European Finance Review 3: 311–317, 1999.

Comment on 'Index Option Pricing Models with Stochastic Volatility and Stochastic Interest Rates'

BENT JESPER CHRISTENSEN
School of Economics and Management, University of Aarhus, Building 350, University Park, DK-8000 Aarhus C, Denmark; E-mail: bjchristensen@econ.au.dk

1. Introduction

The option pricing model has been the work horse in the development of financial economics ever since the seminal papers by Black and Scholes (1973) and Merton (1973). Here, the arbitrage free prices of equity options were derived in closed form. Apart from the price of the underlying asset and contractual terms (strike price and expiration date), the resulting formula only depends on two parameters of the economic environment, both assumed constant over the life of the option: The rate of interest and the volatility of the underlying asset. Subsequent empirical research has shown that the central Black-Scholes option pricing formula (henceforth the BS formula) does not perform equally well during all time periods. Thus, Rubinstein (1994) finds that the implied volatility "smile" is more prominent after the October 1987 stock market crash than before. Thus, the implied volatilities lining up observed option prices with their BS values vary more with strike price in recent periods, suggesting that the BS formula performs best before the crash. Christensen and Prabhala (1998) find that BS implied volatility of one month at-the-money S&P 100 (OEX) call options forecast the subsequently realized index return volatility over the life of the option well after the crash, but less so before, suggesting that the BS formula performs best after the crash. Clearly, considerable modelling effort is required to explain both the time series (forecasting) and cross-sectional (across strikes) evidence before and after the crash.

2. Stochastic Volatility and Stochastic Interest Rates

Option pricing models allowing for stochastic volatility, stochastic interest rates and jumps have been estimated by Bakshi, Cao and Chen (1997), among others. Jiang and van der Sluis (2000) (henceforth JS) rule out jumps and consider the bivariate discrete time stochastic volatility (SV) model

$$y_{s,t} = \sigma_{s,t}\varepsilon_{s,t}, \tag{1}$$

$$y_{r,t} = \sigma_{r,t}\varepsilon_{r,t}, \tag{2}$$

where $y_{s,t}$ and $y_{r,t}$ are the de-meaned log differences of equity index levels and interest rates at time t, and the shocks $\varepsilon_{s,t}$ and $\varepsilon_{r,t}$ are jointly normally distributed and i.i.d. over time, with contemporaneous correlation λ_1. The BS model would imply $\sigma_{s,t}$ constant and $\sigma_{r,t} = 0$. Instead, JS let the log volatilities follow mean reverting Gaussian processes,

$$\log \sigma^2_{s,t+1} = w_s + \gamma_s \log \sigma^2_{s,t} + \sigma_s \eta_{s,t}, \tag{3}$$

$$\log \sigma^2_{r,t+1} = w_r + \gamma_r \log \sigma^2_{r,t} + \sigma_r \eta_{r,t}, \tag{4}$$

and impose the structure

$$\begin{pmatrix} \varepsilon_{s,t} \\ \varepsilon_{r,t} \\ \eta_{s,t} \\ \eta_{r,t} \end{pmatrix} \backsim IIN \left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \lambda_1 & \lambda_2 & \lambda_1\lambda_3 \\ \lambda_1 & 1 & \lambda_1\lambda_2 & \lambda_3 \\ \lambda_2 & \lambda_1\lambda_2 & 1 & \lambda_1\lambda_2\lambda_3 \\ \lambda_1\lambda_3 & \lambda_3 & \lambda_1\lambda_2\lambda_3 & 1 \end{pmatrix} \right), \tag{5}$$

where $|\lambda_i| \leq 1$. Thus, (3) relaxes the BS constant volatility assumption and indeed allows for genuine stochastic volatility, i.e., instead of specifying e.g.,

$$\sigma_{s,t} = \sigma\left(S_t, t\right), \tag{6}$$

where S_t is the index level and $\sigma(\cdot)$ a known deterministic function, a separate shock $\eta_{s,t}$ not necessarily perfectly correlated with $\varepsilon_{s,t}$ is entered. This immediately leaves the resulting model incomplete, thus precluding exact hedging and arbitrage pricing, and calling for specification of economic risk premia for pricing purposes. Even without dropping the complete market assumption, the simpler specification (6) with a decreasing function $\sigma(\cdot)$ would have given a leverage effect, i.e., drops in asset prices lead to increased return volatility. In (1)–(3), this is captured if $\lambda_2 < 0$, i.e., negative $\varepsilon_{s,t}$ (drop in S_t) is associated with positive $\eta_{s,t}$ (increased $\sigma^2_{s,t+1}$).

Specification (2) relaxes the BS constant interest rate assumption. Previous research has considered stochastic processes for the interest rate, i.e. included a shock such as $\varepsilon_{r,t}$ in (2). However, JS go one step further and even allow the interest rate to follow an SV process. The cost is loss of parsimony, i.e., $\eta_{r,t}$ in (4) is one more shock in the model.

It is hard to say whether (5) is exactly the structure to start with. The assumption is that 6 correlations are explained by 3 parameters λ_i. It would be nice to know how many of the 6 correlations are jointly identified in the data. If this is more than 3, a formal test of the restrictions imposed in (5) should be carried out. Were this test to reject, then all the inferences would be called in question.

The structure (5) does have the feature that the test for complete independence between the index and interest rate processes reduces to a test of the simple hypothesis $\lambda_1 = 0$ (under which 4 of the 6 correlations in (5) vanish). This hypothesis cannot be rejected by JS. More generally, they conclude that interest rates matter little for option prices. In particular, this conclusion is not changed by allowing the extra SV term $\eta_{r,t}$ in (4). This would suggest that modelling efforts should focus on the equity side.

3. Estimation of the Stochastic Volatility Model

The unknown structural parameters in (1)–(5) are $(w_i, \gamma_i, \sigma_i)$, $i = r, s$, and λ_i, $i = 1, 2, 3$, i.e., 9 parameters that are fit by the efficient method of moments (EMM) of Gallant and Tauchen (1996) using index return and interest rate data. This means that first, a parameter vector of much higher dimension is estimated by maximizing a pseudo-likelihood function (an auxiliary model) based on a semi-nonparametric (SNP) density with a MEGARCH leading term. Next, the structural parameters are estimated by choosing them such that simulated data from (1)–(5) are similar to the real data, in the sense that at the maximum pseudo-likelihood estimator from the real data, the pseudo-score vector based on a long simulated series (roughly the expected pseudo-score) is as close to zero as possible in the norm given by its inverse dispersion matrix estimated in the real data.

At least six issues arise, regarding this procedure. First, the efficiency claim depends crucially on the central Assumption 3 in Gallant and Long (1997), which imposes that the true (but unknown) score function may be approximated by a polynomial. This would seem hard to verify in SV models, and as noted by JS, Andersen and Lund (1997) and Gallant, Hsieh and Tauchen (1997) adopt this assumption in such models without explicit proof. This leaves room for interesting future work in the area, perhaps giving explicit proofs in simpler models.

Secondly, the de-meaning leading to $y_{s,t}$ and $y_{r,t}$ by regression techniques actually implies that 3 mean parameters have been estimated before the EMM stage. As noted by JS, this pre-whitening approach is used also by Harvey and Shephard (1996) and Sandmann and Koopman (1998). However, joint estimation of all 12 parameters (pre-whitening and structural) would be more efficient. The situation is similar to APT testing, where the two-pass procedure of estimating betas from time series of returns, then estimating risk premia by regressing returns on betas cross-sectionally is inefficient when factors are observed with error (Christensen (1994)). The general approximate exogeneity (local cut) conditions for separate inference (Christensen and Kiefer (1994, 2000)) are not satisfied in the SV and APT models.

Third, the reported EMM standard errors do not reflect the sampling error involved in the pre-whitening stage and hence in the calculated values $y_{s,t}$ and $y_{r,t}$. Methods for correcting this along the lines of Murphy and Topel (1985) should be developed.

Fourth, in view of the efficiency issues and the standard error issue, and particularly since the referenced theory is asymptotic in nature, a sampling experiment would be of great interest, including for determining how well the asymptotics work for the sample size considered.

Fifth, certain parameters in the auxiliary SNP/MEGARCH model correspond to including the cross-terms $\log \sigma_{r,t}^2$ on the RHS of (3) and $\log \sigma_{s,t}^2$ on the RHS of (4). These additional parameters are actually significant, but this does not lead JS to expand the structural model (1)–(5) accordingly. Marginal significance and a Bayesian information criterion are cited, but the approach is nervewrecking. The point of the EMM is that the next round estimates are consistent and more efficient than the estimates in the first (auxiliary) round. Strong significance and hence the need for the indicated expansion of (3)–(4) might have been found in the second (EMM) round, had these additional parameters not been dropped based on the inefficient auxiliary model fit.

Finally, the entire approach is based on estimation of the structural parameters in (1)–(5) from index and interest rate data, without regard to option price data. Whether this is the strategy of choice is a pressing issue, particularly since the ultimate interest is focused on option pricing implications, and even more so because option price data are the data most sensitive to – and hence most informative about – volatility parameters, and thus all 9 parameters in (1)–(5).

4. Empirical Option Pricing

To put things in perspective, the idea is to fit a model to the return and interest series in the first (EMM) stage, as described above (this stage includes an auxiliary and an EMM round of estimation), then in the second stage see whether the options are in fact priced in accordance with this model, for suitably selected risk premia. The risk premia are estimated by choosing them to minimize the sum of squared differences between observed option prices and the theoretical values that should prevail for given risk premia, treating the parameters estimated in (1)–(5) as known. If the options are priced in accordance with the fitted model from the first (EMM) stage, then this is seen as consistent with rational behavior on the part of option market participants and hence market efficiency.

Specifically, the price of a European call option may be written as

$$
\begin{aligned}
&C\left(S_t, r_t, \sigma_{s,t}, \sigma_{r,t}, \theta; T, X\right) \\
&\quad = E\left(e^{-\int_t^T r_u d_u} \max\{S_T - X, 0\} \,|S_t, r_t, \sigma_{s,t}, \sigma_{r,t}\right), \qquad (7)
\end{aligned}
$$

where the expectation is under the risk-neutral measure setting the expected return on S equal to r and allowing risk premia in the drifts of r_t, $\sigma_{s,t}$ and $\sigma_{r,t}$. JS impose that the risk premia in the drifts of r_t and $\sigma_{r,t}$ are zero. Of course, this restriction is not necessary in general. Further, the risk premium in the drift of $\sigma_{s,t}$ is θ in (7). The Hull and White (1987) hypothesis $\lambda_2 = 0$, i.e., symmetry (no leverage effect) in asset returns, allows reducing (7) to an expected BS formula. Further reductions would be $\lambda_3 = 0$ (symmetry in interest rates), $\sigma_r = 0$ (deterministic interest rates), and $\theta = \sigma_s = 0$, giving the BS formula itself. The second stage minimization problem is

$$\min \sum_i (C_i - C(S, r, \widetilde{\sigma}_s, \widetilde{\sigma}_r, \theta; T_i, X_i))^2, \tag{8}$$

where at a given point in time option prices C_i associated with strike prices X_i and terms to expiration T_i are observed. Data also include the contemporaneous index level S and interest rate r.

Here, four issues arise. First, JS only minimize with respect to the risk premium θ in (8), not the unobserved contemporaneous volatilities $\widetilde{\sigma}_s$ and $\widetilde{\sigma}_r$. Instead, reprojected values from the first stage are inserted.

Secondly, after carrying out the minimization for a given day, the resulting value of θ is used for calculating the average percentage option pricing bias the following day. A fully implicit procedure would back out θ, $\widetilde{\sigma}_r$ and $\widetilde{\sigma}_s$ simultaneously, for the day for which the bias is being calculated, using neither reprojection nor out-of-sample calculations. Reprojection in itself is somewhat unnatural in this context. JS consider

$$\log y_{s,t}^2 = \log \sigma_{s,t}^2 + \log \varepsilon_{s,t}^2, \tag{9}$$

based on (1), and

$$\log \sigma_{s,t}^2 = \alpha_0 + \sum_i \log y_{s,t-i-1}^2 + \log u_{s,t}^2 \tag{10}$$

as an approximation to (1), (3), and similarly for the interest rates. Thus, lagged squared return data are used in (10), instead of the lagged squared volatility in (3), and the approximate filtering system (9)–(10) is used to calculate reprojections $E(\log \sigma_{s,t}^2 | y_{s,t}, y_{s,t-1}, \cdots)$. Since this is the system actually employed for volatility calculation, why not use this as the auxiliary model already in the first (EMM) stage, instead of the SNP/MEGARCH? This would allow a focus on the volatilities, i.e. the quantities required for the second stage. The relative merits of these alternative procedures (fully implicit, partially implicit as in JS, with reprojections from a different model than the one used in EMM, and partially implicit with reprojections from the auxiliary model from EMM) are not well understood.

Third, it may be awkward to compare biases across models when the minimization is not designed to minimize percentage bias (but squared errors).

Finally, the minimization and bias comparison could have been done using option prices either in raw levels or in BS implied volatility terms (i.e., both C_i and C in (8) could be subjected to the inverse BS transform). The latter provides for more stable data, but the former is where the money is.

5. Discussion and Conclusion

Unfortunately, JS find evidence against the model (1)–(5) in the first stage, thus in principle leaving the second stage redundant. Their second stage result is that options are not priced in accordance with the first stage model, but this is now hard to interpret: The first stage model is rejected in the EMM test of overidentifying restrictions, so the finding that options are not priced in accordance with this already rejected model is highly indirect evidence, to say the least, on market efficiency.

The second stage test is designed to look across strikes to pick up whether the SV features are sufficient to explain the smile. Here, it would have been nice to see a figure showing how much of the smile remains in implicit volatilities backed out from the SV formula, i.e., inverting the SV option pricing formula (7) instead of the BS formula. Such a figure would yield more direct evidence than comparison of figures showing BS implied volatilities from real and simulated data. Furthermore, as noted earlier, in addition to the smile/cross-section evidence, models should explain the dynamic/forecasting evidence before and after the 1987 stock market crash. Certainly, it would have been nice to see results excluding the period around the crash (the JS data span the period 1980–1995), since the results may be driven by this. In the literature on the term structure of interest rates, the importance of consistency between cross-sectional (yield curve) and dynamic (stochastic evolution) properties has been clearly recognized (Björk and Christensen (1999)). If the dynamic dimension were incorporated, including squared forecasting errors from the relation between implied and subsequently realized return volatility in the criterion function (8), a more powerful second stage test would result, perhaps revealing more clearly that options indeed are not priced according to the model already rejected in stage one.

Again, testing for the three correlation restrictions embodied in (5) might actually be one route toward a more general model that would not be rejected in the first stage. Otherwise, leaving the Gaussian framework might be the alternative. The latter conclusion seems likely and would be consistent with Dumas, Fleming and Whaley (1998), where (6) was considered in detail and found insufficient to explain data, and with Bakshi, Cao and Chen (1997), where both stochastic interest rates and stochastic jumps (in addition to stochastic volatility) were found empirically important for option pricing.

Still, it should be useful for a wide audience to inspect the large amount of work carried out by JS on the Gaussian case. Given the complexity of the area (models, data dimensions, the computational burden, criterion functions, including hedging performance (not considered by JS)), the detailed study design is key in future research. Hopefully, the article in conjunction with the above comments should be helpful in this direction.

References

Andersen, T. G. and Lund, J. (1997), 'Estimating continuous time stochastic volatility models of the short term interest rate', *Journal of Econometrics* **77**, 343–377.

Bakshi, G., Cao, C. and Chen, Z. (1997), 'Empirical performance of alternative option pricing models', *Journal of Finance* **52**, 2003–2049.

Björk, T. and Christensen, B. J. (1999), 'Interest rate dynamics and consistent forward rate curves', *Mathematical Finance* **9**, 323–348.

Black, F. and Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**, 637–654.

Christensen, B. J. (1994), 'Efficiency gains in beta-pricing models', *Mathematical Finance* **4**, 143–154.

Christensen, B. J. and Kiefer, N. M. (1994), 'Local cuts and separate inference', *Scandinavian Journal of Statistics* **21**, 389–407.

Christensen, B. J. and Kiefer, N. M. (2000), 'Panel data, local cuts, and orthogeodesic models', *Bernoulli* **6**(4), 1–12.

Christensen, B. J. and Prabhala, N. R. (1998), 'The relation between implied and realized volatility', *Journal of Financial Economics* **50**, 125–150.

Dumas, B., Fleming, J. and Whaley, R. (1998), 'Implied volatility functions: Empirical tests', *Journal of Finance* **53**, 2059–2106.

Gallant, A. R. and Long, J. R. (1997), 'Estimating stochastic differential equations efficiently by minimum chi-square', *Biometrika* **84**, 125–141.

Gallant, A. R., Hsieh, D. A. and Tauchen, G. E. (1997), 'On fitting a recalcitrant series: The pound/dollar exchange rate, 1974-1983', in W. A. Barnett, J. Powell and G. E. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Fifth International Symposium in Economic Theory and Econometrics, Cambridge University Press, Cambridge, pp. 199–240.

Gallant, A. R. and Tauchen, G. E. (1996), 'Which moments to match?', *Econometric Theory* **12**, 657–681.

Harvey, A. C. and Shephard, N. G. (1996), 'Estimation of an asymmetric stochastic volatility model for asset returns', *Journal of Economics and Business Statistics* **14**, 429–434.

Hull, J. and White, A. (1987), 'The pricing of options on assets with stochastic volatilities', *Journal of Finance* **42**, 281–300.

Jiang, G. J. and van der Sluis, P. J. (2000), 'Index option pricing models with stochastic volatility and stochastic interest rates', *European Finance Review* **3**, 273–310.

Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**, 141–183.

Murphy, K. M. and Topel, R. (1985), 'Estimation and inference in two-step econometric models', *Journal of Business and Economic Statistics* **3**, 370–379.

Rubinstein, M. (1994), 'Implied binomial trees', *Journal of Finance* **49**, 771–818.

Sandmann, G. and Koopman, S. J. (1998), 'Estimation of stochastic volatility models via Monte Carlo Maximum Likelihood', *Journal of Econometrics* **87**, 271–301.