Methodol Comput Appl Probab (2015) 17:823–831
DOI 10.1007/s11009-014-9409-4

Estimating Parametric Models of Probability Distributions

Dilip B. Madan

Received: 16 September 2013 / Revised: 13 March 2014 /
Accepted: 17 April 2014 / Published online: 11 May 2014

Abstract Noting that risk neutral distributions are estimated by minimizing the squared deviations between market and model option prices we consider using option payoff moments in estimating distributional parameters from a sample of observations. It is observed, in particular when compared to maximum likelihood estimation, that digital option payoff moments yield the lowest chisquare statistics for a test of uniformity for data transformed to the unit interval by the estimated distribution function.

Keywords Moment estimators · Digital options · Variance gamma · Double gamma model

AMS 2000 Subject Classifications 62-07 · 62P25

1 Introduction

Parametric models of probability densities have been estimated from data on option prices by least squares minimization of the difference between model and market price for well over twenty years. By way of examples we cite Bakshi, Cao and Chen (1997), Bates (1991), Carr, Geman, Madan and Yor (2002), Heston (1993), and Madan, Carr and Chang (1998). When one has time series data on returns the recommended estimation method is that of maximum likelihood, see Fisher (1922). This is clearly the preferred method when data are drawn from the said distribution but many advantages are known to remain when this is not the case White (1982). For complex models much effort is expended in finding or approximating the likelihood function, A ït Sahalia (1999, 2002). Essentially all the methods may be viewed as applications of moment methods (Hansen 1982), with the differences lying in the moments being selected. Gallant and Tauchen (1996) advocate the use of the score function.

D. B. Madan (✉)
Robert H. Smith School of Business, University of Maryland, College Park, MD 20742, USA
e-mail: dbm@rhsmith.umd.edu

Though one may not estimate option pricing models using maximum likelihood, lacking a set of data possibly drawn from the proposed distribution, one may evaluate expectations of call payoffs from data on observed outcomes and then apply option pricing methods to estimate parameters. These are moment matching methods for particular moments mimicking option payoffs. Other moment choices include skewness, kurtosis or the characteristic function (Feuerverger and Mureika 1977).

Once a model has been estimated for a distribution function one may evaluate the estimated distribution function on the data and, presuming the data come from this estimated distribution function, the sequence of evaluations should be uniformly distributed on the unit interval. The different estimation methods may be compared by testing this sequence of distribution function evaluations for uniformity, using for example, a chisquare statistic. One could instead estimate parameters by minimizing the chisquare statistic, but our interest is in selecting moments for moment matching equations that yield the lower values for the chisquare statistic.

Given that we never know whether any real data comes from a particular model we study and report on properties of estimators when the data is known not to come from the proposed model. Of course we shall use particular models to generate some data samples but as we suppose the researcher does not know this model, we cannot employ methods like the correct maximum likelihood or related EM approaches known to be superior or relevant for the particular data generating scheme. The data generation scheme is known to us but not to the researcher who is estimating the wrong model. The researcher also knows he has the wrong model and expects it to be rejected, and is not interested in finding the right model as there may not be one for the data at hand that can be ascertained in the time available. We expect and the resulting chisquare statistics do reject the proposed model, but by construction this is not an observation of interest. The interest is in the the type of moments that yield the lower and hence better chisquare statistics, thereby providing the more acceptable estimated though incorrect model, on this metric. Properties of parameters in small or large samples are also not of interest given that the model is known to be wrong. The interest is just not in the model. One is merely trying to get the best estimates of probabilities from a selection of proposed wrong models, recognizing that the true data generation mechanism lies in all probability beyond our tractable analytical modeling capabilities.

We shall argue in favor of the possibility of the superiority, at least for some examples, for using digital moment estimation relative to maximum likelihood or other methods. Since we do not wish to prove this proposition but merely to point to the possibility, a few examples suffice in the tradition of counterexamples. Other researchers in practical situations are merely being pointed to consider the relevance of digital moment estimation in their particular context.

For the proposed model we work with the two parameter Gaussian model for real valued data and the two parameter gamma distribution for the data on the positive half line. As there are two parameters to be estimated, we employ two moments for the purpose. One of our estimators is the maximum likelihood estimator, henceforth MLE. The second estimator we call the digital moment estimator, DME, that finds parameters to match two quantiles. In addition we consider CME where the moment is the expectation of a call option type payoff, for strikes taken at particular quantiles. We also consider an exponential call payoff at the same strikes $ECME$.

We find that DME consistently yields the lowest chisquare statistic even when the true distribution sought has been exponentially tilted. We conjecture that this is perhaps due to the fact that the digital payoff is a fairly robust bounded moment. This leads us to suggest that distributions should probably not be estimated by maximum likelihood as classically advocated but one may employ instead a digital moment estimator. The number of moments may be adjusted to match or dominate the number of parameters.

The outline of the rest of the paper is as follows. Section 2 reports on results for the Gaussian model. Section 3 reports on results for the gamma model on the upper half line. Section 4 concludes.

2 The Gaussian Model

The Gaussian model postulates that the underlying probability distribution is the normal distribution with mean μ and variance σ^2. For a sample of data x_t, $t = 1, \cdots, T$ possibly drawn from this distribution the MLE estimators are also the mean and variance moment estimates with

$$\begin{aligned}\widehat{\mu}_m &= \frac{1}{T}\sum_t x_t \\ \widehat{\sigma}_m^2 &= \frac{1}{T}\sum_t (x_t - \widehat{\mu})^2 .\end{aligned}$$

The DME estimators define strikes $k_1 < k_2$ for quantiles $a_1 < a_2$ by

$$\frac{1}{T}\sum_t \mathbf{1}_{x_t \le k_i} = a_i, \quad i = 1, 2.$$

The DME parameter estimates $\widehat{\mu}_d, \widehat{\sigma}_d$ solve the equations

$$N\left(\frac{k_i - \widehat{\mu}_d}{\widehat{\sigma}_d}\right) = a_i,$$

where $N(x)$ is the standard normal distribution function. It follows that

$$\begin{aligned}\widehat{\sigma}_d &= \frac{k_2 - k_1}{N^{-1}(a_2) - N^{-1}(a_1)} \\ \widehat{\mu}_d &= k_1 - \widehat{\sigma}_d N^{-1}(a_1).\end{aligned}$$

The CME parameter estimates $\widehat{\mu}_c, \widehat{\sigma}_c$ solve for

$$c_i = \frac{1}{T}\sum_t (x_t - k_i)^+ ,$$

the equations

$$\begin{aligned}c_i &= \int_{k_i}^{\infty} (x - k_i)\frac{1}{\widehat{\sigma}_c\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x - \widehat{\mu}_c)^2}{\widehat{\sigma}_c^2}\right) dx \\ &= \int_{(k_i - \widehat{\mu}_c)/\widehat{\sigma}_c}^{\infty} (\widehat{\mu}_c + \widehat{\sigma}_c z - k_i) n(z) dz \\ &= \frac{\widehat{\sigma}_c}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{k_i - \widehat{\mu}_c}{\widehat{\sigma}_c}\right)^2\right) + (\widehat{\mu}_c - k_i)\, N\left(\frac{\widehat{\mu}_c - k_i}{\widehat{\sigma}_c}\right).\end{aligned}$$

Similarly the $ECME$ parameter estimates $\widehat{\mu}_e, \widehat{\sigma}_e$ solve for

$$e_i = \frac{1}{T}\sum_t \left(e^{x_t} - e^{k_i}\right)^+$$

the equations

$$\begin{aligned} e_i &= N(d_1) - e^{k_i}N(d_2) \\ d_1 &= \frac{\widehat{\mu}_e}{\widehat{\sigma}_e} + \widehat{\sigma}_e \\ d_2 &= \frac{\widehat{\mu}_e}{\widehat{\sigma}_e} \end{aligned}$$

that follow from the Black, Merton and Scholes option pricing model for a drift of $\widehat{\mu}_e + \widehat{\sigma}_e^2/2$ and a volatility of $\widehat{\sigma}_e$.

We present two sets of results on the Gaussian model. First we generate a thousand sets of samples each with 10,000 points drawn from a centered variance gamma model (Madan and Seneta 1990; Madan, Carr and Chang 1998). The centered variance gamma distribution has three parameters σ, ν, θ that may be sampled by first sampling a gamma variable g with shape and scale parameter equal to $1/\nu$ and the variance gamma variable X is then, for Z a standard normal variable,

$$X = \theta(g-1) + \sigma\sqrt{g}Z.$$

When $\theta = \nu = 0$ the centered variance gamma reduces to a normally distributed variable with zero mean and variance σ^2. The parameter θ calibrates skewness while ν is a measure of excess kurtosis. By taking $\nu = .15$ and $\theta = -.03$ we sample from a distribution that is not Gaussian, for which we estimate mean and variance by MLE, DME and CME. The quantiles used were .25 and .75.

For each of the thousand samples of 10,000 points we transform the data to variables

$$u_t^w = N\left(\frac{x_t - \widehat{\mu}_w}{\widehat{\sigma}_w}\right), \quad t = 1, \cdots, 10000, \quad w = m, d, c.$$

For 20 uniformly constructed cells on the unit interval we evaluate three chisquare statistics to test for the uniformity of $u_t^w, w = m, d, c$. This is done for each of a thousand samples with 10,000 points. Figure 1 presents a graph of the resulting chisquare statistics. We observe that the lowest chisquare statistics are associated with the DME estimates, followed by CME and MLE.

The second set of results we report for the Gaussian model incorporate the exponential call payoffs or the estimator $ECME$ with parameter estimates $\widehat{\mu}_e, \widehat{\sigma}_e$ where we now evaluate for uniformity of the exponentially tilted distribution function defined by

$$G(y) = \int_{-\infty}^{y} \frac{e^x g(x)dx}{\int_{-\infty}^{\infty} e^x g(x)dx}$$

where we take $g(x)$ to be

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

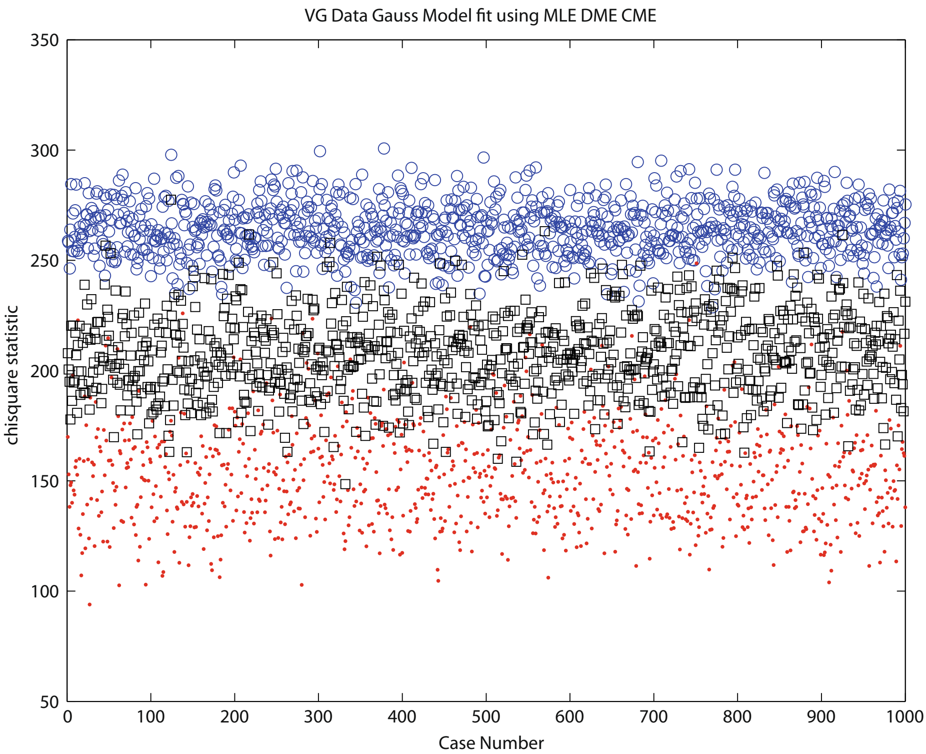

Fig. 1 Chisquare statistics for MLE, DME and CME estimates of the Gaussian model on VG data transformed to the unit interval by the estimated distribution function. The MLE statistics are represented by circles, CME by squares and DME by dots

We may evaluate

$$\begin{aligned} G(y) &= \int_{-\infty}^{\frac{y-\mu}{\sigma}} \frac{e^{\mu+\sigma z}}{e^{\mu+\sigma^2/2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\ &= \int_{-\infty}^{\frac{y-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\sigma)^2}{2}} dz \\ &= N\left(\frac{y-\mu}{\sigma} - \sigma\right). \end{aligned} \tag{1}$$

We may see from Eq. 1 that if x_t is normally distributed with mean μ and variance σ^2 then

$$G\left(x_t + \sigma^2\right) = N\left(\frac{x_t - \mu}{\sigma}\right)$$

is a uniformly distributed variate. Exponential tilting of a normal random variable is just another normal variate with a mean shifted by σ^2 . Evaluating $G(x_t)$ in place of $G(x_t + \sigma^2)$ instead is another model error and here we are just observing how much the departure from uniformity is being affected for each of the estimators by another error specification or wrong model construction.

We present in Fig. 2 the chisquare statistics for the MLE, DME, CME and ECME estimators evaluating the uniformity of the data transformed by the distribution function of Eq. 1. The data is from a variance gamma variable with the same parametric settings as before. We observe again that the lowest chisquare statistics are those of DME followed now by ECME, MLE and CME.

3 The Gamma Model

On the positive half line we consider the gamma distribution as the proposed model with scale parameter c and shape parameter γ with density

$$f(x) = \frac{c^{\gamma}}{\Gamma(\gamma)} x^{\gamma-1} e^{-cx}, \quad x > 0.$$

We now compare the DME, CME and MLE estimators for data generated from a richer distribution than that of the postulated gamma model. In this regard consider the gamma distribution as a model for data generated from a double gamma model where the shape parameter of the second gamma distribution is itself sampled as affine in a first gamma

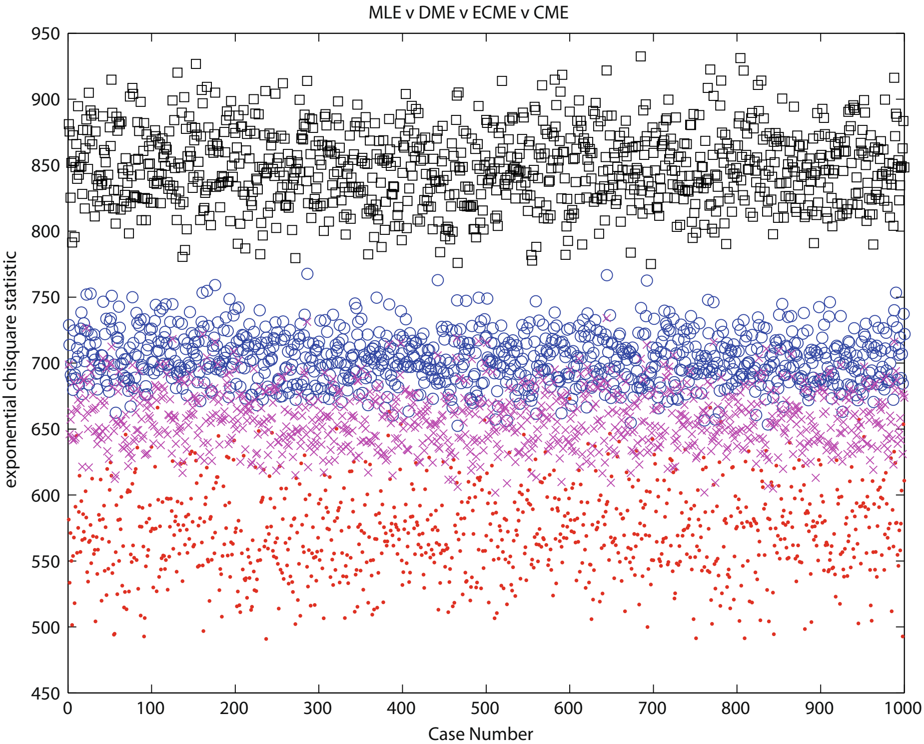

Fig. 2 Chisquare statistics for MLE, DME, CME and ECME estimators testing for uniformity of data transformed by a tilted distribution function. MLE is represented by circles, DME by dots, CME by squares and ECME by stars

variable. The joint density for the pair v, x is

$$k(v,x) = \frac{\gamma^{\gamma} v^{\gamma-1} e^{-\gamma v}}{\Gamma(\gamma)} \frac{\alpha^{\beta+\delta v} x^{\beta+\delta v-1} e^{-\alpha x}}{\Gamma(\beta+\delta v)}$$

and the marginal density for x is

$$h(x) = \int_0^{\infty} k(v,x) dv$$

where we fit a gamma distribution to the data on x.

For the MLE estimator we maximize over parameters $\widehat{c}_m, \widehat{\gamma}_m$ the log likelihood for data $x_t, t = 1, \cdots, T$, given by

$$\mathcal{L}((x_t, t = 1, \cdots, T), \widehat{c}_m, \widehat{\gamma}_m) = T\widehat{\gamma}_m \ln(\widehat{c}_m) - T \ln\Gamma(\widehat{\gamma}_m) + (\widehat{\gamma}_m - 1)\sum_t \ln(x_t) - \widehat{c}_m \sum_t x_t$$

For the DME estimates $\widehat{c}_d, \widehat{\gamma}_d$ we solve the equations for k_i being the a_i quantiles

$$\frac{1}{\Gamma(\widehat{\gamma}_d)} \int_0^{\widehat{c}_d k_i} u^{\widehat{\gamma}_d - 1} e^{-u} du = a_i$$

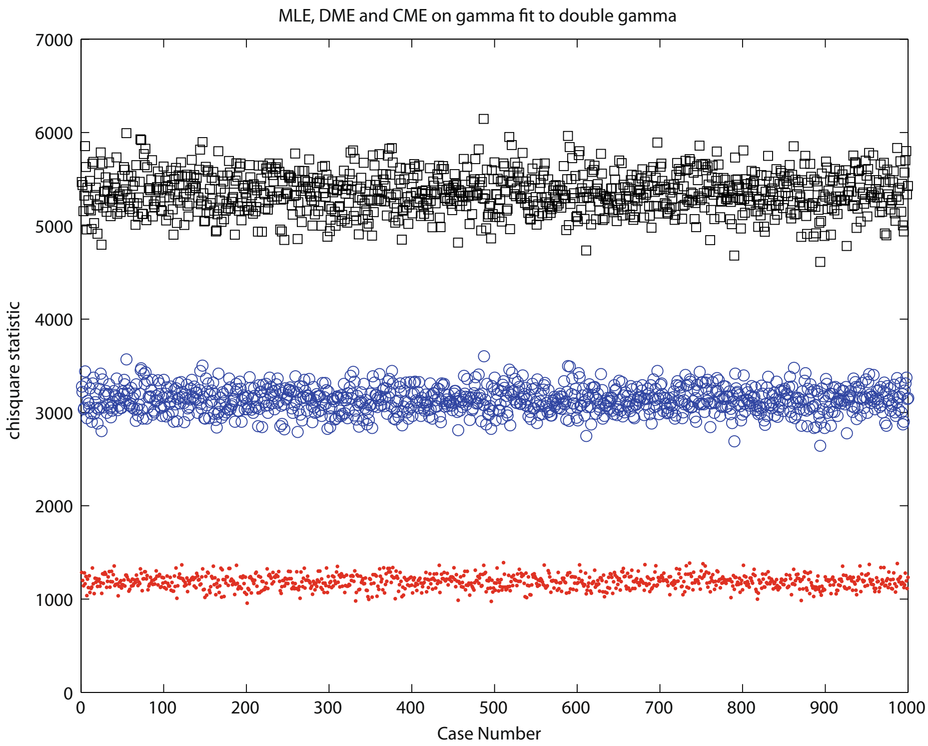

Fig. 3 Chisquare statistics for the gamma distribution model fit to double gamma data using MLE, DME and CME. MLE is represented by circles, CME is presented bu squares and DME are shown by dots

For the CME estimates $\widehat{c}_c, \widehat{\gamma}_c$ we solve the equations

$$c_i = \frac{\widehat{\gamma}_c}{\widehat{c}_c \Gamma(\widehat{\gamma}_c)} \int_0^{\widehat{c}_c k_i} u^{\widehat{\gamma}_c} e^{-u} du - k_i \int_0^{\widehat{c}_c k_i} u^{\widehat{\gamma}_c - 1} e^{-u} du$$
$$c_i = \frac{1}{T} \sum_t (x_t - k_i)^+ .$$

We generated data from the double gamma model with parameters that were used in fitting VIX options of

$$\gamma = 0.1$$
$$\alpha = 0.01$$
$$\beta = 3$$
$$\delta = 3$$

We then generated a thousand sets of 10,000 points from this data and fit a gamma distribution using MLE, DME and CME. The quantiles used were .25 and .75. We then transformed the data by the estimated distribution function to compute chisquare statistics for a test of unifromity for the transformed data. The result is displayed in Fig. 3. We again observe that the DME estimators yield the lowest chisquare statistics followed by MLE and CME.

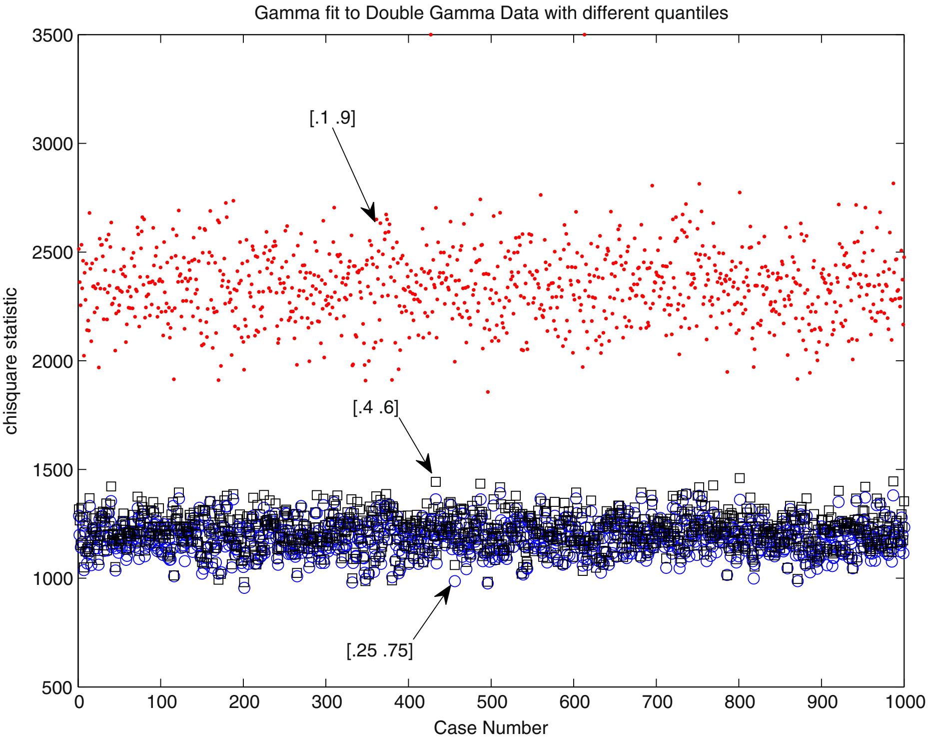

Fig. 4 Chisquare statistics for the gamma distribution fit to data from the double gamma model for DME at quantiles [.4,.6], [.25,.57] and [.1,.9]. The first two quantile sets are represented by squares and circles respectively while the last is represented by dots

Finally with a view to addressing the sensitivity of DME with respect to the choice of the quantiles used we present in Fig. 4 the gamma fit to the double gamma using digitals at [0.4 0.6], [0.1 0.9] along with [0.25 0.75]. We observe that the quantiles [.25, .75] and [.4, .6] perform equally well and significantly better than [.1, .9].

4 Conclusion

Following the practice in estimating risk neutral distributions of minimizing by least squares the deviation between market and model prices for options we consider the use of option payoff moments in estimating parameters of distributions from a sample of data on observations from a possibly related distribution. It is observed that digital option payoff moments yield the lowest chisquare statistics on a test of uniformity for data transformed to the unit interval by the estimated distribution function. For a two parameter proposed model, the use of the first and thrid quartiles as moments appears to be adequate.

References

Aït Sahalia Y (1999) Transition densities for interest rate and other nonlinear diffusions. J Finance 54:1361–1395

Aït Sahalia Y (2002) Maximum-likelihood estimation of discretely-sampled diffusions: a closed-form approximationapproach. Econometrica 70:223–262

Bakshi G, Cao C, Chen Z (1997) Empirical performance of alternative option pricing models. J Finance 52:2003–2049

Bates D (1991) The crash of '87: was it expected? The evidence from options markets. J Finance 46:1009–1044

Carr P, Geman H, Madan DB, Yor M (2002) The fine structure of asset returns: an empirical investigation. J Bus 75:305–332

Fisher RA (1922) On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London. Series A 222, pp. 309-368

Gallant AR, Tauchen G (1996) Which moments to match? Econ Theory 12:657–661

Hansen LP (1982) Large sample properties of generalized method of moments estimators. Econometrica 50:1029–1054

Heston S (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. Rev Financ Stud 6:327–343

Feuerverger A, Mureika RA (1977) The empirical characteristic function and its applications. Annals Stat 5:88–97

Madan DB, Seneta E (1990) The variance gamma (VG) model for share market returns. J Bus 63:511–524

Madan DB, Carr P, Chang E (1998) The variance gamma process and option pricing. Eur Finance Rev 2:79–105

White H (1982) Maximum likelihood estimation of misspecified models. Econometrica 50:1–25