Statistical Science
2005, Vol. 20, No. 4, 351–357
DOI 10.1214/088342305000000421

Rejoinder: A Selective Overview of Nonparametric Methods in Financial Econometrics

Jianqing Fan

I am very grateful to the Executive Editor, Edward George, for organizing this stimulating discussion. I would like to take this opportunity to thank Professors Peter Phillips, Jun Yu, Michael Sørensen, Per Mykland and Lan Zhang for their insightful and stimulating comments, touching both practical, methodological and theoretical aspects of financial econometrics and their applications in asset pricing, portfolio allocation and risk management. They have made valuable contributions to the understanding of various financial econometric problems.

The last two decades have witnessed an explosion of developments of data-analytic techniques in statistical modeling and analysis of complex systems. At the same time, statistical techniques have been widely employed to confront various complex problems arising from financial and economic activities. While the discipline has grown rapidly over the last two decades and has rich and challenging statistical problems, the number of statisticians involved in studying financial econometric problems is still limited. In comparison with statisticians working on problems in biological sciences and medicine, the group working on financial and econometric problems is dismally small. It is my hope that this article will provide statisticians with quick access to some important and interesting problems in financial econometrics and to catalyze the romance between statistics and finance. A similar effort was made by Cai and Hong [12], where various aspects of nonparametric methods in continuous-time finance are reviewed. It is my intention to connect financial econometric problems as closely to statistical problems as possible so that familiar statistical tools can be employed. With this in mind, I sometimes oversimplify the problems and techniques so that key features can be highlighted.

Jianqing Fan is Professor, Benheim Center of Finance and Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey 08544, USA (e-mail: jqfan@princeton.edu).

I am fully aware that financial econometrics has grown into a vast discipline itself and that it is impossible for me to provide an overview within a reasonable length. Therefore, I greatly appreciate what all discussants have done to expand the scope of discussion and provide additional references. They have also posed open statistical problems for handling nonstationary and/or non-Markovian data with or without market noise. In addition, statistical issues on various versions of capital asset pricing models and their related stochastic discount models [15, 19], the efficient market hypothesis [44] and risk management [17, 45] have barely been discussed. These reflect the vibrant intersection of the interfaces between statistics and finance. I will make some further efforts in outlining econometric problems where statistics plays an important role after brief response to the issues raised by the discussants.

1. BIASES IN STATISTICAL ESTIMATION

The contributions by Professors Phillips, Yu and Sørensen address the bias issues on the estimation of parameters in diffusion processes. Professors Phillips and Yu further translate the bias of diffusion parameter estimation into those of pricing errors of bonds and bond derivatives. Their results are very illuminating and illustrate the importance of estimation bias in financial asset pricing. Their results can be understood as follows. Suppose that the price of a financial asset depends on certain parameters θ (the speed of the reversion κ in their illustrative example). Let us denote it by $p(\theta)$, which can be in one case the price of a bond and in another case the prices of derivatives of a bond. The value of the asset is now estimated by $p(\hat\theta)$ with $\hat\theta$ being estimated from empirical data. When $\hat\theta$ is overestimated (say), which shifts the whole distribution of $\hat\theta$ to the left, the distribution of $p(\hat\theta)$ will also be shifted, depending on the sensitivity of p to θ. The sensitivity is much larger for bond derivatives when κ is close to zero (see Figure 2 of [46]), and hence the pricing errors are much larger. On the other hand, as the distribution of κ is shifted to the left, from Figure 2 of [46], both prices of bonds and their derivatives get smaller and so does the variance of pricing errors. Simulation studies in [46] suggest that these two effects cancel each other out in terms of mean square error.

I agree with Phillips and Yu's observation that discretization is not the main source of biases for many reasonable financial applications. Finite-sample estimation bias can be more severe. This partially answers the question raised by Professor Sørensen. On the other hand, his comments give theoretical insights into the bias due to discretization. For financial applications (such as modeling short-term rates) when the data are collected at monthly frequency, the bias $\{1-\exp(-\kappa\Delta)\}/\Delta-\kappa=-0.0019$ and -0.00042, respectively, for $\kappa=0.21459$ used in Figure 3 of [34] and for $\kappa=0.1$ used in the discussion by Phillips and Yu. For weekly data, using the parameter $\kappa=0.0446$ cited in [14], the discretization bias is merely 9.2×10^{-5}.

For other types of applications, such as climatology, Professor Sørensen is right that the bias due to discretization can sometimes be substantial. It is both theoretically elegant and practically viable to have methods that work well for all situations. The quasi-maximum likelihood methods and their modifications discussed by Professor Sørensen are attractive alternatives. As he pointed out, analytical solutions are rare and computation algorithms are required. This increases the chance of numerical instability in practical implementations. The problem can be attenuated with the estimates based on the Euler approximation as an initial value. The martingale method is a generalization of his quasi-maximum likelihood estimator, which aims at improving efficiency by suitable choice of weighting functions a_j. However, unless the conditional density has multiplicative score functions, the estimation equations will not be efficient. This explains the observation made by Professor Sørensen that the methods based on martingale estimating functions are usually not efficient for low frequency data. The above discussion tends to suggest that when the Euler approximation is reasonable, the resulting estimates tend to have smaller variances.

In addition to the discretization bias and finite sample estimation bias, there is model specification bias. This can be serious in many applications. In the example given by Professors Phillips and Yu, the modeling errors do not have any serious adverse effects on pricing bonds and their derivatives. However, we should be wary of generalizing this statement. Indeed, for the model parameters given in the discussion by Phillips and Yu, the transition density of the CIR model has a noncentral χ^2-distributions with degrees of freedom 80, which is close to the normal transition density given by the Vasicek model. Therefore, the model is not very seriously misspecified.

Nonparametric methods reduce model specification errors by either global modeling such as spline methods or local approximations. This reduces significantly the possibility of specification errors. Since nonparametric methods are somewhat crude and often used as model diagnostic and exploration tools, simple and quick methods serve many practical purposes. For example, in time domain smoothing, the bandwidth h is always an order of magnitude larger than the sampling frequency Δ. Therefore, the approximation errors due to discretization are really negligible. Similarly, for many realistic problems, the function approximation errors can be an order of magnitude larger than discretization errors. Hence, discretization errors are often not a main source of errors in nonparametric inference.

2. HIGH-FREQUENCY DATA

Professors Mykland, Zhang, Phillips and Jun address statistical issues for high-frequency data. I greatly appreciate their insightful comments and their elaborations on the importance and applications of the subject. Thanks to the advances in modern trading technology, the availability of high-frequency data over the last decade has significantly increased. Research in this area has advanced very rapidly lately. I would like to thank Professors Mykland and Zhang for their comprehensive overview on this active research area.

With high-frequency data, discretization errors have significantly been reduced. Nonparametric methods become even more important for this type of large sample problem. The connections between the realized volatility and the time-inhomogeneous model can simply be made as follows. Consider a subfamily of models of (8) in [34],

$$dX_t=\alpha_t\,dt+\sigma_t\,dW_t.$$

For high-frequency data the sampling interval is very small. For the sampling frequency of a minute, $\Delta=1/(252*24*60)\approx2.756\times10^{-6}$. Hence, standardized residuals in Section 2.5 of [34] become $E_t=\Delta^{-1/2}(X_{t+\Delta}-X_t)$ and the local constant estimate of the spot volatility reduces to

$$\hat{\sigma}^2_{j\Delta}=\sum_{i=-\infty}^{j-1}w_{j-i}E^2_{i\Delta},$$

where $\{w_i\}$ are the weights induced by a kernel function satisfying $\sum_{i=1}^{\infty} w_i = 1$. Now, for the weights with a bounded support, the quadratic variation of the process or integrated volatility $\int_t^T \sigma_t^2\, dt$ is naturally estimated by $\Delta \sum_{i=t/\Delta}^{T/\Delta-1} \hat{\sigma}_i^2$, which is simply

$$\sum_{i=t/\Delta}^{T/\Delta-1} \{X_{i\Delta} - X_{(i-1)\Delta}\}^2.$$

This shows that our nonparametric estimation of the integrated volatility for high-frequency data is indeed the same as the realized volatility.

As suggested by Professors Mykland, Zhang, Phillips and Yu, the applications of realized volatilities are not without difficulties. Market microstructure noises emerge at such a fine frequency of observation and market prices can contain multiple jumps due to the flux of information during a trading session. Figure 1 in the discussion by Mykland and Zhang demonstrates convincingly the existence of the market microstructure noise. Aït-Sahalia, Mykland and Zhang [1] and Zhang, Mykland and Aït-Sahalia [50] give comprehensive accounts of this under the assumption that the observed prices are the true ones contaminated with random noise of market microstructure: $Y_t = X_t + \varepsilon_t$. However, they do not take into account that the price processes $\{X_t\}$ may contain jumps in addition to random noises. An effort in this direction has been made recently by Fan and Wang [38] using wavelet techniques.

3. ESTIMATING COVARIANCE MATRICES

Covariance matrices play an important role in risk management and asset allocation. They are featured prominently in many financial econometrics problems. For example, the smallest and largest eigenvalues are related to the minimum and the maximum of the volatility of portfolios and their corresponding eigenvectors are related to portfolio allocation. See [40] for applications of covariance matrices to portfolio selection and [43] for their applications to other scientific problems. There are a couple of approaches to these kinds of problems, depending on the size of the covariance matrices. I hereby give a brief overview and address some of the open challenges.

The simplest estimate of a covariance matrix is probably the sample covariance matrix of the log-returns of p assets over a period of n days prior to the current time t. This is indeed a nonparametric estimation of the covariance matrix localizing in time and has been studied in multivariate analysis when p is finite and the underlying model is correct, that is, the covariance matrix remains the same in the n days prior to time t. See, for example, [26, 27, 47]. However, the impact of the biases in nonparametric methods on the estimation of eigenvalues and eigenvectors has not yet been thoroughly investigated.

The sample covariance matrices can be augmented by using the information from the state domain, which is an extension of the method discussed in Section 3.6 of [34] and allows us to use the historical information. This is particularly useful for estimating the covariance matrices of bonds with different maturities. Useful parametric models such as affine models have been popularly used in interest rate modeling. See, for example, [20, 24, 23]. Nonparametric methods provide useful alternatives to estimating the covariance matrices and to validating parametric models. A naive extension involves high-dimensional smoothing in the state domain. But this can be avoided by localizing only on the yields of a few bonds with intermediate length of maturity.

Another class of techniques is to use a form of GARCH model [28] to estimate covariance matrices. As noted in [30], the number of parameters grows rapidly with the dimensionality p. Various efforts have been made to reduce the complexity of the models. These include constant conditional correlation multivariate GARCH models [10], vectorized multivariate GARCH models [11], dynamic conditional correlation models [29, 31], orthogonal GARCH models [2], generalized orthogonal GARCH models [48] and conditionally uncorrelated component models [37]. For a survey, see [8].

In portfolio allocation and risk management, the number of stocks p can be well in the order of hundreds, which is typically in the same order as the sample size n. The sample covariance matrix may not be a good estimator of the population one. The estimated variance of a portfolio based on the sample covariance may far exceed the true one. The estimation errors can accumulate quickly when p grows with n. Indeed, Johnstone [43] shows that the largest eigenvalue of the covariance matrix is far larger than the population one. There are many studies on the behavior of random matrices when the dimensionality p grows with n. See, for example, [5, 22, 21, 49]. For a survey, see [4].

Estimating covariance matrices for large p is intrinsically challenging. For example, when $p = 200$, there are more than 20,000 free parameters. Yet, the available sample size is usually in the order of hundreds or a few thousand. Longer time series (larger n) will increase modeling biases. Without imposing structures on the covariance matrices, they are hard to estimate. Thanks to the multi-factor models (see Chapter 6 of [13]), if a few factors can capture completely the cross-sectional risks, the number of parameters can be significantly reduced. For example, using the Fama–French three-factor models [32, 33], there are $4p$ instead of $p(p+1)/2$ parameters. Natural questions arise with this structured estimate of the covariance matrix, how large p can be such that the estimation error in the covariance matrix is negligible in asset allocation and risk management. The problems of this kind are interesting and remain open.

Another possible approach to the estimation of covariance matrices is to use a model selection approach. First of all, according to Chapter 3 of [39], the Cholesky decomposition admits nice autoregressive interpretation. We may reasonably assume that the elements in the Cholesky decomposition of the covariance matrix are sparse. Hence, the penalized likelihood method [3, 35, 42] can be employed to select and estimate nonsparse elements. The sampling property of such a method remains unknown. Its impact on portfolio allocation and risk management needs to be studied.

4. STATISTICS IN DERIVATIVE PRICING

Over last three decades, option pricing has witnessed an explosion of new models that extend the original work of Black and Scholes [9]. Empirically pricing financial derivatives is innately related to statistical regression problems. This is well documented in papers such as [6, 7, 15, 16, 25, 41]. See also a brief review given by Cai and Hong [12]. For a given stochastic model with given structural parameters under the risk-neutral measure, the prices of European options can be determined, which are simply the discounted expected payoffs under the risk-neutral measure. Bakshi, Cao and Chen [6] give the analytic formulas of option prices for five commonly used stochastic models, including the stochastic-volatility random-jump model. They then estimate the risk-neutral parameters by minimizing the discrepancies between the observed prices and the theoretical ones. With estimated risk-neutral parameters, option prices with different characteristics can be evaluated. They conduct a comprehensive study of the relative merits of competing option pricing models by computing pricing errors for new options. Dumas, Fleming and Whaley [25] model the implied volatility function by a quadratic function of the strike price and time to maturity and determine these parameters by minimizing pricing errors. Based on the analytic formula of Bakshi, Cao and Chen [6] for option price under the stochastic volatility models, Chernov and Ghysels [16] estimate the risk neutral parameters by integrating information from both historical data and risk-neutral data implied by observed option prices. Instead of using continuous-time diffusion models, Heston and Nandi [41] assume that the stock prices under the risk-neutral world follow a GARCH model and derive a closed form for European options. They determine the structural parameters by minimizing the discrepancy between the empirical and theoretical option prices. Barone-Adesi, Engle and Mancini [7] estimate risk-neutral parameters by integrating the information from both historical data and option prices. Christoffersen and Jakobs [18] expand the flexility of the model by introducing long- and short-run volatility components.

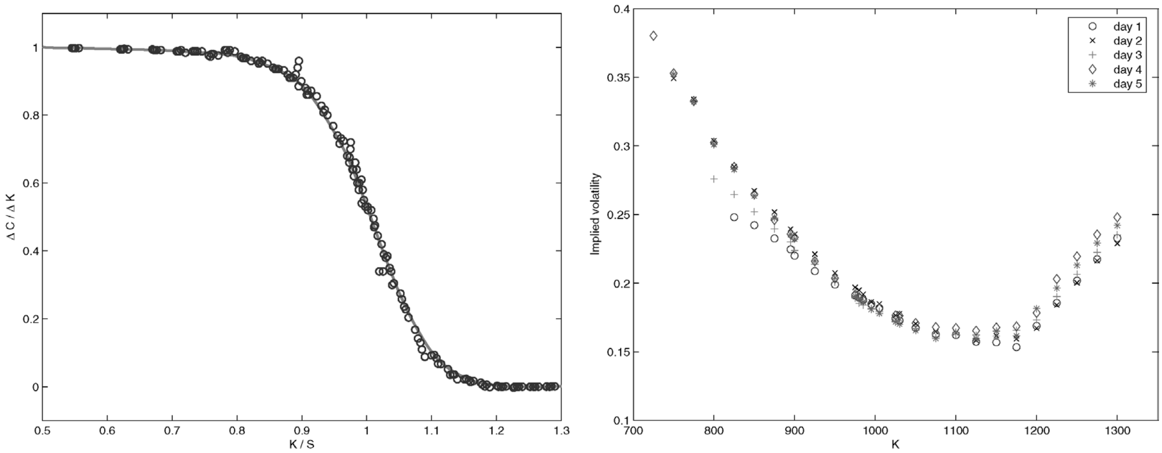
The above approaches can be summarized as follows. Using the notation in Section 4.1 of [34], the theoretical option price with option characteristics $(S_i, K_i, T_i, r_i, \delta_i)$ is governed by a parametric form $C(S_i, K_i, T_i, r_i, \delta_i, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of structural parameters of the stock price dynamics under the risk-neutral measure. The form depends on the underlying parameters of the stochastic model. This can be in one case a stochastic volatility model and in another case a GARCH model. The parameters are then determined by minimizing

$$\sum_{i=1}^{n} \{C_i - C(S_i, K_i, T_i, r_i, \delta_i, \boldsymbol{\theta})\}^2$$

or similar discrepancy measures. The success of a method depends critically on the correctness of model assumptions under the risk-neutral measure. Since these assumptions are not on the physical measure, they are hard to verify. This is why so many parametric models have been introduced. Their efforts can be regarded as searching an appropriate parametric form $C(\cdot; \boldsymbol{\theta})$ to better fit the option data. Nonparametric methods in Section 4.1 provide a viable alternative for this purpose. They can be combined with parametric approaches to improve the accuracy of pricing.

As an illustration, let us consider the options with fixed $(S_i, T_i, r_i, \delta_i)$ so that their prices are only a function of K or equivalently a function of the moneyness $m = K/S$,

$$C = \exp(-rT) \int_K^\infty (x - K) f^*(x)\, dx.$$



(a) (b)

FIG. 1. (a) *Scatterplot of the response variable computed based on option prices with consecutive strike price against the moneyness.* (b) *The implied volatilities of the options during the period July 7–11, 2003.*

Denoting $D = \exp(rT)C/S$ and letting $\bar{F}^*(x) = 1 - F^*(x) = \int_x^\infty f^*(y)\,dy$ be the survival function, then by integration by parts,

$$D = -S^{-1}\int_K^\infty (x-K)\,d\bar{F}^*(x) = S^{-1}\int_K^\infty \bar{F}^*(x)\,dx.$$

By a change of variable, we have

$$D = \int_m^\infty \bar{F}(u)\,du,$$

where $F(u) = F^*(Su)$ is the state price distribution in the normalized scale (the stock price is normalized to \$1). Let us write explicitly $D(m)$ to stress the dependence of discounted option price on the moneyness m. Then

$$\begin{aligned}\frac{D(m_1) - D(m_2)}{m_2 - m_1} &= (m_2 - m_1)^{-1}\int_{m_1}^{m_2} \bar{F}(u)\,du \\ &= \bar{F}\left(\frac{m_2 + m_1}{2}\right) + O\big((m_2 - m_1)^2\big).\end{aligned}$$

Assume that the moneyness $m_i = K_i/S_t$ has already been ordered for N_t options with strike prices $\{K_i, i = 1, \ldots, N_t\}$ traded at time t. Let $x_i = (m_i + m_{i+1})/2 = (K_i + K_{i+1})/(2S)$ and y_i be the observed value of $\frac{D(m_i) - D(m_{i+1})}{m_{i+1} - m_i}$, namely,

$$y_i = \exp(r_t T_t)\{C_i - C_{i+1}\}/\{K_{i+1} - K_i\}, \qquad i = 1, \ldots, N_t - 1,$$

where r_t, T_t and S_t are, respectively, the risk-free interest rate, time to maturity and spot stock price at time t, and C_{i+1} and C_i are the option prices at time t associated with strike prices K_{i+1} and K_i. Then, estimating the state price distribution becomes a familiar nonparametric regression problem,

$$y_i \approx \bar{F}(x_i) + \varepsilon_i.$$

In the above equation, the dependence on t is suppressed. Figure 1(a) shows the scatterplot of the pairs (x_i, y_i) based on the closing call option prices (average of bid-ask prices) of the Standard and Poor's 500 index with maturity of $T_t = 75 - t$ days on the week of July 7 to July 11, 2003 ($t = 0, \ldots, 4$). The implied volatility curve is given in Figure 1(b). It is not a constant and provides stark evidence against the Black–Scholes formula.

The waterfall shape of the regression curve is very clear. The naive applications of nonparametric techniques will incur large approximation biases resulting in systematic pricing errors. One possible improvement is to use a parametric method such as the ad-hoc Black–Scholes model of Dumas, Fleming and Whaley [25] to estimate the main shape of the regression function and then use a nonparametric method to estimate the difference. This kind of idea has been investigated by Fan and Mancini [36]. When we aggregate the data in the week of July 7 to July 11, 2003, the times to maturity T_t vary slightly. Semiparametric techniques can be used to adjust for this effect. Similarly to many practical problems, we always have side information available that can be incorporated into modeling and analysis of the data. This reinforces the claim that pricing financial derivatives is fundamentally a statistical problem where statisticians can play an important role.

ACKNOWLEDGMENTS

This research was supported in part by NSF Grant DMS-03-55179. The author acknowledges gratefully various discussions with Dr. Loriano Mancini and would like to thank him for providing Figure 1 presented in this discussion.

REFERENCES

[1] AÏT-SAHALIA, Y., MYKLAND, P. and ZHANG, L. (2005). How often to sample a continuous-time process in the presence of maket microstructure noise. *Review of Financial Studies* **18** 351–416.

[2] ALEXANDER, C. (2001). Orthogonal GARCH. In *Mastering Risk* **2** (C. Alexander, ed.) 21–38. Financial Times Prentice Hall, London.

[3] ANTONIADIS, A. and FAN, J. (2001). Regularization of wavelet approximations (with discussion). *J. Amer. Statist. Assoc.* **96** 939–967.

[4] BAI, Z. D. (1999). Methodologies in spectral analysis of large-dimensional random matrices. A review (with discussion). *Statist. Sinica* **9** 611–677.

[5] BAI, Z. D. and SILVERSTEIN, J. W. (1999). Exact separation of eigenvalues of large-dimensional sample covariance matrices. *Ann. Probab.* **27** 1536–1555.

[6] BAKSHI, G., CAO, C. and CHEN, Z. (1997). Empirical performance of alternative option pricing models. *J. Finance* **52** 2003–2049.

[7] BARONE-ADESI, G., ENGLE, R. F. and MANCINI, L. (2004). GARCH options in incomplete markets. NCCR-FinRisk Working Paper 155, Univ. Zurich.

[8] BAUWENS, L., LAURENT, S. and ROMBOUTS, J. V. K. (2003). Multivariate GARCH models: A survey. CORE Discussion Paper 2003/31.

[9] BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *J. Political Economy* **81** 637–654.

[10] BOLLERSLEV, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics* **72** 498–505.

[11] BOLLERSLEV, T., ENGLE, R. and WOOLDRIDGE, J. (1988). A capital asset pricing model with time-varying covariances. *J. Political Economy* **96** 116–131.

[12] CAI, Z. and HONG, Y. (2003). Nonparametric methods in continuous-time finance: A selective review. In *Recent Advances and Trends in Nonparametric Statistics* (M. G. Akritas and D. N. Politis, eds.) 283–302. North-Holland, Amsterdam.

[13] CAMPBELL, J. Y., LO, A. W. and MACKINLAY, A. C. (1997). *The Econometrics of Financial Markets*. Princeton Univ. Press.

[14] CHAPMAN, D. A. and PEARSON, N. D. (2000). Is the short rate drift actually nonlinear? *J. Finance* **55** 355–388.

[15] CHEN, X. and LUDVIGSON, S. (2005). Land of Addicts? An empirical investigation of habit-based asset pricing model. Unpublished manuscript.

[16] CHERNOV, M. and GHYSELS, E. (2000). A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *J. Financial Economics* **56** 407–458.

[17] CHRISTOFFERSEN, P. F. (2003). *Elements of Financial Risk Management*. Academic Press, San Diego.

[18] CHRISTOFFERSEN, P. F. and JAKOBS, K. (2005). Option valuation with long-run and short-run volatility components. Unpublished manuscript.

[19] COCHRANE, J. H. (2001). *Asset Pricing*. Princeton Univ. Press.

[20] DAI, Q. and SINGLETON, K. (2000). Specification analysis of affine term structure models. *J. Finance* **55** 1943–1978.

[21] DETTE, H. (2002). Strong approximation of eigenvalues of large dimensional Wishart matrices by roots of generalized Laguerre polynomials. *J. Approx. Theory* **118** 290–304.

[22] DIACONIS, P. and SHAHSHAHANI, M. (1994). On the eigenvalues of random matrices. *J. Appl. Probab.* **31A** 49–61.

[23] DUFFEE, G. R. (2002). Term premia and interest rate forecasts in affine models. *J. Finance* **57** 405–443.

[24] DUFFIE, D. and KAN, R. (1996). A yield-factor model of interest rates. *Math. Finance* **6** 379–406.

[25] DUMAS, B., FLEMING, J. and WHALEY, R. E. (1998). Implied volatility functions: Empirical tests. *J. Finance* **53** 2059–2106.

[26] EATON, M. L. and TYLER, D. E. (1991). On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix. *Ann. Statist.* **19** 260–271.

[27] EATON, M. L. and TYLER, D. (1994). The asymptotic distribution of singular values with applications to canonical correlations and correspondence analysis. *J. Multivariate Anal.* **50** 238–264.

[28] ENGLE, R. F., ed. (1995). *ARCH, Selected Readings*. Oxford Univ. Press.

[29] ENGLE, R. F. (2002). Dynamic conditional correlation—a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *J. Bus. Econom. Statist.* **20** 339–350.

[30] ENGLE, R. F. and KRONER, K. F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory* **11** 122–150.

[31] ENGLE, R. F. and SHEPPARD, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. Preprint.

[32] FAMA, E. and FRENCH, K. (1992). The cross-section of expected stock returns. *J. Finance* **47** 427–465.

[33] FAMA, E. and FRENCH, K. (1993). Common risk factors in the returns on stocks and bonds. *J. Financial Economics* **33** 3–56.

[34] FAN, J. (2005). A selective overview of nonparametric methods in financial econometrics. *Statist. Sci.* **20** 317–337.

[35] FAN, J. and LI, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360.

[36] FAN, J. and MANCINI, L. (2005). Derivative pricing using semiparametric techniques. Unpublished manuscript.

[37] FAN, J., WANG, M. and YAO, Q. (2005). Modelling multivariate volatilities via conditionally uncorrelated components. Unpublished manuscript.

[38] FAN, J. and WANG, Y. (2005). Multiscale jump and volatility analysis for high-frequency financial data. Unpublished manuscript.

[39] FAN, J. and YAO, Q. (2003). *Nonlinear Time Series*: *Nonparametric and Parametric Methods*. Springer, New York.

[40] GOLDFARB, D. and IYENGAR, G. (2003). Robust portfolio selection problems. *Math. Oper. Res.* **28** 1–37.

[41] HESTON, S. L. and NANDI, S. (2000). A closed-form GARCH option valuation model. *Review of Financial Studies* **13** 585–625.

[42] HUANG, J. Z., LIU, N. and POURAHMADI, M. (2005). Covariance selection and estimation via penalized normal likelihood. Unpublished manuscript.

[43] JOHNSTONE, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *Ann. Statist.* **29** 295–327.

[44] LO, A. (2000). Finance: A selective survey. *J. Amer. Statist. Assoc.* **95** 629–635.

[45] MOFFATT, H. K. (2002). *Risk Management*: *Value at Risk and Beyond*. Cambridge Univ. Press.

[46] PHILLIPS, P. C. B. and YU, J. (2005). Jackknifing bond option prices. *Review of Financial Studies* **18** 707–742.

[47] TYLER, D. E. (1981). Asymptotic inference for eigenvectors. *Ann. Statist.* **9** 725–736.

[48] VAN DER WEIDE, R. (2002). GO-GARCH: A multivariate generalized orthogonal GARCH model. *J. Appl. Econometrics* **17** 549–564.

[49] WIEAND, K. (2002). Eigenvalue distributions of random unitary matrices. *Probab. Theory Related Fields* **123** 202–224.

[50] ZHANG, L., MYKLAND, P. and AÏT-SAHALIA, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100** 1394–1411.