

# Density Forecasting: A Survey<sup>1</sup>

Anthony S. Tay\* and Kenneth F. Wallis<sup>†</sup>

\*Department of Economics  
National University of Singapore  
10 Kent Ridge Crescent  
Singapore 119260

<sup>†</sup> Department of Economics  
University of Warwick  
Coventry CV4 7AL  
England

## Abstract:

A density forecast of the realization of a random variable at some future time is an estimate of the probability distribution of the possible future values of that variable. This article presents a selective survey of applications of density forecasting in macroeconomics and finance, and discusses some issues concerning the production, presentation and evaluation of density forecasts.

---

<sup>1</sup> The helpful comments and suggestions of Frank Diebold, Stewart Hodges and two anonymous referees are gratefully acknowledged. Responsibility for errors remains with the authors.

## 1. INTRODUCTION

A density forecast of the realization of a random variable at some future time is an estimate of the probability distribution of the possible future values of that variable. It thus provides a complete description of the uncertainty associated with a prediction, and stands in contrast to a point forecast, which by itself contains no description of the associated uncertainty. Intermediate between these two extremes is a prediction interval, which specifies the probability that the actual outcome will fall within a stated interval. To report a prediction interval represents a first response by point forecasters to criticism of their silence on the subject of uncertainty, and this practice is becoming more common in macroeconomic forecasting. This field also contains several examples of density forecasts, and these feature as a more direct object of attention in the fields of finance and risk management. This article presents a selective survey of the use of density forecasts, and discusses some issues concerning their construction, presentation and evaluation.

A density forecast is of course implicit in the standard construction and interpretation of a symmetric prediction interval. This is usually given as a band of plus or minus one or two standard errors around the point forecast, and the associated statement of a probability of 68 or 95 per cent respectively rests on an assumption of a normal or Gaussian distribution. Sometimes the use of Student's *t*-distribution is suggested, having in mind that the variance of the forecast error must be estimated, or the departure from normality caused by the more general parameter estimation problem may be acknowledged, nevertheless such calculations typically rest on a model with normally distributed innovations. These methods seem to assume that the normal distribution is "too good not to be true", in the classic phrase of Anscombe (1967). This was a position of which he was critical, however, instead advocating the use of non-normal

distributions: the complete sentence reads “the disposition of present-day statistical theorists to suppose that all “error” distributions are exactly normal can be ascribed to their ontological perception that normality is too good not to be true” (Anscombe, 1967, p.16). His advice has been taken to heart in much of the density forecasting literature, where a wide range of functional forms and nonparametric empirical methods appears, as discussed below.

A density forecasting problem can be placed in a decision theoretic environment by a direct generalization of the point forecasting problem, as described by Diebold, Gunther and Tay (1998). Suppose that the forecast user has a loss function  $L(a(p(y)), y)$  which depends on the action  $a(\cdot)$  chosen in the light of the density forecast  $p(y)$  and on the realization of the forecast variable. Choosing an action  $a^*$  to minimize expected loss calculated as if the density forecast  $p(y)$  is correct implies that

$$a^*(p(y)) = \underset{a \in A}{\operatorname{argmin}} \int L(a, y) p(y) dy$$

where  $A$  denotes all possible choices that the forecast user might make. The choice  $a^*$  incurs loss  $L(a^*, y)$ , which is a random variable whose expected value with respect to the true density  $f(y)$  is

$$E[L(a^*(y))] = \int L(a^*(y)) f(y) dy,$$

and a “better” density forecast is one with lower expected loss. However there is even less literature on explicit decision problems and their associated loss functions for density forecasts than for point forecasts, where it is in any event very scarce. Here the standard treatment takes the action to be the choice of a point forecast  $\hat{y}$ , with the loss function a simple non-negative function of the error  $(y \& \hat{y})$ , taking the value zero for zero error. Different loss functions may lead to different optimal point forecasts, particularly if  $f(y)$  is not symmetric. A corresponding

body of results is only just beginning to emerge in the density forecasting case. The problem of evaluating density forecasts *ex post* is considered below, with some discussion of the problem of evaluation under general loss functions, following discussion of a range of applications in macroeconomics and finance.

The general setting of this article is the time series forecasting problem, rather than other prediction problems discussed by Aitchison and Dunsmore (1975), for example, under the heading of statistical prediction analysis. In both cases an essential feature of the characterization is the pair of experiments  $e$  and  $f$ . “From the information which we gain from a performance of  $e$ , the *informative* experiment, we wish to make some reasoned statement concerning the performance of  $f$ , the *future* experiment” (1975, p.1). The experiments are linked by the indexing parameter of the probability models for  $e$  and  $f$ , and a feature of all the problems Aitchison and Dunsmore consider is that, for given index, the experiments  $e$  and  $f$  are independent. This excludes the time series forecasting problem, in which  $e$  records a realization to date of some stochastic process of which  $f$  is a continuation. Our focus on the time series forecasting problem in turn tends to exclude a range of prediction problems where Bayesian techniques have found application.

The article proceeds as follows. Section 2 discusses the use of density forecasts in macroeconomics and finance. Section 3 considers some issues concerning the interpretation and presentation of density forecasts, and the available techniques for evaluating density forecasts are reviewed in Section 4. Section 5 contains concluding comments.

## **2. DENSITY FORECASTS IN MACROECONOMICS AND FINANCE**

### **2.1 Applications in macroeconomics**

The longest-running series of density forecasts in macroeconomics dates back to 1968, when the Business and Economic Statistics Section of the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER) jointly initiated a quarterly survey of macroeconomic forecasters in the United States, known as the ASA-NBER survey; Zarnowitz (1969) describes its original objectives. Later the Federal Reserve Bank of Philadelphia assumed responsibility for the survey, and the name was changed to the Survey of Professional Forecasters (see Croushore, 1993). Most of the questions ask for point forecasts, for a range of variables and forecast horizons, but density forecasts are also requested for inflation and output growth. In each case a number of intervals, or bins, in which the future value of the variable might fall are provided, and each forecaster is asked to report their associated forecast probabilities. The density forecast is thus represented as a histogram, on a preassigned grid. These are averaged over respondents, and the mean probability distributions are published. A recent example of a density forecast of inflation is shown in Table 1.

[ insert Table 1 here ]

There are numerous interesting issues associated with the compilation and reporting of such forecasts, including the combination of individual responses by simple averaging and the potential usefulness of measures such as the standard errors of the mean probabilities. These are familiar topics in point forecasting, whose extension to density forecasting has scarcely begun to be considered.

In the United Kingdom the history is much shorter. When the Treasury established its Panel of Independent Forecasters (the “seven wise men”) in late 1992, one of the present authors suggested that the panel members be asked to report density forecasts for inflation and growth,

using the same questions as the US Survey of Professional Forecasters, in addition to their point forecasts. It was some time before this suggestion was adopted, and only one set of density forecasts was published before the panel, by now the “six wise people”, was dissolved on the change of government in May 1997. In February 1996 the Bank of England had begun to publish average survey responses of density forecasts of inflation in its quarterly *Inflation Report*, along the lines of the US survey, but in contrast the panel published the individual responses; in any case this was not difficult, since there were only six. These are shown, for one-year-ahead inflation (in the Retail Prices Index excluding mortgage interest payments, RPIX), in Table 2.

[ insert Table 2 here ]

It is seen that the first five density forecasts are much less dispersed than the sixth, and round numbers catch the eye, with modal probabilities 55, 50, 40, 40, 35 per cent, coincidentally declining with the alphabet. Such rounding is commonly observed in reported probabilities with a strong subjective element. In contrast, the probabilities in the sixth column of Table 2 are known to be calculated assuming a normal distribution. The forecast is based on that published in the *National Institute Economic Review* (Martin Weale being the Institute’s Director), which since February 1996 has included a density forecast of inflation and growth alongside its long-established macroeconomic point forecasts. The forecasts are based on a macroeconomic model, subject to the forecaster’s residual adjustments, as usual, and analysis of past forecast errors cannot reject the hypothesis that they are normally distributed (Poulizac *et al.*, 1996). This analysis also cannot reject the hypothesis that past forecasts are unbiased, so the normal density forecast is centred on the point forecast, with variance equal to that of the errors in past forecasts at the same horizon (over 1983-95). Historical forecast errors are an appropriate starting point for this calculation since they incorporate a range of possible sources of error including model error, but projecting the variance forward is itself a forecasting problem and, as with point

forecasts, failures may occur due to structural breaks. In this case an obvious change is the introduction of a monetary policy regime of inflation targeting, so the past may not be a good guide to future behaviour, as anticipated by Poulizac *et al.* In these circumstances the first five forecasters are clearly much less uncertain about future inflation, which in the event stood at 2.7 per cent at the end of 1997.

An alternative way of calibrating a model-based density forecast is through stochastic simulation of the model in the new policy environment, as discussed in a companion article on the National Institute model by Blake (1996), although this does not include the effects of model uncertainty. Stochastic simulation methods are required to calculate empirical distributions of estimated outcomes from large-scale macroeconomic models, because these are typically nonlinear in variables and analytic methods are not available. The underlying pseudo-random error terms are usually generated assuming a normal distribution, a test of the normality of residuals being a standard diagnostic test during the specification of the econometric equations. Often these equations are log-linear, or include more complex forms such as the constant elasticity of substitution (CES) function, whereas definitions and accounting identities are linear, involving products or ratios if variables appear in both real and nominal terms. The result is that the predictive densities are non-normal but of unknown form, and they are simply reported graphically or numerically. Even with linear models stochastic simulation methods may be advantageous whenever it is desired to take account of uncertainty due not only to the model's random error terms but also to coefficient estimation error, errors in projecting exogenous variables, and so forth. For a vector autoregressive model of interest rates, money, prices and output Kling and Bessler (1989) present density forecasts calculated by stochastic simulation with respect to model errors and sampling errors in regression coefficients, also taking account of nonconstant residual variances. We return to this example in Section 4 below.

A density forecast of inflation represented analytically by a specific non-normal probability distribution has been published by the Bank of England in its quarterly *Inflation Report* since February 1996. The chosen distribution is the two-piece normal distribution, whose probability density function, with parameters  $\mu$ ,  $s_1$  and  $s_2$ , is

$$f(y) = \begin{cases} A \exp[-(y - \mu)^2/2s_1^2] & y \leq \mu \\ A \exp[-(y - \mu)^2/2s_2^2] & y > \mu \end{cases}$$

where  $A = (\sqrt{2p}(s_1 + s_2)/2)^{-1}$  (John, 1982; Johnson, Kotz and Balakrishnan, 1994). It is formed by taking two halves of normal distributions with parameters  $(\mu, s_1)$  and  $(\mu, s_2)$  respectively and scaling them to give the common value  $f(\mu)$  as above. The illustration in Figure 1 uses the parameter values corresponding to the eight-quarter-ahead forecast published in the August 1997 *Inflation Report*, also used for illustrative purposes by Britton *et al.* (1998) and Wallis (1999). With  $s_2 > s_1$  this has positive skewness, which can be seen algebraically by noting that the mean is

$$E(Y) = \mu + \sqrt{\frac{2}{p}}(s_2 - s_1)$$

and so exceeds the mode,  $\mu$ , if  $s_2 > s_1$ , or that the third moment about the mean is a positive multiple of  $(s_2 - s_1)$ . The coefficient of kurtosis exceeds 3 whenever  $s_1 \neq s_2$ , so the distribution is leptokurtic. It is a convenient way of representing departures from the symmetry of the normal distribution since probability calculations can still be based on standard normal tables, with suitable scaling.



[ Insert Figure 1 here]

The density forecast describes the subjective assessment of inflationary pressures by the Bank's Monetary Policy Committee, and although the prevailing level of uncertainty is initially assessed with reference to past forecast errors, the final calibration of the distribution represents the Committee's judgement, the degree of skewness in particular exhibiting their view of the balance of risks on the "upside" and "downside" of the forecast (Britton *et al.*, 1998). This approach to the construction of a density forecast is closer to Bayesian elicitation methods, although the extent to which the Committee's procedures reflect the common approach in the general field of probability assessment of basing elicitation procedures on quantiles is not known. It nevertheless stands in contrast to the National Institute's approach of superimposing a data-based normal distribution on a point forecast. It should be noted, however, that the conventional null hypothesis of normality would be unlikely to be rejected in this approach if the degree of asymmetry in the example in Figure 1 were to apply in practice. For these parameter values, one would expect to require a sample of almost 200 observations before the conventional asymptotic  $\chi^2(2)$  statistic would reject normality at the 5 per cent level; even the more powerful likelihood ratio test given by Kimber (1985) for the null hypothesis that  $s_1 = s_2$  in the two-piece normal distribution would be expected to require a sample of 115 to reject in this case. Samples of this size are rare in macroeconomic forecasting, and in any event there is no presumption in the Bank of England's forecasts that the balance of risks is constant over time.

In addition to constructing macroeconomic density forecasts to assist in setting their policy instruments, monetary policy makers also use financial market forecasts, to which we now turn. In particular density forecasts of future interest rates and exchange rates can be extracted from sets of prices of traded options and futures contracts as discussed below. These then assist the interpretation of market sentiment and provide a check on the credibility of policy.

## 2.2 Applications in finance

In finance, much effort has been put into generating forecasts with a complete characterization of the uncertainty associated with asset and portfolio returns, recognizing that the normal distribution is inadequate for this purpose. Tests of normality typically rely on third and fourth moments, rejecting the null hypothesis of normality if there is significant skewness and/or excess kurtosis, and many empirical studies have found non-normal higher moments in the (unconditional) distributions of stock returns, interest rates, and other financial data series. An early example is Fama (1965), who reports evidence of excess kurtosis in the unconditional distribution of daily returns of stocks listed in the Dow Jones Industrial Average. Many studies have since attempted to model this excess kurtosis, and also asymmetries in the distribution of asset returns, although the early studies tended not to deal with forecasts *per se*. The models proposed considered returns to be iid processes and should be viewed as attempts to explain the unconditional distribution of asset returns rather than to develop forecasting models, nevertheless they do reflect an increasing awareness that such asymmetries should carry over to forecasts of asset returns.

There are several motivations for an interest in more complete and accurate probability statements. A leading example is the issue of risk management, with which the concept of density forecasting is intimately related. The business of risk management has in recent years developed into an important industry, with density forecasts regularly issued by J.P.Morgan, Reuters and Bloomberg, among others. The basic idea of such systems is to allow the user to generate density forecasts of the change in the value of customized portfolios over a particular holding period. The focus of these forecast densities is typically the  $n$ th percentile of the distribution, embodied in a measure commonly known as Value-at-Risk (VaR); the portfolio is forecasted to lose a value greater or equal to its VaR over the specified period with probability

n/100. It is clear that departures from normality in the portfolio returns will adversely affect the usefulness of VaR estimates if the assumption of normality is inappropriately used in generating the forecast. In particular, the presence of excess kurtosis implies that the Gaussian-based VaR of a portfolio will be underestimated, and much effort has gone into developing accurate density forecasts that can overcome this difficulty in the context of VaR analysis. Such measurements are also of concern to regulators, the Capital Adequacy Directive of the Bank for International Settlements, for example, requiring a bank to hold risk capital adequate to cover losses on its trading portfolio over a ten-day holding period on 99 per cent of occasions. This example indicates that it may be only a single quantile of the distribution that is of interest, moreover one located in its extreme tail, and some methods, discussed below, focus directly on these aspects of the problem. The presence of non-normal higher moments in financial data also has implications for numerous other issues in addition to risk management, including asset pricing, portfolio diversification and options pricing. For example, Simkowitz and Beedles (1978) demonstrate that the degree of skewness preference among investors will affect the extent to which they diversify. The higher the degree of skewness, the fewer assets the investor would hold since diversification reduces the amount of skewness in the portfolio. Cotner (1991) provides some evidence that asymmetries affect options prices. The presence of skew in asset returns will also play an important role in developing hedging strategies.

Density forecasting in finance can be viewed as beginning with the literature that aims to model and forecast volatility. The ARCH model of Engle (1982) models the conditional variance as a linear function of square of past observations and conditional variances, and thus delivers forecasts with time varying conditional variances. A Generalized ARCH(1,1) process (with zero mean) can be written as

$$y_t = e_t h_t^{1/2}$$

$$h_t = \omega + \alpha e_{t-1}^2 + \beta h_{t-1}$$

where the standardized residual  $e_t$  is distributed identically and independently with some zero-mean unit-variance density  $f(e_t)$ . ARCH models imply larger kurtosis in the unconditional distribution than the normal distribution, although the commonly used conditionally Gaussian varieties deliver symmetric density forecasts, which is unsatisfactory given evidence of asymmetries in asset returns. The excess kurtosis generated by conditionally Gaussian models was also found to be insufficient to explain the degree of leptokurtosis in many financial time series. The literature on ARCH models and its applications is enormous, and we do not attempt to review that literature here, nor the related literature on stochastic volatility models. For a good review of ARCH models, see Bollerslev, Engle and Nelson (1994) and, in a more general context, Pagan (1996). Our focus on density forecasting directs attention towards higher-order moments.

GARCH models can produce skewed and leptokurtic conditional forecasts - many do - and they do so by incorporating skew and kurtosis directly into the distribution of the standardized residuals of GARCH processes. A leading example is the t-distributed GARCH model of Bollerslev (1987) which directly incorporates excess kurtosis in the conditional forecasts by specifying a Student's t-distribution for the standardized residuals. Some applications have attempted to model the densities of the standardized residuals by using asymptotic expansions, instead of assuming fully parameterized functional forms. Lee and Tse (1991) model the standardized residuals as a Gram-Charlier type distribution; an application to the 1- and 2-month interbank rates of the Singapore Asian Dollar Market reveals evidence of excess kurtosis in the residuals of an ARCH-M regression, but not of skew. Baillie and Bollerslev (1992) use asymptotic expansions to obtain multi-step-ahead predictions from a GARCH process, which are non-normal even if the one-step-ahead conditional prediction

densities are normal; they approximate the quantiles of these forecast densities using Cornish-Fisher asymptotic expansions. Another approach to producing density forecasts, along similar lines, uses nonparametric procedures to model the density of the standardized residual. An example is the semiparametric ARCH approach of Engle and Gonzalez-Rivera (1991) which specifies the variable of interest as a GARCH process and exploits the fact that quasi-maximum likelihood estimation can deliver consistent estimates of the GARCH parameters. The standardized residuals from this step are then estimated nonparametrically using the discrete maximum penalized likelihood estimation technique. An application to daily stock returns for the sample period 1962 to 1988 showed that skewness is an important feature of stock returns, while an analysis of daily returns to the £/US\$ exchange rate produced evidence of less pronounced skew. Other nonparametric approaches, including bootstrap methods, can be combined with parametric GARCH models in a similar manner.

The idea of accommodating higher moment effects by developing sophisticated distributional specifications, parametric or otherwise, for the standardized residuals of a GARCH model nevertheless does not allow the higher moments to be time-varying in general. For instance, the forecasts delivered by a conditional normal GARCH model will always deliver a conditional density forecast with kurtosis equal to 3. Asymmetric GARCH models, such as the Exponential GARCH model of Nelson (1991), allow negative and positive shocks to affect the conditional variance asymmetrically, but still deliver density forecasts that are symmetric if the distribution of the standardized residuals is so specified. The possibility of forecasting higher moments is something worth exploring, and while there are some studies that pertain to the forecastability of higher moments, the evidence is mixed. Singleton and Wingender (1986) find considerable skewness in the distribution of the monthly returns of common stocks from 1961 to 1980, but also find that this skewness did not persist, that is, stocks that displayed skewness

in one period did not always remain skewed in future periods, and using the Spearman rank order correlation test they provide evidence that the skew in the distribution of stock returns is not predictable. However, they base their study on the standardized third central moment. This measure is known to be very sensitive to outliers, so the results they obtain may be due to a few anomalous observations. Alles and Kling (1994) consider the distributions of daily returns on the NYSE (Jul 1962 to Dec 1989), AMEX (Jul 1962 to Dec 1989) and NASDAQ (Dec 1972 to Dec 1989) market indices. Across four-year-long sample periods, they find that the skew varies in direction and size from period to period. No attempt is made to relate current skew to the returns' observed history, although there is evidence that the skew in the distribution of the returns varies according to the business cycle. Their analysis is also based on the standardized third central moment. An analysis of bond returns (market value-weighted indices of long and medium term treasury bonds) shows that the degree of skew in these series is generally less than for stock indices.

Rare examples of the modelling of the dependence of higher-order moments on the past are the studies by Gallant, Hsieh and Tauchen (1991) and Hansen (1994), again from contrasting points of view. Gallant, Hsieh and Tauchen apply their semi-nonparametric approach to density estimation, based on a series expansion about the Gaussian density which in effect makes the density a polynomial in the past history of returns; the application deals with daily pound/dollar exchange rates. Hansen notes that the method needs very large data sets to achieve good precision, and prefers a model that is fully - yet in contrast parsimoniously - parameterized. His ARCD (autoregressive conditional density) model uses a novel "skewed Student's  $t$ " conditional distribution with an additional skewness parameter. The density function of the standardized residuals (zero mean, unit variance) is

$$f(e_t | \lambda, \eta) = \begin{cases} bc \left( 1 - \frac{1}{\lambda + 2} \left( \frac{be_t - a}{1 + \lambda} \right)^2 \right)^{(\lambda + 1)/2} & e_t < a/b \\ bc \left( 1 - \frac{1}{\lambda + 2} \left( \frac{be_t - a}{1 + \lambda} \right)^2 \right)^{(\lambda + 1)/2} & e_t \geq a/b \end{cases}$$

where  $2 < \lambda < 4$ ,  $0 < \eta < 1$  and

$$a = \frac{1}{\lambda} c \left( \frac{\lambda + 2}{\lambda + 1} \right),$$

$$b^2 = 1 - 3\eta^2 + a^2,$$

$$c = \frac{G\left(\frac{\lambda + 1}{2}\right)}{\sqrt{\lambda(\lambda + 2)} G(\lambda/2)}.$$

The skewness and kurtosis (degrees of freedom) parameters,  $\lambda$  and  $\eta$ , along with the other parameters of the model, are allowed to depend on the lagged error terms, that is,  $\lambda_t = \lambda(e_{t-1})$ ,  $\eta_t = \eta(e_{t-1})$ . Application of the model to the monthly excess holding yield on US Treasury securities (allowing only for time varying kurtosis) and to the US\$/Swiss Franc exchange rate (allowing for both time varying kurtosis and skew) produces evidence of time-varying higher moments in the data. It would be of interest to see if other distributions can be effectively applied to the question of the predictability of distributional shape, perhaps using measures of asymmetry and elongation that are known to be more resistant to the influences of outliers.

A recent development is the derivation of density forecasts from the information about market participants' perceptions of the underlying asset price distribution contained in option

market data. Soderlind and Svensson (1997) describe how methods of extracting information about market expectations from asset prices for monetary policy purposes have developed from the estimation of expected means of future interest rates and exchange rates from forward rates to estimation of their complete densities from traded options prices. The risk-management motivation noted above has also driven applications in finance, as deeper markets and new instruments have provided relevant data. The starting point is the derivation by Breeden and Litzenberger (1978) of a relationship between the risk-neutral density of the price of the underlying asset and the pricing function of European (exercised only at the maturity date) call options. For example, the classic Black-Scholes option pricing model implies a lognormal risk-neutral density function. Once markets were established and became sufficiently active, empirical applications followed. Thus Fackler and King's (1990) analysis of implied density forecasts of US agricultural commodity prices followed the initiation of trading in commodity option futures in 1984, while Jackwerth and Rubinstein's (1996) data on options on the S & P 500 index start on 2 April 1986, the date the Chicago Board Options Exchange switched from American (exercised at any time up to the maturity date) to European options. Bahra (1996, 1997) analyses the risk-neutral density functions of short-term sterling and Deutschmark interest rates implied by options traded on the London International Financial Futures and Options Exchange (LIFFE), which launched options trading in 1985.

Observed differences between market prices and the prices implied by the Black-Scholes model have led to variations in the model's assumptions and the use of distributions other than the lognormal for the underlying asset price. In particular, the estimated volatility of the underlying asset's return varies with the exercise or strike price, contrary to the Black-Scholes model - the so-called "volatility smile" - which calls for distributions with fatter tails than the lognormal and possibly a different degree of skewness. More general functional forms that have



been employed include a mixture of lognormals (Bahra, 1996; Melick and Thomas, 1997; Soderlind and Svensson, 1997) and the generalized beta (Aparicio and Hodges, 1998); Ait-Sahalia and Lo (1998) develop a nonparametric approach to estimating the density. Jackwerth (1999) reviews the various methods, noting that the results tend to be rather similar unless only a very few option prices are available. It should also be noted, as several authors do, that the risk-neutral densities coincide with the “true” densities only in the absence of risk premia, and Soderlind (1999) considers how measures of risk premia might be used to adjust estimated risk-neutral densities.

In the VaR context, as in macroeconomics, stochastic simulation methods are a means of generating non-normal density forecasts. This approach develops a density forecast by simulating possible future paths of the various components of the portfolio of interest, based on models of those components with parameters estimated from past realizations of the corresponding variable. These various sample paths are combined into a sample path of the portfolio, and a density forecast is then obtained from repeated simulations. Early implementations of this approach assumed price changes to be iid Gaussian, so that the portfolio return is also normally distributed. Later examples employ mixtures, jump diffusions, and models with conditional dynamics in the second moments, which would generate non-normal distributions. Duffie and Pan (1997) and Dowd (1998) present comprehensive summaries of the issues and methods in generating VaRs. The focus of attention in VaR analysis is often an extreme quantile of the distribution. However, the methods described above usually perform better near the centre of the data than at the extreme tails, where data are scarce. This has motivated the introduction of methods based on extreme value theory to improve the estimation of the tail of the distribution (see McNeil and Frey, 1998, and references therein). The application of this theory in risk management adds a practical concern to the more theoretical

interest in the shape of the tails of the density, which is related to the question of the existence of moments (see Pagan, 1996, for example). Extreme value theory often begins with the assumption that the data are iid and the distribution has Pareto-type tails, with distribution function  $F(y) = 1 - ky^{-\alpha}$  above a certain threshold, and attention focuses on the estimation of the tail index  $\alpha$ . It can be argued that the development of extreme value methods beyond the limitations of these two assumptions is important for the theory's relevance to financial applications (see, for example, Diebold, Schuermann and Stroughair, 1998), and McNeil and Frey (1998) extend existing methods by considering the conditional distribution of asset returns and the generalized Pareto distribution function for the upper tail of the standardized residuals.

This is

$$F(y) = \begin{cases} 1 - (1 - y/\beta)^{1/\alpha} & \alpha > 0 \\ 1 - \exp(-y/\beta) & \alpha = 0 \end{cases}$$

where  $\beta > 0$ , with support  $y \geq 0$  when  $\alpha \geq 0$  and  $0 \leq y \leq \beta/\alpha$  when  $\alpha < 0$ . The case  $\alpha > 0$  corresponds to heavy-tailed distributions whose tails decay like power functions, including the Pareto, Student's  $t$ , Cauchy and Burr distributions, while  $\alpha = 0$  corresponds to the distributions such as the normal, lognormal, exponential and gamma, whose tails decay exponentially. The case  $\alpha < 0$  includes distributions like the uniform and beta, with finite upper bound.

Quantile regression methods represent an alternative approach to the estimation of conditional quantile functions. Koenker and Zhao (1996) extend the quantile regression ideas of Koenker and Bassett (1978) to the ARCH setting, noting that quantiles are readily interpretable in semiparametric location-scale models and are easier to estimate robustly than moments. Thus, moving outside the Gaussian context, they focus on models for conditional

scale rather than conditional variance. In an earlier application Granger, White and Kamstra (1989) use quantile regression to combine several quantile estimates, in the course of constructing prediction intervals which vary over time thanks to ARCH effects. In a general time-series forecasting context, Taylor and Bunn (1999) apply quantile regression techniques to obtain quantiles of multi-step-ahead forecast densities as functions of the lead time.

Finally, we note another forecasting methodology that is related to density forecasting, namely the predictive likelihood approach (see, for example Bjørnstad 1990). Here a likelihood function

$$L(y_{t+h}, \theta | y_t, y_{t+1}, \dots)$$

is formulated from a fully specified parametric model, treating the forecast variable at the horizon of interest as a parameter to be estimated. The other parameters of the model, denoted here as  $\theta$ , are considered to be nuisance parameters, and are removed from the likelihood using a variety of methods to obtain the predictive likelihood

$$L(y_{t+h} | y_t, y_{t+1}, \dots).$$

Some forecasters normalize the predictive likelihood, which is then interpreted as an estimate of the density of  $y_{t+h}$  conditioned on its past history.

### 3. PRESENTATION OF DENSITY FORECASTS

No matter how density forecasts are generated, it is important to pay attention to their presentation. The way density forecasts are presented and communicated may hide some features of the forecasts, and emphasize others. Inappropriate presentation of forecasts will reduce the forecasts' usefulness for certain forecast users, and worse, may be misleading or misinterpreted, resulting in improper use of the forecasts. Presentation of forecasts tends, however, to be given limited coverage in discussions of forecasting, and we highlight some aspects of this important

issue.

In forecasting, different users have different concerns (or loss functions, formally) and so may be interested in different aspects of the forecast. In the context of point forecasts, for instance, it is known that if the relevant loss function is symmetric then the optimal point predictor is the conditional mean, but that this is not true for general loss functions, as noted by Granger (1969) and Christoffersen and Diebold (1997). Producers of forecasts need to understand the concerns of the target audience (the particular set of forecast users) and present the appropriate point forecast, and a user with asymmetric loss who bases a decision on a conditional mean point forecast will make a sub-optimal decision. In principle the use of density forecasts overcomes this problem: density forecasts are relevant to all forecast users since they are a complete description of the uncertainty associated with the forecast of a variable. In practice, however, the question of how to present density forecasts without losing vital information remains.

The notion of finding the most appropriate way of presenting density forecasts is very much in the spirit of exploratory data analysis where the concern is how best to summarize graphically the information contained in a sample of data, see Tukey (1977). In forecasting, the same problem arises because different features of the forecast are of interest to different users, and the presentation of the forecast should highlight the appropriate features. For example, if a forecast user is interested in possible asymmetries in the future distribution of the forecast variable, a plot of the density estimate with an appropriate symmetric density superimposed may be more helpful than the presentation of the forecast by itself.

It is sometimes feasible to present the analytical form of the density forecasts, but these are available only if standard distributions are used, and in any event the features of the forecast may not be immediately obvious from algebraic expressions. A more common way of presenting

density forecasts is by plotting the density estimate. This would usually be the case for forecasts obtained from semiparametric approaches to density forecasting discussed above, or if the forecast is obtained via simulation methods and the density estimate computed using kernel methods such as those described by Silverman (1986). Graphical presentations are useful - asymmetries in the forecasts are often easily picked out, and the normal distribution (the symmetric distribution taken by convention as having “neutral elongation”) can be imposed to highlight the presence of asymmetries or excess kurtosis in the data to the user. It is often helpful to discretize the density, perhaps by presenting it as a histogram, graphically or in tabular form, indeed this is how density forecasts are elicited in the Survey of Professional Forecasters, as noted above: Table 1 gives an example. When a complete density is available, the conventional discretisation is based on quantiles, so that prediction intervals are reported, again as a graph or a table, for a regular sequence of coverage probabilities, in contrast to the reporting of probabilities for a regular sequence of intervals, as in the histogram case. On the other hand a particular interval may be a focus of attention. In a monetary policy regime of inflation targeting, for example, the objective of policy is sometimes expressed as a target range for inflation, whereupon it is of interest to report the forecast probability that the future outcome will fall in the target range.

In real-time forecasting, a sequence of forecasts for a number of future periods from a fixed origin (the “present”) is often presented as a time-series plot. The central forecast may be shown as a continuation of the plot of actual data recently observed, and limits may be attached, either as standard error bands or quantiles, becoming wider as the forecast horizon increases. As Thompson and Miller (1986) note, “typically forecasts and limits are graphed as dark lines on a white background, which tends to make the point forecast the focal point of the display”. They argue for and illustrate the use of selective shading of quantiles, as “a deliberate attempt to draw

attention away from point forecasts and toward the *uncertainty* in forecasting” (1986, p.431, emphasis in original). In presenting its density forecasts of inflation the Bank of England takes this argument a stage further, by replacing a point forecast by a central 10 per cent interval. The alternative presentation of a recent Bank forecast by Wallis (1999), based on conventional percentiles, is shown in Figure 2. The density forecast is represented graphically as a set of prediction intervals covering 10, 20, ..., 90 per cent of the probability distribution, of lighter shades for the outer bands; equivalently the boundaries of the bands are the 5<sup>th</sup>, 10<sup>th</sup>, ..., 95<sup>th</sup> percentiles (excluding the 50<sup>th</sup>). This is done for inflation forecasts one to eight quarters ahead, and since the dispersion increases and the intervals “fan out” as the forecast horizon increases, the result has become known as the “fan chart”.

The Bank of England’s preferred presentation of the fan chart is based on the shortest intervals for the assigned probabilities, which differ from the “central” intervals used in Figure 2 whenever the density is asymmetric: in particular they converge on the mode, rather than the median, as the coverage is reduced. Moreover the tail probabilities for the shortest intervals are not only unequal, they are not reported, leaving open the possibility of misinterpretation by a user who assumes them to be equal. Comparison of the two can be illuminated by reference to a loss function, which provides a rare example of the use of a loss function beyond the point prediction context, albeit to the interval prediction problem rather than the complete density forecast. To construct an optimal or minimum cost prediction interval (a,b) it is first assumed that there is a cost proportional to the length of the interval which is incurred irrespective of the outcome, and the distinction between the two cases arises from the assumption about the additional cost associated with a mistake, that is, the interval not containing the outcome. If this has an all-or-nothing form, being zero if the interval contains the outcome and a positive constant otherwise, then the optimal interval satisfies  $f(a) = f(b)$ , and this “equal height” property is also a property

of the interval with shortest length  $b-a$  for given coverage. If, however, the cost of a mistake is proportional to the amount by which the outcome lies outside the interval, then the optimal interval is the central interval with equal tail probabilities. (See Wallis, 1999, for derivations.) We find the all-or-nothing loss function's indifference to the actual magnitude of error unrealistic. A preference for the alternative is implicit in the practice of the overwhelming majority of statisticians of summarizing densities by presenting selected percentiles.

While a density forecast can be seen as an acknowledgment of the uncertainty in a point forecast, it is itself uncertain, and this second level of uncertainty is of more than casual interest if the density forecast is the direct object of attention, as in several of the finance applications discussed above. How this might be described and reported is beginning to receive attention. For parametric approaches that treat the underlying distributional assumptions as correct the effect of parameter estimation error can be estimated via an appropriate covariance matrix and reported as a point-by-point confidence interval for the density forecast: Soderlind and Svensson (1997) plot an example for an options-based implied risk-neutral density. For approaches based on simulation methods, not only the point estimate of the density but also a sampling standard error or robust alternative measure of dispersion could likewise be reported.

Finally, we note that the problem of reporting results and communicating with "remote clients" has been a concern of Bayesian statisticians at least since Hildreth (1963) characterized users in this way. With the increasing use of simulation methods in economic forecasting, many forecasts will be available in the form of a sample from the predictive distribution, and these simulations could be used by various forecast users to generate estimates of specific features appropriate to their individual purposes. In this connection, the proposal by Geweke (1997) to exploit modern computation, communication and information storage and retrieval capabilities to facilitate Bayesian communication also merit attention by density forecasters.

#### 4. EVALUATION AND CALIBRATION

Given a series of forecasts over a period of time, we consider the question of how to assess forecasting performance *ex post*. Evaluation of the quality of the forecasts may be of interest for its own sake, or may be explicitly directed towards improvement of future performance. For point forecasts, there is a large literature on the *ex post* evaluation of *ex ante* forecasts, and a range of techniques has been developed, recently surveyed by Wallis (1995) and Diebold and Lopez (1996). The evaluation of interval forecasts has a much newer literature (Christoffersen, 1998; Taylor, 1999), as does the evaluation of density forecasts, which is our present concern.

One strand of the literature on point forecasts comprises descriptive accounts of forecasts and forecast errors in specific episodes of particular interest, often business cycle turning points. For example Wallis (1989) reviews several accounts of macroeconomic forecast performance during the recessions of the 1970s. It is striking that many studies of the performance of options-based densities adopt the same approach, describing the behaviour of the implied probability distributions before and after such events as the US stock market crash of October 1987 (Jackwerth and Rubinstein, 1996), the Persian Gulf crisis of 1990-91 (Melick and Thomas, 1997), the crisis in the European exchange rate mechanism around 16 September 1992 - “black Wednesday” - (Malz, 1996) together with the following month’s announcement of a new monetary policy framework in the United Kingdom (Soderlind, 1999) and, at a more mundane level, announcements of economic news and shifts in official interest rates (Bahra, 1996). Agreement with other sources of information validates the estimated distributions - they “are consistent with the market commentary at the time” (Melick and Thomas) or are “validated by recent market developments” (Bahra) - and they can also supplement accounts of the events as they unfolded - “the market clearly believed that UK monetary policy was in big trouble”; later



“this took the market by some surprise” (Soderlind). Interesting, indeed entertaining, as these accounts are, they provide little systematic information on such questions as the comparative performance of different models or methods, and one suspects that here, as in other areas of forecasting, they will gradually be replaced by more formal analyses.

It is often argued that forecasts should be evaluated in an explicit decision context, that is, in terms of the economic consequences that would have resulted from using the forecasts to solve a sequence of decision problems. The incorporation of a specific loss function into the evaluation process would focus attention on the features of interest to the forecast user, perhaps also showing the optimality of a particular forecast. In macroeconomic forecasting this does not happen, given the difficulty of specifying a realistic loss function and the absence of a well-defined user. In finance there is usually a more obvious profit-and-loss criterion, and there is a long tradition of forecast evaluation in the context of investment performance. This extends to volatility models (West, Edison and Cho, 1993, for example) but not yet to density forecasts. Here there are relatively few results based on explicit loss functions, as noted above. The basic result that a “correct” forecast is optimal regardless of the form of the loss function is extended from point forecasts to event probability forecasts by Granger and Pesaran (1996) and to density forecasts by Diebold, Gunther and Tay (1998). The latter authors also show that there is no ranking of sub-optimal density forecasts that holds for all loss functions. The problem of the choice of forecast would require the use of loss functions defined over the distance between forecast and actual densities. Instead the general objective of density forecasters is to get close to the correct density in some sense, and practical evaluations are based on the same idea.

Statistical evaluations of real-time density forecasts have recently begun to appear, although the key device, the probability integral transform, has a long history. The literature usually cites Rosenblatt (1952) for the basic result, and the approach features in several

expositions from different points of view, such as Dawid (1984) and Cooke (1991). For a sample of  $n$  one-step-ahead forecasts and the corresponding outcomes, the probability integral transform of the realized variables with respect to the forecast densities is

$$z_t = \int_0^{y_t} p_t(u) du$$

$$z_t \sim P_t(y_t), \quad t = 1, \dots, n,$$

where  $P_t(\cdot)$  is the forecast distribution function. If  $P_t(\cdot)$  is “correct”, then the  $z_t$  are independent uniform  $U[0,1]$  variates. The independence property is obvious in the case of iid forecasts, and also extends to the case of time dependent density forecasts, as when the forecasts comprise a sequence of conditional densities, provided that the forecasts are based on an information set that contains the past history of the forecast variable (Diebold, Gunther and Tay, 1998). The question then is whether the  $z_t$  sequence “looks like” a random sample from  $U[0,1]$  (Dawid, 1984, quotation marks in original), that is, whether the forecasts are “well calibrated”. Deviations from uniform iid will indicate that the forecasts have failed to capture some aspect of the underlying data generating process. Serial correlation in the  $z_t$  sequence (in squares, third powers, etc) would indicate poorly modelled dynamics, whereas non-uniformity may indicate improper distributional assumptions, or poorly captured dynamics, or both. Diebold, Hahn and Tay (1999) show that if the true conditional density belongs to a location-scale family, and the forecaster issues location-scale density forecasts with correctly specified location and scale, the  $z_t$  sequence will continue to bear the iid property but will no longer be uniform if the wrong location-scale density is used. Extensions to multi-step-ahead forecasts are awaited.

The uniformity of the probability integral transforms is usually evaluated by plotting the empirical distribution function and comparing with a 45 degree line, whereas some authors find an estimate of the density easier to check visually for departures from uniformity. Formal tests

of goodness of fit can be employed, such as that based on Kolmogorov's  $D_n$ - statistic, the maximum absolute difference between the empirical distribution function and the null hypothesis uniform distribution function. The distribution theory for  $D_n$  rests on an assumption of random sampling, however, whereas the hypothesis of interest is the joint hypothesis of iid uniformity, and little is known about the impact on critical values of  $D_n$  of departures from independence. Diebold, Gunther and Tay (1998) also consider the conditional dynamics of the density forecasts by examining the correlograms of the levels and powers of the probability integral transforms, and present an application to density forecasts of daily S & P 500 returns. Diebold, Tay and Wallis (1999) add resampling procedures to the toolkit in the course of an evaluation of the Survey of Professional Forecasters' density forecasts of inflation. These authors supplement formal tests with a graphical approach, in an exploratory spirit, so that the evaluation process might be informative about the direction in which a forecasting model could be improved. Extensions to the multivariate case are considered by Diebold, Hahn and Tay (1998) and Clements and Smith (1998). Both approaches are based on a decomposition of multivariate forecasts into univariate conditionals. The first paper evaluates a bivariate forecasting model of high-frequency exchange rates, and the latter paper uses density forecast performance to compare linear models with nonlinear forecasting models of output growth and unemployment.

Clements and Smith use forecast densities to discriminate between competing models, in place of standard measures such as mean square forecast errors. To see how forecast densities can discriminate between competing models, consider the simple example of comparing two zero-mean volatility models differing only in distributional assumptions. The conditional mean forecasts generated by both models are identical (zero), as are their mean square forecast errors. The forecast densities, on the other hand, would be different in general, and the probability integral transform would highlight these differences. Clements and Smith find that while

standard means of comparing models suggest that nonlinear models offer no improvement over linear models, evaluation techniques that consider the entire forecast density are able to discriminate between the two: in their example the nonlinear models they consider (self-exciting threshold autoregressive models) outperform the linear models, being better able to predict higher order moments.

If a given model or systematic forecasting method is found to produce poorly calibrated forecasts over an early time period, it may be possible to use the results of this initial evaluation to improve subsequent forecasts. Dawid (1984) speaks of “tuning” the system to provide better calibrated forecasts; other authors speak of “de-biasing” or “recalibrating” the later forecasts. Given an estimate  $\hat{Q}(z)$  of an empirical distribution function whose graph does not look like a 45 degree line, the recalibrated forecast distribution function is

$$P^*(y) = \hat{Q}[P(y)].$$

Specifying  $\hat{Q}(\cdot)$  is not straightforward, and Kling and Bessler (1989) estimate it as a piecewise linear function in the course of a recalibration of their VAR-based density forecasts of interest rates, money, prices and output. Fackler and King (1990) fit a beta distribution to the empirical distribution of the  $z$ 's to recalibrate option-based density forecasts of prices in agricultural commodity markets, initially based on a lognormal assumption. The beta distribution contains the uniform distribution as a special case, and so admits the possibility of a likelihood ratio test. Diebold, Hahn and Tay (1998) again provide a multivariate extension and recalibrate a series of bivariate Gaussian density forecasts of high frequency exchange rate returns based on an exponential smoothing approach to estimation of the conditional variances and covariances of the returns. These examples all illustrate a further way of generalizing distributional assumptions away from normality.

## 5. CONCLUSION

The interest in and use of density forecasts has increased in recent years. As in other areas of quantitative economics, advances in statistical methodology, the greater availability of relevant data, and increases in computing power have all played a part. Public discussion of macroeconomic point forecasts too often treats them as exact, and to acknowledge explicitly that they are not, perhaps by publishing a density forecast, can only improve the policy debate. In finance density forecasts are used more directly for specific assessments of risk, where mistakes have obvious commercial consequences. In our discussion of the construction, presentation and evaluation of density forecasts the need to keep in mind the loss function of the forecast user is always present in principle, but in practice this rarely features explicitly in the literature. Perhaps this is one direction in which to look for future developments. The commercial consequences of risk assessments can be expressed as a loss function, whose formulation might be extended to illuminate the specification and evaluation of density forecasts. Even when this is done, however, it will be important to have methods of communication that can accommodate users with a variety of loss functions.

Other future developments in density forecasting might be identified with reference to the literature on point forecasts. Both forecasts are usually model-based, and model improvements are often motivated by the need to improve forecasts; a particular issue identified above relates to the predictability of higher moments. The evaluation of density forecasts is at a comparatively rudimentary stage, and issues that merit attention are the optimal revision of fixed-event forecasts, the comparative evaluation of forecasts, and evaluation of multi-step-ahead density forecasts. Combinations of forecasts sometimes form part of comparative evaluations and sometimes are of interest for their own sake, and the literature on combining point forecasts that springs from Bates and Granger (1969) has reached prediction intervals but not yet density forecasts. Finally we note that, while a density forecast provides a representation of the

uncertainty in a point forecast, its own uncertainty should also be acknowledged and quantified. How this second level of uncertainty might in turn be described and the description subsequently evaluated has scarcely begun to be considered. A future survey of this area has much to look forward to.

## References

- Ait-Sahalia Y, Lo A. 1998. Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance* **53**: 499-547.
- Aitchison J, Dunsmore IR. 1975. *Statistical Prediction Analysis*. Cambridge University Press: Cambridge.
- Alles LA, Kling JL. 1994. Regularities in the variation of skewness in asset returns. *Journal of Financial Research* **17**: 427-438.
- Anscombe FJ. 1967. Topics in the investigation of linear relations fitted by the method of least squares (with discussion). *Journal of the Royal Statistical Society B* **29**:1-52.
- Aparicio SD, Hodges SD. 1998. Implied risk-neutral distribution: a comparison of estimation methods. Preprint 98/95, Financial Options Research Centre, University of Warwick.
- Bahra B. 1996. Probability distributions of future asset prices implied by option prices. *Bank of England Quarterly Bulletin* **36**: 299-311.
- Bahra B. 1997. Implied risk-neutral probability density functions from option prices: theory and application. Working Paper No.66, Bank of England: London.
- Baillie RT, Bollerslev T. 1992. Prediction in dynamic models with time dependent conditional variances. *Journal of Econometrics* **52**: 91-114.
- Bates JM, Granger CWJ. 1969. The combination of forecasts. *Operations Research Quarterly* **20**: 451-468.
- Bjørnstad JF. 1990. Predictive likelihood: a review (with discussion). *Statistical Science* **5**: 242-265.
- Blake AP. 1996. Forecast error bounds by stochastic simulation. *National Institute Economic Review* No.156: 72-79.
- Bollerslev T. 1987. A conditionally heteroskedastic time series model for speculative prices

- and rates of return. *Review of Economics and Statistics* **69**: 542-546.
- Bollerslev T, Engle RF, Nelson D. 1994. ARCH Models In *Handbook of Econometrics 4*, Engle RF, McFadden D. (eds); Elsevier: Amsterdam.
- Breeden DT, Litzenberger RH. 1978. Prices of state-contingent claims implicit in options prices. *Journal of Business* **51**: 621-651.
- Britton E, Fisher PG, Whitley JD. 1998. The *Inflation Report* projections: understanding the fan chart. *Bank of England Quarterly Bulletin* **38**: 30-37.
- Christoffersen PF. 1998. Evaluating interval forecasts. *International Economic Review* **39**: 841-862.
- Christoffersen PF, Diebold FX. 1997. Optimal prediction under asymmetric loss. *Econometric Theory* **13**: 808-817.
- Clements MP, Smith J. 1998. Evaluating the forecast densities of linear and non-linear models: applications to output growth and unemployment. *Journal of Forecasting*, forthcoming.
- Cooke RM. 1991. *Experts in Uncertainty: Opinion and Subjective Probability in Science*. Oxford University Press: Oxford.
- Cotner JS. 1991. Index option pricing: do investors pay for skewness. *Journal of Futures Markets* **11**: 1-8.
- Croushore D. 1993. Introducing: The Survey of Professional Forecasters. *Business Review*, Federal Reserve Bank of Philadelphia November/December 1993: 3-15.
- Dawid AP. 1984. Statistical theory, the prequential approach. *Journal of the Royal Statistical Society A*, **147**: 278-290.
- Diebold FX, Lopez JA. 1996. Forecast evaluation and combination. In *Handbook of Statistics 14: Statistical Methods in Finance*, Maddala GS, Rao CR (eds); North-



Holland: Amsterdam.

Diebold FX, Gunther TA, Tay AS. 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* **39**: 863-883.

Diebold FX, Schuermann T, Stroughair JD. 1998. Pitfalls and opportunities in the use of extreme value theory in risk management. In *Advances in Computational Finance*, Refenes AN, Burgess AN, Moody JD (eds); Kluwer: Amsterdam.

Diebold FX, Tay AS, Wallis KF. 1999. Evaluating density forecasts of inflation: the Survey of Professional Forecasters. In *Cointegration, Causality, and Forecasting: Festschrift in Honour of Clive W.J. Granger*, Engle RF, White H (eds); Oxford University Press: Oxford.

Diebold FX, Hahn J, Tay AS. 1999. Multivariate density forecast evaluation and calibration in financial risk management: high-frequency returns on foreign exchange. *Review of Economics and Statistics*, forthcoming.

Dowd K. 1998. *Beyond Value at Risk: The New Science of Risk Management*. John Wiley: Chichester.

Duffie D, Pan J. 1997. An overview of Value at Risk. *Journal of Derivatives* **4**: 7-49.

Engle RF. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**: 987-1008.

Engle RF, González-Rivera G. 1991. Semiparametric ARCH models. *Journal of Business and Economic Statistics* **9**: 345-359.

Fama EF. 1965. The behavior of stock market prices. *Journal of Business* **38**: 34-105.

Fackler PL, King RP. 1990. Calibration of options-based probability assessments in agricultural commodity markets. *American Journal of Agricultural Economics* **72**: 73-83.

- Gallant R, Hsieh DA, Tauchen G. 1991. On fitting a recalcitrant series:the pound/dollar exchange rate 1974-83. In *Nonparametric and Semiparametric Methods in Econometrics and Statistics: Proceedings of the Fifth International Symposium in Economic Theory and Econometrics*, Barnett WA, Powell J, Tauchen G (eds); Cambridge University Press: Cambridge.
- Geweke J. 1997. Posterior simulators in econometrics. In *Advances in Economics and Econometrics: Theory and Applications Vol III*, Kreps DM, Wallis KF (eds); Cambridge University Press: Cambridge.
- Granger CWJ. 1969. Prediction with a generalized cost of error function. *Operational Research Quarterly* **20**: 199-207.
- Granger CWJ, Pesaran MH. 1996. A decision theoretic approach to forecast evaluation. Discussion Paper 96-23, Department of Economics, University of California, San Diego.
- Granger CWJ, White H, Kamstra M. 1989. Interval forecasting: an analysis based on ARCH-quantile estimators. *Journal of Econometrics* **40**: 87-96.
- Hansen BE. 1994. Autoregressive conditional density estimation. *International Economic Review* **35**: 705-730.
- Hildreth C. 1963. Bayesian statisticians and remote clients. *Econometrica* **31**:422-438.
- Jackwerth JC. 1999. Option implied risk-neutral distributions and implied binomial trees: a literature review. *Journal of Derivatives*, forthcoming.
- Jackwerth JC, Rubinstein M. 1996. Recovering probability distributions from option prices. *Journal of Finance* **51**: 1611-1631.
- John S. 1982. The three-parameter two-piece normal family of distributions and its fitting. *Communications in Statistics - Theory and Methods* **11**: 879-885.

- Johnson N, Kotz S, Balakrishnan N. 1994. *Continuous Univariate Distributions*, 2nd edn vol.1. Wiley: New York.
- Kimber AC. 1985. Methods for the two-piece normal distribution. *Communications in Statistics - Theory and Methods*, **14**: 235-245.
- Kling JL, Bessler DA. 1989. Calibration-based predictive distributions: an application of prequential analysis to interest rates, money, prices and output. *Journal of Business* **62**: 477 - 499.
- Koenker R, Bassett G. 1978. Regression quantiles. *Econometrica* **46**:33-50.
- Koenker R, Zhao Q. 1996. Conditional quantile estimation and inference for ARCH models. *Econometric Theory* **12**: 793-813.
- Lee TKY, Tse YK. 1991. Term structure of interest rates in the Singapore Asian dollar market. *Journal of Applied Econometrics* **6** :143-152.
- Malz AM. 1996. Using option prices to estimate realignment probabilities in the European monetary system: the case of sterling-mark. *Journal of International Money and Finance* **15**: 717-748
- McNeil AJ, Frey R. 1998. Estimation of time-related risk measures for heteroscedastic financial time series: an extreme value approach. Preprint, Department of Mathematics, ETH, Zurich.

- Melick WR, Thomas CP. 1997. Recovering an asset's implied PDF from option prices: an application to crude oil during the Gulf crisis. *Journal of Financial and Quantitative Analysis* **32**: 91-115.
- Nelson DB. 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**: 347-370.
- Pagan AR. 1996. The econometrics of financial markets. *Journal of Empirical Finance* **3**: 15-102.
- Poulizac D, Weale MR, Young G. 1996. The performance of National Institute economic forecasts. *National Institute Economic Review* No.156: 55-62.
- Rosenblatt M. 1952. Remarks on a multivariate transformation. *Annals of Mathematical Statistics* **23**: 470-472.
- Simkowitz MA, Beedles WL. 1978. Diversification in a three-moment world. *Journal of Finance* **33**: 288-292.
- Singleton JF, Wingender J. 1986. Skewness persistence in common stock returns. *Journal of Financial and Quantitative Analysis* **21**: 335-341.
- Silverman BW. 1986. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: New York.
- Soderlind P. 1999. Market expectations in the UK before and after the ERM crisis. *Economica*, forthcoming.
- Soderlind P, Svensson LEO. 1997. New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics* **40**: 383-429.
- Taylor JW. 1999. Evaluating volatility and interval forecasts. *Journal of Forecasting* **18**: 111-128.
- Taylor JW, Bunn DW. 1999. A quantile regression approach to generating prediction

- intervals. *Management Science* **45**: 225-237.
- Thompson PA, Miller RB. 1986. Sampling the future: a Bayesian approach to forecasting from univariate time series models. *Journal of Business and Economic Statistics* **4**: 427-436.
- Tukey JW. 1977. *Exploratory Data Analysis*. Addison-Wesley: Reading, Mass.
- Wallis KF. 1989. Macroeconomic forecasting: a survey. *Economic Journal* **99**: 28-61.
- Wallis KF. 1995. Large-scale macroeconometric modeling. In *Handbook of Applied Econometrics*, Pesaran MH, Wickens MR (eds); Blackwell: Oxford.
- Wallis KF. 1999. Asymmetric density forecasts of inflation and the Bank of England's fan chart. *National Institute Economic Review*, No.167: 106-112.
- West KD, Edison HJ, Cho D. 1993. A utility-based comparison of some models of exchange rate volatility. *Journal of International Economics* **35**: 23-45.
- Zarnowitz V. 1969. The new ASA-NBER survey of forecasts by economic statisticians. *American Statistician* **23**:12-16.

**Table 1. Density forecasts of US inflation**  
Mean probability (of 28 forecasters) attached to possible percent changes in GDP price index, 1998-99

Inflation rate (per cent)	Probability (per cent)
8.0 or more	0.07
7.0 to 7.9	0.11
6.0 to 6.9	0.14
5.0 to 5.9	0.25
4.0 to 4.9	1.21
3.0 to 3.9	8.96
2.0 to 2.9	20
1.0 to 1.9	49.54
0.0 to 0.9	17.89
Will decline	1.82

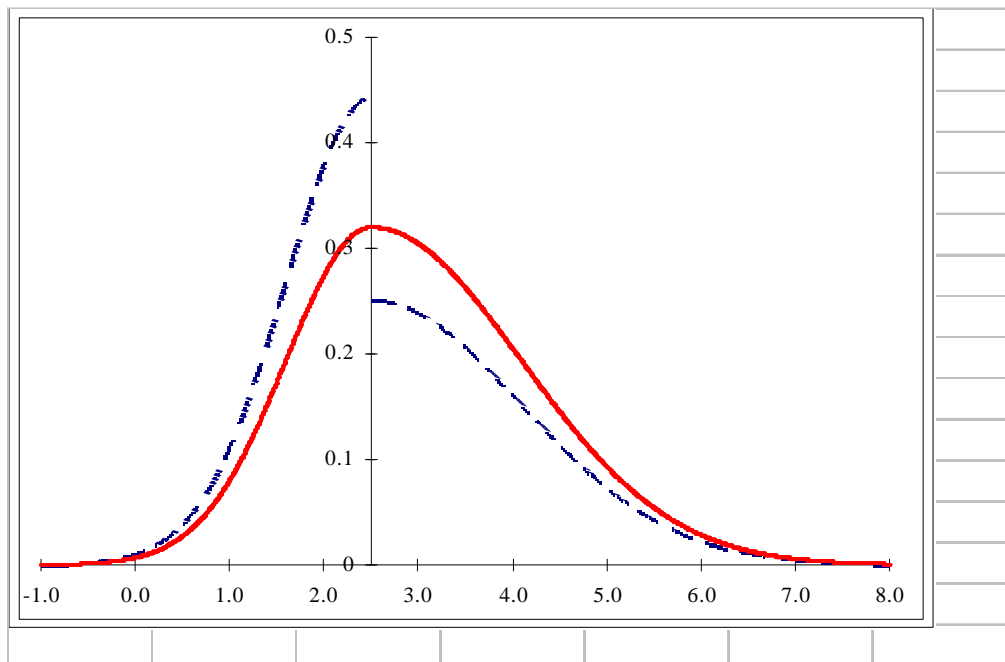
Source: Philadelphia Fed., *Survey of Professional Forecasters*, 20 November 1998  
(<http://www.phil.frb.org/econ/spf/spfq498.pdf>)

**Table 2. Density forecasts of UK inflation**  
Probability of indicated annual RPIX inflation rate, percent, 1997Q4

	KB	TC	GD	PM	BR	MW
8.0 or more	0	0	0	0	0	1
7.0 to 7.9	0	0	0	0	0	1
6.0 to 6.9	1	0	1	0	0	4
5.0 to 5.9	4	2	2	0	1	8
4.0 to 4.9	8	7	8	0	6	13
3.0 to 3.9	22	32	16	5	22	18
2.0 to 2.9	55	50	40	30	35	19
1.0 to 1.9	10	9	25	40	26	16
0.0 to 0.9	0	0	6	20	9	11
Will decline	0	0	2	5	1	9

Source: Panel of Independent Forecasters, November 1996 Report (London: HM Treasury).  
The panel members are Kate Barker, Tim Congdon, Gavin Davies, Patrick Minford,  
Bridget Rosewell and Martin Weale

**Figure 1** The probability density function of the two-piece normal distribution



dashed line : two halves of normal distributions with  $m= 2.5$ ,  
 $s_1 = 0.902$  (left) and  $s_2 = 1.592$  (right)  
solid line : the two-piece normal distribution



**Figure 2** The August 1997 *Inflation Report* forecast

