Jacek Osiewalski* , Mateusz Pipień**

BAYESIAN PRICING OF AN EUROPEAN CALL OPTION USING A GARCH MODEL WITH ASYMMETRIES

Summary. In this paper option pricing is treated as an application of Bayesian predictive analysis. The distribution of the discounted payoff, induced by the predictive density of future observables, is the basis for direct option pricing, as in Bauwens and Lubrano (1997). We also consider another, more eclectic approach to option pricing, where the predictive distribution of the Black-Scholes value is used (with volatility measured by the conditional standard deviation at time of maturity).

We use a model framework that allows for two types of asymmetry in GARCH processes: skewed *t* conditional densities and different reactions of conditional scale to positive/negative shocks. Our skewed *t*-GARCH(1, 1) model is used to describe daily changes of the Warsaw Stock Exchange Index (WIG) from 4.01.1995 till 8.02.2002. The data till 28.09.2001 are used to obtain the posterior and predictive distributions, and to illustrate Bayesian option pricing for the remaining period.

Keywords: Bayesian inference, financial econometrics, derivative pricing, volatility models, forecasting

1. INTRODUCTION

The paper presents two Bayesian approaches to option pricing; both use univariate daily time series of a basic financial instrument. The distribution of the payoff function, induced by the predictive density of future observables, is the basis for direct pricing. This first approach, proposed by Bauwens and Lubrano (1997), relies only on the statistical model for discrete observations. We also consider another (more eclectic) approach to option pricing, where the predictive distribution of the famous Black-Scholes value is calculated using the same discrete statistical model as in the direct approach.

*Prof. zw. dr hab., Katedra Ekonometrii, Akademia Ekonomiczna w Krakowie.
**Dr, adiunkt, Katedra Ekonometrii, Akademia Ekonomiczna w Krakowie.

Our model framework generalises univariate AR – t-GARCH specifications from previous Bayesian studies; see Kleibergen and van Dijk (1993), Bauwens and Lubrano (1997, 1998), Bauwens, Lubrano and Richard (1999). We allow for two types of asymmetry in GARCH processes. The first one means different reactions of conditional scale to positive and negative shocks and is modelled in the way proposed by Glosten, Jagannathan and Runkle (1993), and used by Bauwens and Lubrano (1997, 1998). The other type is asymmetry of conditional distributions, modelled through skewed Student t family that has been defined in a general multivariate framework by Fernàndez, Osiewalski and Steel (1995). Univariate skewed t family was analysed by Fernàndez and Steel (1998), who used it, however, in simpler models, without any ARCH-type structure. Osiewalski and Pipień (1999) adopted the skewed t distribution for error terms in GARCH processes, thus generalising and unifying the works by Bauwens and Lubrano (1998) and Fernàndez and Steel (1998). By introducing the free mode parameter of the skewed t conditional sampling distribution of a GARCH(1, 1) process, Osiewalski and Pipień (2000) proposed a formal statistical approach to the derivation and testing of the GARCH-in-Mean (GARCH-M) effect, first considered by Engle, Lilien and Robins (1987). Osiewalski and Pipień (2003) used this approach for a general GARCH(p, q) process, considering the choice of (p, q) through posterior odds. Here, again, we restrict considerations to the most important case, the GARCH(1, 1) specification.

Our Bayesian model, defined in section 2, is used in section 3 to describe daily changes of the Warsaw Stock Exchange Index (WIG) from 04.01.1995 till 28.09.2001. Section 4 focuses on option pricing treated as one of the main applications of Bayesian predictive analysis. Note that Osiewalski and Pipień (2003) also considered Bayesian option pricing; they presented results based on the WIG data from the very beginning of the Warsaw Stock Exchange (April 1991) till the end of 2000.

2. THE BAYESIAN MODEL FOR DISCRETE OBSERVATIONS

Let x_t denote the price of an asset or an exchange rate or, as in our application, the value of a stock market index (at time t). For x_t we assume an AR(2) process with asymmetric GARCH(1, 1) errors. In terms of growth or return rates (expressed in percentage points) $y_t = 100 \Delta \ln x_t = 100 \ln (x_t/x_{t-1})$, our model can be written as

$$y_t = \delta + \rho(y_{t-1} - \delta) + \delta_1 \ln x_{t-1} + \varepsilon_t, \quad t = 1, ..., T + k \qquad (1)$$

where T observations are used in estimation, k is the forecasting horizon, $\varepsilon_t = z_t\sqrt{h_t}$,

$$h_t = a_0 + a_1\varepsilon_{t-1}^2 I(\varepsilon_{t-1} < 0) + a_1^+ \varepsilon_{t-1}^2 I(\varepsilon_{t-1} \geqslant 0) + b_1 h_{t-1} \qquad (2)$$

and we treat h_0 as an additional parameter. Our specification of asymmetric reactions to positive and negative shocks, given in (2), follows Glosten, Jagannathan and Runkle (1993); it nests the basic symmetric ARCH and GARCH processes introduced by Engle (1982) and Bollerslev (1986).

We assume that z_t are independent skewed t random variables with $v > 2$ degrees of freedom, mode $\zeta \in (-\infty, +\infty)$, unit precision and asymmetry parameter $\gamma > 0$. The density is:

$$f_{skS}(z|v, \zeta, 1, \gamma) = \frac{2\Gamma((v+1)/2)}{(\gamma + \gamma^{-1})\Gamma(v/2)\sqrt{\pi v}} \cdot$$

$$\cdot [1 + (z-\zeta)^2 v^{-1}\{\gamma^2 I_{(-\infty, 0)}(z-\zeta) + \gamma^{-2} I_{(0, +\infty)}(z-\zeta)\}]^{-(v+1)/2} \qquad (3)$$

Note that $\gamma = 1$ corresponds to the usual symmetric Student t distribution. In general, γ^2 measures the degree of distributional asymmetry, as it is the ratio of the probability masses on the right and on the left side of the mode of z_t:

$$\gamma^2 = \frac{\Pr(z_t \geqslant \zeta)}{\Pr(z_t < \zeta)}.$$

The moments of z_t and the conditional moments of ε_t (given the past of the process, ψ_{t-1}, and the parameters) take the form:

$$E(z_t|v, \zeta, \gamma) = \zeta + \frac{(\gamma^2 - \gamma^{-2})2v\Gamma((v+1)/2)}{(\gamma + \gamma^{-1})(v-1)\Gamma(v/2)\sqrt{\pi v}} = \zeta + \tau(\gamma, v)$$

$$E(z_t^2|v, \zeta, \gamma) = \frac{(\gamma^3 - \gamma^{-3})v}{(\gamma + \gamma^{-1})(v-2)} + 2\zeta\tau(\gamma, v) + \zeta^2 \qquad (4)$$

$$E(\varepsilon_t|\psi_{t-1}, v, \zeta, \gamma) = \sqrt{h_t}E(z_t) = \sqrt{h_t}(\zeta + \tau(\gamma, v))$$

$$Var(\varepsilon_t|\psi_{t-1}, v, \zeta, \gamma) = \sigma_t^2 = h_t\left(\frac{(\gamma^3 + \gamma^{-3})v}{(\gamma + \gamma^{-1})(v-2)} - \tau^2(\gamma, v)\right) = h_t d(\gamma, v)$$

see Fernàndez and Steel (1998) and Osiewalski and Pipień (1999, 2000, 2003). Remark that $\tau(\gamma, v) = 0$ iff $\gamma = 1$, and τ measures the effect of distributional asymmetry on the mean of z_t.

The conditional distribution of y_t (given the past of the process, ψ_{t-1}, and the parameters) is skewed Student t with $v > 2$ degrees of freedom, asymmetry parameter $\gamma > 0$, mode $\mu_t = \delta + \rho(y_{t-1} - \delta) + \delta_1 \ln x_{t-1} + \zeta\sqrt{h_t}$, and precision h_t^{-1}, where inverse precision h_t is defined in (2). The density function is

$$p(y_t|\psi_{t-1}, \theta) = f_{sks}(y_t|v, \mu_t, h_t^{-1}, \gamma) = \frac{2\Gamma((v+1)/2)}{(\gamma + \gamma^{-1})\Gamma(v/2)\sqrt{\pi v}} h_t^{-1/2} \cdot \tag{5}$$

$$\cdot [1 + (y_t - \mu_t)^2 (v h_t)^{-1}\{\gamma^2 I_{(-\infty, 0)}(y_t - \mu_t) + \gamma^{-2} I_{[0, +\infty)}(y_t - \mu_t)\}]^{-(v+1)/2}$$

where $\theta = (\zeta, v, \delta, \delta_1, \rho, \gamma, h_0, a_0, a_1, a_1^+, b_1)'$ groups all the parameters. The sampling model is represented by the following joint density function of T observed and k forecasted values:

$$p(y, y_f|\theta) = \prod_{t=1}^{T+k} p(y_t|\psi_{t-1}, \theta) = \prod_{t=1}^{T+k} f_{sks}(y_t|v, \mu_t, h_t^{-1}, \gamma).$$

This density conditions on some initial observations, which are not shown in our notation.

Our assumptions lead to a GARCH-M representation of (1), namely:

$$y_t = \delta + \rho(y_{t-1} - \delta) + \delta_1 \ln x_{t-1} + \varphi\sqrt{h_t} + u_t, \quad E(u_t) = 0 \tag{6}$$

where $u_t = \varepsilon_t - E(\varepsilon_t|\psi_{t-1}, v, \zeta, \gamma) = \varepsilon_t - \varphi\sqrt{h_t}$ and the GARCH-M effect is measured by the following function of basic parameters:

$$\varphi = E(z_t|v, \zeta, \gamma) = \zeta + \tau(\gamma, v).$$

Note that φ can be zero (no GARCH-M effect) iff $\zeta = -\tau(\gamma, v)$, i.e. when the free location parameter ζ compensates the effect of asymmetry on the mean of the process (as measured by τ). It should be stressed that the stochastic specification in (5) enables modelling statistical sources of the GARCH-M effect (distributional asymmetry, non-zero location ζ) and making formal inference both on the total magnitude of this effect and on its components. Formal testing of this effect amounts to testing $\varphi = 0$, a complicated non-linear restriction on ζ, v and γ. Our GARCH-M parameterisation differs from the usual one. If we rewrite (6) in terms of the conditional sampling standard deviation $\sigma_t = \sqrt{h_t} d(\gamma, v)$, we obtain:

$$y_t = \delta + \rho(y_{t-1} - \delta) + \delta_1 \ln x_{t-1} + \lambda\sigma_t + u_t \tag{7}$$

where $\lambda = \lambda(\zeta, \gamma, v) = \varphi/\sqrt{d(\gamma, v)}$ and (7) is closer to the original parameterisation of the GARCH-M effect as in Engle, Lilien and Robins (1987); see also Bollerslev, Chou and Kroner (1992) and Shepard (1995). Of course, (6) is more general than (7) since the former does not require the existence of conditional variance (it needs only $v > 1$). However, assuming $v > 2$ we follow the most popular way of modelling risk through the (conditional) standard deviation. Note that in our formulation ε_{t-j}, and not $u_{t-j} = (\varepsilon_{t-j} - \lambda\sigma_{t-j})$, enter the GARCH equation (2).

Note that the AR(2) model, as in Bauwens and Lubrano (1998) and Bauwens, Lubrano and Richard (1999), enables to make inference on the presence of a unit root in $\ln x_t$. If $\delta_1 = 0$ then (1) reduces to an AR(1) process for $y_t = 100\Delta \ln x_t$, i.e. a unit root process for $\ln x_t$.

Our Bayesian model is defined through the joint density of all the observables and unknown parameters:

$$p(y, y_f, \theta) = p(y, y_f | \theta)p(\theta),$$

where $p(\theta)$ is the marginal (or prior) density of the parameters. The prior density in our AR(2)-GARCH(1, 1) model is:

$$p(\theta) = p(\zeta)p(v)p(\delta)p(\delta_1)p(\rho)p,(\gamma)p(h_0)p(a_0)p(a_1)p(a_1^+)p(b_1),$$

where $p(\zeta)$ and $p(\delta)$ are standard normal, $p(\delta_1)$ is normal with mean 0 and standard deviation 0.1, $p(\rho)$ is uniform on the interval (-1, 1), $p(\gamma)$ is log – standard normal (truncated at 0.5 and 2), $p(v)$ is exponential with mean 10, truncated at 2 as $v > 2$, $p(h_0)$ and $p(a_0)$ are exponential with mean 1, and $p(a_1)$, $p(a_1^+)$, $p(b_1)$ are uniform on the unit interval. These assumptions reflect rather weak prior knowledge about the parameters. In the case when $\delta_1 = 0$, $\gamma = 1$ and $\zeta = 0$, δ is the expected (systematic) daily growth or return rate (in percentage points); its most likely value is zero and the standardised normal distribution is spread enough to represent little prior knowledge. The normal prior for δ_1 is located around $\delta_1 = 0$, reflecting prior beliefs in the unit root model for $\ln x_t$. However, other processes are not ruled out *a priori*. The uniform density for ρ refers only to the usual restriction $\rho \in (-1, 1)$. Our prior distribution of γ assumes that not more than 80% of the sampling probability mass can be located on one side of the mode and that symmetry is the most likely situation. The priors for the remaining parameters are very flat. In particular, we expect low values of v, but situations close to conditional normality ($v > 30$) are still possible. Our joint prior distribution is proper, thus leading to the well-defined posterior distribution. Note that the use of the improper uniform prior for the degrees of freedom parameter would preclude the existence of the posterior distribution; see Bauwens and Lubrano (1998).

Fig. 1. The log of Warsaw Stock Exchange Index (WIG), 04.01.1995–08.02.2002
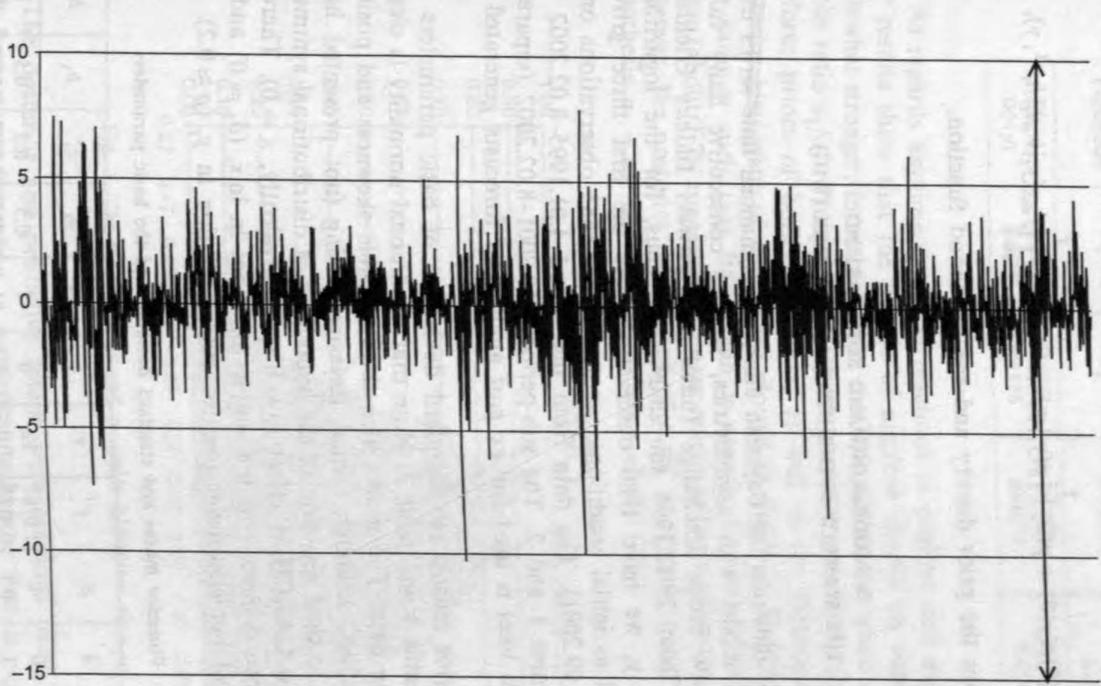
Fig. 2. Daily growth rates of WIG, 04.01.1995–08.02.2002

The prior distribution for our basic parameters induces prior probability distributions for all their measurable functions. In particular, Osiewalski and Pipień (2003) present histograms of the implied prior distributions for τ and φ.

The posterior distribution of θ has a density of the form

$$p(\theta|y) \propto p(\theta)p(y|\theta) = p(\theta)\prod_{t=1}^{T} p(y_t|\psi_{t-1},\theta) = p(\theta)\prod_{t=1}^{T} f_{sks}(y_t|v,\mu_t,h_t^{-1},\gamma),$$

which combines the prior density and the likelihood function.

3. MODELLING AND FORECASTING
THE WARSAW STOCK EXCHANGE INDEX (WIG)

In order to illustrate the Bayesian analysis of financial time series using our GARCH model with asymmetries, we use consecutive daily values of the Warsaw Stock Exchange Index (Warszawski Indeks Giełdowy – WIG, x_t) from 29.12.1994 till 28.09.2001. Thus, for the logarithmic growth rates y_t we have 1686 observations. As the first three growth rates are used as initial conditions, we model $T = 1683$ observations on y_t (4.01.1995–28.09.2001). The data from the period 4.01.1995–8.02.2002 are plotted in figures 1 and 2. The sub-period 1.10.2001–8.02.2002 (separated by the vertical line) is used for *ex post* analysis of forecasts generated by our model.

The posterior means and standard deviations of basic parameters are presented in table 1 and table 2. Note that conditional normality is clearly rejected by the data. The results also show possible skewness and positive GARCH-M effect, although formal Bayesian testing (not presented here) does not lead to clear rejection of the hypothesis of distributional symmetry ($\gamma = 1$) and no GARCH-M effect ($\varphi = 0$ or, equivalently, $\lambda = 0$). There is enough evidence in favour of the simple unit root in $\ln x_t$ ($\delta_1 = 0$) and in favour of small but significant positive autocorrelation in y_t ($\rho \approx 0.2$).

Table 1. Posterior means and standard deviations of the basic parameters

Para-meter	ρ	δ	δ_1	γ	v	ζ	a_0	a_1	a_1^+	b_1	b_0
$E(\cdot)$	0	0	0	1.0767	12	0	1	0.5	0.5	0.5	1
$D(\cdot)$	0.5773	1	0.1	0.4115	10	1	1	0.2828	0.2828	0.2828	1
$E(\cdot|y)$	0.1842	-0.1657	-0.0137	0.9523	8.4123	0.2074	0.1857	0.1772	0.1038	0.7541	1.3911
$D(\cdot|y)$	0.0259	0.3088	0.0957	0.0325	1.6915	0.1176	0.0640	0.0426	0.0243	0.0489	1.2058

Table 2. Posterior means and standard deviations of GARCH-in-Mean parameters

Parameter	τ	ζ	$\varphi = \tau + \zeta$	$\lambda = \dfrac{\varphi}{\sqrt{d(y, v)}}$
$E(\cdot\|y)$	-0.0875	0.2074	0.1199	0.1037
$D(\cdot\|y)$	0.0607	0.1176	0.1014	0.0876

As regards asymmetry of the reactions to positive and negative shocks, our results show that the impact of negative shocks on volatility may be somewhat stronger; formal inference can be based on the posterior distribution of the ratio a_1/a_1^+. The prior density of this ratio, induced by the independent uniform priors of both parameters, as well as the (appropriately scaled) posterior histogram of a_1/a_1^+, are plotted in figure 3. The prior density of this ratio is such that its median is exactly 1. The posterior median is larger (about 1.7), suggesting stronger reactions to negative shocks ($a_1 > a_1^+$). However, the posterior distribution of a_1/a_1^+ is spread enough to make the hypothesis $a_1 = a_1^+$ not unlikely in view of the data.
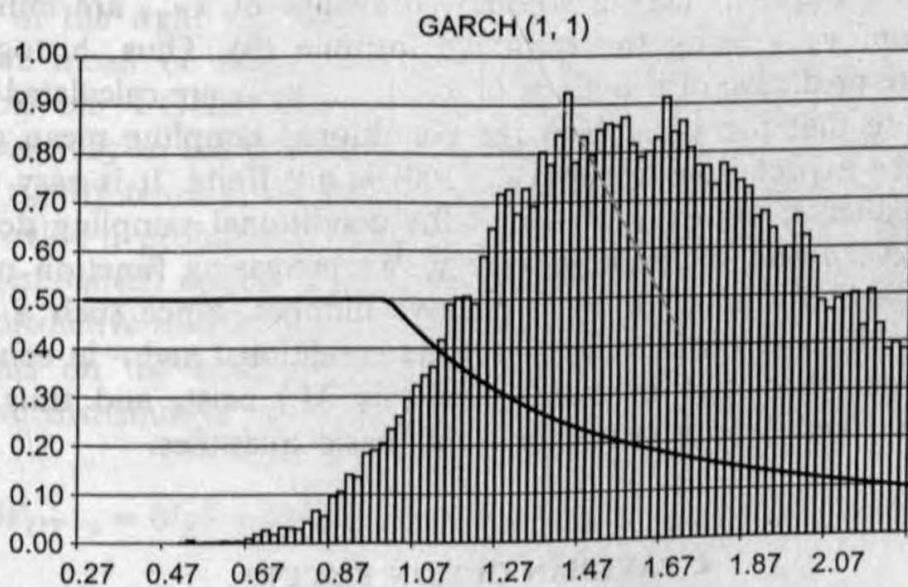

Fig. 3. The prior density and posterior histogram of a_1/a_1^+

The posterior and predictive results were obtained using Monte Carlo with Importance Sampling, a numerical technique that is very stable and efficient in our application, where the parameter space is 11-dimensional. The out-of-sample predictive density of k future return rates is obtained through averaging the sampling predictive density over the parameter space, with the use of the posterior density as the weight function:

$$p(y_f|y) = \int_\Theta p(y_f|y, \theta)p(\theta|y)d\theta \tag{8}$$

where

$$p(y_f|y, \theta) = \prod_{t=T+1}^{T+k} p(y_t|\psi_{t-1}, \theta) = \prod_{t=T+1}^{T+k} f_{skS}(y_t|v, \mu_t, h_t^{-1}, \gamma).$$

The numerical approximation we base our results on requires simulating future values of the series from their sampling predictive density. For a given parameter vector, drawn according to the importance function $r(\theta)$, a multivariate Student t density with 3 degrees of freedom, we recursively generate y_{T+j}, given y_{T+j-1} ($j = 1, ..., k$) from the skewed t distribution (5). This is done by drawing a value $\dot{y}_{T+j}$ from the appropriate symmetric Student t distribution and re-scaling it on the basis of its sign ($y_{T+j} = \gamma\dot{y}_{T+j}$ if $\dot{y}_{T+j} \geqslant 0$ and $y_{T+j} = \gamma^{-1}\dot{y}_{T+j}$ otherwise). It is also of interest to forecast future levels of the series, i.e.

$$x_{T+j} = x_{T+j-1}\exp\{y_{T+j}/100\}, \quad \text{for } j = 1, ..., k \tag{9}$$

In our Monte Carlo numerical strategy, drawings of x_{T+j} are immediately obtained from y_{T+j} using the recursive formula (9). Thus, histograms of the univariate predictive distributions of $x_{T+1}, ..., x_{T+k}$ are calculated as a by-product. Note that for x_{T+j}, even the conditional sampling mean does not exist since the expectation of $\exp(y_{T+j}/100)$ is not finite. It is easy to prove that the product of $\exp(n \cdot y_t/100)$ and the conditional sampling density (5) is, for positive n and sufficiently large y_t, an increasing function of y_t and can be bounded from below by a positive number. Since such a function is not integrable in any interval $(d, +\infty)$, no conditional and – in consequence – no marginal predictive moments $E(x_{T+j}^n|y, M_i)$ exist, and thus we can only compute and report histograms and basic quantiles.

4. BAYESIAN OPTION PRICING

4.1. Direct Approach Using the Predictive Distribution of Discounted Payoff

One of the main applications of the Bayesian predictive analysis is option pricing as discussed by Bauwens and Lubrano (1997). Assume that the analysed stock exchange index is a tradable security and its price at time t is equal to its numerical value x_t. Consider a hypothetical European

call option evaluated at time T (the last period in the observed series), s units of time before maturity. The payoff function is $(x_{T+s} - K)^+ = \max\{x_{T+s} - K, 0\}$, where x_{T+s} is the price of the underlying security at maturity (no dividend being paid) and K is the strike (exercise price). The discounted (present value) payoff considered at time T is

$$W_{T|T+s} = e^{-rs}(x_{T+s} - K)^+,$$

where r is the interest rate (assumed known for the sake of simplicity). Of course, this discounted payoff is a random variable as a measurable function of x_{T+s}, which is random. Direct option pricing would use the assumed statistical model for x_t (and the data) to construct the predictive distribution of $W_{T|T+s}$ and would not refer to any other theoretical specifications or tools (like the famous Black-Scholes formula, based on different stochastic assumptions).

While the predictive distribution of $W_{T|T+s}$ is censored at zero (and thus has a probability mass at this point), its right tail is essentially the same as the right tail of the underlying distribution of x_{T+s}. As Osiewalski and Pipień (2003) already discussed, in the case of t-GARCH models the conditional sampling mean $E(x_{T+s}|\psi_{T+s-1}, \theta)$ is infinite due to non-integrability in the right tail. Thus, the corresponding sampling mean and the predictive mean of $W_{T|T+s}$ are also infinite. Hence, direct option pricing based on our t-GARCH(1, 1) model cannot use this infinite predictive mean, contrary to the proposal by Bauwens and Lubrano (1997).[1] However, as they rightly remark, the main advantage of the Bayesian direct approach is to provide a probability distribution with respect to which any observed (or contemplated) option price can be assessed. This benchmark is provided by the predictive distribution of the discounted payoff. The latter distribution conditions on the observed data and the assumed statistical model. The predictive distribution of $W_{T|T+s}$ consists of a point mass at zero:

$$\Pr\{W_{T|T+s} = 0|y\} = \Pr\{x_{T+s} \leqslant K|y\} = \Pr\{0.01 \sum_{j=1}^{s} y_{T+1} \leqslant \ln(K/x_T)|y\}$$

(10)

and a continuous part for $W_{T|T+s} \in (0, +\infty)$, defined by the density function $p(W_{T|T+s}|y)$ that is obtained from the truncated density $p(x_{T+s}|y)I(x_{T+s} > K)$. As Bauwens and Lubrano (1997) write, the predictive distribution of

---

[1] Bauwens and Lubrano's (1997) proposal to take the predictive mean of the discounted payoff can be useful in the context of stochastic volatility (SV) models (see Pajor, 2003), where normality of the conditional distribution of return rates is often acceptable – in contrast to GARCH models, which usually need heavier tails.

$W_{T|T+s}$ can be used by market participants who wish to compare the model predictions to potential prices on the market or to other predictions. Also, a t-GARCH specification can be used to provide an "objective" option price $C_{T|T+s}$. As the predictive mean of $W_{T|T+s}$ is infinite, and thus can hardly be considered a reasonable price, one can consider the predictive median defined by the conditions

$$\Pr\{W_{T|T+s} \geqslant C_{T|T+s}|y\} \geqslant 0.5 \quad \text{and} \quad \Pr\{W_{T|T+s} \leqslant C_{T|T+s}|y\} \geqslant 0.5.$$

Of course, this suggestion leads to $C_{T|T+s} = 0$ when the point mass at zero, given by (10), is at least 0.5. It means that the "objective" price of an option, which with probability at least 0.5 would not be exercised ($x_{T+s} \leqslant K$), is zero (except possibly for some fixed operation cost that is not considered at this level of reasoning). However, if the probability that the option will be exercised ($x_{T+s} > K$) is greater than 0.5, the option is priced at a uniquely defined level $C_{T|T+s} > 0$.

In our empirical illustration we set the strike at the last observed value of the stock index ($K = x_T = 11816.43$). We set the interest rate at $r = 0.14/360$ (14% on annual base), which is roughly the average interest rate on 3 month deposits, paid by Polish commercial banks in the first half of 2001. At $t = T$ we consider hypothetical European call options of different maturity (s): 30, 60 or 90 working days. Table 3 shows the predictive probability that the option will not be exercised (i.e., the point mass at 0) and the quartiles: $Q_1, Q_2 = C_{T|T+s}$ (median) and Q_3 of the predictive distribution of the discounted payoff $W_{T|T+s}$.

Table 3. The point mass at 0 and quartiles of the predictive distribution of $W_{T|T+s}$

Time to maturity (s)	30	60	90
$\Pr\{W_{T\backslash T+s} = 0|y\} = \Pr\{x_{T+s} \leqslant K|y\}$	0.500	0.501	0.497
$Q_1(W_{T\backslash T+s}|y)$	0.0	0.0	0.0
$Q_2(W_{T\backslash T+s}|y)$	0.0	0.0	0.0
$Q_3(W_{T\backslash T+s}|y)$	1105.3	2139.6	3227.3

Figure 4 shows histograms of the marginal predictive distributions of the value of WIG at time $T + s$, x_{T+s}, for $s = 30$, $s = 60$, and $s = 90$, as well as the conditional distributions of $W_{T|T+s}$ given that the option is exercised ($W_{T|T+s} > 0$, i.e. $x_{T+s} > K$); remind that the latter distribution is continuous. Since the posterior (or predictive) probability that the option will be exercised is approximately 0.5 for all three values of s, the predictive probability that $W_{T|T+s} > 0$ is also 0.5 and hence the predictive median of

$W_{T|T+s}$ is 0^2. Thus the vertical line in the right panels of figure 4 represents only the median of the continuous part of the distribution of $W_{T|T+s}$. This median increases quickly with s due to the heavier and heavier right tail and increased spread of the predictive distribution.

Note that the assumed exercise price ($K = 11816.43$), represented in the left panels of figure 4 by the light square, is always lower than the true value of WIG at time $T + s$, indicated by the dark square. Thus our hypothetical options would have been exercised. It is important to stress that the true values of x_{T+s} ($s = 30, 60, 90$) lie in the areas of high predictive density, indicating good forecasting properties of our GARCH specification even for relatively distant future periods. However, this is achieved at the cost of huge *ex ante* uncertainty.

4.2. Using the Black-Scholes Formula

Although the famous Black-Scholes formula for option pricing relies on very different assumptions than GARCH models for discrete time-series data, it is sometimes used with the volatility predicted on the basis of a GARCH specification for the observed return series. The Black-Scholes (BS) formula

$$BS_{T|T+s} = x_T N(d_1) - K e^{-rs} N(d_2), \tag{11}$$

where $N(.)$ is the cumulative distribution function of the standard normal distribution and

$$d_1 = \frac{\ln(x_T/K) + (r + 0.5\sigma^2) \cdot s}{\sigma\sqrt{s}}, \quad d_2 = d_1 - \sigma\sqrt{s},$$

can be viewed as a known function of the unknown volatility parameter σ. In the GARCH framework, σ can be interpreted as the conditional standard deviation of $0.01 y_{T+s}$, the return rate at time $T + s$. Hence σ is given by the equation

$$\sigma = 0.01\sqrt{\text{Var}(y_{T+s}|\psi_{T+s-1}, \theta)} = 0.01\sqrt{h_{T+s} d(\gamma, v)}.$$

The marginal posterior (or predictive) distributions of this quantity (with $s = 30, 60, 90$) are presented in figure 5. The predictive character of $p(\sigma|y)$ is due to the dependence of σ (for $s > 1$) on future (unobserved at time T)

---

[2] This is very different from the results presented by Osiewalski and Pipień (2003), obtained for different periods. Their option prices were positive.

$p(x_{T+30} \mid y)$

$p(W_{T|T+30} \mid y, x_{T+30} > K)$

$p(x_{T+60} \mid y)$

$p(W_{T/T+60} \mid y, x_{T+60} > K)$
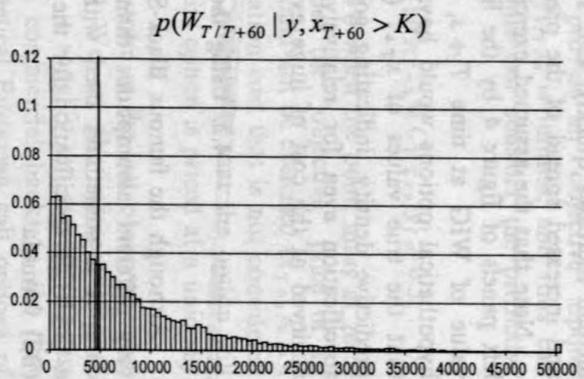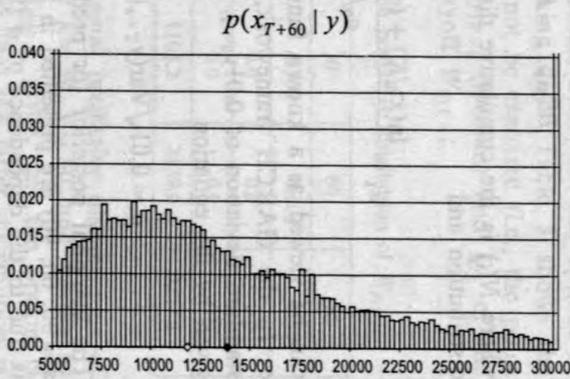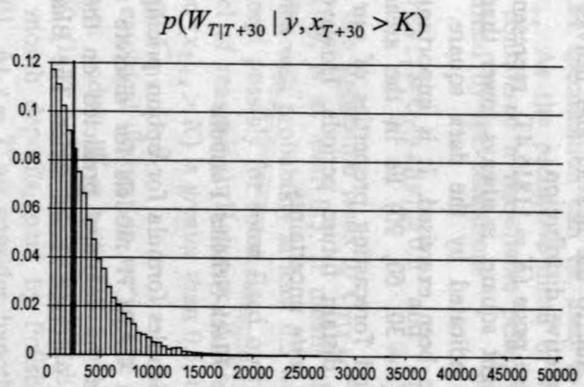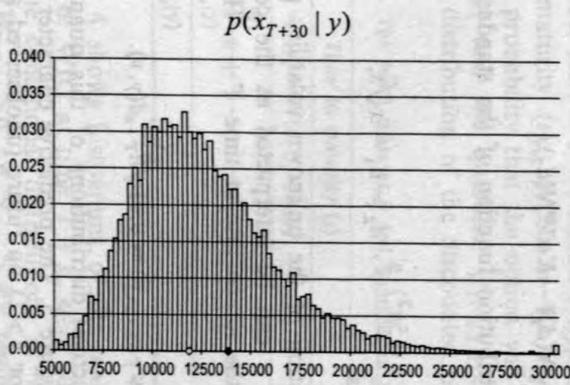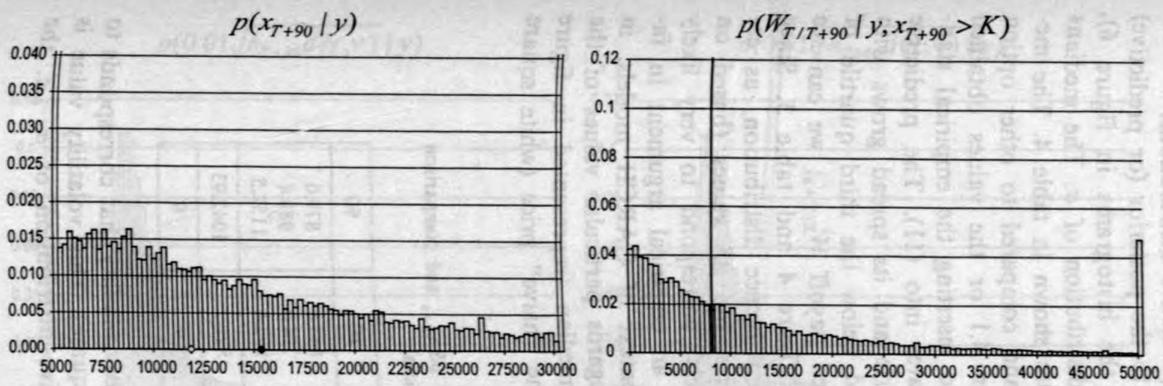
Fig. 4. Predictive histograms of x_{T+s} and the continuous part of $W_{T|T+s}$ ($s = 30$, 60, 90)

values of $y_{T+1}, ..., y_{T+s-1}$ through the future error terms $\varepsilon_{T+1}, ..., \varepsilon_{T+s-1}$; see (1)–(2). In our Monte Carlo computations future growth (or return) rates are simulated using their conditional predictive distributions.

The Bayesian approach naturally leads to the posterior (or predictive) distribution of $BS_{T|T+s}$ (represented by the histograms in figure 6), a continuous distribution induced by the distribution of σ. The medians and inter-quartile ranges of $p(BS_{T|T+s}|y)$ are shown in table 4. The medians can be taken as BS option prices and compared to other option prices, like $C_{T|T+s}$ proposed in subsection 4.1 or the values obtained using a "naive" approach that amounts to inserting the empirical standard deviation of the observed return rates into (11). The predictive distribution of $BS_{T|T+s}$ is highly asymmetric and its spread grows with s. Most of its probability mass is located below the third quartile of the predictive distribution of the discounted payoff $W_{T|T+s}$; we can see that comparing figure 6 and table 4 to figure 4 and table 3. Since we treat the distribution of $W_{T|T+s}$ as a reference distribution, as explained in subsection 4.1, we conclude that the BS values (based on different assumptions than GARCH models) correspond to very likely values of the payoff function. This gives an empirical argument in favour of using the BS formula in the context of GARCH models, in spite of theoretical incompatibilities. As regards particular values of the random variable $BS_{T|T+s}$, its predictive median (represented in figure 6 by the vertical line) is quite close to the "naive" price (white square in figure 6).

Table 4. The predictive characteristics of $BS_{T|T+s}$ and comparison to other option prices

Time to maturity (s)	30	60	90
$Q_1(BS_{T|T+s}|y)$	450.7	680.8	879,0
$Q_2(BS_{T|T+s}|y)$	518.5	774.0	988.4
$Q_3(BS_{T|T+s}|y)$	611.1	902.4	1152.5
Naive pricing	553.2	821.2	1042.93
$Q_2(W_{T|T+s}|y)$	0	0	0

Finally, figure 5 also shows the BS implied volatility that corresponds to the predictive median of $BS_{T|T+s}$ (white square). This volatility value is approximately the median of the predictive distribution of σ, i.e. the distribution presented in figure 6.

$$p(0.01\sqrt{h_{T+30}}d(\gamma,\nu)\mid y)$$

$$p(0.01\sqrt{h_{T+60}}d(\gamma,\nu)\mid y)$$

$$p(0.01\sqrt{h_{T+90}}d(\gamma,\nu)\mid y)$$
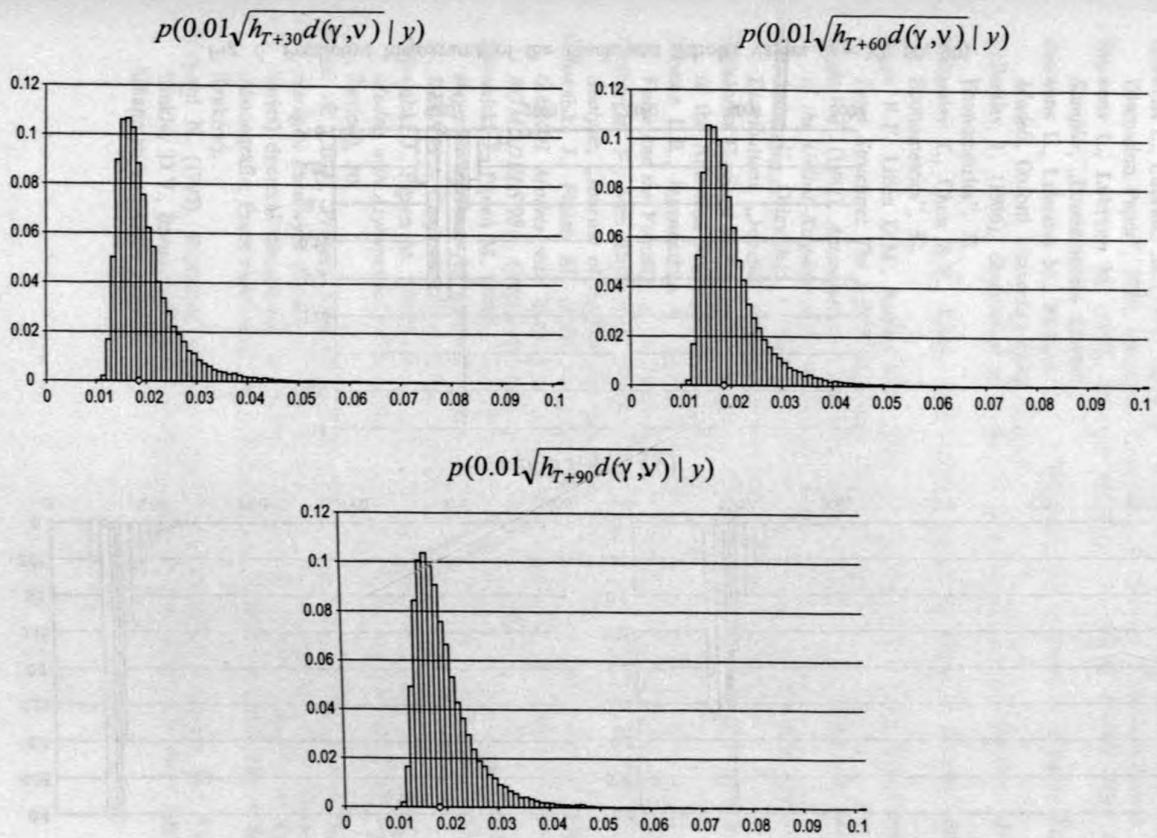
Fig. 5. Predictive histograms of the volatility parameter σ at time $T+s$ ($s = 30,\ 60,\ 90$)

$p(BS_{T|T+30} \mid y)$

$p(BS_{T|T+60} \mid y)$

$p(BS_{T|T+90} \mid y)$
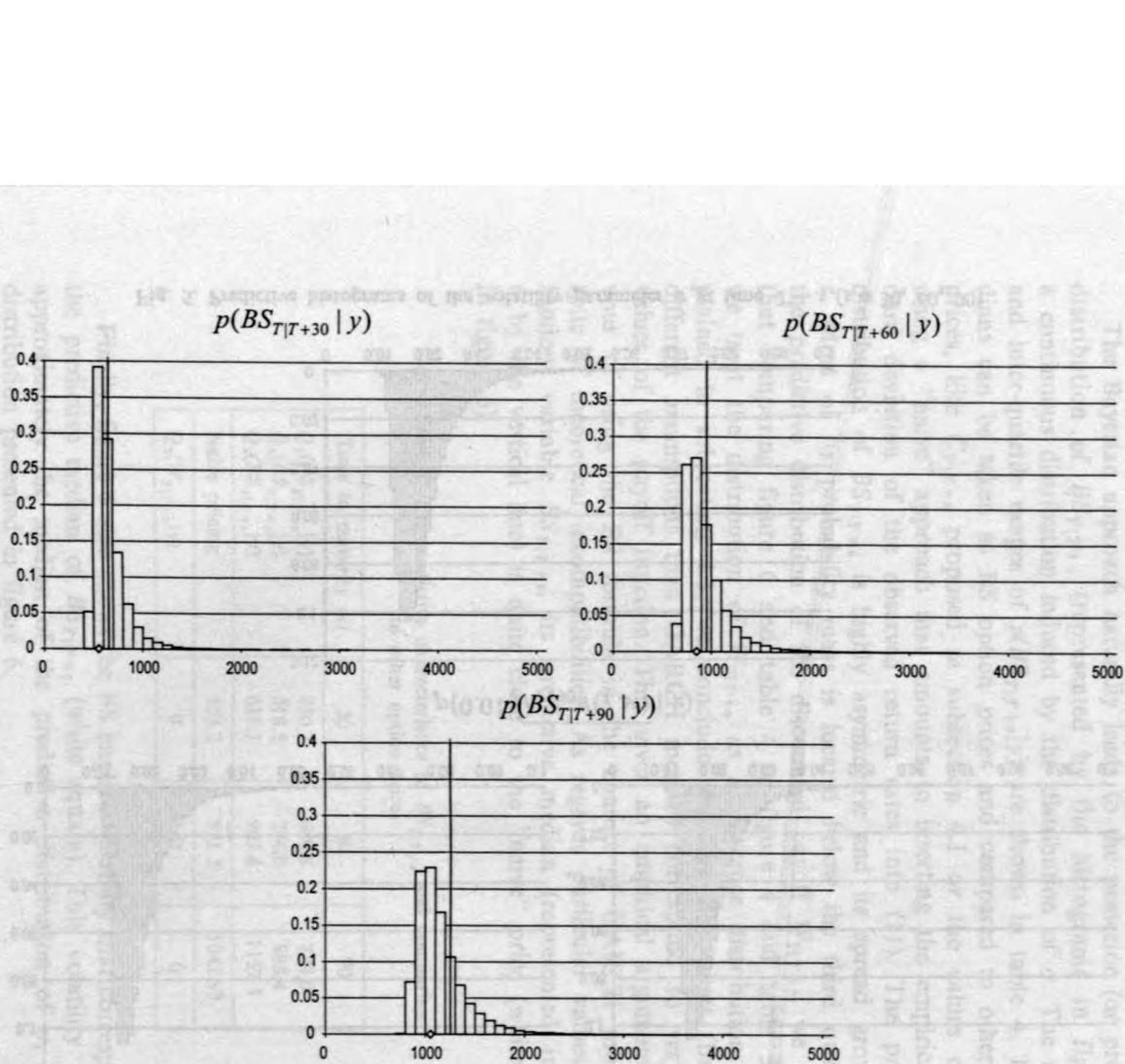
Fig. 6. Predictive histograms of the Black and Scholes values ($s = 30, 60, 90$)

REFERENCES

Bauwens L., Lubrano M. (1997), *Bayesian Option Pricing Using Asymmetric GARCH*, „CORE Discussion Paper", 9759, Université Catholique de Louvain, Louvain-la-Neuve.

Bauwens L., Lubrano M. (1998), *Bayesian Inference on GARCH Models Using the Gibbs Sampler*, „Econometrics Journal", 1.

Bauwens L., Lubrano M., Richard J.-F. (1999), *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, Oxford.

Bollerslev T. (1986), *Generalised Autoregressive Conditional Heteroscedasticity*, „Journal of Econometrics", 31.

Bollerslev T., Chou R.Y., Kroner K.F. (1992), *ARCH Modelling in Finance*, „Journal of Econometrics", 52.

Engle R.F, Lilien D.M., Robins R.P. (1987), *Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model*, „Econometrica" 55.

Engle R.F. (1982), *Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation*, „Econometrica", 50.

Fernández C., Osiewalski J., Steel M.F.J. (1995), *Modelling and Inference with v-Spherical Distributions*, „Journal of the American Statistical Association", 90.

Fernández C., Steel M.F.J. (1998), *On Bayesian Modelling of Fat Tails and Skewness*, „Journal of the American Statistical Association" 93.

Glosten L.R., Jagannathan R., Runkle D.E. (1993), *On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks*, „Journal of Finance", 48.

Kleibergen F., Van Dijk H.K. (1993), *Non-Stationarity in GARCH Models: A Bayesian Analysis*, „Journal of Applied Econometrics", 8.

Osiewalski J., Pipień M. (1999), *Bayesian Forecasting of Foreign Exchange Rates Using GARCH Models with Skewed t Conditional Distributions*, [in:] Welfe W. (ed.), *MACROMODELS'98 – Conference Proceedings*, Absolwent, Łódź.

Osiewalski J., Pipień M. (2000), *GARCH-in-Mean through Skewed t Conditional Distributions: Bayesian Inference for Exchange Rates*, [in:] Welfe W., Wdowiński P. (eds) *MACROMODELS'99 – Conference Proceedings*, Absolwent, Łódź.

Osiewalski J., Pipień M. (2003), *Bayesian Analysis and Option Pricing in Univariate GARCH Models with Asymmetries and GARCH-in-Mean Effects*, „Przegląd Statystyczny" (Statistical Review), 50.

Pajor A. (2003), *Procesy zmienności stochastycznej SV w bayesowskiej analizie finansowych szeregów czasowych* (Stochastic Volatility Processes in Bayesian Analysis of Financial Time Series) doctoral dissertation (in Polish), published by Cracow University of Economics (Monografie: Prace Doktorskie, nr 2, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków).

Shepard N. (1995), *Statistical Aspects of ARCH and Stochastic Volatility*, [in:] Cox D.R., Hinkley D.V., Barndorf-Nielsen O. (eds), *Likelihood, Time. Series and Other Applications*, Chapman and Hall, London.

Jacek Osiewalski, Mateusz Pipień

BAYESOWSKA WYCENA EUROPEJSKIEJ OPCJI KUPNA Z WYKORZYSTANIEM MODELU GARCH Z ASYMETRIAMI

Streszczenie

W prezentowanym artykule wycena opcji jest traktowana jako jedno z zastosowań bayesowskiej analizy predyktywnej. Rozkład wartości zdyskontowanej wypłaty, indukowany przez gęstość predyktywną przyszłych stóp zwrotu, jest podstawą bezpośredniej wyceny opcji (zob. Bauwens, Lubrano, 1997). Rozważamy też bardziej eklektyczne podejście, wykorzystujące rozkład predyktywny formuły Blacka i Scholesa (ze zmiennością określoną jako warunkowe odchylenie standardowe w momencie realizacji opcji).

Przyjmujemy ramy modelowe, które uwzględniają dwa rodzaje asymetrii w procesach GARCH: skośne rozkłady warunkowe (typu t-Studenta) oraz zróżnicowane reakcje wariancji warunkowej na szoki dodatnie lub ujemne. Model: skośny t-GARCH(1, 1) jest stosowany do opisu dziennej zmienności Warszawskiego Indeksu Giełdowego (WIG) od 4.01.1995 r. do 8.02.2002 r. Dane do 28.09.2001 wykorzystujemy do budowy rozkładów *a posteriori* i predyktywnego oraz do ilustracji bayesowskiej wyceny opcji na pozostały okres.