SCHOOL OF ECONOMICS AND MANAGEMENT
FACULTY OF SOCIAL SCIENCES
AARHUS UNIVERSITY

CREATES Research Paper 2010-10

Stochastic Volatility

Torben G. Andersen and Luca Benzoni

School of Economics and Management
Aarhus University
Bartholins Allé 10, Building 1322, DK-8000 Aarhus C
Denmark

Stochastic Volatility[1]

Torben G. Andersen[a] and Luca Benzoni[b]

[a] Kellogg School of Management, Northwestern University, Evanston, IL; NBER, Cambridge, MA; and CREATES, Aarhus, Denmark.
[b] Federal Reserve Bank of Chicago, Chicago, Illinois, USA.

Abstract: We give an overview of a broad class of models designed to capture stochastic volatility in financial markets, with illustrations of the scope of application of these models to practical finance problems. In a broad sense, this model class includes GARCH, but we focus on a narrower set of specifications in which volatility follows its own random process and is therefore a latent factor. These stochastic volatility specifications fit naturally in the continuous-time finance paradigm, and therefore serve as a prominent tool for a wide range of pricing and hedging applications. Moreover, the continuous-time paradigm of financial economics is naturally linked with the theory of volatility modeling and forecasting, and in particular with the practice of constructing ex-post volatility measures from high-frequency intraday data (realized volatility). One drawback is that in this setting volatility is not measurable with respect to observable information, and this feature complicates estimation and inference. Further, the presence of an additional state variable—volatility—renders the model less tractable from an analytic perspective. New estimation methods, combined with model restrictions that allow for closed-form solutions, make it possible to address these challenges while keeping the model consistent with the main properties of the data.

Key words: Stochastic Volatility, Realized Volatility, Implied Volatility, Options, Volatility Smirk, Volatility Smile, Dynamic Term Structure Models, Affine Models

JEL Classification Numbers: E43, G12

I. Introduction

The label Stochastic Volatility is applied in two distinct ways in the literature. For one, it is used to signify that the (absolute) size of the innovations of a time series displays random fluctuations over time. Descriptive models of financial time series almost invariably embed this feature nowadays as asset return series tend to display alternating quiet and turbulent periods of varying length and intensity. To distinguish this feature from models that operate with an a priori known or deterministic path for the volatility process, the random evolution of the conditional return variance is termed stochastic volatility. The simplest case of deterministic volatility is the constant variance assumption

[1]Original version: June 15, 2008; lightly revised February 2010. Chapter prepared for the Encyclopedia of Complexity and System Science, Springer. We are grateful to Olena Chyruk, Bruce Mizrach (the Editor), and Neil Shephard for helpful comments and suggestions. Of course, all errors remain our sole responsibility. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago or the Federal Reserve System. The work of Andersen is supported by a grant from the NSF to the NBER and support from CREATES funded by the Danish National Research Foundation. The most recent version of this paper is at: http://ssrn.com/abstract=1076672

invoked in, e.g., the Black and Scholes [63] framework. Another example is modeling the variance purely as a given function of calendar time, allowing only for effects such as time-of-year (seasonals), day-of-week (institutional and announcement driven) or time-of-day (diurnal effects due to, e.g., market microstructure features). Any model not falling within this class is then a stochastic volatility model. For example, in the one-factor continuous-time Cox, Ingersoll, and Ross [113] (CIR) model the (stochastic) level of the short term interest rate governs the dynamics of the (instantaneous) drift and diffusion term of all zero-coupon yields. Likewise, in GARCH models the past return innovations govern the one-period ahead conditional mean and variance. In both models, the volatility is known, or deterministic, at a given point in time, but the random evolution of the processes renders volatility stochastic for any horizon beyond the present period.

The second notion of stochastic volatility, which we adopt henceforth, refers to models in which the return variation dynamics is subject to an unobserved random shock so that the volatility is inherently latent. That is, the current volatility state is not known for sure, conditional on the true data generating process and the past history of all available discretely sampled data. Since the CIR and GARCH models described above do render the current (conditional) volatility known, they are not stochastic volatility models in this sense. In order to make the distinction clear cut, we follow Andersen [10] and label this second, more restrictive, set *genuine* stochastic volatility (SV) models.

There are two main advantages to focusing on SV models. First, much asset pricing theory is built on continuous-time models. Within this class, SV models tend to fit more naturally with a wide array of applications, including the pricing of currencies, options, and other derivatives, as well as the modeling of the term structure of interest rates. Second, the increasing use of high-frequency intraday data for construction of so-called realized volatility measures is also starting to push the GARCH models out of the limelight as the realized volatility approach is naturally linked to the continuous-time SV framework of financial economics.

One drawback is that volatility is not measurable with respect to observable (past) information in the SV setting. Accordingly, an estimate of the current volatility state must be filtered out from a noisy environment and the estimate will change as future observations become available. Hence, in-sample estimation typically involves smoothing techniques, not just filtering. In contrast, the conditional variance in GARCH is observable given past information, which renders (quasi-)maximum likelihood techniques for inference quite straightforward while smoothing techniques have no role. Thus, GARCH models are easier to estimate and practitioners often rely on them for time-series forecasts of volatility. However, the development of powerful method of simulated moments, Markov Chain Monte Carlo (MCMC) and other simulation based procedures for estimation and forecasting of SV models may well render them competitive with ARCH over time on that dimension.

Direct indications of the relations between SV and GARCH models are evident in the sequence of papers by Dan Nelson and Dean Foster exploring the SV diffusion limits of ARCH models as the case of continuous sampling is approached, see, e.g., Nelson and Foster [219]. Moreover, as explained in further detail in the estimation section below, it can be useful to summarize the dynamic features of asset returns by tractable pseudo-likelihood scores obtained from GARCH-style models when performing simulation based inference for SV models. As such, the SV and GARCH frameworks are closely related and should be viewed as complements. Despite these connections we focus, for the sake of brevity, almost exclusively on SV models and refer the interested reader to the GARCH chapter for further

information.

The literature on SV models is vast and rapidly growing, and excellent surveys are available, e.g., Ghysels et al. [158] and Shephard [239, 240]. Consequently, we focus on providing an overview of the main approaches with illustrations of the scope for applications of these models to practical finance problems.

II. Model Specification

The original econometric studies of SV models were invariably cast in discrete time and they were quite similar in structure to ARCH models, although endowed with a more explicit structural interpretation. Recent work in the area has been mostly directly towards a continuous time setting and motivated by the typical specifications in financial economics. This section briefly reviews the two alternative approaches to specification of SV models.

II.1. Discrete-Time SV Models and the Mixture-of-Distributions Hypothesis

Asset pricing theory contends that financial asset prices reflect the discounted value of future expected cash flows, implying that all news relevant for either discount rates or cash flows should induce a shift in market prices. Since economic news items appear almost continuously in real time, this perspective rationalizes the ever-changing nature of prices observed in financial markets. The process linking news arrivals to price changes may be complex, but if it is stationary in the statistical sense it will nonetheless produce a robust theoretical association between news arrivals, market activity and return volatility. In fact, if the number of news arrival is very large, standard central limit theory will tend to imply that asset returns are approximately normally distributed *conditional* on the news count. More generally, variables such as the trading volume, the number of transactions or the number of price quotes are also naturally related to the intensity of the information flow. This line of reasoning has motivated specifications such as,

$$y_t|s_t \rightsquigarrow N(\mu_y s_t, \sigma_y^2 s_t) \tag{1}$$

where y_t is an "activity" variable related to the information flow, s_t is a positive intensity process reflecting the rate of news arrivals, μ_y represents the mean response to an information event, and σ_y is a pure scaling parameter.

This is a normal mixture model, where the s_t process governs or "mixes" the scale of the distribution across the periods. If s_t is constant, this is simply an i.i.d. Gaussian process for returns and possible other related variables. However, this is clearly at odds with the empirical evidence for, e.g., return volatility and trading volume. Therefore, s_t is typically stipulated to follow a separate stochastic process with random innovations. Hence, each period the return series is subject to two separate shocks, namely the usual idiosyncratic error term associated with the (normal) return distribution, but also a shock to the variance or volatility process, s_t. This endows the return process with genuine stochastic volatility, reflecting the random intensity of news arrivals. Moreover, it is typically assumed that only returns, transactions and quotes are observable, but not the actual value of s_t itself, implying that σ_y cannot be separately identified. Hence, we simply fix this parameter at unity.

The time variation in the information flow series induces a fat-tailed unconditional distribution, consistent with stylized facts for financial return and, e.g., trading volume series. Intuitively, days with a lot of news display more rapid price fluctuations and trading activity than days with a low news count. In addition, if the s_t process is positively correlated, then shocks to the conditional mean and variance processes for y_t will be persistent. This is consistent with the observed clustering in financial markets, where return volatility and trading activity are contemporaneously correlated and each display pronounced positive serial dependence.

The inherent randomness and unobserved nature of the news arrival process, even during period t, renders the true mean and variance series latent. This property is the major difference with the GARCH model class, in which the one-step-ahead conditional mean and variance are a known function of observed variables at time $t-1$. Accordingly, for genuine SV models, we must distinguish the full, but infeasible, information set ($s_t \in \mathcal{F}_t$) and the observable information set ($s_t \notin \mathcal{I}_t$). This basic latency of the mixing variable (state vector) of the SV model complicates inference and forecasting procedures as discussed below.

For short horizon returns, μ_y is nearly negligible and can reasonably be ignored or simply fixed at a small constant value, and the series can then be demeaned. This simplification produces the following return (innovation) model,

$$r_t = \sqrt{s_t} z_t \,, \tag{2}$$

Where z_t is an i.i.d. standard normal variable, implying a simple normal-mixture representation,

$$r_t | s_t \rightsquigarrow N(0, s_t) \,. \tag{3}$$

Univariate return models of the form (3) as well as multivariate systems including a return variable along with other related market activity variables, such as the transactions count, the quote intensity or the aggregate trading volume, stem from the Mixture-of-Distributions Hypothesis (MDH).

Actual implementation of the MDH hinges on a particular representation of the information-arrival process s_t. Clark [102] uses trading volume as a proxy for the activity variable, a choice motivated by the high contemporaneous correlation between return volatility and volume. Tauchen and Pitts [247] follow a structural approach to characterize the joint distribution of the daily return and volume relation governed by the underlying latent information flow s_t. However, both these models assume temporal independence of the information flow, thus failing to capture the clustering in these series. Partly in response, Gallant et al. [153] examine the joint conditional return-volume distribution without imposing any structural MDH restrictions. Nonetheless, many of the original discrete-time SV specifications are compatible with the MDH framework, including Taylor [249],[2] who proposes an autoregressive parameterization of the latent log-volatility (or information flow) variable

$$\log(s_{t+1}) = \eta_0 + \eta_1 \log(s_t) + u_t \,,\ u_t \rightsquigarrow i.i.d(0, \sigma_u^2) \,, \tag{4}$$

where the error term, u_t, may be correlated with the disturbance term, z_t, in the return equation (2) so that $\rho = \mathrm{corr}(u_t, z_t) \neq 0$. If $\rho < 0$, downward movements in asset prices result in higher future

[2]Discrete-time SV models go farther back in time, at least to the early paper by Rosenberg [232] recently reprinted in Shephard [240].

volatility as also predicted by the so-called 'leverage effect' in the exponential GARCH, or EGARCH, form of Nelson [218] and the asymmetric GARCH model of Glosten et al. [160].

Early tests of the MDH include Lamoureux and Lastrapes [194] and Richardson and Smith [231]. Subsequently, Andersen [11] studies a modified version of the MDH that provides a much improved fit to the data. Further refinements of the MDH specification have been pursued by, e.g., Liesenfeld [198, 199] and Bollerslev and Jubinsky [67]. Among the first empirical studies of the related approach of stochastic time changes are Ané and Geman [29], who focus on stock returns, and Conley et al. [109], who focus on the short-term risk-free interest rate.

II.2. Continuous-Time Stochastic Volatility Models

Asset returns typically contain a predictable component, which compensates the investor for the risk of holding the security, and an unobservable shock term, which cannot be predicted using current available information. The conditional asset return variance pertains to the variability of the unobservable shock term. As such, over a non-infinitesimal horizon it is necessary to first specify the conditional mean return (e.g., through an asset pricing model) in order to identify the conditional return variation. In contrast, over an infinitesimal time interval this is not necessary because the requirement that market prices do not admit arbitrage opportunities implies that return innovations are an order of magnitude larger than the mean return. This result has important implications for the approach we use to model and measure volatility in continuous time.

Consider an asset with log-price process $\{p(t)\,, t \in [0\,, T]\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$. Following Andersen et al. [19] we define the continuously compounded asset return over a time interval from $t-h$ to t, $0 \leq h \leq t \leq T$, to be

$$r(t,h) = p(t) - p(t-h)\,. \tag{5}$$

A special case of (5) is the cumulative return up to time t, which we denote $r(t) \equiv r(t,t) = p(t) - p(0)$, $0 \leq t \leq T$. Assume the asset trades in a frictionless market void of arbitrage opportunities and the number of potential discontinuities (jumps) in the price process per unit time is finite. Then the log-price process p is a semi-martingale (e.g., Back [33]) and therefore the cumulative return $r(t)$ admits the decomposition (e.g., Protter [229])

$$r(t) = \mu(t) + M^C(t) + M^J(t)\,, \tag{6}$$

where $\mu(t)$ is a predictable and finite variation process, $M^c(t)$ a continuous-path infinite-variation martingale, and $M^J(t)$ is a compensated finite activity jump martingale. Over a discrete time interval the decomposition (6) becomes

$$r(t,h) = \mu(t,h) + M^C(t,h) + M^J(t,h)\,, \tag{7}$$

where $\mu(t,h) = \mu(t) - \mu(t-h)$, $M^C(t,h) = M^C(t) - M^C(t-h)$, and $M^J(t,h) = M^J(t) - M^J(t-h)$.

Denote now with $[r,r]$ the quadratic variation of the semi-martingale process r, where (Protter [229])

$$[r,r]_t = r(t)^2 - 2\int r(s-)dr(s)\,, \tag{8}$$

and $r(t-) = \lim_{s \uparrow t} r(s)$. If the finite variation process μ is continuous, then its quadratic variation is identically zero and the predictable component μ in decomposition (7) does not affect the quadratic variation of the return r. Thus, we obtain an expression for the quadratic return variation over the time interval from $t-h$ to t, $0 \leq h \leq t \leq T$ (e.g., Andersen et al. [21] and Barndorff-Nielsen and Shephard [51, 52]):

$$\begin{aligned} QV(t,h) = [r,r]_t - [r,r]_{t-h} &= [M^C, M^C]_t - [M^C, M^C]_{t-h} + \sum_{t-h<s\leq t} \Delta M^2(s) \\ &= [M^C, M^C]_t - [M^C, M^C]_{t-h} + \sum_{t-h<s\leq t} \Delta r^2(s)\,. \end{aligned} \tag{9}$$

Most continuous-time models for asset returns can be cast within the general setting of equation (7), and equation (9) provides a framework to study the model-implied return variance. For instance, the Black and Scholes [63] model is a special case of the setting described by equation (7) in which the conditional mean process μ is constant, the continuous martingale M^C is a standard Brownian motion process, and the jump martingale M^J is identically zero:

$$dp(t) = \mu dt + \sigma dW(t)\,. \tag{10}$$

In this case, the quadratic return variation over the time interval from $t-h$ to t, $0 \leq h \leq t \leq T$, simplifies to

$$QV(t,h) = \int_{t-h}^{t} \sigma^2 ds = \sigma^2\, h\,, \tag{11}$$

that is, return volatility is constant over any time interval of length h.

A second notable example is the jump-diffusion model of Merton [214],

$$dp(t) = (\mu - \lambda\overline{\xi}\,)dt + \sigma dW(t) + \xi(t)\, dq_t\,, \tag{12}$$

where q is a Poisson process uncorrelated with W and governed by the constant jump intensity λ, i.e., $\text{Prob}(dq_t = 1) = \lambda\, dt$. The scaling factor $\xi(t)$ denotes the magnitude of the jump in the return process if a jump occurs at time t. It is assumed to be normally distributed,

$$\xi(t) \rightsquigarrow N(\,\overline{\xi}, \sigma_\xi^2\,)\,. \tag{13}$$

In this case, the quadratic return variation process over the time interval from $t-h$ to t, $0 \leq h \leq t \leq T$ becomes

$$QV(t,h) = \int_{t-h}^{t} \sigma^2 ds + \sum_{t-h\leq s\leq t} J(s)^2 = \sigma^2\, h + \sum_{t-h\leq s\leq t} J(s)^2\,, \tag{14}$$

where $J(t) \equiv \xi(t) dq(t)$ is non-zero only if a jump actually occurs.

Finally, a broad class of stochastic volatility models is defined by

$$dp(t) = \mu(t)dt + \sigma(t)dW(t) + \xi(t)\, dq_t\,, \tag{15}$$

where q is a constant-intensity Poisson process with log-normal jump amplitude (13). Equation (15) is also a special case of (7) and the associated quadratic return variation over the time interval from

$t-h$ to t, $0 \leq h \leq t \leq T$, is

$$\begin{aligned} QV(t,h) &= \int_{t-h}^{t} \sigma(s)^2 ds + \sum_{t-h \leq s \leq t} J(s)^2 \\ &\equiv IV(t,h) + \sum_{t-h \leq s \leq t} J(s)^2 . \end{aligned} \tag{16}$$

As in the general case of equation (9), equation (16) identifies the contribution of diffusive volatility, termed 'integrated variance' (IV), and cumulative squared jumps to the total quadratic variation.

Early applications typically ignored jumps and focused exclusively on the integrated variance component. For instance, IV plays a key role in Hull and White's [174] SV option pricing model, which we discuss in Section IV.1 below along with other option pricing applications. For illustration, we focus here on the SV model specification by Wiggins [256]:

$$dp(t) = \mu dt + \sigma(t)\, dW_p(t) \tag{17}$$

$$d\sigma(t) = f(\sigma(t))dt + \eta\, \sigma(t)\, dW_\sigma(t)\,, \tag{18}$$

where the innovations to the return dp and volatility σ, W_p and W_σ, are standard Brownian motions. If we define $y = \log(\sigma)$ and apply Itô's formula we obtain

$$dy(t) = d\log(\sigma(t)) = \left[-\frac{1}{2}\eta^2 + \frac{f(\sigma(t))}{\sigma(t)} \right] dt + \eta\, dW_\sigma(t)\,. \tag{19}$$

Wiggins approximates the drift term $f(\sigma(t)) \approx \{\alpha + \kappa[\log(\overline{\sigma}) - \log(\sigma(t))]\}\sigma(t)$. Substitution in equation (19) yields

$$d\log(\sigma(t)) = [\,\overline{\alpha} - \kappa \log(\sigma(t))\,]\, dt + \eta\, dW_\sigma(t)\,, \tag{20}$$

where $\overline{\alpha} = \alpha + \kappa \log(\overline{\sigma}) - \frac{1}{2}\eta^2$. Thus, the logarithmic standard deviation process in Wiggins has diffusion dynamics similar in spirit to Taylor's discrete time AR(1) model for the logarithmic information process, equation (4). As in Taylor's model, negative correlation between return and volatility innovations, $\rho = \mathrm{corr}(W_p, W_\sigma) < 0$, generates an asymmetric response of volatility to return shocks similar to the leverage effect in discrete-time EGARCH models.

More recently, several authors have imposed restrictions on the continuous-time SV jump-diffusion (15) that render the model more tractable while remaining consistent with the empirical features of the data. We return to these models in Section IV.1 below.

III. Realized Volatility

Model-free measures of return variation constructed only from concurrent return realizations have been considered at least since Merton [215]. French et al. [148] construct monthly historical volatility estimates from daily return observations. More recently, the increased availability of transaction data has made it possible to refine early measures of historical volatility into the notion of 'realized volatility,' which is endowed with a formal theoretical justification as an estimator of the quadratic return variation as first noted in Andersen and Bollerslev [18]. The realized volatility of an asset

return r over the time interval from $t-h$ to t is

$$RV(t,h;n) = \sum_{i=1}^{n} r\left(t-h+\frac{ih}{n}, \frac{h}{n}\right)^2 . \tag{21}$$

Semi-martingale theory ensures that the realized volatility measure RV converges to the return quadratic variation QV, previously defined in equation (9), when the sampling frequency n increases. We point the interested reader to, e.g., Andersen et al. [19] to find formal arguments in support of this claim. Here we convey intuition for this result by considering the special case in which the asset return follows a continuous-time diffusion without jumps,

$$dp(t) = \mu(t)dt + \sigma(t)\, dW(t)\,. \tag{22}$$

As in equation (21), consider a partition of the $[t-h,t]$ interval with mesh h/n. A discretization of the diffusion (22) over a sub-interval from $\left(t-h+\frac{(i-1)h}{n}\right)$ to $\left(t-h+\frac{ih}{n}\right)$, $i=1,\ldots,n$, yields

$$r\left(t-h+\frac{ih}{n}, \frac{h}{n}\right) \approx \mu\left(t-h+\frac{(i-1)h}{n}\right)\frac{h}{n} + \sigma\left(t-h+\frac{(i-1)h}{n}\right)\Delta W\left(t-h+\frac{ih}{n}\right), \tag{23}$$

where $\Delta W\left(t-h+\frac{ih}{n}\right) = W\left(t-h+\frac{ih}{n}\right) - W\left(t-h+\frac{(i-1)h}{n}\right)$.

Suppressing time indices, the squared return r^2 over the time interval of length h/n is therefore:

$$r^2 = \mu^2\left(\frac{h}{n}\right)^2 + 2\,\mu\,\sigma\,\Delta W\,\left(\frac{h}{n}\right) + \sigma^2(\Delta W)^2\,. \tag{24}$$

As $n\to\infty$ the first two terms vanish at a rate higher than the last one. In particular, to a first order approximation the squared return equals the squared return innovation and therefore the squared return conditional mean and variance are

$$\mathrm{E}\big[\, r^2|\mathcal{F}_t\,\big] \;\approx\; \sigma^2\,\frac{h}{n} \tag{25}$$

$$\mathrm{Var}\big[\, r^2|\mathcal{F}_t\,\big] \;\approx\; 2\,\sigma^4\left(\frac{h}{n}\right)^2 . \tag{26}$$

The no-arbitrage condition implies that return innovations are serially uncorrelated. Thus, summing over $i=1,\ldots,n$ we obtain

$$\begin{aligned}
\mathrm{E}\big[RV(t,h,n)|\,\mathcal{F}_t\,\big] = \sum_{i=1}^{n}\mathrm{E}\left[\, r\left(t-h+\frac{ih}{n},\frac{h}{n}\right)^2 |\,\mathcal{F}_t\,\right] &\approx \sum_{i=1}^{n}\sigma\left(t-h+\frac{(i-1)h}{n}\right)^2\frac{h}{n} \\
&\approx \int_{t-h}^{t}\sigma(s)^2 ds \qquad (27)\\
\mathrm{Var}\big[RV(t,h,n)|\,\mathcal{F}_t\,\big] = \sum_{i=1}^{n}\mathrm{Var}\left[\, r\left(t-h+\frac{ih}{n},\frac{h}{n}\right)^2 |\,\mathcal{F}_t\,\right] &\approx \sum_{i=1}^{n} 2\,\sigma\left(t-h+\frac{(i-1)h}{n}\right)^4\left(\frac{h}{n}\right)^2 \\
&\approx 2\left(\frac{h}{n}\right)\int_{t-h}^{t}\sigma(s)^4 ds\,. \qquad (28)
\end{aligned}$$

Equation (27) illustrates that realized volatility is an unbiased estimator of the return quadratic variation, while equation (28) shows that the estimator is consistent as its variance shrinks to zero

when we increase the sampling frequency n and keep the time interval h fixed. Taken together, these results suggest that RV is a powerful and model-free measure of the return quadratic variation. Effectively, RV gives practical empirical content to the latent volatility state variable underlying the models previously discussed in Section II.2.

Two issues complicate the practical application of the convergence results illustrated in equations (27) and (28). First, a continuum of instantaneous return observations must be used for the conditional variance in equation (28) to vanish. In practice, only a discrete price record is observed, and thus an inevitable discretization error is present. Barndorff-Nielsen and Shephard [52] develop an asymptotic theory to assess the effect of this error on the RV estimate (see also Meddahi [209]). Second, market microstructure effects (e.g., price discreteness, bid-ask spread positioning due to dealer inventory control, and bid-ask bounce) contaminate the return observations, especially at the ultra-high frequency. These effects tend to generate spurious correlations in the return series which can be partially eliminated by filtering the data prior to forming the RV estimates. However, this strategy is not a panacea and much current work studies the optimal sampling scheme and the construction of improved realized volatility in the presence of microstructure noise. This growing literature is surveyed by Hansen and Lunde [165], Bandi and Russell [46], McAleer and Medeiros [205], and Andersen and Benzoni [14]. Recent notable contributions to this literature include Bandi and Russell [45], Barndorff-Nielsen et al. [49], Diebold and Strasser [121], and Zhang, Mykland, and Aït-Sahalia [262]. Related, there is the issue of how to construct RV measures when the market is rather illiquid. One approach is to use a lower sampling frequency and focus on longer-horizon RV measure. Alternatively the literature has explored volatility measures that are more robust to situations in which the noise-to-signal ratio is high, e.g., Alizadeh et al. [8], Brandt and Diebold [72], Brandt and Jones [73], Gallant et al. [151], Garman and Klass [157], Parkinson [221], Schwert [237], and Yang and Zhang [259] consider the high-low price range measure. Dobrev [122] generalizes the range estimator to high-frequency data and shows its link with RV measures.

Equations (27) and (28) also underscore an important difference between RV and other volatility measures. RV is an ex-post model-free estimate of the quadratic variation process. This is in contrast to ex-ante measures which attempt to forecast future quadratic variation using information up to current time. The latter class includes parametric GARCH-type volatility forecasts as well as forecasts built from stochastic volatility models through, e.g., the Kalman filter (e.g., Harvey and Shephard [168], Harvey et al. [167]), the particle filter (e.g., Johannes and Polson [185, 186]) or the reprojection method (e.g., Gallant and Long [152] and Gallant and Tauchen [155]).

More recently, other studies have pursued more direct time-series modeling of volatility to obtain alternative ex-ante forecasts. For instance, Andersen et al. [21] follow an ARMA-style approach, extended to allow for long memory features, to model the logarithmic foreign exchange rate realized volatility. They find the fit to dominate that of traditional GARCH-type models estimated from daily data. In a related development, Andersen, Bollerslev, and Meddahi [24, 25] exploit the general class of Eigenfunction Stochastic Volatility (ESV) models introduced by Meddahi [208] to provide optimal analytic forecast formulas for realized volatility as a function of past realized volatility. Other scholars have pursued more general model specifications to improve forecasting performance. Ghysels et al. [159] consider Mixed Data Sampling (MIDAS) regressions that use a combination of volatility measures estimates at different frequencies and horizons. Related, Engle and Gallo [137] exploit the information

in different volatility measures, captured by a multivariate extension of the multiplicative error model suggested by Engle [136], to predict multi-step volatility. Finally, Andersen et al. [20] build on the Heterogeneous AutoRegressive (HAR) model by Barndorff-Nielsen and Shephard [50] and Corsi [110] and propose a HAR-RV component-based regression to forecast the h-steps ahead quadratic variation:

$$RV(t+h,h) = \beta_0 + \beta_D RV(t,1) + \beta_W RV(t,5) + \beta_M RV(t,21) + \varepsilon(t+h)\,. \tag{29}$$

Here the lagged volatility components $RV(t,1)$, $RV(t,5)$, and $RV(t,21)$ combine to provide a parsimonious approximation to the long-memory type behavior of the realized volatility series, which has been documented in several studies (e.g., Andersen et al. [19]). Simple OLS estimation yields consistent estimates for the coefficients in the regression (29), which can be used to forecast volatility out of sample.

As mentioned previously, the convergence results illustrated in equations (27) and (28) stem from the theory of semi-martingales under conditions more general than those underlying the continuous-time diffusion in equation (22). For instance, these results are robust to the presence of discontinuities in the return path as in the jump-diffusion SV model (15). In this case the realized volatility measure (21) still converges to the return quadratic variation, which is now the sum of the diffusive integrated volatility IV and the cumulative squared jump component:

$$QV(t,h) = IV(t,h) + \sum_{t-h\leq s\leq t} J(s)^2\,. \tag{30}$$

The decomposition in equation (30) motivates the quest for separate estimates of the two quadratic variation components, IV and squared jumps. This is a fruitful exercise in forecasting applications, since separate estimation of the two components increases predictive accuracy (e.g., Andersen et al. [20]). Further, this decomposition is relevant for derivatives pricing, e.g., options are highly sensitive to jumps as well as large moves in volatility (e.g., Pan [220] and Eraker [141]).

A consistent estimate of integrated volatility is the k-skip bipower variation, BV (e.g., Barndorff-Nielsen and Shephard [53]),

$$BV(t,h;k,n) = \frac{\pi}{2}\sum_{i=k+1}^{n} |\, r\left(t-h+\frac{ih}{n},\frac{h}{n}\right)|\,|\, r\left(t-h+\frac{(i-k)h}{n},\frac{h}{n}\right)|\,. \tag{31}$$

Liu and Maheu [202] and Forsberg and Ghysels [147] show that realized power variation, which is robust to the presence of jumps, can improve volatility forecasts. A well-known special case of (31) is the 'realized bipower variation,' which has $k=1$ and is denoted $BV(t,h;n) \equiv BV(t,h;1,n)$. We can combine bipower variation with the realized volatility RV to obtain a consistent estimate of the squared jump component, i.e.,

$$RV(t,h;n) - BV(t,h;n) \underset{n\to\infty}{\longrightarrow} QV(t,h) - IV(t,h) = \sum_{t-h\leq s\leq t} J(s)^2\,. \tag{32}$$

The result in equation (32) are useful to design tests for the presence of jumps in volatility, e.g., Andersen et al. [20], Barndorff-Nielsen and Shephard [53, 54], Huang and Tauchen [172], and Mizrach [217]. More recently, alternative approaches to test for jumps have been developed by Aït-Sahalia and Jacod [6], Andersen et al. [23], Lee and Mykland [195], and Zhang [261].

IV. Applications

The power of the continuous-time paradigm has been evident ever since the work by Merton [212] on intertemporal portfolio choice, Black and Scholes [63] on option pricing, and Vasicek [255] on bond valuation. However, the idea of casting these problems in a continuous-time diffusion context goes back all the way to the work in 1900 by Bachelier [32].

Merton [213] develops a continuous-time general-equilibrium intertemporal asset pricing model which is later extended by Cox et al. [112] to a production economy. Because of its flexibility and analytical tractability, the Cox et al. [112] framework has become a key tool used in several financial applications, including the valuation of options and other derivative securities, the modeling of the term structure of risk-free interest rates, the pricing of foreign currencies and defaultable bonds.

Volatility has played a central role in these applications. For instance, an option's payoff is non-linear in the price of the underlying asset and this feature renders the option value highly sensitive to the volatility of underlying returns. Further, derivatives markets have grown rapidly in size and complexity and financial institutions have been facing the challenge to manage intricate portfolios exposed to multiple risk sources. Risk management of these sophisticated positions hinges on volatility modeling. More recently, the markets have responded to the increasing hedging demands of investors by offering a menu of new products including, e.g., volatility swaps and derivatives on implied volatility indices like the VIX. These innovations have spurred an even more pressing need to accurately measure and forecast volatility in financial markets.

Research has responded to these market developments. We next provide a brief illustrative overview of the recent literature dealing with option pricing and term structure modeling, with an emphasis on the role that volatility modeling has played in these two key applications.

IV.1. Options

Rubinstein [233] and Bates [55], among others, note that prior to the 1987 market crash the Black and Scholes [63] (BS) formula priced option contracts quite accurately whereas after the crash it has been systematically underpricing out-of-the-money equity-index put contracts. This feature is evident from Figure 1, which is constructed from options on the S&P 500 futures. It shows the implied volatility function for near-maturity contracts traded both before and after October 19, 1987 ('Black Monday'). The mild u-shaped pattern prevailing in the pre-crash implied volatilities is labeled a 'volatility smile,' in contrast to the asymmetric post-1987 'volatility smirk.' Importantly, while the steepness and level of the implied volatility curve fluctuate day to day depending on market conditions, the curve has been asymmetric and upward sloping ever since 1987, so the smirk remains in place to the current date, e.g., Benzoni et al. [60]. In contrast, before the crash the implied volatility curve was invariably flat or mildly u-shaped as documented in, e.g., Bates [57]. Finally, we note that the post-1987 asymmetric smirk for index options contrasts sharply with the pattern for individual equity options, which possess flat or mildly u-shaped implied volatility curves (e.g., Bakshi et al. [37] and Bollen and Whaley [65]).

Given the failures of the BS formula, much research has gone into relaxing the underlying assumptions. A natural starting point is to allow volatility to evolve randomly, inspiring numerous studies that examine the option pricing implications of SV models. The list of early contributions includes Hull and White [174], Johnson and Shanno [187], Melino and Turnbull [211], Scott [238], Stein [244],

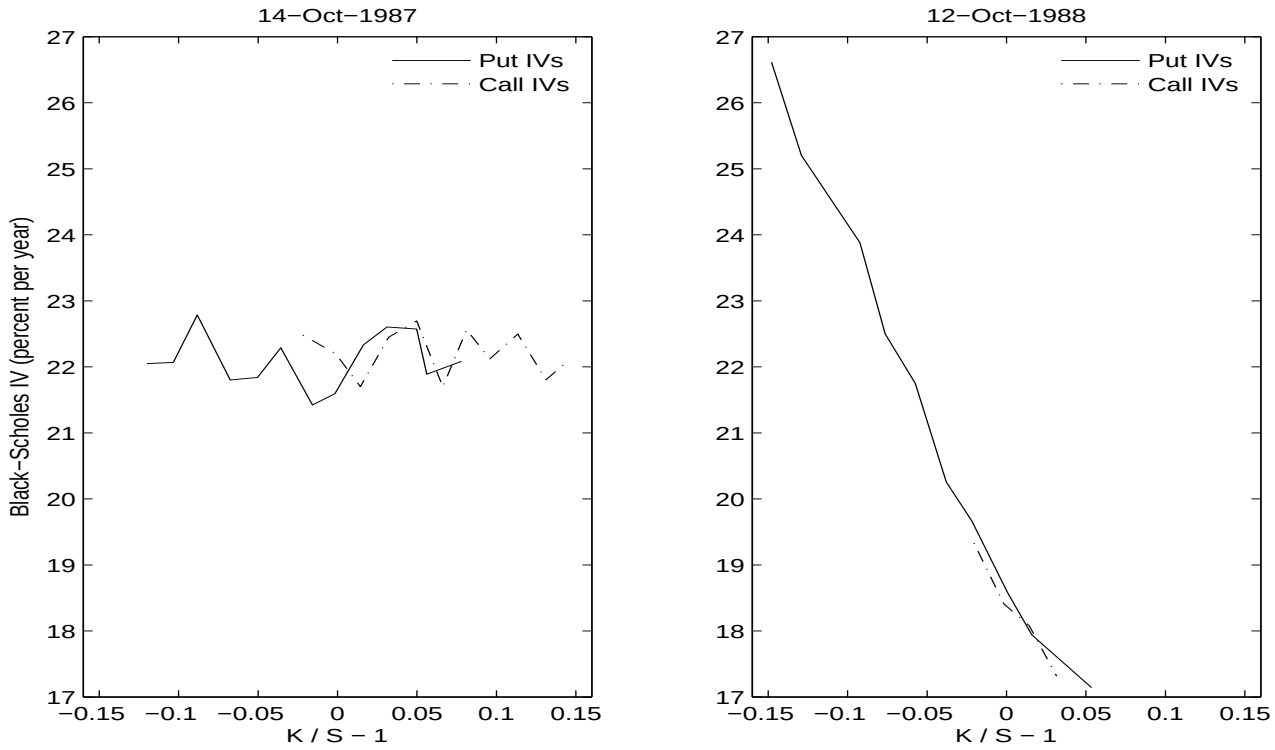

Figure 1: Pre- and Post-1987 Crash Implied Volatilities. The plots depict Black-Scholes implied volatilities computed from near-maturity options on the S&P500 futures on October 14, 1986 (the week before the 1987 market crash) and a year later.

Stein and Stein [245], and Wiggins [256]. Here we focus in particular on the Hull and White [174] model,

$$dp(t) = \mu_p\, dt + \sqrt{V(t)}\, dW_p(t) \tag{33}$$

$$\frac{dV(t)}{V(t)} = \mu_V\, dt + \sigma_V\, dW_V(t)\,, \tag{34}$$

where W_p and W_V are standard Brownian motions. In general, shocks to returns and volatility may be (negatively) correlated, however for tractability Hull and White assume $\rho = \mathrm{corr}(dW_p, dW_V) = 0$. Under this assumption they show that, in a risk-neutral world, the premium C^{HW} on a European call option is the Black and Scholes price C^{BS} evaluated at the average integrated variance $\overline{V}$,

$$\overline{V} = \frac{1}{T-t}\int_t^T V(s)ds\,, \tag{35}$$

integrated over the distribution $h(\overline{V}|V(t))$ of $\overline{V}$:

$$C^{HW}(\,p(t), V(t)\,) = \int C^{BS}(\overline{V})h(\overline{V}|V(t))d\overline{V}\,. \tag{36}$$

The early efforts to identify a more realistic probabilistic model for the underlying return were slowed by the analytical and computational complexity of the option pricing problem. Unlike the BS

setting, the early SV specifications do not admit closed-form solutions. Thus, the evaluation of the option price requires time-consuming computations through, e.g., simulation methods or numerical solution of the pricing partial differential equation by finite difference methods. Further, the presence of a latent factor, volatility, and the lack of closed-form expressions for the likelihood function complicate the estimation problem.

Consequently, much effort has gone into developing restrictions for the distribution of the underlying return process that allow for (semi) closed-form solutions and are consistent with the empirical properties of the data. The 'affine' class of continuous-time models has proven particularly useful in providing a flexible, yet analytically tractable, setting. Roughly speaking, the defining feature of affine jump-diffusions is that the drift term, the conditional covariance term, and the jump intensity are all a linear-plus-constant (affine) function of the state vector. The Vasicek [255] bond valuation model and the Cox et al. [112] intertemporal asset pricing model provide powerful examples of the advantages of the affine paradigm.

To illustrate the progress in option pricing applications built on affine models, consider the return dynamics

$$dp(t) = \mu\, dt + \sqrt{V(t)}\, dW_p(t) + \xi_p(t)\, dq(t) \tag{37}$$

$$dV(t) = \kappa(\overline{V} - V(t))\, dt + \sigma_V \sqrt{V(t)}\, dW_V(t) + \xi_V(t)\, dq(t)\,, \tag{38}$$

where W_p and W_V are standard Brownian motions with non-zero correlation $\rho = \mathrm{corr}(dW_p, dW_V)$, q is a Poisson process, uncorrelated with W_p and W_V, with jump intensity

$$\lambda(t) = \lambda_0 + \lambda_1 V(t)\,, \tag{39}$$

that is, $\mathrm{Prob}(dq_t = 1) = \lambda(t)\, dt$. The jump amplitudes variables ξ_p and ξ_V have distributions

$$\xi_V(t) \rightsquigarrow \exp(\overline{\xi}_V) \tag{40}$$

$$\xi_p(t)\,|\,\xi_V(t) \rightsquigarrow N(\overline{\xi}_p + \rho_\xi\, \xi_V(t), \sigma_p^2)\,. \tag{41}$$

Here volatility is not only stochastic but also subject to jumps which occur simultaneously with jumps in the underlying return process. The Black and Scholes model is a special case of (37)-(41) for constant volatility, $V(t) = \sigma^2, 0 \le t \le T$, and no jumps, $\lambda(t) = 0, 0 \le t \le T$. The Merton [214] model arises from (37)-(41) if volatility is constant but we allow for jumps in returns.

More recently, Heston [170] has considered a special case of (37)-(41) with stochastic volatility but without jumps. Using transform methods he derives a European option pricing formula which may be evaluated readily through simple numerical integration. His SV model has GARCH-type features, in that the variance is persistent and mean reverts at a rate κ to the long-run mean $\overline{V}$. Compared to Hull and White's [174] setting, Heston's model allows for shocks to returns and volatility to be negatively correlated, i.e., $\rho < 0$, which creates a leverage-type effect and skews the return distribution. This feature is consistent with the properties of equity index returns. Further, a fatter left tail in the return distribution results in a higher cost for crash insurance and therefore makes out-of-the-money put options more expensive. This is qualitatively consistent with the patterns in implied volatilities observed after the 1987 market crash and discussed above.

Bates [56] has subsequently extended Heston's approach to allow for jumps in returns and using similar transform methods he has obtained a semi-closed form solution for the option price. The addition of jumps provides a more realistic description of equity returns and has important option pricing implications. With diffusive shocks (e.g., stochastic volatility) alone a large drop in the value of the underlying asset over a short time span is very unlikely whereas a market crash is always possible as long as large negative jumps can occur. This feature increases the value of a short-dated put option, which offers downside protection to a long position in the underlying asset.

Finally, Duffie et al. [130] have introduced a general model with jumps to volatility which embeds the dynamics (37)-(41). In model (37)-(41), the likelihood of a jump to occur increases when volatility is high ($\lambda_1 > 0$) and a jump in returns is accompanied by an outburst of volatility. This is consistent with what is typically observed during times of market stress. As in the Heston case, variance is persistent with a mean reversion coefficient κ towards its *diffusive* long-run mean $\overline{V}$, while the total long-run variance mean is the sum of the diffusive and jump components. In the special case of constant jump intensity, i.e., $\lambda_1 = 0$, the total long-run mean is $\overline{V}+\overline{\xi}_V\lambda_0/\kappa$. The jump term $(\xi_V(t)\, dq(t))$ fattens the right tail of the variance distribution, which induces leptokurtosis in the return distribution. Two effects generate asymmetrically distributed returns. The first channel is the diffusive leverage effect, i.e., $\rho < 0$, the second is the correlation between the volatility and the jump amplitude of returns generated through the coefficient ρ_ξ. Taken together, these effects increase model-implied option prices and help produce a realistic volatility smirk.

Several empirical studies rely on models of the form (37)-(41) in option-pricing applications. For instance, Bates [56] uses Deutsche Mark options to estimate a model with stochastic volatility and constant-intensity jumps to returns, while Bates [57] fits a jump-diffusion model with two SV factors to options on S&P 500 futures. In the latter case, the two SV factors combine to help capture features of the long-run memory in volatility while retaining the analytical tractability of the affine setting (see, e.g., Christoffersen et al. [101] for another model with similar features). Alternative approaches to model long memory in continuous-time SV models rely on the fractional Brownian motion process, e.g., Comte and Renault [108] and Comte et al. [107], while Breidt et al. [76], Harvey [166] and Deo et al. [118] consider discrete-time SV models (see Hurvich et al. [175] for a review). Bakshi et al. [34, 37] estimate a model similar to the one introduced by Bates [56] using S&P 500 options.

Other scholars rely on underlying asset return data alone for estimation. For instance, Andersen et al. [15] and Chernov et al. [95] use equity-index returns to estimate jump-diffusion SV models within and outside the affine (37)-(41) class. Eraker et al. [142] extend this analysis and fit a model that includes constant-intensity jumps to returns and volatility.

Finally, another stream of work examines the empirical implications of SV jump-diffusions using a joint sample of S&P 500 options and index returns. For example, Benzoni [59], Chernov and Ghysels [93], and Jones [189] estimate different flavors of the SV model without jumps. Pan [220] fits a model that has jumps in returns with time-varying intensity, while Eraker [141] extends Pan's work by adding jumps in volatility.

Overall, this literature has established that the SV jump-diffusion model dramatically improves the fit of underlying index returns and options prices compared to the Black and Scholes model. Stochastic volatility alone has a first-order effect and jumps further enhance model performance by generating fatter tails in the return distribution and reducing the pricing error for short-dated options.

The benefits of the SV setting are also significant in hedging applications.

Another aspect related to the specification of SV models concerns the pricing of volatility and jump risks. Stochastic volatility and jumps are sources of uncertainty. It is an empirical issue to determine whether investors demand to be compensated for bearing such risks and, if so, what the magnitude of the risk premium is. To examine this issue it is useful to write model (37)-(41) in so-called risk-neutral form. It is common to assume that the volatility risk premium is proportional to the instantaneous variance, $\eta(t) = \eta_V V(t)$. Further, the adjustment for jump risk is accomplished by assuming that the amplitude $\tilde{\xi}_p(t)$ of jumps to returns has mean $\overline{\tilde{\xi}}_p = \overline{\xi}_p + \eta_p$. These specifications are consistent with an arbitrage-free economy. More general specifications can also be supported in a general equilibrium setting, e.g., a risk adjustment may apply to the jump intensity $\lambda(t)$. However, the coefficients associated to these risk adjustments are difficult to estimate and to facilitate identification they typically are fixed at zero. Incorporating such risk premia in model (37)-(41) yields the following risk-neutral return dynamics (e.g., Pan [220] and Eraker [141]):

$$dp(t) = (r - \mu^*)\,dt + \sqrt{V(t)}\,d\widetilde{W}_p(t) + \tilde{\xi}_p(t)\,dq(t) \tag{42}$$

$$dV(t) = [\,\kappa(\overline{V} - V(t)) + \eta_V\,V(t)\,]\,dt + \sigma_V\,\sqrt{V(t)}\,d\widetilde{W}_V(t) + \xi_V(t)\,dq(t)\,, \tag{43}$$

where r is the risk-free rate, μ^* a jump compensator term, $\widetilde{W}_p$ and $\widetilde{W}_V$ are standard Brownian motions under this so-called $\mathcal{Q}$ measure, and the risk-adjusted jump amplitude variable $\tilde{\xi}_p$ is assumed to follow the distribution,

$$\tilde{\xi}_p(t)\,|\,\xi_V(t) \rightsquigarrow N(\overline{\tilde{\xi}}_p + \rho_\xi\,\xi_V(t), \sigma_p^2)\,. \tag{44}$$

Several studies estimate the risk-adjustment coefficients η_V and η_p for different specifications of model (37)-(44); see, e.g., Benzoni [59], Broadie et al. [78], Chernov and Ghysels [93], Eraker [141], Jones [189], and Pan [220]. It is found that investors demand compensation for volatility and jump risks and these risk premia are important for the pricing of index options. This evidence is reinforced by other studies examining the pricing of volatility risk using less structured but equally compelling procedures. For instance, Coval and Shumway [111] find that the returns on zero-beta index option straddles (i.e., combinations of calls and puts that have offsetting covariances with the index) are significantly lower than the risk-free return. This evidence suggests that in addition to market risk at least a second factor (likely, volatility) is priced in the index option market. Similar conclusions are obtained by Bakshi and Kapadia [36], Buraschi and Jackwerth [79], and Broadie et al. [78].

IV.2. Risk-Free Bonds and their Derivatives

The market for (essentially) risk-free Treasury bonds is liquid across a wide maturity spectrum. It turns out that no-arbitrage restrictions constrain the allowable dynamics in the cross-section of bond yields. Much work has gone into the development of tractable dynamic term structure models capable of capturing the salient time-series properties of interest rates while respecting such cross-sectional no-arbitrage conditions. The class of so-called 'affine' dynamic term structure models provides a flexible and arbitrage-free, yet analytically tractable, setting for capturing the dynamics of the term structure of interest rates. Following Duffie and Kan [129], Dai and Singleton [114, 115], and Piazzesi [227], the short term interest rate, $y_0(t)$, is an affine (i.e., linear-plus-constant) function of a vector of state

variables, $X(t) = \{x_i(t),\ i = 1, \ldots, N\}$:

$$y_0(t) = \delta_0 + \sum_{i=1}^{N} \delta_i\, x_i(t) = \delta_0 + \delta_X' X(t)\,, \tag{45}$$

where the state-vector X has risk-neutral dynamics

$$dX(t) = \tilde{\mathcal{K}}(\tilde{\Theta} - X(t))dt + \Sigma\sqrt{S(t)}d\widetilde{W}(t)\,. \tag{46}$$

In equation (46), $\widetilde{W}$ is an N-dimensional Brownian motion under the so-called $\mathcal{Q}$-measure, $\tilde{\mathcal{K}}$ and $\tilde{\Theta}$ are $N \times N$ matrices, and $S(t)$ is a diagonal matrix with the ith diagonal element given by $[S(t)]_{ii} = \alpha_i + \beta_i' X(t)$. Within this setting, the time-t price of a zero-coupon bond with time-to-maturity τ is given by

$$P(t, \tau) = e^{A(\tau) - B(\tau)' X(t)}\,, \tag{47}$$

where the functions $A(\tau)$ and $B(\tau)$ solve a system of ordinary differential equations (ODEs); see, e.g., Duffie and Kan [129]. Semi-closed form solutions are also available for bond derivatives, e.g., bond options as well as caps and floors (e.g., Duffie et al. [130]).

In empirical applications it is important to also establish the evolution of the state vector X under the physical probability measure $\mathcal{P}$, which is linked to the $\mathcal{Q}$-dynamics (46) through a market price of risk, $\Lambda(t)$. Following Dai and Singleton [114] the market price of risk is often given by

$$\Lambda(t) = \sqrt{S(t)}\lambda\,, \tag{48}$$

where λ is an $N \times 1$ vector of constants. More recently, Duffee [127] proposed a broader 'essentially affine' class, which retains the tractability of standard models but, in contrast to the specification in equation (48), allows compensation for interest rate risk to vary independently of interest rate volatility. This additional flexibility proves useful in forecasting future yields. Subsequent generalization are in Duarte [124] and Cheridito et al. [92].

Litterman and Scheinkman [201] demonstrate that virtually all variation in U.S. Treasury rates is captured by three factors, interpreted as changes in 'level,' 'steepness,' and 'curvature.' Consistent with this evidence, much of the term-structure literature has focused on three-factor models. One problem with these models, however, is that the factors are latent variables void of immediate economic interpretation. Therefore, it is challenging to impose appropriate identifying conditions for the model coefficients and in particular to find the ideal representation for the 'most flexible' model, i.e., the model with the highest number of identifiable coefficients. Dai and Singleton [114] conduct an extensive specification analysis of multi-factor affine term structure models. They classify these models into subfamilies according to the number of (independent linear combination of) state variables that determine the conditional variance matrix of the state vector. Within each subfamily, they proceed to identify the models that lead to well-defined bond prices (a condition they label 'admissibility') and among the admissible specifications they identify a 'maximal' model that nests econometrically all others in the subfamily. Joslin [190] builds on Dai and Singleton's [114] work by pursuing identification through a normalization of the drift term in the state vector dynamics (instead of the diffusion term, as in Dai and Singleton [114]). Duffie and Kan [129] follow an alternative approach to obtain an identifiable model by rotating from a set of latent state variables to a set of observable zero-coupon

yields. Collin-Dufresne et al. [104] build on the insights of both Dai and Singleton [114] and Duffie and Kan [129]. They perform a rotation of the state vector into a vector that contains the first few components in the Taylor series expansion of the yield curve around a maturity of zero and their quadratic variation. One advantage is that the elements of the rotated state vector have an intuitive and unique economic interpretation (such as level, slope, and curvature of the yield curve) and therefore the model coefficients in this representation are identifiable. Further, it is easy to construct a model-independent proxy for the rotated state vector, which facilitates model estimation as well as interpretation of the estimated coefficients across models and sample periods.

This discussion underscores an important feature of affine term structure models. The dependence of the conditional factor variance $S(t)$ on one or more of the elements in X introduces stochastic volatility in the yields. However, when a square-root factor is present parametric restrictions (admissibility conditions) need to be imposed so that the conditional variance $S(t)$ is positive over the range of X. These restrictions affect the correlations among the factors which, in turn, tend to worsen the cross-sectional fit of the model. Specifically, CIR models in which $S(t)$ depends on all the elements of X require the conditional correlation among the factors to be zero, while the admissibility conditions imposed on the matrix $\mathcal{K}$ renders the unconditional correlations non-negative. These restrictions are not supported by the data. In contrast, constant-volatility Gaussian models with no square-root factors do not restrict the signs and magnitude of the conditional and unconditional correlations among the factors but they do, of course, not accommodate the pronounced and persistent volatility fluctuations observed in bond yields. The class of models introduced by Dai and Singleton [114] falls between these two extremes. By including both Gaussian and square-root factors they allow for time-varying conditional volatilities of the state variables and yet they do not constrain the signs of some of their correlations. This flexibility helps to address the trade off between generating realistic correlations among the factors while capturing the time-series properties of the yields' volatility.

A related aspect of (unconstrained) affine models concerns the dual role that square-root factors play in driving the time-series properties of yields' volatility and the term structure of yields. Specifically, the time-t yield $y_\tau(t)$ on a zero-coupon bond with time-to-maturity τ is given by

$$P(t,\tau) = e^{-\tau\, y_\tau(t)}\,. \tag{49}$$

Thus, we have

$$y_\tau(t) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)'}{\tau}\, X(t)\,. \tag{50}$$

It is typically assumed that the B matrix has full rank and therefore equation (50) provides a direct link between the state-vector $X(t)$ and the term-structure of bond yields. Further, Itô's Lemma implies that the yield y_τ also follows a diffusion process:

$$dy_\tau(t) = \mu_{y_\tau}(X(t),t)\,dt + \frac{B(\tau)'}{\tau}\,\Sigma\,\sqrt{S(t)}d\widetilde{W}(t)\,. \tag{51}$$

Consequently, the (instantaneous) quadratic variation of the yield given as the squared yield volatility coefficient for y_τ is

$$V_{y_\tau}(t) = \frac{B(\tau)'}{\tau}\,\Sigma\,S(t)\,\Sigma'\,\frac{B(\tau)}{\tau}\,. \tag{52}$$

The elements of the $S(t)$ matrix are affine in the state vector $X(t)$, i.e., $[S(t)]_{ii} = \alpha_i + \beta_i' X(t)$. Further, invoking the full rank condition on $B(\tau)$, equation (50) implies that each state variable in the vector $X(t)$ is an affine function of the bond yields $Y(t) = \{y_{\tau_j}(t),\ j = 1, \ldots, J\}$. Thus, for any τ there is a set of constants $a_{\tau,j}$, $j = 0, \ldots, J$, so that

$$V_{y_\tau}(t) = a_{\tau,0} + \sum_{j=1}^{J} a_{\tau,j}\, y_{\tau_j}(t)\,. \tag{53}$$

Hence, the current quadratic yield variation for bonds at any maturity is a linear combination of the term structure of yields. Therefore, the market is complete, i.e., volatility is perfectly spanned by a portfolio of bonds.

Collin-Dufresne and Goldstein [103] note that this spanning condition is unnecessarily restrictive and propose conditions which ensures that volatility no longer directly enters the main bond pricing equation (47). This restriction, which they term 'unspanned stochastic volatility' (USV), effectively breaks the link between the yields' quadratic variation and the level of the term structure by imposing a reduced rank condition on the $B(\tau)$ matrix. Further, since their model is a special (nested) case of the affine class it retains the analytical tractability of the affine model class. Recently Joslin [190] has derived more general conditions for affine term structure models to exhibit USV. His restrictions also produce a market incompleteness (i.e., volatility cannot be hedged using a portfolio of bonds) but do not constrain the degree of mean reversion of the other state variables so that his specification allows for more flexibility in capturing the persistence in interest rate series. (See also the USV conditions in the work by Trolle and Schwartz [253]).

There is conflicting evidence on the volatility spanning condition in fixed income markets. Collin-Dufresne and Goldstein (2002) find that swap rates have limited explanatory power for returns on at-the-money 'straddles,' i.e., portfolios mainly exposed to volatility risk. Similar findings are in Heidari and Wu [169], who show that the common factors in LIBOR and swap rates explain only a limited part of the variation in the swaption implied volatilities. Moreover, Li and Zhao [197] conclude that some of the most sophisticated multi-factor dynamic term structure models have serious difficulties in hedging caps and cap straddles, even though they capture bond yields well. In contrast, Fan et al. [143] argue that swaptions and even swaption straddles can be well hedged with LIBOR bonds alone, supporting the notion that bond markets are complete.

More recently other studies have examined several versions of the USV restriction, again coming to different conclusions. A direct comparison of these results, however, is complicated by differences in the model specification, the estimation method, and the data and sample period used in the estimation. Collin-Dufresne et al. [105] consider swap rates data and fit the model using a Bayesian Markov Chain Monte Carlo method. They find that a standard three-factor model generates a time series for the variance state variable that is essentially unrelated to GARCH estimates of the quadratic variation of the spot rate process or to implied variances from options, while a four-factor USV model generates both realistic volatility estimates and a good cross-sectional fit. In contrast, Jacobs and Karoui [177] consider a longer data set of U.S. Treasury yields and pursue quasi-maximum likelihood estimation. They find the correlation between model-implied and GARCH volatility estimates to be high. However, when estimating the model with a shorter sample of swap rates, they find such correlations to be small or negative. Thompson [250] explicitly tests the Collin-Dufresne and Goldstein [103] USV restriction

and rejects it using swap rates data. Bikbov and Chernov [62], Han [164], Jarrow et al. [182], Joslin [191], and Trolle and Schwartz [254] rely on data sets of derivatives prices and underlying interest rates to better identify the volatility dynamics.

Andersen and Benzoni [12] directly relate model-free realized volatility measures (constructed from high-frequency U.S. Treasury data) to the cross-section of contemporaneous bond yields. They find that the explanatory power of such regressions is very limited, which indicates that volatility is not spanned by a portfolio of bonds. The evidence in Andersen and Benzoni [12] is consistent with the USV models of Collin-Dufresne et al. [105] and Joslin [190], as well as with a model that embeds weak dependence between the yields and volatility as in Joslin [191]. Moreover, Duarte [125] argues that the effects of mortgage-backed security hedging activity affects both the interest rate volatility implied by options and the actual interest rate volatility. This evidence suggests that variables that are not in the span of the term structure of yields and forward rates contribute to explain volatility in fixed income markets. Also related, Wright and Zhou [258] find that adding a measure of market jump volatility risk to a regression of excess bond returns on the term structure of forward rates nearly doubles the R^2 of the regression. Taken together, these findings suggest more generally that genuine SV models are critical for appropriately capturing the dynamic evolution of the term structure.

V. Estimation Methods

There are a very large number of alternative approaches to estimation and inference for parametric SV models and we abstain from a thorough review. Instead, we point to the basic challenges that exist for different types of specifications, how some of these were addressed in the early literature and finally provide examples of methods that have been used extensively in recent years. Our exposition continues to focus on applications to equity returns, interest rates, and associated derivatives.

Many of the original SV models were cast in discrete time, inspired by the popular GARCH paradigm. In that case, the distinct challenge for SV models is the presence of a strongly persistent latent state variable. However, more theoretically oriented models, focusing on derivatives applications, were often formulated in continuous time. Hence, it is natural that the econometrically-oriented literature has moved in this direction in recent years as well. This development provides an added complication as the continuous-time parameters must be estimated from discrete return data and without direct observations on volatility. For illustration, consider a fully parametric continuous-time SV model for the asset return r with conditional variance V and coefficient vector Ψ. Most methods to estimate Ψ rely on the conditional density f for the data generating process,

$$f(r(t), V(t) \,|\, \mathcal{I}(t-1), \Psi) = f_{r|V}(r(t) \,|\, V(t), \mathcal{I}(t-1), \Psi) \times f_V(V(t) \,|\, \mathcal{I}(t-1), \Psi)\,, \tag{54}$$

where $\mathcal{I}(t-1)$ is the available information set at time $t-1$. The main complications are readily identified. First, analytic expressions for the discrete-time transition (conditional) density, f, or the discrete-time moments implied by the data generating process operating in continuous time, are often unavailable. Second, volatility is latent in SV models, so that even if a closed-form expression for f is known, direct evaluation of the above expression is infeasible due to the absence of explicit volatility measures. The marginal likelihood with respect to the observable return process alone is obtained

by integrating over all possible paths for the volatility process, but this integral has a dimension corresponding to sample size, rendering the approach infeasible in general.

Similar issues are present when estimating continuous-time dynamic term structure models. Following Piazzesi [228], a change of variable gives the conditional density for a zero-coupon yield y on a bond with time to maturity τ:

$$f(y_\tau(t)\,|\,\mathcal{I}(t-1),\Psi) = f_X(g(y_\tau(t),\Psi)\,|\,\mathcal{I}(t-1),\Psi) \times |\bigtriangledown_y g(y_\tau(t),\Psi)|\,. \tag{55}$$

Here the latent state vector X has conditional density f_X, the function $g(\cdot,\Psi)$ maps the observable yield y into X, $X(t) = g(y_\tau(t),\Psi)$, and $\bigtriangledown_y g(y_\tau(t),\Psi)$ is the Jacobian determinant of the transformation. Unfortunately, analytic expressions for the conditional density f_X are known only in some special cases. Further, the mapping $X(t) = g(y_\tau(t),\Psi)$ holds only if the model provides an exact fit to the yields, while in practice different sources of error (e.g., model mis-specification, microstructure effects, measurement errors) inject a considerable degree of noise into this otherwise deterministic linkage (for correct model specification) between the state vector and the yields. As such, a good measure of X might not be available to evaluate the conditional density (55).

V.1. Estimation via Discrete-Time Model Specification or Approximation

The first empirical studies have estimated discrete-time SV models via a (Generalized) Method of Moments procedure by matching a number of theoretical and sample moments, e.g., Chan et al. [89], Ho et al. [171], Longstaff and Schwartz [204], and Melino and Turnbull [211]. These models were either explicitly cast in discrete time or were seen as approximate versions of the continuous-time process of interest. Similarly, several authors estimate diffusive affine dynamic term structure models by approximating the continuous-time dynamics with a discrete-time process. If the error terms are stipulated to be normally distributed, the transition density of the discretized process is multivariate normal and computation of unconditional moments then only requires knowledge of the first two moments of the state vector. This result facilitates quasi-maximum likelihood estimation. In evaluating the likelihood function, some studies suggest using closed-form expressions for the first two moments of the continuous-time process instead of the moments of the discretized process (e.g., Fisher and Gilles [145] and Duffee [127]), thus avoiding the associated discretization bias. This approach typically requires some knowledge of the state of the system which may be obtained, imperfectly, through matching the system, given the estimated parameter vector, to a set of observed zero-coupon yields to infer the state vector X. A modern alternative is to use the so-called particle filter as an efficient filtering procedure for the unobserved state variables given the estimated parameter vector. We provide more detailed accounts of both of these procedures later in this section.

Finally, a number of authors develop a simulated maximum likelihood method that exploit the specific structure of the discrete-time SV model. Early examples are Danielsson and Richard [117] and Danielsson [116] who exploit the Accelerated Gaussian Importance Sampler for efficient Monte Carlo evaluation of the likelihood. Subsequent improvements were provided by Fridman and Harris [149] and Liesenfeld and Richard [200], with the latter relying on Efficient Importance Sampling (EIS). In a second step, EIS can also be used for filtering the latent volatility state vector. In general, these inference techniques provide quite impressive efficiency but the methodology is not always easy

to generalize beyond the structure of the basic discrete-time SV asset return model. We discuss the general inference problem for continuous-time SV models for which the lack of a closed-form expression for the transition density is an additional complicating factor in a later section.

V.2. Filtering the Latent State Variable Directly During Estimation

Some early studies focused on direct ways to extract estimates of the latent volatility state variable in discrete-time SV asset return models. The initial approach was based on quasi-maximum likelihood (QML) methods exploiting the Kalman filter. This method requires a transformation of the SV model to a linear state-space form. For instance, Harvey and Shephard [168] consider a version of the Taylor's [249] discrete-time SV model,

$$\begin{aligned} p(t) &= p(t-1) + \beta + \sqrt{V(t)}\varepsilon(t) && (56) \\ \log(V(t)) &= \alpha + \phi \log(V(t-1)) + \eta(t)\,, && (57) \end{aligned}$$

where p is the logarithmic price, ε is a zero-mean error term with unit variance, and η is an independently-distributed error term with zero mean and variance σ_η^2.

Define $y(t) = p(t) - p(t-1) - \beta$, square the observations in equation (56), and take logarithms to obtain the *measurement equation*,

$$\ell(t) = \omega + h(t) + \xi(t)\,, \tag{58}$$

where $\ell(t) \equiv \log y(t)^2$, $h(t) \equiv \log(V(t))$. Further, ξ is a zero-mean disturbance term given by $\xi(t) = \log(\varepsilon(t)^2) - \mathrm{E}[\log(\varepsilon(t)^2)]$, $\omega = \log(\sigma^2) + \mathrm{E}[\log(\varepsilon(t)^2)]$, and σ is a scale constant which subsumes the effect of the drift term α in equation (57). The autoregression (57) yields the *transition equation*,

$$h(t) = \phi h(t-1) + \eta(t)\,, \tag{59}$$

Taken together, equations (58) and (59) are the linear state-space transformation of the SV model (56)-(57). If the joint distribution of ε and η is symmetric, i.e., $f(\varepsilon, \eta) = f(-\varepsilon, -\eta)$, then the disturbance terms in the state-space form are uncorrelated even if η and ε are not. A possible dependence between ε and η allows the model to pick up some of the asymmetric behavior often observed in stock returns. Projection of $[\,h(t) - \mathrm{E}_{t-1}h(t)\,]$ over $[\,\ell(t) - \mathrm{E}_{t-1}\ell(t)\,]$ yields the Kalman filter estimate of the latent (logarithmic) variance process:

$$\mathrm{E}_t\, h(t) = \mathrm{E}_{t-1}h(t) + \frac{\mathrm{E}\{\,[\,h(t) - \mathrm{E}_{t-1}h(t)\,] \times [\,\ell(t) - \mathrm{E}_{t-1}\ell(t)\,]\,\}}{\mathrm{E}\{\,[\,\ell(t) - \mathrm{E}_{t-1}\ell(t)\,]^2\}} \times [\,\ell(t) - \mathrm{E}_{t-1}\ell(t)\,]\,, \tag{60}$$

where the conditional expectations $\mathrm{E}_{t-1}\ell(t)$ and $\mathrm{E}_{t-1}h(t)$ are given by:

$$\begin{aligned} \mathrm{E}_{t-1}\ell(t) &= \omega + \mathrm{E}_{t-1}h(t) && (61) \\ \mathrm{E}_{t-1}h(t) &= \phi\, \mathrm{E}_{t-1}\, h(t-1)\,. && (62) \end{aligned}$$

To start the recursion (60)-(62), the initial value $\mathrm{E}_0\, h(0)$ is fixed at the long-run mean $\log(\overline{V})$.

Harvey and Shephard [168] estimate the model coefficients via quasi-maximum likelihood, i.e. by treating the errors ξ and η as though they were normal and maximizing the prediction-error decomposition form of the likelihood function obtained via the Kalman filter. Inference is valid as

long as the standard errors are appropriately adjusted. In their application they rely on daily returns on the value-weighted U.S. market index over 1967-1987 and daily returns for 30 individual stocks over 1974-1983. Harvey et al. [167] pursue a similar approach to fit a multivariate SV model to a sample of four exchange rate series from 1981 to 1985. One major drawback of the Kalman filter approach is that the finite sample properties can be quite poor because the error term, ξ, is highly non-Gaussian, see, e.g., Andersen, Chung, and Sørensen [27]. The method may be extended to accommodate various generalizations including long memory persistence in volatility as detailed in Ghysels, Harvey, and Renault [158].

A related literature, often exploited in multivariate settings, specifies latent GARCH-style dynamics for a state vector which governs the systematic evolution of a higher dimensional set of asset returns. An early representative of these specifications is in Diebold and Nerlove [120], who exploit the Kalman filter for estimation, while Fiorentina et al. [144] provide a likelihood-based estimation procedure using MCMC techniques. We later review the MCMC approach in some detail.

The state-space form is also useful to characterize the dynamics of interest rates. Following, e.g., Piazzesi [227], for a discrete-time dynamic term structure model the measurement and transition equations are

$$y_\tau(t) = -\frac{A(\tau)}{\tau} + \frac{B(\tau)'}{\tau} X(t) + \xi_\tau(t) \tag{63}$$

$$X(t) = \mu + \Phi X(t-1) + \Sigma\sqrt{S(t)}\varepsilon(t)\,, \tag{64}$$

where $S(t)$ is a matrix whose elements are affine functions of the state vector X, and A and B solve a system of difference equations. When all the yields are observed with error (i.e., $\xi_\tau \neq 0\ \forall\tau$, $0 \leq \tau \leq T$), QML estimation of the system (63)-(64) via the extended Kalman filter method yields an estimate of the coefficient vector. Applications of this approach for the U.S. term structure data include Campbell and Viceira [81], Gong and Remolona [161], and Pennacchi [225]. The extended Kalman filter involves a linear approximation of the relation between the observed data and the state variables, and the associated approximation error will produce biased estimates. Christoffersen et al. [99] raise this concern and recommend the use of the so-called unscented Kalman filter for estimation of systems in which the relation between data and state variables is highly non-linear, e.g., options data.

V.3. Methods Accommodating the Lack of a Closed-Form Transition Density

We have so far mostly discussed estimation techniques for models with either a known transition density or one that is approximated by a discrete-time system. However, the majority of empirically-relevant continuous-time models do not possess explicit transition densities and alternative approaches are necessary. This problem leads us naturally towards the large statistics and econometric literature on estimation of diffusions from discretely-observed data. The vast majority of these studies assume that all relevant variables are observed so the latent volatility or yield curve state variables, integral to SV models, are not accounted for. Nonetheless, it may be feasible to extract the requisite estimates of the state variable by alternate means, thus restoring the feasibility, albeit not efficiency, of the basic approach. Since the literature is large and not directly geared towards genuine SV models, we focus on methods that have seen use in applications involving latent state variables.

A popular approach is to invert the map between the state vector and a subset of the observables assuming that the model prices specific securities exactly. In applications to equity markets this is done, e.g., by assuming that one option contract is priced without error, which implies a specific value (estimate) of the variance process given the model parameters Ψ. For instance, Pan [220] follows this approach in her study of S&P 500 options and returns, which we review in more detail in Section V.5. In applications to fixed income markets it is likewise stipulated that certain bonds are priced without error, i.e., in equation (63) the error term $\xi_{\tau_i}(t)$ is fixed at zero for a set of maturities $\tau_1, \ldots, \tau_N$, where N matches the dimension of the state vector X. This approach yields an estimate for the latent variables through the inverse-map $X(t) = g(y_\tau(t), \Psi)$.

One criticism of the state vector inversion procedure is that it requires ad hoc assumptions regarding the choice of the securities that are error-free (those used to compute model-implied measures of the state vector) vis-a-vis those observed with error (used either for estimation or to assess model performance in an 'out-of-sample' cross-sectional check). In fact, the extracted state vector can be quite sensitive to the choice of derivatives (or yields) used. Nevertheless, this approach has intuitive appeal. Model-implied measures of the state vector, in combination with a closed-form expression for the conditional density (55), allow for efficient estimation of the coefficient vector Ψ via maximum likelihood. Analytic expressions for f_X in equation (55) exist in a limited number of cases. For instance, if X is Gaussian then f_X is multivariate normal, while if X follows a square-root process then f_X can be expressed in terms of the modified Bessel function (e.g., Cox et al. [113]). Different flavors of these continuous-time models are estimated in, e.g., Chen and Scott [91], Collin-Dufresne and Solnik [106], Duffie and Singleton [132], Jagannathan et al. [181], and Pearson and Sun [223]. In more general cases, including affine processes that combine Gaussian and square-root state variables, closed-form expressions for f_X are no longer available. In the rest of this section we briefly review different methods to overcome this problem. The interested reader may consult, e.g., Piazzesi [227] for more details.

Lo [203] warns that the common approach of estimating parameters of an Itô process by applying maximum likelihood to a discretization of the stochastic differential equation yields inconsistent estimators. In contrast, he characterizes the likelihood function as a solution to a partial differential equation. The method is very general, e.g., it applies not only to continuous-time diffusions but also to jump processes. In practice, however, analytic solutions to the partial differential equations (via, e.g., Fourier transforms) are available only for a small class of models so computationally-intensive methods (e.g., finite differencing or simulations) are generally required to solve the problem. This is a severe limitation in the case of multivariate systems like SV models.

For general Markov processes, where the above solution is infeasible, a variety of procedures have been advocated in recent years. Three excellent surveys provide different perspectives on the issue. Aït-Sahalia, Hansen, and Scheinkman [5] discuss operator methods and mention the potential of applying a time deformation technique to account for genuine SV features of the process, as in Conley, Hansen, Luttmer, and Scheinkman [109]. In addition, the Aït-Sahalia [3, 4] closed-form polynomial expansions for discretely-sampled diffusions are reviewed along with the Schaumburg [235] extension to a general class of Markov processes with Lévy-type generators. Meanwhile, Bibby, Jacobsen, and Sørensen [61] survey the extensive statistics literature on estimating functions for diffusion-type models and Bandi and Phillips [42] explicitly consider dealing with nonstationary processes (see also the work of Bandi

[39], Bandi and Nguyen [41], and Bandi and Phillips [43, 44]).

The characteristic function based inference technique has been particularly widely adopted due to the natural fit with the exponentially affine model class which provides essentially closed-form solutions for many pricing applications. Consequently, we dedicate a separate section to this approach.

V.3.1. Characteristic Functions

Singleton [242] proposes to exploit the information contained in the conditional characteristic function of the state vector X,

$$\phi(iu, X(t), \Psi) = \mathrm{E}\big[\, e^{iu'X(t+1)} \big| X(t) \big]\,, \tag{65}$$

to pursue maximum likelihood estimation of affine term structure models. In equation (65) we highlight the dependence of the characteristic function on the unknown parameter vector Ψ. When X is an affine (jump-)diffusion process, ϕ has the exponential affine form,

$$\phi(iu, X(t), \Psi) = e^{\alpha_t(u) + \beta_t(u)'X(t)}\,, \tag{66}$$

where the functions α and β solve a system of ODEs. Therefore, the transition density f_X is known explicitly up to an inverse-Fourier transformation of the characteristic function (65),

$$f_X(X(t+1)\big|X(t);\Psi) = \frac{1}{\pi^N} \int_{\mathrm{R}_+^N} \mathrm{Re}\big[\, e^{-iu'X(t+1)}\, \phi(iu, X(t), \Psi)\,\big]\, du\,. \tag{67}$$

Singleton shows that Gauss-Legendre quadrature with a relatively small number of quadrature points allows to accurately evaluate the integral in equation (67) when X is univariate. As such, the method readily delivers efficient estimates of the parameter vector, Ψ, subject to an auxiliary assumption, namely that the state vector may be extracted by assuming that a pre-specified set of security prices is observed without error while the remainder have non-trivial error terms.

When X is multivariate the Fourier inversion in equation (67) is computationally more demanding. Thus, when estimating multi-dimensional systems Singleton suggests focusing on the conditional density function of the individual elements of X, but conditioned on the full state vector,

$$f_{X_j}(X_j(t+1)|X(t);\Psi) = \frac{1}{2\pi} \int_{\mathrm{R}} e^{-i\omega \mathbf{I}_j' X(t+1)} \phi(i\omega\, \mathbf{I}_j, X(t), \Psi) d\omega\,, \tag{68}$$

where the vector $\mathbf{I}_j$ has 1 in the jth element and zero elsewhere so that the jth element of X is $X_j(t+1) = \mathbf{I}_j' X(t+1)$. Maximization of the likelihood function obtained from f_{X_j}, for a fixed j, will often suffice to obtain a consistent estimate of Ψ. Exploiting more than one of the conditional densities (68) will result in more efficient Ψ estimate. For instance, the scores of multiple univariate log-likelihood functions, stacked in a vector, yield moment conditions that allow for generalized method of moment (GMM) estimation of the system. Alternatively, Joslin [191] proposes a change-of-measure transformation which reduces the oscillatory behavior of the integrand in equation (67). When using this transformation, Gauss-Hermite quadrature more readily provides a solution to the integral in (67) even if the state vector X is multi-dimensional, thus facilitating full ML estimation of the system.

Related, several studies have pursued GMM estimation of affine processes using characteristic functions. Definition (65) yields the moment condition

$$\mathrm{E}\big[(\phi(iu, X(t), \Psi) - e^{iu'X(t+1)}) z(u, X(t))\big] = 0\,, \tag{69}$$

where X is an N-dimensional (jump-)diffusion, $u \in \mathrm{R}^N$, and z is an instrument function. When X is affine, the characteristic function takes the exponential form (66). Different choices of u and z yield a set of moment conditions that can be used for GMM estimation and inference. Singleton [242] derives the optimal instrument in terms of delivering efficient estimates. Carrasco et al. [86] approximate the optimal instrument with a set of basis functions that do not require the knowledge of the conditional likelihood function, thus avoiding one of the assumptions invoked by Singleton. Further, they build on Carrasco and Florens [87] to implement estimation using a continuum of moment conditions, which yields maximum-likelihood efficiency. Other applications of GMM-characteristic function methods to affine (jump-) diffusions for equity index returns are in Chacko and Viceira [88] and Jiang and Knight [183].

In some cases the lack of closed-form expressions for the moment condition in equation (69) can hinder GMM estimation. In these cases the expectation in equation (69) can be evaluated by Monte Carlo integration. This is accomplished by simulating a long sample from the discretized process for a given value of the coefficient vector Ψ. The parameter Ψ is then estimated via the simulated method of moments (SMM) of McFadden [206] and Duffie and Singleton [131]. Singleton [242] proposes SMM characteristic function estimators that exploit the special structure of affine term structure models.

V.4. Efficient Estimation of General Continuous-Time Processes

A number of recent approaches offer excellent flexibility in terms of avoiding approximations to the continuous-time model-implied transition density while still facilitating efficient estimation of the evolution of the latent state vector for the system.

V.4.1. Maximum Likelihood with Characteristic Functions

Bates [58] develops a filtration-based maximum likelihood estimation method for affine processes. His approach relies on Bayes' rule to recursively update the joint characteristic function of latent variables and data conditional on past data. He then obtains the transition density by Fourier inversion of the updated characteristic function.

Denote with $y(t)$ and $X(t)$ the time-t values of the observable variable and the state vector, respectively, and let $Y(t) \equiv \{y(1), \ldots, y(t)\}$ be the data observed up to time t. Consider the case in which the characteristic function of $z(t+1) \equiv (y(t+1), X(t+1))$ conditional on $z(t) \equiv (y(t), X(t))$, is an exponential affine function of $X(t)$:

$$\begin{aligned} \phi(is, iu, z(t), \Psi) &= \mathrm{E}\big[\, e^{is' y(t+1) + iu' X(t+1)} \big| z(t) \,\big] \\ &= e^{\alpha(is,iu,y(t)) + \beta(is,iu,y(t))' X(t)}, \end{aligned} \tag{70}$$

Next, determine the value of the characteristic function conditional on the observed data $Y(t)$:

$$\begin{aligned} \phi(is, iu, Y(t), \Psi) &= \mathrm{E}\Big[\mathrm{E}\big[\, e^{is' y(t+1) + iu' X(t+1)} \big| z(t) \,\big] \Big| \, Y(t) \Big] \\ &= \mathrm{E}\Big[e^{\alpha(is,iu,y(t)) + \beta(is,iu,y(t))' X(t)} \,\Big|\, Y(t) \Big] \\ &= e^{\alpha(is,iu,y(t))}\, \psi(\beta(is, iu, y(t)), Y(t), \Psi)\,, \end{aligned} \tag{71}$$

where $\psi(iu, Y(t), \Psi) \equiv \mathrm{E}\left[e^{iu'X(t)} \,\middle|\, Y(t)\right]$ denotes the (marginal) characteristic function for the state vector conditional on the observed data. Fourier inversion then yields the conditional density for the observation $y(t+1)$ conditional on $Y(t)$:

$$f_y(y(t+1) \,|\, Y(t); \Psi) = \frac{1}{2\pi} \int_{\mathrm{R}} e^{-is'y(t+1)} \phi(is, 0, Y(t), \Psi) ds \,. \tag{72}$$

The next step updates the characteristic function ψ (Bartlett [48]):

$$\psi(iu, Y(t+1), \Psi) = \frac{1}{2\pi f_y(y(t+1) \,|\, Y(t); \Psi)} \int_{\mathrm{R}} e^{-is'y(t+1)} \phi(is, iu, Y(t), \Psi) ds \,. \tag{73}$$

To start the recursion, Bates initializes ψ at the unconditional characteristic function of the latent variable X. The log-likelihood function is then given by

$$\log \mathcal{L}(Y(T); \Psi) = \log(f_y(y(1); \Psi) + \sum_{t=2}^{T} \log(f_y(y(t) \,|\, Y(t-1); \Psi)) \,. \tag{74}$$

A nice feature is that the method provides a natural solution to the filtering problem. The filtered estimate of the latent state X and its variance are computed from the first and second derivatives of the moment generating function $\psi(u, Y(t); \Psi)$ in equation (73), evaluated at $u = 0$:

$$\mathrm{E}[\, X(t+1) \,|\, Y(t+1); \Psi \,] = \frac{1}{2\pi f_y(y(t+1) \,|\, Y(t); \Psi)} \int_{\mathrm{R}} e^{-is'y(t+1)} \phi_u(is, 0, Y(t); \Psi) ds \tag{75}$$

$$\begin{aligned} \mathrm{Var}[\, X(t+1) \,|\, Y(t+1); \Psi \,] &= \frac{1}{2\pi f_y(y(t+1) \,|\, Y(t); \Psi)} \int_{\mathrm{R}} e^{-is'y(t+1)} \phi_{uu}(is, 0, Y(t); \Psi) ds \\ &\quad - \{\mathrm{E}[\, X(t+1) \,|\, Y(t+1) \,]\}^2 \,. \end{aligned} \tag{76}$$

A drawback is that at each step t of the iteration the method requires storage of the entire characteristic function $\psi(iu, Y(t); \Psi)$. To deal with this issue Bates recommends to approximate the true ψ with the characteristic function of a variable with a two-parameter distribution. The choice of the distribution depends on the X-dynamics while the two parameters of the distribution are determined by the conditional mean $\mathrm{E}[\, X(t+1) \,|\, Y(t+1); \Psi \,]$ and variance $\mathrm{Var}[\, X(t+1) \,|\, Y(t+1); \Psi \,]$ given in equations (75)-(76).

In his application Bates finds that the method is successful in estimating different flavors of the SV jump-diffusion for a univariate series of daily 1953-1996 S&P 500 returns. In particular, he shows that the method obtains estimates that are equally, if not more, efficient compared to the efficient method of moments and Markov Chain Monte Carlo methods described below. Extensions of the method to multivariate processes are theoretically possible, but they require numerical integration of multi-dimensional functions, which is computationally demanding.

V.4.2. Simulated Maximum Likelihood

In Section VI.2 we discussed methods for simulated ML estimation and inference in discrete-time SV models. Pedersen [224] and Santa-Clara [234] independently develop a simulated maximum likelihood (SML) method to estimate continuous-time diffusion models. They divide each interval in between

two consecutive data points X_{t+1} and X_t into M sub-intervals of length $\Delta = 1/M$ and they discretize the X proces using the Euler scheme,

$$X_{t+(i+1)\Delta} = X_{t+i\Delta} + \mu(X_{t+i\Delta})\Delta + \Sigma(X_{t+i\Delta})\sqrt{\Delta}\,\varepsilon_{t+(i+1)\Delta}\,, \quad i = 0, \ldots, M-1\,, \tag{77}$$

where μ and Σ are the drift and diffusion terms of the X process and ε is multivariate normal with mean zero and identity variance matrix. The transition density of the discretized process is multivariate normal with mean μ and variance matrix $\Sigma\Sigma'$. As Δ goes to zero, this density converges to that of the continuous-time process X. Thus, the transition density from X_t to X_{t+1} is given by

$$f_X(X_{t+1}|X_t;\Psi) = \int f_X(X_{t+1}|X_{t+1-\Delta};\Psi) f_X(X_{t+1-\Delta}|X_t;\Psi) dX_{t+1-\Delta}\,. \tag{78}$$

For sufficiently small values of Δ the first term in the integrand, $f_X(X_{t+1}|X_{t+1-\Delta};\Psi)$, is approximated by the transition density of the discretized process, while the second term, $f_X(X_{t+1-\Delta}|X_t;\Psi)$, is a multi-step-ahead transition density that can be computed from the recursion from X_t to $X_{t+1-\Delta}$. Writing the right-hand side of equation (78) as a conditional expectation yields

$$f_X(X_{t+1}|X_t;\Psi) = E_{X_{t+1-\Delta}|X_t}\left[\, f_X(X_{t+1}|X_{t+1-\Delta};\Psi)\,\right]. \tag{79}$$

The expectation in equation (79) can be computed by Monte Carlo integration over a large number of paths for the process X, simulated via the Euler scheme (77). As Δ vanishes, the Euler scheme is consistent. Thus, when the size of the simulated sample increases the sample average of the function f_X, evaluated at the random draws of $X_{t+1-\Delta}$, converges to the true transition density. Application of the principles in Bladt and Sørensen [64] may well be useful in enhancing the efficiency of the simulation scheme and hence the actual efficiency of the inference procedure in practice.

Brandt and Santa-Clara [75] apply the SML method to estimate a continuous-time model of the joint dynamics of interest rates in two countries and the exchange rate between the two currencies. Piazzesi [228] extends the SML approach for jump-diffusion processes with time-varying jump intensity. She considers a high-frequency policy rule based on yield curve information and an arbitrage-free bond market and estimates the model using 1994-1998 data on the Federal Reserve target rate, the six-month LIBOR rate, and swap yields.

An important issue is how to initialize any unobserved component of the state vector, $X(t)$, such as the volatility state at each observation to provide a starting point for next Monte Carlo integration step. This may be remedied through application of the particle filter, as mentioned earlier and discussed below in connection with MCMC estimation. Another possibility is, as also indicated previously, to extract the state variable through inversion from derivatives prices or yields assumed observed without pricing errors.

V.4.3. Indirect inference

There are also other method-of-moments strategies to estimate finitely-sampled continuous-time processes of a general type. One prominent approach approximates the unknown transition density for the continuous-time process with the density of a semi-nonparametric (SNP) auxiliary model. Then one can use the score function of the auxiliary model to form moment conditions for the parameter

vector Ψ of the continuous-time model. This approach yields the efficient method of moments estimator (EMM) of Gallant and Tauchen [154], Gallant et al. [150], and Gallant and Long [152], and the indirect inference estimator of Gouriéroux et al. [162] and Smith [243].

To fix ideas, suppose that the conditional density for a continuous-time return process r (the 'structural' model) is unknown. We intend to approximate the unknown density with a discrete-time model (the 'auxiliary' model) that is tractable and yet sufficiently flexible to accommodate the systematic features of the actual data sample well. A parsimonious auxiliary density for r embeds ARMA and EGARCH leading terms to capture the conditional mean and variance dynamics. There may be residual excess skewness and kurtosis that elude the ARMA and EGARCH forms. As such, the auxiliary density is rescaled using a nonparametric polynomial expansion of order K, which yields

$$g_K(r(t)|x(t);\xi) = \left(\nu + (1-\nu) \times \frac{[P_K(z(t),x(t))]^2}{\int_{\mathrm{R}}[P_K(z(t),x(t))]^2\phi(u)du}\right)\frac{\phi(z(t))}{\sqrt{h}(t)}\,, \tag{80}$$

where ν is a small constant, $\phi(.)$ is the standard normal density, $x(t)$ contains lagged return observations, and

$$z(t) = \frac{r(t)-\mu(t)}{\sqrt{h(t)}}\,, \tag{81}$$

$$\mu(t) = \phi_0 + c\,h(t) + \sum_{i=1}^{s}\phi_i r(t-1) + \sum_{i=1}^{u}\delta_i\varepsilon(t-1)\,, \tag{82}$$

$$\begin{aligned}\log h(t) &= \omega + \sum_{i=1}^{p}\beta_i \log h(t-1) + \\ &\quad (1+\alpha_1 L + ... + \alpha_q L^q)\,[\,\theta_1 z(t-1) + \theta_2\,(b(z(t-1)) - \sqrt{2/\pi})\,]\,,\end{aligned} \tag{83}$$

$$P_K(z,x) = \sum_{i=0}^{K_z} a_i(x) z^i = \sum_{i=0}^{K_z}\left(\sum_{|j|=0}^{K_x} a_{ij}x^j\right) z^i\,, \qquad a_{00} = 1\,, \tag{84}$$

Here j is a multi-index vector, $x^j \equiv (x_1^{j_1},\ldots,x_M^{j_M})$, and $|j\,| \equiv \sum_{m=1}^{M} j_m$. The term $b(z)$ is a smooth (twice-differentiable) function that closely approximates the absolute value operator in the EGARCH variance equation.

In practice, the representation of P_K is given by Hermite orthogonal polynomials. When the order K of the expansion increases, the auxiliary density will approximate the data arbitrarily well. If the structural model is indeed the true data generating process, then the auxiliary density will converge to that of the structural model. For a given K, the QML estimator $\hat{\xi}$ for the auxiliary model coefficient satisfies the score condition

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial \log g_K(r(t)\,|\,x(t);\hat{\xi})}{\partial\xi} = 0\,. \tag{85}$$

Suppose now that the structural model is correct and Ψ_0 is the true value of its coefficient vector. Consider a series $\{r(t;\Psi), x(t;\Psi)\}$, $t = 1,\ldots,\mathcal{T}(T)$, simulated from the structural model. Then we expect that the score condition (85) holds when evaluated by averaging over the simulated returns rather than over the actual data:

$$m_{\mathcal{T}(T)}(\Psi_0,\hat{\xi}) = \frac{1}{\mathcal{T}(T)}\sum_{t=1}^{\mathcal{T}(T)}\frac{\partial \log g_K(r(t,\Psi_0)\,|\,x(t,\Psi_0);\hat{\xi})}{\partial\xi} \approx 0. \tag{86}$$

When T and $\mathcal{T}(T)$ tend to infinity, condition (86) holds exactly.

Gallant and Tauchen [154] propose the EMM estimator $\hat{\Psi}$ defined via

$$\hat{\Psi} = \arg\min_{\Psi} \; m_{\mathcal{T}(T)}(\Psi, \hat{\xi})' \, \hat{W}_T \, m_{\mathcal{T}(T)}(\Psi, \hat{\xi}) \,, \tag{87}$$

where the weighting matrix $\hat{W}_T$ is a consistent estimate of the inverse asymptotic covariance matrix of the auxiliary score function, e.g., the inverse outer product of the SNP gradient:

$$\hat{W}_T^{-1} = \frac{1}{T} \sum_{t=1}^{T} \left[\frac{\partial \log g_K(r(t) \,|\, x(t); \hat{\xi})}{\partial \xi} \right] \left[\frac{\partial \log g_K(r(t) \,|\, x(t); \hat{\xi})}{\partial \xi} \right]' . \tag{88}$$

An important advantages of the technique is that EMM estimates achieve the same degree of efficiency as the ML procedure, when the score of the auxiliary model asymptotically spans the score of the true model. It also delivers powerful specification diagnostics that provide guidance in the model selection. Gallant and Tauchen [154] show that the EMM estimator is asymptotically normal. Further, under the assumption that the structural model is correctly specified, they derive a χ^2 statistic for the test of over-identifying restrictions. Gallant et al. [150] normalize the vector $m_{\mathcal{T}(T)}(\hat{\Psi}, \hat{\xi})$ by its standard error to obtain a vector of score t-ratios. The significance of the individual score elements is often informative of the source of model mis-specification, with the usual caveat that failure to capture one characteristic of the data may result in the significance of a moment condition that pertains to a coefficient not directly related to that characteristic (due to correlation in the moment conditions). Finally, EMM provides a straightforward solution to the problem of filtering and forecasting the latent return variance process V, i.e., determining the conditional densities $f(V(t) \,|\, x(t), \Psi)$ and $f(V(t+j) \,|\, x(t), \Psi)$, $j \geq 0$. This is accomplished through the *reprojection* method discussed in, e.g., Gallant and Long [152] and Gallant and Tauchen [155]. In applications to dynamic term structure models, the same method yields filtered and forecasted values for the latent state variables.

The reprojection method assumes that the coefficient vector Ψ is known. In practice, Ψ is fixed at the EMM estimate $\hat{\Psi}$. Then one simulates a sample of returns and latent variables from the structural model and fits the auxiliary model on the simulated data. This is equivalent to the first step of the EMM procedure except that, in the reprojection step, we fit the auxiliary model assuming the structural model is correct, rather than using actual data. The conditional density of the auxiliary model, estimated under the null, approximates the unknown density of the structural model:

$$g_K(r(t+j)|x(t); \tilde{\xi}) \approx f(r(t+j)|x(t); \hat{\Psi}), \; j \geq 0 \,, \tag{89}$$

where $\tilde{\xi}$ is the QML estimate of the auxiliary model coefficients obtained by fitting the model on simulated data. This approach yields filtered estimates and forecasts for the conditional mean and variance of the return via

$$E\big[r(t+j) \,|\, x(t); \hat{\Psi}\,\big] \;=\; \int y \, g_K(y|x(t); \tilde{\xi}) \, dy \,, \tag{90}$$

$$Var\big[r(t+j) \,|\, x(t); \hat{\Psi}\,\big] \;=\; \int \Big(y - E\big[r(t+j) \,|\, x(t); \hat{\Psi}\,\big]\Big)^2 \, g_K(y|x(t); \tilde{\xi}) \, dy \,. \tag{91}$$

An alternative approach consists in fitting an auxiliary model for the latent variable (e.g., the return conditional variance) as a function of current and lagged returns. It is straightforward to estimate

such model using data on the latent variable and the associated returns simulated from the structural model with the EMM coefficient $\hat{\Psi}$. Also in this case the auxiliary model density approximates the true one, i.e.,

$$g_K^V(V(t+j)|x(t);\tilde{\xi}) \approx f^V(V(t+j)|x(t);\hat{\psi}),\ j \geq 0\,. \tag{92}$$

This approach yields a forecast for the conditional variance process,

$$E\big[V(t+j)\,|\,x(t);\hat{\Psi}\,\big] \;=\; \int v\, g_K^V(v|x(t);\tilde{\xi})\, dv\,. \tag{93}$$

In sum, reprojection is a simulation approach to implement a non-linear Kalman-filter-type technique, which yields effective forecasts for the unobservable state vector.

The indirect inference estimator by Gouriéroux et al. [162] and Smith [243] is closely related to the EMM estimator. Indirect inference exploits that the following two quantities should be close when the structural model is correct and the data are simulated at the true parameter Ψ_0: (i) the QML estimator $\hat{\xi}$ for the auxiliary model computed from actual data; (ii) the QML estimator $\hat{\xi}(\Psi)$ for the auxiliary model fitted on simulations from the structural model. Minimizing the distance between $\hat{\xi}$ and $\hat{\xi}(\Psi)$ in an appropriate metric yields the indirect inference estimator for Ψ. Similar to EMM, asymptotic normality holds and a χ^2 test for over-identifying restrictions is available. However, the indirect inference approach is computationally more demanding, because finding the value of Ψ that minimizes the distance function requires re-estimating the auxiliary model on a different simulated sample for each iteration of the optimization routine. EMM does not have this drawback, since the EMM objective function is evaluated at the same fitted score at each iteration. Nonetheless, there may well be circumstances where particular auxiliary models are of primary economic interest and estimation based on the corresponding moment conditions may serve as a useful diagnostic tool for model performance in such directions.

Several studies have used EMM to fit continuous-time SV jump-diffusion models for equity index returns, e.g., Andersen et al. [15], Benzoni [59], Chernov and Ghysels [93], and Chernov et al. [94, 95]. Andersen and Lund [28] and Andersen et al. [16] use EMM to estimate SV jump-diffusion models for the short-term interest rate. Ahn et al. [1, 2], Brandt and Chapman [71], and Dai and Singleton [114] fit different flavors of multi-factor dynamic term structure models. Andersen et al. [27] document the small-sample properties of the efficient method of moments estimator for stationary processes, while Duffee and Stanton [128] study its properties for near unit-root processes.

A. Ronald Gallant and George E. Tauchen at Duke University have prepared well-documented general-purpose EMM and SNP packages, available for download at the web address `ftp.econ.duke.edu` in the directories `pub/get/emm` and `pub/get/snp`. In applications it is often useful to customize the SNP density to allow for a more parsimonious fit of the data under investigation. For instance, Andersen et al. [15, 16], Andersen and Lund [28], and Benzoni [59] rely on the SNP density (80)-(84).

V.4.4. Markov Chain Monte Carlo

The MCMC method provides a Bayesian solution to the inference problem for a dynamic asset pricing model. The approach treats the model coefficient Ψ as well as the vector of latent state variables X as random variables and computes the posterior distribution $f(\Psi, X|Y)$, conditional on certain

observable variables Y, predicted by the model. The setting is sufficiently general to deal with a wide range of situations. For instance, X and Y can be the (latent) volatility and (observable) return processes as is the case of an SV model for asset returns. Or X and Y can be the latent state vector and observable yields in a dynamic term structure model.

The posterior distribution $f(\Psi, X|Y)$ is the main tool to draw inference not only on the coefficient Ψ but also on the latent vector X. Since $f(\Psi, X|Y)$ is unknown in closed-form in relevant applications, MCMC relies on a simulation (a Markov Chain) from the conditional density $f(\Psi, X|Y)$ to compute mode, mean, and standard deviations for the model coefficients and state variables via the Monte Carlo method.

The posterior $f(\Psi, X|Y)$ is analytically untractable and extremely high-dimensional, so that simulation directly from $f(\Psi, X|Y)$ is typically infeasible. The MCMC approach hinges on the Clifford-Hammersley theorem, which determines conditions under which the posterior $f(\Psi, X|Y)$ is uniquely determined by the marginal posterior distributions $f(\Psi|X, Y)$ and $f(X|\Psi, Y)$. In turn, the posteriors $f(\Psi|X, Y)$ and $f(X|\Psi, Y)$ are determined by a set or univariate posterior distributions. Specifically, denote with $\Psi(i)$ the ith element of the coefficient Ψ, $i = 1, \ldots, K$, and with $\Psi(-i)$ the vector consisting of all elements in Ψ except for the ith one. Similarly denote with $X(t)$ the tth row of the state vector, $t = 1, \ldots, T$, and with $X(-t)$ the rest of the vector. Then the Clifford-Hammersley theorem allows to characterize the posterior $f(\Psi, X|Y)$ via $K + T$ univariate posteriors,

$$f(\Psi(i)|\Psi(-i), X, Y), \qquad i = 1, \ldots, K \tag{94}$$

$$f(X(t)|X(-t), \Psi, Y), \qquad t = 1, \ldots, T\,. \tag{95}$$

The construction of the Markov Chain relies on the so-called Gibbs sampler. The first step of the algorithm consists in choosing initial values for the coefficient and the state, Ψ_0 and X_0. When (one of or both) the multi-dimensional posteriors are tractable, the Gibbs sampler generates values Ψ_1 and X_1 directly from $f(\Psi|X, Y)$ and $f(X|\Psi, Y)$. Alternatively, each element of Ψ_1 and X_1 is drawn from the univariate posteriors (94)-(95). Some of these posteriors may also be analytically intractable or efficient algorithms to draw from these posteriors may not exist. In such cases the Metropolis-Hastings algorithm ensures that the simulated sample is consistent with the posterior target distribution. Metropolis-Hastings sampling consists of an accept-reject procedure of the draws from a 'proposal' or 'candidate' tractable density, which is used to approximate the unknown posterior (see, e.g., Johannes and Polson [186]).

Subsequent iterations of Gibbs sampling, possibly in combination with the Metropolis-Hastings sampling, yield a series of 'sweeps' $\{\Psi_s, X_s\}$, $s = 1, \ldots, S$, with limiting distribution $f(\Psi, X|Y)$. A long number of sweeps may be necessary to 'span' the whole posterior distribution and obtain convergence due to the serial dependence of subsequent draws of coefficients and state variables. When the algorithm has converged, additional simulations provide a sample from the joint posterior distribution.

The MCMC approach has several advantages. First, the inference automatically accounts for parameter uncertainty. Further, the Markov Chain provides a direct and elegant solution to the *smoothing* problem, i.e., the problem of determining the posterior distribution for the state vector X conditional on the entire data sample, $f(X(t)\,|\,Y(1), \ldots, Y(T), \Psi)$, $t = 1, \ldots, T$. The limitation on the approach is largely that efficient sampling schemes for the posterior distribution must be constructed

for each specific problem at hand which by nature is case specific and potentially cumbersome or inefficient. Nonetheless, following the development of more general simulation algorithms, the method has proven flexible for efficient estimation of a broad class of important models.

One drawback is that MCMC does not deliver an immediate solution to the *filtering* problem, i.e., determining $f(X(t) \,|\, Y(1), \ldots, Y(t), \Psi)$, and the *forecasting* problem, i.e., determining $f(X(t+j) \,|\, Y(1), \ldots, Y(t), \Psi)$, $j > 0$. However, recent research is overcoming this limitation through the use of the 'particle filter.' Bayes rule implies

$$f(X(t+1) \,|\, Y(1), \ldots, Y(t+1), \Psi) \propto f(Y(t+1) \,|\, X(t+1), \Psi)\, f(X(t+1) \,|\, Y(1), \ldots, Y(t), \Psi)\,, \quad (96)$$

where the symbol $\propto$ denotes 'proportional to.' The first density on the right-hand side of equation (96) is determined by the SV model and it is often known in closed form. In contrast, the second density at the far-right end of the equation is given by an integral that involves the unknown filtering density at the prior period, $f(X(t) \,|\, Y(1), \ldots, Y(t), \Psi)$:

$$f(X(t+1) \,|\, Y(1), \ldots, Y(t), \Psi) = \int f(X(t+1) \,|\, X(t), \Psi)\, f(X(t) \,|\, Y(1), \ldots, Y(t), \Psi)\, dX(t)\,. \quad (97)$$

The particle method relies on simulations to construct a finite set of weights $w^i(t)$ and particles $X^i(t)$, $i = 1, \ldots, N$, that approximate the unknown density with a finite sum,

$$f(X(t) \,|\, Y(1), \ldots, Y(t), \Psi) \approx \sum_{i=1}^{N} w^i(t) \delta_{X^i(t)}\,. \quad (98)$$

where the Dirac function $\delta_{X^i(t)}$ assigns mass one to the particle $X^i(t)$. Once the set of weights and particles are determined, it is possible to re-sample from the discretized distribution. This step yields a simulated sample $\{X^s(t)\}_{s=1}^{S}$ which can be used to evaluate the density in equation (97) via Monte Carlo integration:

$$f(X(t+1) \,|\, Y(1), \ldots, Y(t), \Psi) \approx \frac{1}{S} \sum_{s=1}^{S} f(X(t+1) \,|\, X^s(t), \Psi)\,. \quad (99)$$

Equation (99) solves the forecasting problem while combining formulas (96) and (99) solves the filtering problem. The challenge in practical application of the particle filter is to identify an accurate and efficient algorithm to construct the set of particles and weights. We point the interested reader to Kim et al. [193], Pitt and Shephard [226], and Johannes and Polson [186] for a discussion on how to approach this problem.

The usefulness of the MCMC method to solve the inference problem for SV models has been evident since the early work by Jacquier et al. [179], who develop an MCMC algorithm for the logarithmic SV model. Jacquier et al. [180] provide extensions to correlated and non-normal error distributions. Kim et al. [193] and Chib et al. [96] develop simulation-based methods to solve the filtering problem, while Chib et al. [97] use the MCMC approach to estimate a multivariate SV model. Elerian et al. [135] and Eraker [140] discuss how to extend the MCMC inference method to a continuous-time setting. Eraker [140] uses the MCMC approach to estimate an SV diffusion process for interest rates, while Jones [188] estimates a continuous-time model for the spot rate with non-linear drift function. Eraker et al. [142]

estimate an SV jump-diffusion process using data on S&P 500 return while Eraker [141] estimates a similar model using joint data on options and underlying S&P 500 returns. Li et al. [196] allow for Lévy-type jumps in their model. Collin-Dufresne et al. [104] use the MCMC approach to estimate multi-factor affine dynamic term structure model using swap rates data. Johannes and Polson [185] give a comprehensive survey of the still ongoing research on the use of the MCMC approach in the general nonlinear jump-diffusion SV setting.

V.5. Estimation from Option Data

Options' payoffs are non-linear functions of the underlying security price. This feature renders options highly sensitive to jumps in the underlying price and to return volatility, which makes option data particularly useful to identify return dynamics. Therefore, several studies have taken advantage of the information contained in option prices, possibly in combination with underlying return data, to estimate SV models with or without discontinuities in returns and volatility.

Applications to derivatives data typically require a model for the pricing errors. A common approach is to posit that the market price of an option, O^*, normalized by the underlying observed security price S^*, is the sum of the normalized model-implied option price, O/S^*, and a disturbance term ε (e.g., Renault [230]):

$$\frac{O^*}{S^*} = \frac{O(S^*, V, K, \tau, \Psi)}{S^*} + \varepsilon\,, \tag{100}$$

where V is the latent volatility state, K is the option strike price, τ is time to maturity, and Ψ is the vector with the model coefficients. A pricing error ε could arise for several reasons, including measurement error (e.g., price discreteness), asynchroneity between the derivatives and underlying price observations, microstructure effects, and perhaps most importantly specification error. The structure imposed on ε depends on the choice of a specific 'loss function' used for estimation (e.g., Christoffersen and Jacobs [98]). Several studies have estimated the coefficient vector Ψ by minimizing the sum of the squared option pricing errors normalized by the underlying price S^*, as in equation (100). Others have focused on either squared dollar pricing errors, or squared errors normalized by the options market price (instead of S^*). The latter approach has the advantage that a \$1 error on an expensive in-the-money option carries less weight than the same error on a cheaper out-of-the-money contract. The drawback is that giving a lot of weight to the pricing errors on short-maturity deep-out-of-the-money options could bias the estimation results. Finally, the common practice of expressing option prices in terms of their Black-Scholes implied volatilities has inspired other scholars to minimize the deviations between Black-Scholes implied volatilities inferred from model and market prices (e.g., Mizrach [216]). An alternative course is to form a moment-based loss function and follow a GMM- or SMM-type approach to estimate Ψ. To this end moment conditions stem from distributional assumptions on the pricing error ε (e.g., $\mathrm{E}[\varepsilon] = 0$) or from the scores of a reduced-form model that approximates the data.

In estimating the model, some researchers have opted to use a panel of options consisting of contracts with multiple strikes and maturities across dates in the sample period. This choice brings a wealth of information on the cross-sectional and term-structure properties of the implied volatility smirk into the analysis. Others rely on only one option price observation per time period, which shifts the focus to the time-series dimension of the data. Some studies re-estimate the model on a daily

basis rather than seeking a single point estimate for the coefficient Ψ across the entire sample period. This ad hoc approach produces smaller in-sample pricing errors, which can be useful to practitioners, but at the cost of concealing specification flaws by over-fitting the model, which tends to hurt out-of-sample performance. The different approaches are in part dictated by the intended use of the estimated system as practitioners often are concerned with market making and short-term hedging while academics tend to value stable relations that may form the basis for consistent modeling of the dominant features of the system over time.

Early contributions focus on loss functions based on the sum of squared option pricing errors and rely entirely on option data for estimation. This approach typically yields an estimate of the model coefficient Ψ that embeds an adjustment for risk, i.e., return and volatility dynamics are identified under the risk-neutral rather than the physical probability measure. For instance, Bates [56] considers an SV jump-diffusion model for Deutsche Mark foreign currency options and estimates its coefficient vector Ψ via nonlinear generalized least squares of the normalized pricing errors with daily option data from January 1984 to June 1991. A similar approach is followed by Bates [57] who fits an SV model with two latent volatility factors and jumps using daily data on options on the S&P 500 futures from January 1988 to December 1993. Bakshi et al. [34] focus on the pricing and hedging of daily S&P 500 index options from June 1988 to May 1991. In their application they re-calibrate the model on a daily basis by minimizing the sum of the squared dollar pricing errors across options with different maturities and strikes. Huang and Wu [173] explore the pricing implications of the time-changed Lévy process by Carr and Wu [84] for daily S&P 500 index options from April 1999 to May 2000. Their Lévy return process allows for discontinuities that exhibit higher jump frequencies compared to the finite-intensity Poisson jump processes in equations (37)-(41). Further, their model allows for a random time change, i.e., a monotonic transformation of the time variable which generates SV in the diffusion and jump components of returns. In contrast, Bakshi et al. [35] fit an SV jump-diffusion model by SMM using daily data on long-maturity S&P 500 options (LEAPS).

More recent studies have relied on joint data on S&P 500 option prices and underlying index returns, spanning different periods, to estimate the model. This approach forces the same model to price securities in two different markets and relies on information from the derivatives and underlying securities to better pin down model coefficients and risk premia. For instance, Eraker [141] and Jones [189] fit different flavors of the SV model (with and without jumps, respectively) by MCMC. Pan [220] follows a GMM approach to estimate an SV jump-diffusion model using weekly data. She relies on a single at-the-money option price observation each week, which identifies the level of the latent volatility state variable (i.e., at each date she fixes the error term ε at zero and solves equation (100) for V). Aït-Sahalia and Kimmel [7] apply Aït-Sahalia's [4] method to approximate the likelihood function for a joint sample of options and underlying prices. Chernov and Ghysels [93] and Benzoni [59] obtain moment conditions from the scores of a SNP auxiliary model. Similarly, other recent studies have found it useful to use joint derivatives and interest rate data to fit dynamic term structure models, e.g., Almeida et al. [9], and Bikbov and Chernov [62].

Finally, a different literature has studied the option pricing implications of a model in which asset return volatility is a deterministic function of the asset price and time, e.g., Derman and Kani [119], Dupire [134], Rubinstein [233], and Jackwerth and Rubinstein [176]. Since volatility is not stochastic in this setting, we do not review these models here and point the interested reader to, e.g., Dumas et

al. [133] for an empirical analysis of their performance.

VI. Future Directions

In spite of much progress in our understanding of volatility new challenges lie ahead. In recent years a wide array of volatility-sensitive products has been introduced. The market for these derivatives has rapidly grown in size and complexity. Research faces the challenge to price and hedge these new products. Moreover, the recent developments in model-free volatility modeling have effectively given empirical content to the latent volatility variable, which opens the way for a new class of estimation methods and specification tests for SV systems. Related, improved volatility measures enable us to shed new light on the properties and implications of the volatility risk premium. Finally, more work is needed to better understand the linkage between fluctuations in economic fundamentals and low- and high-frequency volatility movements. We conclude this chapter by briefly reviewing some open issues in these four areas of research.

VI.1. Volatility and Financial Markets Innovation

Volatility is a fundamental input to any financial and real investment decision. Markets have responded to investors' needs by offering an array of volatility-linked instruments. In 1993 the Chicago Board Option Exchange (CBOE) has introduced the VIX index, which measures the market expectations of near-term volatility conveyed by equity-index options. The index was originally computed using the Black-Scholes implied volatilities of eight different S&P 100 option (OEX) series so that, at any given time, it represented the implied volatility of a hypothetical at-the-money OEX option with exactly 30 days to expiration (see Whaley [257]). On September 22, 2003, the CBOE began disseminating price level information using a revised 'model-free' method for the VIX index. The new VIX is given by the price of a portfolio of S&P 500 index options and incorporates information from the volatility smirk by using a wider range of strike prices rather than just at-the-money series (see Britten-Jones and Neuberger [77]). On March 26, 2004, trading in futures on the VIX Index started on the CBOE Futures Exchange (CFE) while on February 24, 2006, options on the VIX began trading on the Chicago Board Options Exchange. These developments have opened the way for investors to trade on option-implied measures of market volatility. The popularity of the VIX prompted the CBOE to introduce similar indices for other markets, e.g., the VXN NASDAQ 100 Volatility Index.

Along the way, a new over-the-counter market for volatility derivatives has also rapidly grown in size and liquidity. Volatility derivatives are contracts whose payments are expressed as functions of realized variance. Popular examples are variance swaps, which at maturity pay the difference between realized variance and a fixed strike price. According to estimates by BNP Paribas reported by the Risk [192] magazine, the daily trading volume for variance swaps on indices reached $4-5 million in vega notional (measured in dollars per volatility point) in 2006, which corresponds to payments in excess of $1 billion per percentage point of volatility on an annual basis (Carr and Lee [82]). Using variance swaps hedge fund managers and proprietary traders can easily place huge bets on market volatility.

Finally, in recent years credit derivatives markets have evolved in complexity and grown in size.

Among the most popular credit derivatives are the credit default swaps (CDS), which provide insurance against the risk of default by a particular company. The buyer of a single-name CDS acquires the right to sell bonds issued by the company at face value when a credit event occurs. Multiple-name contracts can be purchased simultaneously through credit indices. For instance, the CDX indices track the credit spreads for different portfolios of North American companies while the iTraxx Europe indices track the spreads for portfolios of European companies. At the end of 2006 the notional amount of outstanding over-the-counter single- and multi-name CDS contracts stood at $19 and $10 trillion, respectively, according to the September 2007 Bank for International Settlements Quarterly Review.

These market developments have raised new interesting issues for research to tackle. The VIX computations based on the new model-free definition of implied volatility used by the CBOE requires the use of options with strike prices that cover the entire support of the return distribution. In practice, liquid options satisfying this requirement often do not exist and the CBOE implementation introduces random noise and systematic error into the index (Jiang and Tian [184]). Related, the VIX implementation entails a truncation, i.e., the CBOE discards illiquid option prices with strikes lying in the tails of the return distribution. Thus, the notion of the VIX is more directly linked to that of corridor volatility (Andersen and Bondarenko [26]). In sum, robust implementation of a model free measure of implied volatility is still an open area of research. Future developments in this direction will also have important repercussions on the hedging practices for implied-volatility derivatives.

Pricing and hedging of variance derivatives is another active area of research. Variance swaps admit a simple replication strategy via static positions in call and put options on the underlying asset, similar to model-free implied volatility measures (e.g., Britten-Jones and Neuberger [77] and Carr and Madan [83]). In contrast, it is still an open area of research to determine the replication strategy for derivatives whose payoffs are non-linear function of realized variance, e.g., volatility swaps, which pay the square-root of realized variance, or call and put options on realized variance. Carr and Lee [82] is an interesting paper in this direction.

Limited liability gives shareholders the option to default on the firm's debt obligation. Hence, a debt claim has features similar to a short position in a put option. The pricing of corporate debt is therefore sensitive to the volatility of the firms' assets: higher volatility increases the probability of default and therefore reduces the price of debt and increases credit spreads. The insights and techniques developed in the SV literature could prove useful in credit risk modeling and applications (e.g., Jacobs and Li [178], Tauchen and Zhou [248] and Zhang et al. [260]).

VI.2. The Use of Realized Volatility for Estimation of SV Models

Another promising line of research aims at extracting the information in RV measures for the estimation of dynamic asset pricing models. Early work along these lines includes Barndorff-Nielsen and Shephard [51], who decompose RV into actual volatility and realized volatility error. They consider a state-space representation for this decomposition and apply the Kalman filter to estimate different flavors of the SV model. Moreover, Bollerslev and Zhou [68] and Garcia et al. [156] build on the insights of Meddahi [210] to estimate SV diffusion models using conditional moments of integrated volatility. More recently, Todorov [252] generalizes the analysis for the presence of jumps.

Related, recent studies have started to use RV measures to test the implications of models previ-

ously estimated with lower-frequency data. Since RV gives empirical content to the latent quadratic variation process, this approach allows for a direct test of the model-implied restrictions on the latent volatility factor. Recent work along these lines includes Andersen and Benzoni [12], who use model-free RV measures to show that the volatility spanning condition embedded in some affine term structure models is violated in the U.S. Treasury market. Christoffersen et al. [100] note that the Heston square-root SV model implies that the dynamics for the standard deviation process are conditionally Gaussian. They reject this condition by examining the distribution of the changes in the square-root RV measure for S&P 500 returns.

VI.3. Volatility Risk Premium

More work is needed to better understand the link between asset return volatility and model risk premia. Also in this case, RV measures are a fruitful source of information to shed new light on the issue. Among the recent studies that pursue this venue is Bollerslev et al. [66], who exploit the moments of RV and option-implied volatility to gauge a measure of the volatility risk premium. Todorov [251] explores the variance risk premium dynamics using high-frequency S&P 500 index futures data and data on the VIX index. He finds the variance risk premium to vary significantly over time and to increase during periods of high volatility and immediately after big jumps in underlying returns. Carr and Wu [85] provide a broader analysis of the variance risk premium for five equity indices and 35 individual stocks. They find the premium to be large and negative for the indices while it is much smaller for the individual stocks. Further, they also find the premium to increase (in absolute value) with the level of volatility. Additional work on the volatility risk premium embedded in individual stock options is in Bakshi and Kapadia [36], Driessen et al. [123], and Duarte and Jones [126]. Other studies have examined the linkage between volatility risk premia and equity returns (e.g., Bollerslev and Zhou [69]) and hedge-fund performance (e.g., Bondarenko [70]). New research is also examining the pricing of aggregate volatility risk in the cross-section of stock returns. For instance, Ang et al. [30] find that average returns are lower on stocks that have high sensitivities to innovations in aggregate volatility and high idiosyncratic volatility (see also the related work by Ang et al. [31], Bandi et al. [**?**], Chen [90], and Guo et al. [163]). This evidence is consistent with the findings of the empirical option pricing literature, which suggests that there is a negative risk premium for volatility risk. Intuitively, periods of high market volatility are associated to worsened investment opportunities and tend to coincide with negative stock market returns (the so-callled leverage effect). Therefore, investors are willing to pay higher prices (i.e., accept lower expected returns) to hold stocks that do well in high-volatility conditions.

VI.4. Determinants of Volatility

Finally, an important area of future research concerns the linkage between asset return volatility and economic uncertainty. Recent studies have proposed general equilibrium models that produce low-frequency fluctuations in conditional volatility, e.g., Campbell and Cochrane [80], Bansal and Yaron [47], McQueen and Vorkink [207], and Tauchen [246]. Related, Engle and Rangel [139] and Engle et al. [138] link macroeconomic variables and long-run volatility movements. It is still an open issue, however, to determine the process through which news about economic fundamentals are embedded

into prices to generate high-frequency volatility fluctuations. Early research by Schwert [236] and Shiller [241] has concluded that the amplitude of the fluctuations in aggregate stock volatility is difficult to explain using simple models of stock valuation. Further, Schwert [236] notes that while aggregate leverage is significantly correlated with volatility, it explains a relatively small part of the movements in stock volatility. Moreover, he finds little evidence that macroeconomic volatility (measured by inflation and industrial production volatility) helps predict future asset return volatility. Model-free realized volatility measures are a useful tool to further investigate this issue. Recent work in this direction includes Andersen et al. [22] and Andersen and Bollerslev [17], who explore the linkage between news arrivals and exchange rates volatility, and Andersen and Benzoni [13], who investigate the determinants of bond yields volatility in the U.S. Treasury market. Related, Balduzzi et al. [38] and Fleming and Remolona [146] study the reaction of trading volume, bid-ask spread, and price volatility to macroeconomic surprises in the U.S. Treasury market, while Brandt and Kavajecz [74] and Pasquariello and Vega [222] focus instead on the price discovery process and explore the implications of order flow imbalances (excess buying or selling pressure) on day-to-day variation in yields.

VII. References

Primary Literature

1. Ahn DH, Dittmar RF, Gallant AR (2002) Quadratic Term Structure Models: Theory and Evidence. Review of Financial Studies 15:243–288

2. Ahn DH, Dittmar RF, Gallant AR, Gao B (2003) Purebred or hybrid?: Reproducing the volatility in term structure dynamics. Journal of Econometrics 116:147–180

3. Aït-Sahalia Y (2002) Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach. Econometrica 70:223–262

4. Aït-Sahalia Y (2007) Closed-Form Likelihood Expansions for Multivariate Diffusions, forthcoming in the Annals of Statistics. Annals of Statistics, forthcoming

5. Aït-Sahalia Y, Hansen LP, Scheinkman J (2004) Operator Methods for Continuous-Time Markov Processes. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

6. Aït-Sahalia Y, Jacod J (2007) Testing for Jumps in a Discretely Observed Process. Annals of Statistics, forthcoming

7. Aït-Sahalia Y, Kimmel R (2007) Maximum Likelihood Estimation of Stochastic Volatility Models Journal of Financial Economics 83:413–452

8. Alizadeh S, Brandt MW, Diebold FX (2002) Range-Based Estimation of Stochastic Volatility Models. Journal of Finance 57:1047–1091

9. Almeida CIR, Graveline JJ, Joslin S (2006) Do Options Contain Information About Excess Bond Returns?. Working Paper, UMN, Fundação Getulio Vargas, MIT

10. Andersen TG (1994) Stochastic Autoregressive Volatility: A Framework for Volatility Modeling. Mathematical Finance 4:75–102

11. Andersen TG (1996) Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility. Journal of Finance 51:169–204

12. Andersen TG, Benzoni L (2006) Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models. Working Paper, KGSM and Chicago FED

13. Andersen TG, Benzoni L (2007a) The Determinants of Volatility in the U.S. Treasury market. Working Paper, KGSM and Chicago FED

14. Andersen TG, Benzoni L (2007b) Realized Volatility. In: Andersen TG, Davis RA, Kreiss JP, Mikosch T (Eds) Handbook of Financial Time Series, Springer Verlag

15. Andersen TG, Benzoni L, Lund J (2002) An Empirical Investigation of Continuous-Time Equity Return Models. Journal of Finance 57:1239–1284

16. Andersen TG, Benzoni L, Lund J (2004) Stochastic Volatility, Mean Drift and Jumps in the Short Term Interest Rate. Working Paper, Northwestern University, University of Minnesota, and Nykredit Bank

17. Andersen TG, Bollerslev T (1998a) Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies. Journal of Finance 53:219–265

18. Andersen TG, Bollerslev T (1998b) Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts. International Economic Review 39:885–905

19. Andersen TG, Bollerslev T, Diebold FX (2004) Parametric and nonparametric volatility measurement. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

20. Andersen TG, Bollerslev T, Diebold FX (2007) Roughing It Up: Including Jump Components in Measuring, Modeling and Forecasting Asset Return Volatility. Review of Economics and Statistics, forthcoming

21. Andersen TG, Bollerslev T, Diebold FX, Labys P (2003) Modeling and Forecasting Realized Volatility. Econometrica 71:579–625

22. Andersen TG, Bollerslev T, Diebold FX, Vega C (2003) Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange. American Economic Review 93:38–62

23. Andersen TG, Bollerslev T, Dobrev D (2007) No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. Journal of Econometrics 138:125–180

24. Andersen TG, Bollerslev T, Meddahi N (2004) Analytic Evaluation of Volatility Forecasts. International Economic Review 45:1079–1110

25. Andersen TG, Bollerslev T, Meddahi N (2005) Correcting the Errors: Volatility Forecast Evaluation Using High-Frequency Data and Realized Volatilities. Econometrica 73:279–296

26. Andersen TG, Bondarenko O (2007) Construction and Interpretation of Model-Free Implied Volatility. Working Paper, KGSM and UIC

27. Andersen TG, Chung HJ, Sørensen BE (1999) Efficient Method of Moments Estimation of a Stochastic Volatility Model: A Monte Carlo Study, Journal of Econometrics 91:61–87

28. Andersen TG, Lund J (1997) Estimating continuous-time stochastic volatility models of the short term interest rate diffusion. Journal of Econometrics 77:343–377

29. Ané T, Geman H (2000) Order Flow, Transaction Clock, and Normality of Asset Returns. Journal of Finance 55:2259–2284

30. Ang A, Hodrick RJ, Xing Y, Zhang X (2006) The Cross-Section of Volatility and Expected Returns. Journal of Finance 51:259–299

31. Ang A, Hodrick RJ, Xing Y, Zhang X (2008) High idiosyncratic volatility and low returns: international and further U.S. evidence. Journal of Financial Economics, forthcoming

32. Bachelier L (1900) Théorie de la Spéculation. Annales de l'École Normale Supérieure 3, Gauthier-Villars, Paris, France. English translation in Cootner PH (Ed) (1964) The Random Character of Stock Market Prices, MIT Press, Cambridge, Massachusetts, U.S.

33. Back K (1991) Asset Prices for General Processes. Journal of Mathematical Economics 20:371–395

34. Bakshi G, Cao C, Chen Z (1997) Empirical Performance of Alternative Option Pricing Models. Journal of Finance 52:2003–2049

35. Bakshi G, Cao C, Chen Z (2002) Pricing and hedging long-term options. Journal of Econometrics 94:277–318

36. Bakshi G, Kapadia N (2003) Delta-Hedged Gains and the Negative Market Volatility Risk Premium. Review of Financial Studies 16:527–566

37. Bakshi G, Kapadia N, Madan D (2003) Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. Review of Financial Studies 16:101–143

38. Balduzzi P, Elton EJ, Green TC (2001) Economic News and Bond Prices: Evidence from the U.S. Treasury Market. Journal of Financial and Quantitative Analysis 36:523–543

39. Bandi FM (2002) Short-term interest rate dynamics: a spatial approach. Journal of Financial Economics 65:73–110

40. Bandi F, Moise CE, Russell JR (2008) Market volatility, market frictions, and the cross section of stock returns. Working Paper, University of Chicago and Case Western Reserve University.

41. Bandi FM, Nguyen T (2003) On the functional estimation of jump-diffusion models. Journal of Econometrics 116:293–328

42. Bandi FM, Phillips PCB (2002) Nonstationary Continuous-Time Processes. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

43. Bandi FM, Phillips PCB (2003) Fully nonparametric estimation of scalar diffusion models. Econometrica 71:241–283

44. Bandi FM, Phillips PCB (2007) A simple approach to the parametric estimation of potentially nonstationary diffusions. Journal of Econometrics 137:354–395

45. Bandi FM, Russell J (2006a) Separating Microstructure Noise from Volatility. Journal of Financial Economics 79:655–692

46. Bandi FM, Russell J (2006b) Volatility. In: Birge J and Linetsky V (Eds) Handbook of Financial Engineering, Elsevier

47. Bansal R, Yaron A (2004) Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. Journal of Finance 59:1481–1509

48. Bartlett MS (1938) The Characteristic Function of a Conditional Statistic. Journal of the London Mathematical Society 13:63–67

49. Barndorff-Nielsen OE, Hansen P, Lunde A, Shephard N (2007) Designing Realized Kernels to Measure the Ex-Post Variation of Equity Prices in the Presence of Noise. Working Paper, University of Aarhus, Denmark; Stanford University; Nuffield College, Oxford, UK

50. Barndorff-Nielsen OE, Shephard N (2001) Non-Gaussian OrnsteinUhlenbeckbased models and some of their uses in financial economics. Journal of the Royal Statistical Society, Series B 63:167–241

51. Barndorff-Nielsen OE, Shephard N (2002a) Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models. Journal of the Royal Statistical Society, Series B 64:253–280

52. Barndorff-Nielsen OE, Shephard N (2002b) Estimating quadratic variation using realized variance. Journal of Applied Econometrics 17:457–477

53. Barndorff-Nielsen OE, Shephard N (2004) Power and bipower variation with stochastic volatility and jumps. Journal of Financial Econometrics 2:1-37

54. Barndorff-Nielsen OE, Shephard N (2006) Econometrics of testing for jumps in financial economics using bipower variation. Journal of Financial Econometrics 4:1–30

55. Bates DS (1991) The Crash of '87: Was It Expected? The Evidence from Options Markets Journal of Finance 46:1009–1044

56. Bates DS (1996) Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. Review of Financial Studies 9:69–107

57. Bates DS (2000) Post-'87 crash fears in the S&P 500 futures option market. Journal of Econometrics 94:181–238

58. Bates DS (2006) Maximum Likelihood Estimation of Latent Affine Processes. Review of Financial Studies 19:909–965

59. Benzoni L (2002) Pricing Options under Stochastic Volatility: An Empirical Investigation. Working Paper, Chicago FED

60. Benzoni L, Collin-Dufresne P, Goldstein RS (2007) Explaining Pre- and Post-1987 Crash Prices of Equity and Options within a Unified General Equilibrium Framework. Working Paper, Chicago FED, UCB, and UMN

61. Bibby BM, Jacobsen M, Sorensen M (2004) Estimating Functions for Discretely Sampled Diffusion-Type Models. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

62. Bikbov R, Chernov M (2005) Term Structure and Volatility: Lessons from the Eurodollar Markets. Working Paper, LBS, Deutsche Bank

63. Black F, Scholes M (1973) The Pricing of Options and Corporate Liabilities. Journal of Political Economy 81:637–654

64. Bladt M, Sørensen M (2007) Simple Simulation of Diffusion Bridges with Application to Likelihood Inference for Diffusions. Working Paper, University of Copenhagen, Denmark.

65. Bollen NPB, Whaley RE (2004) Does Net Buying Pressure Affect the Shape of Implied Volatility Functions?. Journal of Finance 59:711–753

66. Bollerslev T, Gibson M, Zhou H (2004) Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities. Working Paper, Duke, Federal Reserve Board

67. Bollerslev T, Jubinsky PD (1999) Equity Trading Volume and Volatility: Latent Information Arrivals and Common Long-Run Dependencies. Journal of Business & Economic Statistics 17:9–21

68. Bollerslev T, Zhou H (2002) Estimating stochastic volatility diffusion using conditional moments of integrated volatility. Journal of Econometrics 109:33–65

69. Bollerslev T, Zhou H (2007) Expected Stock Returns and Variance Risk Premia. Working Paper, Duke University and Federal Reserve Board

70. Bondarenko O (2004) Market price of variance risk and performance of hedge funds. Working Paper, UIC

71. Brandt MW, Chapman DA (2003) Comparing Multifactor Models of the Term Structure. Working Paper, Duke and Boston College

72. Brandt MW, Diebold FX (2006) A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations. Journal of Business 79:61–73

73. Brandt MW, Jones CS (2006) Volatility Forecasting with Range-Based EGARCH Models. Journal of Business & Economic Statistics 24:470–486

74. Brandt MW, Kavajecz KA Price Discovery in the U.S. Treasury Market: The Impact of Orderflow and Liquidity on the Yield Curve. Journal of Finance 59:2623–2654

75. Brandt MW, Santa-Clara P (2002) Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. Journal of Financial Economics 63:161–210

76. Breidt FJ, Crato N, de Lima P (1998) The detection and estimation of long memory in stochastic volatility Journal of Econometrics 83:325-348

77. Britten-Jones M, Neuberger A (2000) Option Prices, Implied Price Processes, and Stochastic Volatility. Journal of Finance 55:839–866

78. Broadie M, Chernov M, Johannes MJ (2007) Model Specification and Risk Premia: Evidence from Futures Options. Journal of Finance 62:1453–1490

79. Buraschi A, Jackwerth J (2001) The price of a smile: hedging and spanning in option markets. Review of Financial Studies 14:495–527

80. Campbell JY, Cochrane JH (1999) By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. Journal of Political Economy 107:205–251

81. Campbell JY, Viceira LM (2001) Who Should Buy Long-Term Bonds? American Economic Review 91:99–127

82. Carr P, Lee R (2007) Robust Replication of Volatility Derivatives. Working Paper, NYU and UofC

83. Carr P, Madan D (1998) Towards a theory of volatility trading. In: Jarrow R (Ed) Volatility, Risk Publications

84. Carr P, Wu L (2004) Time-changed Lvy processes and option pricing. Journal of Financial Economics 71:113–141

85. Carr P, Wu L (2007) Variance Risk Premia. Review of Financial Studies, forthcoming

86. Carrasco M, Chernov M, Florens JP, Ghysels E (2007) Efficient estimation of general dynamic models with a continuum of moment conditions. Journal of Econometrics 140:529–573

87. Carrasco M, Florens JP (2000) Generalization of GMM to a continuum of moment conditions. Econometric Theory 16:797–834

88. Chacko G, Viceira LM (2003) Spectral GMM estimation of continuous-time processes. Journal of Econometrics 116:259–292

89. Chan KC, Karolyi GA, Longstaff FA, Sanders AB (1992) An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. The Journal of Finance 47:1209–1227

90. Chen J (2003) Intertemporal CAPM and the Cross-Section of Stock Return. Working Paper, USC

91. Chen RR, Scott L (1993) Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates. Journal of Fixed Income 3:14–31

92. Cheridito P, Filipović D, Kimmel RL (2005) Market Price of Risk Specifications for Affine Models: Theory and Evidence. Journal of Financial Economics, forthcoming

93. Chernov M, Ghysels E (2002) A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. Journal of Financial Economics 56:407–458

94. Chernov M, Gallant AR, Ghysels E, Tauchen G (1999) A New Class of Stochastic Volatility Models with Jumps: Theory and Estimation. Working Paper, London Business School, Duke University, University of Northern Carolina

95. Chernov M, Gallant AR, Ghysels E, Tauchen G (2003) Alternative models for stock price dynamics. Journal of Econometrics 116:225–257

96. Chib S, Nardari F, Shephard N (2002) Markov chain Monte Carlo methods for stochastic volatility models. Journal of Econometrics 108:281–316

97. Chib S, Nardari F, Shephard N (2006) Analysis of high dimensional multivariate stochastic volatility models. Journal of Econometrics 134:341–371

98. Christoffersen PF, Jacobs K (2004) The importance of the loss function in option valuation. Journal of Financial Economics 72:291–318

99. Christoffersen PF, Jacobs K, Karoui L, Mimouni K (2007) Estimating Term Structure Models Using Swap Rates. Working Paper, McGill University

100. Christoffersen PF, Jacobs K, Mimouni K (2006a) Models for S&P 500 Dynamics: Evidence from Realized Volatility, Daily Returns, and Option Prices. Working Paper, Mcgill University

101. Christoffersen PF, Jacobs K, Wang Y (2006b) Option Valuation with Long-run and Short-run Volatility Components Working Paper, Mcgill University

102. Clark PK (1973) A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. Econometrica 41:135–156

103. Collin-Dufresne P, Goldstein RS (2002) Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility. Journal of Finance 57:1685–1730

104. Collin-Dufresne P, Goldstein R, Jones CS (2007a) Identification of Maximal Affine Term Structure Models. Journal of Finance, forthcoming

105. Collin-Dufresne P, Goldstein R, Jones CS (2007b) Can Interest Rate Volatility be Extracted from the Cross Section of Bond Yields? An Investigation of Unspanned Stochastic Volatility. Working Paper, UCB, USC, UMN

106. Collin-Dufresne P, Solnik B (2001) On the Term Structure of Default Premia in the Swap and LIBOR Markets Journal of Finance 56:1095–1115

107. Comte F, Coutin L, Renault E (2003) Affine Fractional Stochastic Volatility Models with Application to Option Pricing. Working Paper, CIRANO

108. Comte F, Renault E (1998) Long memory in continuous-time stochastic volatility models. Mathematical Finance 8:291–323

109. Conley TG, Hansen LP, Luttmer EGJ, Scheinkman JA (1997) Short-term interest rates as subordinated diffusions. Review of Financial Studies 10:525–577

110. Corsi F (2003) A Simple Long Memory Model of realized Volatility. Working paper, University of Southern Switzerland

111. Coval JD, Shumway T (2001) Expected Option Returns. Journal of Finance 56:983–1009

112. Cox JC, Ingersoll JE, Ross SA (1985a) An Intertemporal General Equilibrium Model of Asset Prices. Econometrica 53:363–384

113. Cox JC, Ingersoll JE, Ross SA (1985b) A Theory of the Term Structure of Interest Rates. Econometrica 53:385–407

114. Dai Q, Singleton KJ (2000) Specification Analysis of Affine Term Structure Models. Journal of Finance 55:1943–1978

115. Dai Q, Singleton KJ (2003) Term Structure Dynamics in Theory and Reality. Review of Financial Studies 16:631–678

116. Danielsson, J (1994) Stochastic Volatility in Asset Prices: Estimation by Simulated Maximum Likelihood. Journal of Econometrics 64:375–400

117. Danielsson J, Richard, J-F (1993) Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models. Journal of Applied Econometrics 8:S153–S173

118. Deo, R, Hurvich C (2001) On the Log Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models. Econometric Theory 17:686–710

119. Derman E, Kani I (1994) The volatility smile and its implied tree. Quantitative Strategies Research Notes, Goldman Sachs, New York.

120. Diebold FX, Nerlove M (1989) The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model. Journal of Applied Econometrics 4:1–21

121. Diebold FX, Strasser G (2007) On the Correlation Structure of Microstructure Noise in Theory and Practice. Working Paper, University of Pennsylvania,

122. Dobrev D (2007) Capturing Volatility from Large Price Moves: Generalized Range Theory and Applications. Working Paper, Federal Reserve Board

123. Driessen J, Maenhout P, Vilkov G (2006) Option-Implied Correlations and the Price of Correlation Risk. Working Paper, University of Amsterdam and INSEAD

124. Duarte J (2004) Evaluating An Alternative Risk Preference in Affine Term Structure Models. Review of Financial Studies 17:370–404

125. Duarte J (2007) The Causal Effect of Mortgage Refinancing on Interest-Rate Volatility: Empirical Evidence and Theoretical Implications. Review of Financial Studies, forthcoming

126. Duarte J, Jones CS (2007) The Price of Market Volatility Risk. Working Paper, University of Washington and USC

127. Duffee GR (2002) Term Premia and Interest Rate Forecasts in Affine Models. Journal of Finance 57:405–443

128. Duffee G, Stanton R (2007) Evidence on simulation inference for near unit-root processes with implications for term structure estimation. Journal of Financial Econometrics, forthcoming.

129. Duffie D, Kan R (1996) A yield-factor model of interest rates. Mathematical Finance 6:379–406

130. Duffie D, Pan J, Singleton KJ (2000) Transform Analysis and Asset Pricing for Affine Jump-Diffusions Econometrica 68:1343–1376

131. Duffie D, Singleton KJ (1993) Simulated Moments Estimation of Markov Models of Asset Prices. Econometrica 61:929–952

132. Duffie D, Singleton KJ (1997) An Econometric Model of the Term Structure of Interest-Rate Swap Yields. Journal of Finance 52:1287–1321

133. Dumas B, Fleming J, Whaley RE (1996) Implied Volatility Functions: Empirical Tests. Journal of Finance 53:2059–2106

134. Dupire B (1994) Pricing with a smile. Risk 7:18–20

135. Elerian O, Chib S, Shephard N (2001) Likelihood Inference for Discretely Observed Nonlinear Diffusions. Econometrica 69:959–994

136. Engle RF (2002) New frontiers for ARCH models. Journal of Applied Econometrics 17:425–446

137. Engle RF, Gallo GM (2006) A multiple indicators model for volatility using intra-daily data. Journal of Econometrics 131:3–27

138. Engle RF, Ghysels E, Sohn B (2006) On the Economic Sources of Stock Market Volatility. Working Paper, NYU and UNC

139. Engle RF, Rangel JG (2006) The Spline-GARCH Model for Low Frequency Volatility and Its Global Macroeconomic Causes. Working Paper, NYU

140. Eraker B (2001) MCMC Analysis of Diffusions with Applications to Finance. Journal of Business & Economic Statistics 19:177–191

141. Eraker B (2004) Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices. Journal of Finance 59: 1367–1404

142. Eraker B, Johannes MS, Polson N (2003) The Impact of Jumps in Volatility and Returns. Journal of Finance 58: 1269–1300

143. Fan R, Gupta A, Ritchken P (2003) Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets. Journal of Finance 58:2219–2248

144. Fiorentina G, Sentana E, Shephard N (2004) Likelihood-Based Estimation of Latent Generalized ARCH Structures. Econometrica 72:1481–1517

145. Fisher M, Gilles C (1996) Estimating exponential-affine models of the term structure. Working Paper, Atlanta FED

146. Fleming MJ, Remolona EM (1999) Price Formation and Liquidity in the U.S. Treasury Market: The Response to Public Information. Journal of Finance 54:1901–1915

147. Forsberg L, Ghysels E (2007) Why Do Absolute Returns Predict Volatility So Well? Journal of Financial Econometrics 5:31–67

148. French KR, Schwert GW, Stambaugh RF (1987) Expected stock returns and volatility. Journal of Financial Economics 19:3–29

149. Fridman M, Harris L (1998) A Maximum Likelihood Approach for Non-Gaussian Stochastic Volatility Models. Journal of Business & Economic Statistics 16:284–291

150. Gallant AR, Hsieh DA, Tauchen GE (1997) Estimation of Stochastic Volatility Models with Diagnostics. Journal of Econometrics 81:159–192

151. Gallant AR, Hsu C, Tauchen GE (1999) Using Daily Range Data to Calibrate Volatility Diffusions and Extract the Forward Integrated Variance. Review of Economics and Statistics 81:617–631

152. Gallant AR, Long JR (1997) Estimating stochastic differential equations efficiently by minimum chi-squared. Biometrika 84:125–141

153. Gallant AR, Rossi PE, Tauchen GE (1992) Stock Prices and Volume. Review of Financial Studies 5:199–242

154. Gallant AR, Tauchen GE (1996) Which Moments to Match. Econometric Theory 12:657–681

155. Gallant AR, Tauchen G (1998) Reprojecting Partially Observed Systems With Application to Interest Rate Diffusions. Journal of the American Statistical Association 93:10–24

156. Garcia R, Lewis MA, Pastorello S, Renault E (2001) Estimation of Objective and Risk-neutral Distributions based on Moments of Integrated Volatility Working Paper, Université de Montréal, Banque Nationale du Canada, Università di Bologna, UNC

157. Garman MB, Klass MJ (1980) On the Estimation of Price Volatility From Historical Data. Journal of Business 53:67–78

158. Ghysels E, Harvey AC, Renault E (1996) Stochastic Volatility. In: Maddala GS and Rao CR (Eds) Handbook of Statistics, Volume 14, North Holland, Amsterdam, Holland

159. Ghysels E, Santa-Clara P, Valkanov R (2006) Predicting Volatility: How to Get the Most Out of Returns Data Sampled at Different Frequencies. Journal of Econometrics 131:59–95

160. Glosten LR, Jagannathan R, Runkle D (1993) On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Finance 48:1779–1801

161. Gong FF, Remolona EM (1996) A three-factor econometric model of the U.S. term structure. Working paper, Federal Reserve Bank of New York.

162. Gouriéroux C, Monfort A, Renault E (1993) Indirect Inference. Journal of Applied Econometrics 8:S85–S118

163. Guo H, Neely CJ, Higbee J (2007) Foreign Exchange Volatility Is Priced in Equities. Financial Management, forthcoming

164. Han B (2007) Stochastic Volatilities and Correlations of Bond Yields. Journal of Finance 62:1491–1524

165. Hansen PR, Lunde A (2006) Realized Variance and Market Microstructure Noise. Journal of Business and Economic Statistics 24:127–161

166. Harvey AC (1998) Long memory in stochastic volatility. In: Knight J, Satchell S (Eds) Forecasting Volatility in Financial Markets, Butterworth-Heinemann, London.

167. Harvey AC, Ruiz E, Shephard N (1994) Multivariate Stochastic Variance Models. Review of Economic Studies 61:247–264

168. Harvey AC, Shephard N (1996) Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns. Journal of Business & Economic Statistics 14:429–434

169. Heidari M, Wu L (2003) Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates? Journal of Fixed Income 13:75–86

170. Heston SL (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of Financial Studies 6:327–343

171. Ho M, Perraudin W, Sørensen BE (1996) A Continuous Time Arbitrage Pricing Model with Stochastic Volatility and Jumps. Journal of Business & Economic Statistics 14:31–43

172. Huang X, Tauchen G (2005) The relative contribution of jumps to total price variation. Journal of Financial Econometrics 3:456–499

173. Huang J, Wu L (2004) Specification Analysis of Option Pricing Models Based on Time-Changed Levy Processes. Journal of Finance 59:1405–1440

174. Hull J, White A (1987) The Pricing of Options on Assets with Stochastic Volatilities. Journal of Finance 42:281–300

175. Hurvich CM, Soulier P (2007) Stochastic Volatility Models with Long Memory. In: Andersen TG, Davis RA, Kreiss JP, Mikosch T (Eds) Handbook of Financial Time Series, Springer Verlag.

176. Jackwerth JC, Rubinstein M (1996) Recovering Probability Distributions from Option Prices. Journal of Finance 51:1611–1631

177. Jacobs K, Karoui L (2007) Conditional Volatility in Affine Term Structure Models: Evidence from Treasury and Swap Markets. Working Paper, McGill University

178. Jacobs K, Li X (2007) Modeling the Dynamics of Credit Spreads with Stochastic Volatility. Management Science, forthcoming

179. Jacquier E, Polson NG, Rossi PE (1994) Bayesian Analysis of Stochastic Volatility Models. Journal of Business & Economic Statistics 12:371–389

180. Jacquier E, Polson NG, Rossi PE (2004) Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics 122:185–212

181. Jagannathan R, Kaplin A, Sun S (2003) An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices. Journal of Econometrics 116:113–146

182. Jarrow R, Li H, Zhao F (2007) Interest Rate Caps "Smile" Too! But Can the LIBOR Market Models Capture the Smile? Journal of Finance 62:345–382

183. Jiang GJ, Knight JL (2002) Efficient Estimation of the Continuous Time Stochastic Volatility Model via the Empirical Characteristic Function. Journal of Business & Economic Statistics 20:198–212

184. Jiang GJ, Tian YS (2005) The Model-Free Implied Volatility and Its Information Content. Review of Financial Studies 18:1305–1342.

185. Johannes MS, Polson N (2003) MCMC Methods for Continuous-Time Financial Econometrics. In: Hansen LP, Aït-Sahalia I (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

186. Johannes MS, Polson N (2006) Particle Filtering. In: Andersen TG, Davis RA, Kreiss JP, Mikosch T (Eds) Handbook of Financial Time Series, Springer Verlag.

187. Johnson H, Shanno D (1987) Option Pricing when the Variance Is Changing. Journal of Financial and Quantitative Analysis 22:143–152.

188. Jones CS (2003a) Nonlinear Mean Reversion in the Short-Term Interest Rate Review of Financial Studies 16:793–843

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

190. Joslin S (2006) Can Unspanned Stochastic Volatility Models Explain the Cross Section of Bond Volatilities?. Working Paper, MIT

191. Joslin S (2007) Pricing and Hedging Volatility Risk in Fixed Income Markets. Working Paper, MIT

192. Jung J (2006) Vexed by variance. Risk

193. Kim SN, Shephard N, Chib S (1998) Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. Review of Economic Studies 65:361–393

194. Lamoureux CG, Lastrapes WD (1994) Endogenous Trading Volume and Momentum in Stock-Return Volatility. Journal of Business & Economic Statistics 14:253–260

195. Lee SS, Mykland PA (2006) Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. Working Paper, Georgia Institute of Technology and University of Chicago

196. Li H, Wells MT, Yu CL (2006) A Bayesian Analysis of Return Dynamics with Lévy Jumps Review of Financial Studies, forthcoming

197. Li H, Zhao F (2006) Unspanned Stochastic Volatility: Evidence from Hedging Interest Rate Derivatives. Journal of Finance 61:341–378

198. Liesenfeld R (1998) Dynamic Bivariate Mixture Models: Modeling the Behavior of Prices and Trading Volume. Journal of Business & Economic Statistics, 16, 101–109

199. Liesenfeld R (2001) A Generalized Bivariate Mixture Model for Stock Price Volatility and Trading Volume. Journal of Econometrics 104:141–178

200. Liesenfeld R, Richard, J-F (2003) Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics. Journal of Empirical Finance 10:505–531

201. Litterman R, Scheinkman JA (1991) Common Factors Affecting Bond Returns. Journal of Fixed Income 1:54–61

202. Liu C, Maheu JM (2007) Forecasting Realized Volatility: A Bayesian Model Averaging Approach. Working Paper, University of Toronto

203. Lo AW (1988) Maximum likelihood estimation of generalized Itô processes with discretely-sampled data. Econometric Theory 4:231–247

204. Longstaff FA, Schwartz ES (1992) Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model. Journal of Finance 47:1259–1282

205. McAleer M, Medeiros MC (2007) Realized Volatility: A Review. Econometric Reviews, forthcoming

206. McFadden D (1989) A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. Econometrica 57:995–1026

207. McQueen G, Vorkink K (2004) Whence GARCH? A Preference-Based Explanation for Conditional Volatility. Review of Financial Studies 17:915–949

208. Meddahi N (2001) An Eigenfunction Approach for Volatility Modeling. Working Paper, Imperial College

209. Meddahi N (2002) A Theoretical Comparison Between Integrated and Realized Volatility. Journal of Applied Econometrics 17:475–508

210. Meddahi N (2002) Moments of Continuous Time Stochastic Volatility Models. Working Paper, Imperial College

211. Melino A, Turnbull SM (1990) Pricing foreign currency options with stochastic volatility. Journal of Econometrics 45:239–265

212. Merton RC (1969) Lifetime portfolio selection under uncertainty: the continuous-time case. Review of Economics and Statistics 51:247–257

213. Merton RC (1973) An Intertemporal Capital Asset Pricing Model. Econometrica 41:867–887

214. Merton RC (1976) Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3:125–144

215. Merton RC (1980) On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8:323–361

216. Mizrach B (2006) The Enron Bankruptcy: When Did The Options Market Lose Its Smirk. Review of Quantitative Finance and Accounting 27:365–382

217. Mizrach B (2007) Recovering Probabilistic Information From Options Prices and the Underlying. In Lee C, Lee AC (eds), Handbook of Quantitative Finance, Springer-Verlag, New York

218. Nelson DB (1991) Conditional Heteroskedasticity in Asset Returns: A New Approach. Econometrica 59:347–370

219. Nelson DB, Foster DP (1994) Asymptotic Filtering Theory for Univariate ARCH Models. Econometrica 62:1–41

220. Pan J (2002) The jump-risk premia implicit in options: evidence from an integrated time-series study. Journal of Financial Economics 63:3–50

221. Parkinson M (1980) The Extreme ValueMethod for Estimating the Variance of the Rate of Return. Journal of Business 53:61-65

222. Pasquariello P, Vega C (2006) Informed and Strategic Order Flow in the Bond Markets. Review of Financial Studies, forthcoming

223. Pearson ND, Sun TS (1994) Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll, and Ross Model. Journal of Finance 49:1279–1304

224. Pedersen AR (1995) A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. Scandinavian Journal of Statistics 22:55–71

225. Pennacchi GG (1991) Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data. Review of Financial Studies 4:53–86

226. Pitt MK, Shephard N (1999) Filtering via simulation: auxiliary particle filter. Journal of the American Statistical Association 94:590–599

227. Piazzesi M (2003) Affine Term Structure Models. In: Hansen LP, Aït-Sahalia I (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

228. Piazzesi M (2005) Bond Yields and the Federal Reserve. Journal of Political Economy 113:311–344

229. Protter P (1992) Stochastic Integration and Differential Equations: A New Approach. Springer-Verlag, New York, New York, U.S.

230. Renault E (1997) Econometric models of option pricing errors. In: Kreps D, Wallis K (Eds), Advances in Economics and Econometrics. Seventh World Congress. Cambridge University Press, New York and Melbourne, 223-278.

231. Richardson M, Smith T (1994) A Direct Test of the Mixture of Distributions Hypothesis: Measuring the Daily Flow of Information. Journal of Financial and Quantitative Analysis 29:101–116

232. Rosenberg B (1972). The Behavior of Random Variables with Nonstationary Variance and the Distribution of Security Prices. Working Paper, UCB, Berkeley

233. Rubinstein M (1994) Implied Binomial Trees. Journal of Finance 49:771–818

234. Santa-Clara P (1995) Simulated likelihood estimation of diffusions with an application to the short term interest rate. Ph.D. Dissertation, INSEAD

235. Schaumburg E (2005) Estimation of Markov processes with Levy type generators. Working Paper, KGSM

236. Schwert GW (1989) Why Does Stock Market Volatility Change Over Time? Journal of Finance 44:1115–1153

237. Schwert GW (1990) Stock Volatility and the Crash of '87. Review of Financial Studies 3:77-102

238. Scott LO (1987) Option Pricing when the Variance Changes Randomly: Theory, Estimation and an Application. Journal of Financial and Quantitative Analysis 22:419–438

239. Shephard N (1996) Statistical Aspects of ARCH and Stochastic Volatility Models. In: Cox DR, Hinkley DV, Barndorff-Nielsen OE (Eds) Time Series Models in Econometrics, Finance and Other Fields, Chapman & Hall, London, UK, pp. 1–67

240. Shephard N (2004) Stochastic Volatility: Selected Readings. Oxford University Press, Oxford, UK

241. Shiller RJ (1981) Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? American Economic Review 71:421–436

242. Singleton KJ (2001) Estimation of affine asset pricing models using the empirical characteristic function. Journal of Econometrics 102:111–141

243. Smith, Jr., AA (1993) Estimating Nonlinear Time-series Models using Simulated Vector Autoregressions. Journal of Applied Econometrics 8:S63–S84

244. Stein JC (1989) Overreactions in the Options Market. Journal of Finance 44:1011–1023

245. Stein EM, Stein JC (1991) Stock price distributions with stochastic volatility: an analytic approach. Review of Financial Studies 4:727–752

246. Tauchen GE (2005) Stochastic Volatility in General Equilibrium. Working Paper, Duke

247. Tauchen GE, Pitts M (1983) The Price Variability-Volume Relationship on Speculative Markets. Econometrica 51:485–505

248. Tauchen GE, Zhou H (2007) Realized Jumps on Financial Markets and Predicting Credit Spreads. Working Paper, Duke and Board of Governors

249. Taylor SJ (1986) Modeling Financial Time Series. John Wiley and Sons, Chichester, UK

250. Thompson S (2004) Identifying Term Structure Volatility from the LIBOR-Swap Curve. Working Paper, Harvard University

251. Todorov V (2006a) Variance Risk Premium Dynamics. Working Paper, KGSM

252. Todorov V (2006b) Estimation of Continuous-time Stochastic Volatility Models with Jumps using High-Frequency Data. Working Paper, KGSM

253. Trolle AB, Schwartz ES (2007a) Unspanned stochastic volatility and the pricing of commodity derivatives. Working Paper, Copenhagen Business School and UCLA

254. Trolle AB, Schwartz ES (2007b) A general stochastic volatility model for the pricing of interest rate derivatives. Review of Financial Studies, forthcoming

255. Vasicek OA (1977) An equilibrium characterization of the term structure. Journal of Financial Economics 5:177–188

256. Wiggins JB (1987) Option Values under Stochastic Volatility: Theory and Empirical Estimates. Journal of Financial Economics 19:351–372

257. Whaley RE (1993) Derivatives on Market Volatility: Hedging Tools Long Overdue. Journal of Derivatives 1:71–84

258. Wright J, Zhou H (2007) Bond Risk Premia and Realized Jump Volatility. Working Paper, Board of Governors

259. Yang D, Zhang Q (2000) Drift-Independent Volatility Estimation Based on High, Low, Open, and Close Prices. Journal of Business 73:477-491

260. Zhang BY, Zhou H, Zhu H (2005) Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms. Working Paper, Fitch Ratings, Federal Reserve Board, BIS

261. Zhang L (2007) What you don't know cannot hurt you: On the detection of small jumps. Working Paper, UIC

262. Zhang L, Mykland PA, Aït-Sahalia Y (2005) A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High Frequency Data. Journal of the American Statistical Association 100:1394–1411

Books and Reviews

Asai M, McAleer M, Yu J (2006) Multivariate Stochastic Volatility: A Review. Econometric Reviews 25:145-175

Bates DS (2003) Empirical option pricing: a retrospection, Journal of Econometrics 116:387–404

Campbell JY, Lo AW, MacKinlay AC (1996) The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ, US

Chib S, Omori Y, Asai M (2007) Multivariate Stochastic Volatility. Working Paper, Washington University

Duffie D (2001) Dynamic Asset Pricing Theory. Princeton University Press, Princeton, NJ, US

Gallant AR, Tauchen G (2002) Simulated Score Methods and Indirect Inference for Continuous-time Models. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

Garcia R, Ghysels E, Renault E (2003) The Econometrics of Option Pricing. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

Gouriéroux C, Jasiak J (2001) Financial Econometrics. Princeton University Press, Princeton, NJ, US

Johannes MS, Polson N (2006) Markov Chain Monte Carlo. In: Andersen TG, Davis RA, Kreiss JP, Mikosch T (Eds) Handbook of Financial Time Series, Springer Verlag

Jungbacker B, Koopman SJ (2007) Parameter Estimation and Practical Aspect of Modeling Stochastic Volatility. Working Paper, Vrije Universiteit

Mykland PA (2003) Option Pricing Bounds and Statistical Uncertainty. In: Hansen LP, Aït-Sahalia Y (Eds) Handbook of Financial Econometrics, North-Holland, Amsterdam, Holland, forthcoming

Renault E (2007) Moment-Based Estimation of Stochastic Volatility Models. In: Andersen TG, Davis RA, Kreiss JP, Mikosch T (Eds) Handbook of Financial Time Series, Springer Verlag

Singleton KJ (2006) Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment. Princeton University Press, Princeton, NJ, US

Research Papers 2010

2009-56: Valeri Voev: On the Economic Evaluation of Volatility Forecasts

2009-57: Jesper Rangvid, Maik Schmeling and Andreas Schrimpf: Global Asset Pricing: Is There a Role for Long-run Consumption Risk?

2009-58: Olaf Posch: Risk premia in general equilibrium

2009-59: Christian M. Dahl and Emma M. Iglesias: Modelling the Volatility-Return Trade-off when Volatility may be Nonstationary

2009-60: Ole E. Barndorff-Nielsen, José Manuel Corcuera and Mark Podolskij: Limit theorems for functionals of higher order differences of Brownian semi-stationary processes

2010-01: Anders Bredahl Kock and Timo Teräsvirta: Forecasting with nonlinear time series models

2010-02: Gunnar Bårdsen, Stan Hurn and Zoë McHugh: Asymmetric unemployment rate dynamics in Australia

2010-03: Jesper Rangvid, Maik Schmeling and Andreas Schrimpf: Cash Flow-Predictability: Still Going Strong

2010-04: Helle Bunzel and Walter Enders: The Taylor Rule and "Opportunistic" Monetary Policy

2010-05: Martin M. Andreasen: Non-linear DSGE Models and The Optimized Particle Filter

2010-06: Søren Johansen and Bent Nielsen: Discussion of The Forward Search: Theory and Data Analysis by Anthony C. Atkinson, Marco Riani, and Andrea Ceroli

2010-07: Giuseppe Cavaliere, Anders Rahbek and A.M.Robert Taylor: Bootstrap Sequential Determination of the Co-integration Rank in VAR Models

2010-08: Peter R. Hansen and Asger Lunde: Estimating the Persistence and the Autocorrelation Function of a Time Series that is Measured with Error

2010-09: Tom Engsted, Thomas Q. Pedersen and Carsten Tanggaard: Pitfalls in VAR based return decompositions: A clarification

2010-10: Torben G. Andersen and Luca Benzoni: Stochastic Volatility