1.24</td>
</tr>
<tr>
<td>$E[r_{t+2}^2]$</td>
<td>8 0</td>
<td>-1.95</td>
<td>-1.87</td>
</tr>
<tr>
<td>$E[r_{t+3}^2]$</td>
<td>9 1</td>
<td>0.99</td>
<td>1.04</td>
</tr>
<tr>
<td>$E[r_{t+4}^2]$</td>
<td>10 0</td>
<td>-1.43</td>
<td>-1.38</td>
</tr>
<tr>
<td>$E[r_{t+1} r_{t+4}]$</td>
<td>11 0</td>
<td>1.60</td>
<td>1.32</td>
</tr>
<tr>
<td>$E[r_{t+2} r_{t+4}]$</td>
<td>12 0</td>
<td>-5.63</td>
<td>-5.21</td>
</tr>
<tr>
<td>$E[r_{t+3} r_{t+4}]$</td>
<td>13 1</td>
<td>1.24</td>
<td>1.15</td>
</tr>
<tr>
<td>$E[r_{t+4}^3]$</td>
<td>14 0</td>
<td>1.48</td>
<td>1.37</td>
</tr>
<tr>
<td>$E[r_{t+1}^2 r_{t+4}]$</td>
<td>15 1</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>$E[r_{t+1} r_{t+5}]$</td>
<td>16 0</td>
<td>1.10</td>
<td>1.02</td>
</tr>
<tr>
<td>$E[r_{t+1} r_{t+6}]$</td>
<td>17 0</td>
<td>1.39</td>
<td>1.10</td>
</tr>
<tr>
<td>$E[r_{t+2}^2 r_{t+5}]$</td>
<td>18 0</td>
<td>-8.00</td>
<td>-7.11</td>
</tr>
<tr>
<td>$E[r_{t+2}^2 r_{t+6}]$</td>
<td>19 1</td>
<td>1.39</td>
<td>1.24</td>
</tr>
<tr>
<td>$E[r_{t+3}^2 r_{t+5}]$</td>
<td>20 0</td>
<td>2.31</td>
<td>2.06</td>
</tr>
<tr>
<td>$E[r_{t+3}^2 r_{t+6}]$</td>
<td>21 1</td>
<td>1.02</td>
<td>1.02</td>
</tr>
<tr>
<td>$E[r_{t+1} r_{t+7}]$</td>
<td>22 1</td>
<td>0.76</td>
<td>0.68</td>
</tr>
<tr>
<td>$E[r_{t+2}^3]$</td>
<td>23 0</td>
<td>49.92</td>
<td>38.20</td>
</tr>
<tr>
<td>$E[r_{t+3}^3]$</td>
<td>24 0</td>
<td>-2.95</td>
<td>-2.52</td>
</tr>
<tr>
<td>$E[r_{t+4}^3]$</td>
<td>25 1</td>
<td>0.67</td>
<td>0.58</td>
</tr>
<tr>
<td>$E[r_{t+1} r_{t+7}]$</td>
<td>26 0</td>
<td>-0.73</td>
<td>-0.63</td>
</tr>
<tr>
<td>$E[r_{t+2}^3]$</td>
<td>27 1</td>
<td>0.75</td>
<td>0.73</td>
</tr>
<tr>
<td>$E[r_{t+3}^3]$</td>
<td>28 1</td>
<td>2.17</td>
<td>1.86</td>
</tr>
<tr>
<td>$E[r_{t+4}^3]$</td>
<td>29 0</td>
<td>0.53</td>
<td>0.39</td>
</tr>
<tr>
<td>$E[r_{t+1}^2 r_{t+8}]$</td>
<td>30 0</td>
<td>986.72</td>
<td>810.83</td>
</tr>
<tr>
<td>$E[r_{t+2}^2 r_{t+8}]$</td>
<td>31 1</td>
<td>0.94</td>
<td>0.77</td>
</tr>
<tr>
<td>$E[r_{t+3}^2 r_{t+8}]$</td>
<td>32 0</td>
<td>2.48</td>
<td>2.05</td>
</tr>
<tr>
<td>$E[r_{t+4}^2 r_{t+8}]$</td>
<td>33 1</td>
<td>1.12</td>
<td>1.08</td>
</tr>
<tr>
<td>$E[r_{t+1}^3 r_{t+8}]$</td>
<td>34 1</td>
<td>0.81</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Table 7: Moment Matching for Big Stocks.
The entries of the table are ratios of model unconditional moments to their empirical counterparts, based on parameter estimates of the single factor SVS and the HN-S volatility models on big stocks. To perform the GMM estimation for all the three models, we use the same eleven moment conditions that are all significant for big stocks and market indexes as depicted in Figures 2, 3 and 4. The corresponding moments are identified by a 1 in the third column. Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations.

<table>
<thead>
<tr>
<th>Moment</th>
<th>BG</th>
<th>BN</th>
<th>BV</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E[r_t]$</td>
<td>1 0</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>$E[r_t^2]$</td>
<td>2 1</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>$E[r_t^3]$</td>
<td>3 1</td>
<td>0.75</td>
<td>-0.67</td>
</tr>
<tr>
<td>$E[r_t^4]$</td>
<td>4 1</td>
<td>1.01</td>
<td>0.98</td>
</tr>
<tr>
<td>$E[r_{t+5}]$</td>
<td>5 0</td>
<td>-0.99</td>
<td>-0.47</td>
</tr>
<tr>
<td>$E[r_{t+5}^2]$</td>
<td>6 0</td>
<td>1.80</td>
<td>0.85</td>
</tr>
<tr>
<td>$E[r_{t+5}^3]$</td>
<td>7 0</td>
<td>-0.54</td>
<td>-0.21</td>
</tr>
<tr>
<td>$E[r_{t+5}^4]$</td>
<td>8 0</td>
<td>3.77</td>
<td>1.10</td>
</tr>
<tr>
<td>$E[r_{t+4}]$</td>
<td>9 1</td>
<td>-1.64</td>
<td>-0.77</td>
</tr>
<tr>
<td>$E[r_{t+4}^2]$</td>
<td>10 0</td>
<td>-13.38</td>
<td>-6.28</td>
</tr>
<tr>
<td>$E[r_{t+4}^3]$</td>
<td>11 0</td>
<td>0.97</td>
<td>0.37</td>
</tr>
<tr>
<td>$E[r_{t+4}^4]$</td>
<td>12 0</td>
<td>1.11</td>
<td>1.15</td>
</tr>
<tr>
<td>$E[r_{t+3}]$</td>
<td>13 0</td>
<td>44.70</td>
<td>17.24</td>
</tr>
<tr>
<td>$E[r_{t+3}^2]$</td>
<td>14 0</td>
<td>-2.00</td>
<td>-0.58</td>
</tr>
<tr>
<td>$E[r_{t+3}^3]$</td>
<td>15 0</td>
<td>-2.35</td>
<td>-1.09</td>
</tr>
<tr>
<td>$E[r_{t+3}^4]$</td>
<td>16 0</td>
<td>2.09</td>
<td>0.97</td>
</tr>
<tr>
<td>$E[r_{t+2}]$</td>
<td>17 0</td>
<td>1.12</td>
<td>0.43</td>
</tr>
<tr>
<td>$E[r_{t+2}^2]$</td>
<td>18 0</td>
<td>1.03</td>
<td>1.06</td>
</tr>
<tr>
<td>$E[r_{t+2}^3]$</td>
<td>19 0</td>
<td>-2.66</td>
<td>-1.01</td>
</tr>
<tr>
<td>$E[r_{t+2}^4]$</td>
<td>20 0</td>
<td>-2.11</td>
<td>-0.60</td>
</tr>
<tr>
<td>$E[r_{t+1}]$</td>
<td>21 0</td>
<td>-5.71</td>
<td>-2.62</td>
</tr>
<tr>
<td>$E[r_{t+1}^2]$</td>
<td>22 0</td>
<td>0.64</td>
<td>0.29</td>
</tr>
<tr>
<td>$E[r_{t+1}^3]$</td>
<td>23 0</td>
<td>-1.29</td>
<td>-0.49</td>
</tr>
<tr>
<td>$E[r_{t+1}^4]$</td>
<td>24 0</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>$E[r_{t+0}]$</td>
<td>25 0</td>
<td>114.07</td>
<td>42.94</td>
</tr>
<tr>
<td>$E[r_{t+0}^2]$</td>
<td>26 0</td>
<td>3.99</td>
<td>1.11</td>
</tr>
<tr>
<td>$E[r_{t+0}^3]$</td>
<td>27 0</td>
<td>-1.29</td>
<td>-0.59</td>
</tr>
<tr>
<td>$E[r_{t+0}^4]$</td>
<td>28 0</td>
<td>1.01</td>
<td>0.46</td>
</tr>
<tr>
<td>$E[r_{t+1}^2]$</td>
<td>29 0</td>
<td>-0.42</td>
<td>-0.16</td>
</tr>
<tr>
<td>$E[r_{t+1}^3]$</td>
<td>30 0</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>$E[r_{t+1}^4]$</td>
<td>31 0</td>
<td>3.35</td>
<td>1.24</td>
</tr>
</tbody>
</table>
Table 8: Moment Matching for Market Indexes.
The entries of the table are ratios of model unconditional moments to their empirical counterparts, based on parameter estimates of the single factor SVS and the HN-S volatility models on market indexes. To perform the GMM estimation for all the three models, we use the same eleven moment conditions that are all significant for big stocks and market indexes as depicted in Figures 2, 3 and 4. The corresponding moments are identified by a 1 in the third column. Return data are daily and span the period from January 2, 1990 to December 30, 2005, for a total of 4036 observations.

<table>
<thead>
<tr>
<th>CRSP</th>
<th>S&amp;P500</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E[r_t] )</td>
<td>( \eta \neq 0 )</td>
</tr>
<tr>
<td>( E[r_t^2] )</td>
<td>1 0</td>
</tr>
<tr>
<td>( E[r_t^4] )</td>
<td>2 1</td>
</tr>
<tr>
<td>( E[r_{t+1}^2] )</td>
<td>3 1</td>
</tr>
<tr>
<td>( E[r_{t+1}^3] )</td>
<td>4 1</td>
</tr>
<tr>
<td>( E[r_t r_{t+5}] )</td>
<td>5 0</td>
</tr>
<tr>
<td>( E[r_{t+5}^2] )</td>
<td>6 0</td>
</tr>
<tr>
<td>( E[r_{t+5}^3] )</td>
<td>7 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+5}^2] )</td>
<td>8 0</td>
</tr>
<tr>
<td>( E[r_t^2 r_{t+5}] )</td>
<td>9 1</td>
</tr>
<tr>
<td>( E[r_t^3 r_{t+5}] )</td>
<td>10 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+4}] )</td>
<td>11 0</td>
</tr>
<tr>
<td>( E[r_{t+4}^2] )</td>
<td>12 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+4}^2] )</td>
<td>13 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+4}^3] )</td>
<td>14 0</td>
</tr>
<tr>
<td>( E[r_{t+4}^3 r_t] )</td>
<td>15 1</td>
</tr>
<tr>
<td>( E[r_t r_{t+4}^4] )</td>
<td>16 0</td>
</tr>
<tr>
<td>( E[r_t^2 r_{t+4}] )</td>
<td>17 0</td>
</tr>
<tr>
<td>( E[r_t^3 r_{t+4}] )</td>
<td>18 0</td>
</tr>
<tr>
<td>( E[r_{t+4}^2 r_t] )</td>
<td>19 1</td>
</tr>
<tr>
<td>( E[r_t r_{t+4}^3] )</td>
<td>20 0</td>
</tr>
<tr>
<td>( E[r_{t+4}^3 r_t] )</td>
<td>21 1</td>
</tr>
<tr>
<td>( E[r_t r_{t+4}^4] )</td>
<td>22 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+2}] )</td>
<td>23 0</td>
</tr>
<tr>
<td>( E[r_{t+2}^2] )</td>
<td>24 0</td>
</tr>
<tr>
<td>( E[r_{t+2}^3] )</td>
<td>25 1</td>
</tr>
<tr>
<td>( E[r_{t+2}^4] )</td>
<td>26 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+2}^2] )</td>
<td>27 1</td>
</tr>
<tr>
<td>( E[r_t r_{t+2}^3] )</td>
<td>28 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+1}] )</td>
<td>29 0</td>
</tr>
<tr>
<td>( E[r_t^2 r_{t+1}] )</td>
<td>30 0</td>
</tr>
<tr>
<td>( E[r_t r_{t+1}^2] )</td>
<td>31 1</td>
</tr>
<tr>
<td>( E[r_{t+1}^2 r_t] )</td>
<td>32 0</td>
</tr>
<tr>
<td>( E[r_{t+1}^3 r_t] )</td>
<td>33 1</td>
</tr>
<tr>
<td>( E[r_t r_{t+1}^3] )</td>
<td>34 0</td>
</tr>
</tbody>
</table>
Table 9: Summary Statistics for Strike Price and Maturity Categories.

(a) Summary statistics by moneyness

<table>
<thead>
<tr>
<th>Moneyness</th>
<th>Number of Contracts</th>
<th>Average Call Price</th>
<th>Average IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.95</td>
<td>3192</td>
<td>29.412</td>
<td>0.195</td>
</tr>
<tr>
<td>0.95 to 0.975</td>
<td>2351</td>
<td>32.625</td>
<td>0.194</td>
</tr>
<tr>
<td>0.975 to 1</td>
<td>3851</td>
<td>37.292</td>
<td>0.194</td>
</tr>
<tr>
<td>1 to 1.025</td>
<td>3048</td>
<td>47.294</td>
<td>0.202</td>
</tr>
<tr>
<td>&gt;1.025</td>
<td>3216</td>
<td>81.372</td>
<td>0.232</td>
</tr>
<tr>
<td>All</td>
<td>15658</td>
<td>45.986</td>
<td>0.203</td>
</tr>
</tbody>
</table>

(b) Summary statistics by Maturity

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Number of Contracts</th>
<th>Average Call Price</th>
<th>Average IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2061</td>
<td>31.119</td>
<td>0.207</td>
</tr>
<tr>
<td>2</td>
<td>4931</td>
<td>38.680</td>
<td>0.204</td>
</tr>
<tr>
<td>3</td>
<td>2571</td>
<td>42.018</td>
<td>0.203</td>
</tr>
<tr>
<td>4-6</td>
<td>2974</td>
<td>51.108</td>
<td>0.202</td>
</tr>
<tr>
<td>7-12</td>
<td>3121</td>
<td>65.731</td>
<td>0.201</td>
</tr>
<tr>
<td>All</td>
<td>15658</td>
<td>45.986</td>
<td>0.203</td>
</tr>
</tbody>
</table>

(c) Summary statistics by moneyness and maturities. For each moneyness and strike price category, the first line gives the number of contracts and the second line give the average Implied Volatility

<table>
<thead>
<tr>
<th>Moneyness</th>
<th>Months</th>
<th>&lt;0.95</th>
<th>0.95 to 0.975</th>
<th>0.975 to 1</th>
<th>1 to 1.025</th>
<th>&gt;1.025</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>7</td>
<td>154</td>
<td>658</td>
<td>699</td>
<td>543</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.254</td>
<td>0.201</td>
<td>0.187</td>
<td>0.197</td>
<td>0.245</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>277</td>
<td>738</td>
<td>1389</td>
<td>1201</td>
<td>1326</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.208</td>
<td>0.188</td>
<td>0.189</td>
<td>0.198</td>
<td>0.234</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>439</td>
<td>494</td>
<td>720</td>
<td>454</td>
<td>464</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.198</td>
<td>0.191</td>
<td>0.197</td>
<td>0.207</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td>4-6</td>
<td>928</td>
<td>513</td>
<td>574</td>
<td>406</td>
<td>553</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.193</td>
<td>0.194</td>
<td>0.199</td>
<td>0.207</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>7-12</td>
<td>1541</td>
<td>452</td>
<td>510</td>
<td>288</td>
<td>330</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.193</td>
<td>0.204</td>
<td>0.205</td>
<td>0.213</td>
<td>0.224</td>
</tr>
</tbody>
</table>

Figure 1: Return Series.
Table 10: Estimation of Structural Parameters of Risk-Neutral Processes

<table>
<thead>
<tr>
<th>Param.</th>
<th>SVS1F</th>
<th>SVS1F, η = 0</th>
<th>SVS2F</th>
<th>Param.</th>
<th>HN</th>
<th>CHJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>β₁</td>
<td>-1.808E+01</td>
<td>-1.210E+02</td>
<td>1.348E+02</td>
<td>γ*</td>
<td>2.410E+02</td>
<td>1.223E+06</td>
</tr>
<tr>
<td>η₁</td>
<td>-2.135E-01</td>
<td>0.000E+00</td>
<td>-1.340E-01</td>
<td>ω*</td>
<td>5.450E-13</td>
<td>1.266E-06</td>
</tr>
<tr>
<td>ν₁</td>
<td>1.347E-02</td>
<td>2.115E-01</td>
<td>5.474E-01</td>
<td>β*</td>
<td>7.554E-01</td>
<td>9.730E-01</td>
</tr>
<tr>
<td>α₁</td>
<td>1.177E-04</td>
<td>5.270E-06</td>
<td>1.933E-16</td>
<td>α*</td>
<td>3.592E-06</td>
<td>1.223E-06</td>
</tr>
<tr>
<td>φ₁</td>
<td>9.989E-01</td>
<td>9.824E-01</td>
<td>9.979E-01</td>
<td>η*</td>
<td>0.000E+00</td>
<td>-6.218E-02</td>
</tr>
<tr>
<td>β₂</td>
<td>9.838E-01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>η₂</td>
<td>-1.856E-01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ν₂</td>
<td>1.108E+00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α₂</td>
<td>4.824E-16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>φ₂</td>
<td>9.207E-01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RRMSE | 0.078 | 0.103 | 0.058 | 0.092 | 0.085 |

Table 11: Relative RMSE by Moneyness and Maturity

(a) Moneyness

<table>
<thead>
<tr>
<th></th>
<th>&lt;0.95</th>
<th>0.95 to 0.975</th>
<th>0.975 to 1</th>
<th>1 to 1.025</th>
<th>&gt;1.025</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVS1f</td>
<td>0.142</td>
<td>0.089</td>
<td>0.080</td>
<td>0.062</td>
<td>0.055</td>
<td>0.078</td>
</tr>
<tr>
<td>SVS1f, η = 0</td>
<td>0.215</td>
<td>0.122</td>
<td>0.092</td>
<td>0.072</td>
<td>0.072</td>
<td>0.103</td>
</tr>
<tr>
<td>SVS2f</td>
<td>0.105</td>
<td>0.052</td>
<td>0.057</td>
<td>0.047</td>
<td>0.044</td>
<td>0.058</td>
</tr>
<tr>
<td>HN</td>
<td>0.158</td>
<td>0.129</td>
<td>0.106</td>
<td>0.075</td>
<td>0.057</td>
<td>0.092</td>
</tr>
<tr>
<td>CHJ</td>
<td>0.147</td>
<td>0.112</td>
<td>0.094</td>
<td>0.072</td>
<td>0.056</td>
<td>0.085</td>
</tr>
</tbody>
</table>

(b) Maturity

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4-6</th>
<th>7-12</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVS1f</td>
<td>0.106</td>
<td>0.066</td>
<td>0.048</td>
<td>0.053</td>
<td>0.092</td>
<td>0.078</td>
</tr>
<tr>
<td>SVS1f, η = 0</td>
<td>0.105</td>
<td>0.069</td>
<td>0.072</td>
<td>0.088</td>
<td>0.125</td>
<td>0.103</td>
</tr>
<tr>
<td>SVS2f</td>
<td>0.062</td>
<td>0.038</td>
<td>0.038</td>
<td>0.052</td>
<td>0.070</td>
<td>0.058</td>
</tr>
<tr>
<td>HN</td>
<td>0.118</td>
<td>0.081</td>
<td>0.051</td>
<td>0.060</td>
<td>0.109</td>
<td>0.092</td>
</tr>
<tr>
<td>CHJ</td>
<td>0.116</td>
<td>0.069</td>
<td>0.062</td>
<td>0.057</td>
<td>0.100</td>
<td>0.085</td>
</tr>
</tbody>
</table>
Table 12: Relative Bias by Moneyness and Maturity

<table>
<thead>
<tr>
<th></th>
<th>&lt;0.95</th>
<th>0.95 to 0.975</th>
<th>0.975 to 1</th>
<th>1 to 1.025</th>
<th>&gt;1.025</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVS1f</td>
<td>-0.018</td>
<td>-0.008</td>
<td>0.022</td>
<td>0.018</td>
<td>-0.001</td>
<td>0.005</td>
</tr>
<tr>
<td>SVS1f, η = 0</td>
<td>0.029</td>
<td>-0.020</td>
<td>-0.011</td>
<td>-0.009</td>
<td>0.023</td>
<td>0.006</td>
</tr>
<tr>
<td>SVS2f</td>
<td>-0.009</td>
<td>-0.008</td>
<td>0.014</td>
<td>0.014</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>HN</td>
<td>-0.015</td>
<td>0.026</td>
<td>0.036</td>
<td>0.015</td>
<td>-0.016</td>
<td>0.005</td>
</tr>
<tr>
<td>CHJ</td>
<td>-0.043</td>
<td>0.009</td>
<td>0.025</td>
<td>0.014</td>
<td>-0.003</td>
<td>0.002</td>
</tr>
</tbody>
</table>

(b) Maturity

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4-6</th>
<th>7-12</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVS1f</td>
<td>-0.027</td>
<td>0.015</td>
<td>0.013</td>
<td>0.003</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td>SVS1f, η = 0</td>
<td>-0.065</td>
<td>0.011</td>
<td>0.031</td>
<td>0.020</td>
<td>-0.000</td>
<td>0.006</td>
</tr>
<tr>
<td>SVS2f</td>
<td>0.012</td>
<td>0.007</td>
<td>-0.003</td>
<td>0.005</td>
<td>0.007</td>
<td>0.006</td>
</tr>
<tr>
<td>HN</td>
<td>-0.015</td>
<td>0.020</td>
<td>0.021</td>
<td>-0.003</td>
<td>-0.004</td>
<td>0.005</td>
</tr>
<tr>
<td>CHJ</td>
<td>-0.067</td>
<td>0.007</td>
<td>0.036</td>
<td>0.017</td>
<td>-0.010</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Figure 2: Autocorrelation of Squared Returns.
Figure 3: Cross-Correlations Between Returns and Squared Returns.

Panel A: $\text{Corr}(r_t, r_{t+j}^2)$

Panel B: $\text{Corr}(r_{t+j}^2, r_{t+j})$
Figure 4: Cross-Correlations Between Returns and Cubed Returns.

Panel A: \( \text{Corr} \left( r_t, r_{t+j}^3 \right) \)

Panel B: \( \text{Corr} \left( r_t^3, r_{t+j} \right) \)
Figure 5: Portfolios Volatility and Skewness: Market Indexes

S&P500 Volatility $\eta_1 \neq 0$

CRSP Volatility $\eta_1 \neq 0$

S&P500 Skewness $\eta_1 \neq 0$

CRSP Skewness $\eta_1 \neq 0$

S&P500 Volatility $\eta_1 = 0$

CRSP Volatility $\eta_1 = 0$

S&P500 Skewness $\eta_1 = 0$

CRSP Skewness $\eta_1 = 0$
Figure 6: Portfolios Volatility and Skewness: Small and Big Stocks

SV Volatility $\eta_1 \neq 0$

SV Volatility $\eta_1 = 0$

BG Volatility $\eta_1 \neq 0$

BG Volatility $\eta_1 = 0$

SV Skewness $\eta_1 \neq 0$

SV Skewness $\eta_1 = 0$

BG Skewness $\eta_1 \neq 0$

BG Skewness $\eta_1 = 0$
Figure 7: Risk-Neutral Volatility and Conditional Skewness

![Volatility and Conditional Skewness](image)

Figure 8: Implied BSM volatility by Moneyness, Maturity and Model

![Implied BSM volatility by Moneyness, Maturity and Model](image)
Figure 9: Implied BSM volatility by Maturity, Moneyness and Model

- **Observed**
  - SVS1F, $\eta = 0$
  - SVS2F, $\eta = 0$
  - HN
  - CHJ

### Maturity Bands

- **<1**
- **1 to 2**
- **2 to 3**
- **4 to 6**
- **7 to 12**

#### Moneyness Bands

- **all < 0.95**
- **0.95 to 0.975**
- **0.975 to 1**
- **1 to 1.025**
- **> 1.025**