"Forecasting stochastic Volatility using the Kalman filter: An Application to Canadian Interest Rates and Price-Earnings Ratio"

Théoret, Raymond and Racicot, François-Éric

University of Quebec. Outaouais, University of Quebec. Montreal

December 2010

RESEARCH ARTICLE

AESTIMATIO, THE IEB INTERNATIONAL JOURNAL OF FINANCE, 2010. 1: 28-47

Forecasting Stochastic Volatility Using the Kalman Filter:

Racicot, François-Éric
Théoret, Raymond

▸ SUBMITTED : JANUARY 2010
▸ ACCEPTED : SEPTEMBER 2010

In this paper, we aim at forecasting the stochastic volatility of key financial market variables with the Kalman filter using stochastic models developed by Taylor (1986, 1994) and Nelson (1990). First, we compare a stochastic volatility model relying on the Kalman filter to the conditional volatility estimated with the GARCH model. We apply our models to Canadian short-term interest rates. When comparing the profile of the interest rate stochastic volatility to the conditional one, we find that the omission of a constant term in the stochastic volatility model might have a perverse effect leading to a scaling problem, a problem often overlooked in the literature. Stochastic volatility seems to be a better forecasting tool than GARCH(1,1) since it is less conditioned by autoregressive past information. Second, we filter the S&P500 price-earnings *(P/E)* ratio in order to forecast its value. To make this forecast, we postulate a rational expectations process but our method may accommodate other data generating processes. We find that our forecast is close to a GARCH(1,1) profile.

Stochastic volatility; Kalman filter; P/E ratio forecast; Interest rate forecast.

C13; C19; C49; G12; G31.

Racicot, F. E. ✉ Professor, University of Quebec - Outaouais, UQO, 101 St-Jean-Bosco street, Lucien Brault Building, Gatineau (Hull), Quebec, Canada, J8X 3X7. E-mail : francoiseric.racicot@uqo.ca, Tel. number: (819) 595-3900 ext. 1727.

Théoret, R. Professor, University of Quebec - Montreal, UQAM, 315 Ste-Catherine east, Montreal, Quebec, Canada. H3X 2X2. Professeur associé, UQO. E-mail: theoret.raymond@uqam.ca. Tel. number: (514) 987-3000 ext.4417.

Kalman filter is increasingly used in financial applications (Racicot and Théoret, 2006, 2007a; Andersen and Benzoni, 2010; Racicot and Théoret, 2009, 2010). In this paper, we show how to combine Kalman filter and stochastic models to forecast two key financial variables: stochastic volatility and price/earnings (*P/E* ratio).

In their seminal paper published in 1973, Black and Scholes assume that stock price volatility, which is the underlying security volatility of a call option, is constant. They thus rely on unconditional volatility to formulate their equation. As usually done at this time, they thus choose the stock return standard deviation as an empirical measure of volatility[1]. But thereafter, researchers found that the return volatility was not constant but conditional to the information set available at the computation time. However, there are many ways to compute and forecast conditional volatility. In this paper, Canadian Treasury bills monthly yield and the S&P/TSX return volatilities are estimated using the Kalman filter. In order to show the flexibility of this filtering method, we also use it to forecast the S&P500 *P/E* ratio. This article focuses on stochastic volatility[2], which we compare to the standard GARCH(1,1). We find that these two measures provide similar results but that there may be some differences in a short-term horizon. We also show that the empirical specification of the stochastic volatility is very important and that the omission of some parameters, as often done in theoretical models, may give raise to biased results.

This paper is organized as follows. In section 2, we present the Kalman filter procedure in details. Section 3 provides the forecasting method of stochastic volatility. Section 4 presents the *P/E* forecasting application before concluding in section 5.

The Kalman filter is increasingly used in financial applications. In their famous equation, Black and Scholes (1973) assume that the volatility of the call underlying stock is constant. They thus use the concept of unconditional volatility to formulate their equation. The historical standard deviation of stock returns was then the usual method to measure the volatility empirically.

But thereafter researchers realized that the variance of returns was not constant but conditional to the sample of information available at the moment of its computation. However, there are several methods to make this computation or to forecast conditional

[1] *Nowadays, the practitioner use what is now called the practitioner Black and Scholes (PBS) model, which uses the fitted implied volatility as the stock return standard deviation resulting from a linear regression of the computed implied volatility on a polynomial of the strike price and maturity. For more details, see Rouah and Vainberg (2007) and Racicot and Théoret (2001, 2006, 2007b, 2010).*

[2] *On the concept of stochastic volatility, see Taylor (1986, 1994), Nelson (1990) and Mills (1999).*

volatility. In the next sections, we give an application of the Kalman filter for estimating a stochastic volatility model which we compare to the GARCH(1,1) model.

We observe that these two measures provide long-term similar results but these results might differ in a short horizon. We show that the choice of the stochastic volatility model specification is very important and that the omission of some parameters, as often done in theoretical models, could lead to spurious results. As an application of the procedure of the Kalman filter, we estimate the stochastic volatility of the monthly returns of Canadian Treasury bills and of the daily return on the *SP/TSX* Canadian index. Our simulations are initiated over a monthly sample running from 1941 to 2005, more than sufficient to proceed in our objective which is to compare two modelling techniques of the stochastic and GARCH volatilities. We did not aim here at describing the effect of the 2007-2009 subprime crisis. That is why we selected this period of time for running our simulations.

2.1. An introduction to the Kalman filter procedure[3]

Assume an observable time series $\boldsymbol{y}_t$ represented as a vector $(y_0, y_1, \ldots, y_n)$. This variable may be for instance a financial asset return. It depends on the variable h_t which is unobservable or latent. This variable could be the stochastic volatility of y_t. Since we cannot observe h_t, we have to simulate it. The variance of h_t, denoted by ω_t, is also unobservable. The model can be represented as follows:

$$y_t = \theta_1 + \theta_2 h_t + \varepsilon_t \quad (1)$$

$$h_{t+1} = \theta_3 + \theta_4 h_t + \eta_t \quad (2)$$

where θ_i are the parameters to estimate, ε_t stands for a Gaussian noise whose variance is ν_{1t} and η_t is a Gaussian noise with variance ν_{2t}. Equation (1) is the measurement or observation equation whereas equation (2) is the state or transition equation.

Let us now consider the case of time-variable coefficients. At time $(t\text{-}1)$, estimations of h_{t-1} and of its variance ω_{t-1} as well as coefficients $\theta_{i,t-1}$ are predetermined. At time 0, we must have a preliminary estimation of h_0 and of ω_0. But because these values are unknown, the software EViews, used in this study, put a zero value to h_0 and a high value to ω_0 in order to account for the uncertainty related to the estimation of h_0.

Let us set back to time $(t\text{-}1)$ of the simulation or of the filtering and give the three steps of the procedure followed by the Kalman filter: forecasting, updating and parameter estimation.

[3] *This section is based on James and Webber (2000). A classical reference on Kalman filter is Harvey (1989). For an introduction of the Kalman filter in finance, see Wang (2003), Racicot and Théoret (2006, 2007a, 2008, 2010) and Gregoriou (2009).*

In the first step, we make the following two forecasts: $h_{t|t-1}$, that is the forecast of h_t conditional to the information set at time (t-1), i.e. the expectation of h_t conditional to the available information at time (t-1); $\omega_{t|t-1}$, that is the forecast ω_t conditional to the information set at time (t-1), i.e. the expectation of ω_t conditional to the available information at time (t-1). These forecasts, which are unbiased conditional estimations, are computed as follows:

$$h_{t|t-1} = \theta_{3,\,t-1} + \theta_{4,\,t-1} h_{t-1} \tag{3}$$

$$\omega_{t|t-1} = \theta^2_{4,\,t-1} \omega_{t-1} + \nu_{2,\,t-1} \tag{4}$$

The second step is the updating one. At time t, we have a new observation of y, i.e. y_t. We can thus compute the prediction error υ_t:

$$\upsilon_t = y_t - \theta_{1,\,t-1} - \theta_{2,\,t-1} h_{t|t-1} \tag{5}$$

The variance of υ_t, denoted by ψ_t, is given by:

$$\psi_t = \theta^2_{2,\,t-1} \omega_{t|t-1} + \nu_{1,\,t-1} \tag{6}$$

We use υ_t and ψ_t to update h_t and its variance ω_t as follows

$$h_t = h_{t|t-1} + \frac{\theta_{2,\,t-1} \omega_{t|t-1}\ \upsilon_t}{\psi_t} \tag{7}$$

$$\omega_t = \omega_{t|t-1} + \frac{\theta^2_{2,\,t-1} \omega^2_{t|t-1}}{\psi_t} \tag{8}$$

Equation (7) and (8) are conditionally unbiased and efficient estimators. The Kalman filter is thus optimal because it is the best estimator in the class of linear estimators[4].

The third step deals with parameter estimation. To estimate the parameter θ_i, we use the maximum likelihood method. The log-likelihood function can be written as follows:

$$\ell = -\frac{1}{2} \sum_t log(\psi_t) - \frac{1}{2} \sum_t \frac{\upsilon_t^2}{\psi_t} \tag{9}$$

To complete the procedure, we go to time (t+1) and repeat the three-step procedure up to period n.

[4] *Note that the Kalman filter is not restricted to linear processes.*

3.1. The model

Assume the following differential equation for the logarithm of the stock price (P):

$$d(log(P))=\frac{dP}{P}=\mu dt+\sigma(t)dz_{1t} \tag{10}$$

Its discretization is the following product process:

$$x_t=\mu+\sigma_t U_t \tag{11}$$

where $x_t=\Delta\log(P_t)$ and U_t is a standardized variable[6] such as: $E(U_t)=0$ and $V(U_t)=1$.

The conditional variance of x_t, is equal to:

$$V(x_t|\sigma_t)=V(\mu+\sigma_t U_t)=\sigma_t^2 \tag{12}$$

σ_t is thus the conditional standard deviation of x_t.

The distribution of the conditional volatility σ_t must be specified. According to Mills (1999), a lognormal distribution seems appropriate, i.e.

$$h_t=log(\sigma_t^2)=\gamma_0+\gamma_1 h_{t-1}+\xi_t \tag{13}$$

where $\xi_t \sim N(0, \sigma_\xi^2)$. We can thus rewrite the equation (13) for x_t as follows:

$$x_t=\mu+U_t e^{\frac{h_t}{2}} \tag{14}$$

Mills (1999) assumes that μ is equal to 0 because the daily mean return and the intra-day returns of stocks and currencies is zero on average. In order to linearize equation (14), we square x_t and then we transform it in a logarithmic form. We thus obtain:

$$x_t^2=U_t^2 e^{h_t} \tag{15}$$

$$log(x_t^2)=log(U_t^2)+h_t \tag{16}$$

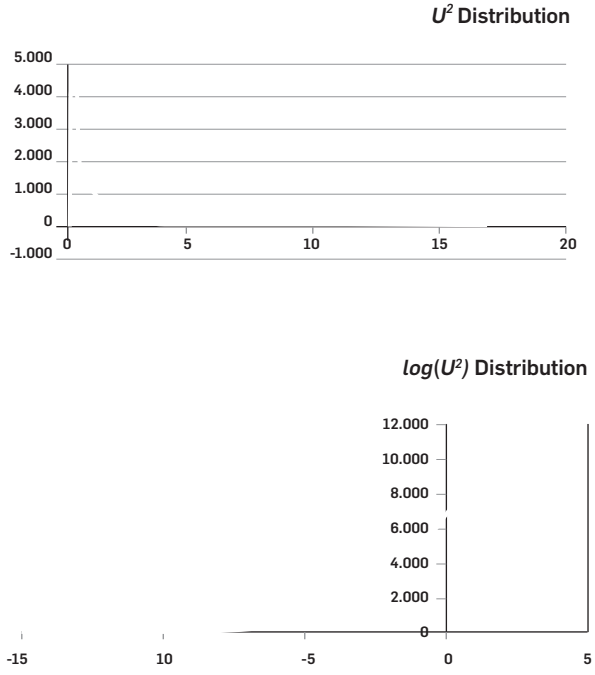
We can elaborate further on the last result since we know that $U_t \sim N(0,1)$. The distribution of $log(U_t^2)$ is therefore known, which is a logarithmic χ^2 distribution with an expectation of -1.27 and a variance of $0.5\pi^2$, which is approximately 4.93.

[5] *This section is based on Mills (1999). For a good reference on the existing models of volatility, see Andersen et al. (2005).*

[6] *Note that* $U_t = (x_t - \mu)/\sigma_t$.

We make a small digression here about the distribution of $log(U_t^2)$. To establish the properties of this distribution, we have generated ten thousands random numbers: $U \sim N(0,1)$. Then we have generated the distribution of U^2 and of $log(U^2)$, which appears at Figure 1.

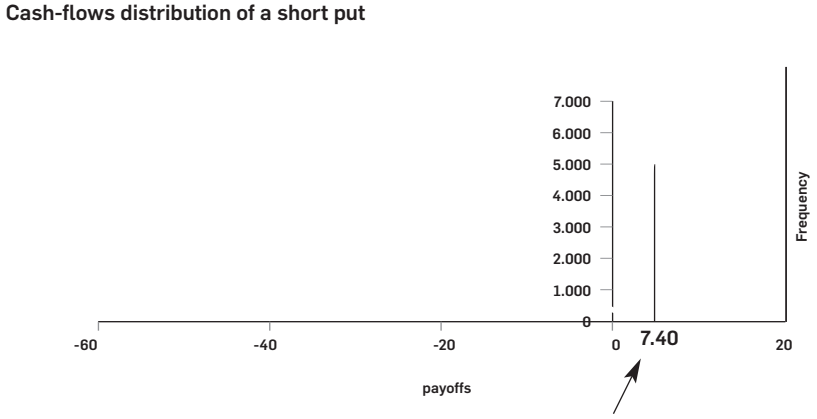
Distributions of U^2 and of $log(U^2)$, $U \sim N(0,1)$



Since $U \sim N(0,1)$, the distribution of U^2 corresponds to a centered χ^2 as shown[7] in Figure I. Furthermore, the distribution of $log(U^2)$ is truncated. It is very similar to the distribution of the payoffs of a short position in a put option. To illustrate that point, we have simulated the payoffs of a short position in an European put option with the following characteristics. The price of the underlying stock is 100$; the strike price is 95$; the option maturity is three months; the risk-free rate is 0%; the volatility is 50%. The resulting price of this put is 7.40$. We have simulated ten thousands payoffs of this put whose distribution is shown at Figure 2. We observe that the distribution of these payoffs is very close to the distribution of $log(U^2)$. This profile of payoffs might be found very frequently in hedge fund returns which have a fat left tail. Incidentally, the payoffs of a short put option are a very good indicator of risk related to adverse rare events. The distribution of $log(U^2)$ seems thus very relevant to capture stock market crashes.

[7] *A centered χ^2 is obtained using Gaussian standard random variables. An uncentered χ^2 distribution obtains when the expectation of the normal random variables used to build this distribution is different from zero.*

Payoffs distribution of a short put whose price is 7.40$



SOURCE: RACICOT AND THÉORET (2006)

Our distribution of $log(U^2)$ has a mean of -1.24 compared to -1.27 for the theoretical mean and a variance of 4.74 compared to 4.93 for the theoretical variance. Even with 10,000 iterations, we cannot replicate perfectly the theoretical moments. This indicates that that the sample must be very large to do so. Furthermore, the simulated distribution has a leptokurtic coefficient equal to 3.62 compared to 3 for the normal distribution and an asymmetry coefficient equal to -1.48 compared to 0 for the normal distribution. These are two adverse risks for the investor who prefers investments having returns whose kurtosis is close to 3 and whose asymmetry is positive. But these two risks can be found in a large number of financial instruments and are thus relevant for the distribution of $log(U^2)$.

To take into account these results, we add and subtract $E\left[log(U_t^2)\right]$ in equation (16)

$$log\left(x_t^2\right)=E\left[log\left(U_t^2\right)\right]+h_t+\left[log\left(U_t^2\right)-E\left[log\left(U_t^2\right)\right]\right] \tag{17}$$

For estimation purpose, we can rewrite equation (17) as follows

$$log\left(x_t^2\right)=\eta_0-1.27+h_t+\varsigma_t \tag{18}$$

where $\varsigma_t=\left[log\left(U_t^2\right)-E\left[log\left(U_t^2\right)\right]\right]$ and η_0 is a constant used to account for the fact that $E\,log(U_t^2)$ is equal to -1.27 only in very large samples as shown when simulating the distribution of $log(U^2)$. By doing so, we resort to a procedure which differs from the one used by the researchers who rely on the Kalman filter to estimate the stochastic volatility. As shown later, adding this constant will give more satisfying results when comparing stochastic volatility to GARCH(1,1) volatility.

ς is an error term that follows a logarithmic χ^2. Its expectation is given by

$$E(\varsigma_t)=E\Big[\,log(U_t^2)-E\big[\,log(U_t^2)\,\big]\Big]=E\big[\,log(U_t^2)\,\big]-E\big[\,log(U_t^2)\,\big]=0 \qquad (19)$$

and its variance is

$$V(\varsigma_t)=E(\varsigma_t^2)=E\Big[\,log(U_t^2)-E\big[\,log(U_t^2)\,\big]\Big]^2=0.5\pi^2=4.93 \qquad (20)$$

Finally, the equation system that we want to estimate is the following

$$log(x_t^2)=\eta_0-1.27+h_t+\varsigma \qquad (21)$$

$$h_t=\gamma_0+\gamma_1 h_{t-1}+\xi_t \qquad (22)$$

Equations (21) and (22) are in an appropriate form to use the Kalman filter presented earlier. Equation (21) is the measurement equation since the variable x_t is observed. Equation (22) is the state equation or the transition equation since h_t, the state variable, is not observed. This equation is simulated with the Kalman filter.

We aim at estimating these two equations using the return on the Canadian Treasury bill. We resort to a time series of monthly returns from 1941 to 2005. This sample seems appropriate to attain our objective, which is to compare two modelling techniques which are stochastic volatility and the GARCH one. This dataset contains sufficient information to proceed with our simulations. We use the EViews software to estimate the parameters of these equations. In the ***Workfile*** windows of EViews, we click on ***Object***, then on ***New Object*** and then we choose in the menu the specification ***SSpace***. Then in the window that appears, we can write the following EViews code which is displayed in Table 1.

EViews specification of a stochastic volatility model*.

```
@signal lnr2=-1.27+H TT+c(1)+[VAR=s2]
@state HTT=c(4)+ c(2)*H TT(-1)+[ename=e1,VAR=exp(c(3))]
@param c(1) 0.01 c(2) 0.9 c(3) 0.1
```

** We added a constant to the measurement equation in order to scale data. We also expressed the data in deviation from the mean because assuming a zero mean is a strong hypothesis for some financial time series.*

In a state model, the measurement equation begins by @signal in the EViews software and in the state equation, by @state. The first command indicates that the dependent variable is observed whereas the second means that the dependent variable is unob-

served and must thus be simulated. In Table 1, the variable *lnr*2 is equal to $log(x_t^2)$, where x_t is yield of the Canadian Treasury bill. Since we cannot assume that the mean of the Treasury bill yield is zero, we have expressed these yields in deviation from the mean in order to implement the equation of x_t.

The variances are expressed in brackets in the EViews command: *VAR*, as shown in Table 1. In the first equation of this table, *s2* is the variance of the innovation. Previously, we have created a scalar *s2* equal to: 0.5*@acos(-1)*@acos(-1) = 4.93, which is the variance associated with the χ^2 distribution of $log(U^2)$. In the equation of the stochastic variance which is the equation of HTT, the variance of the innovation is expressed as an exponential form, more precisely: $VAR = exp(c(3))$, $c(3)$ being a coefficient to be estimated. We have also given seed values to the three coefficients: $c(1)$, $c(2)$, and $c(3)$. The results of the estimation are reported in Table 2.

Estimation of the stochastic volatility model of the yield of the Canadian Treasury Bill.

Sspace: SS01
Method: Maximum likelihood (Marquardt)
Date: 09/13/10 Time: 17:31
Sample: 1941M01 2005M12
Included observations: 780
Convergence achieved after 169 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	19.46189	20793.93	0.000936	0.9993
C(2)	0.998575	0.009191	108.6420	0.0000
C(3)	-4.329276	1.714602	-2.524946	0.0116
C(4)	-0.037478	29.74276	-0.001260	0.9990
	Final State	Root MSE	z-Statistic	Prob.
HTT	-25.43278	0.504542	-50.40764	0.0000
Log likelihood	-1383.653	Akaike info criterion		3.558085
Parameters	4	Schwarz criterion		3.581979
Diffuse priors	0	Hannan-Quinn criter.		3.567275

SOURCE: EVIEWS

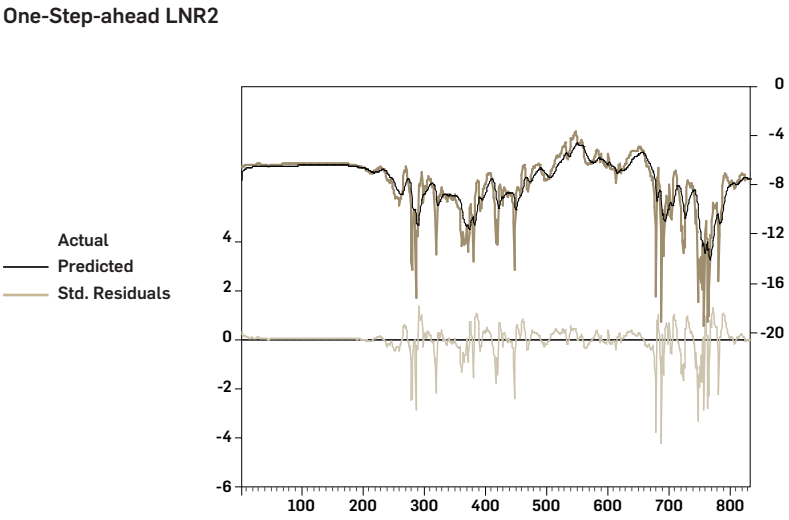
As shown in Table 2, the coefficients $c(1)$ and $c(4)$ are insignificant at the 95% confidence level. The dynamic behaviour of the observed and filtered values of the variable

y might be found in Figure 3. On the other hand, consistent with equation (22), the yield stochastic volatility is equal to

$$\sigma_t = \sqrt{e^{\frac{h_t}{2}}} \tag{23}$$

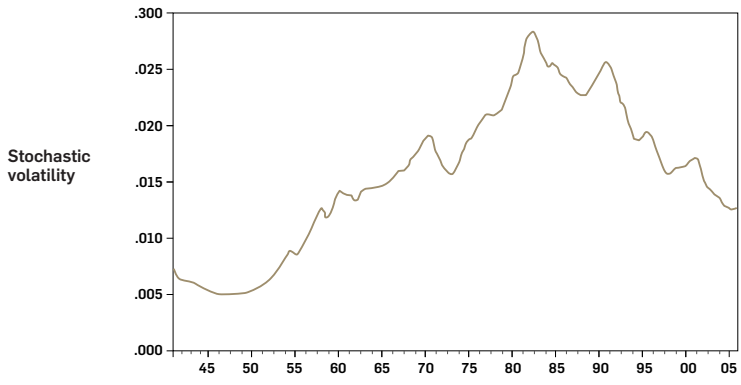
To annualize this standard deviation, we multiply it by $\sqrt{12}$. Figure 4 shows that the yield volatility has achieved its maximum value at the beginning of the 1980s (500th observation) and then had a tendency to decrease progressively.

Observed and estimated values of $log(r_t^2)$.



SOURCE: RACICOT AND THÉORET (2006)

Stochastic volatility of the Canadian T-bill yield.



Practitioners often compare their stochastic volatility models with a standard conditional model like the GARCH(1,1). We have thus applied this basic model to the Canadian T-bill yield:

$$Y_t = c + \xi_t \tag{24}$$

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \xi_{t-1}^2 \tag{25}$$

where c is a constant, $\xi_t = \varepsilon_t \sqrt{h_t}$ with $\varepsilon\ N(0,1)$ and h_t is the conditional variance. Nelson (1990) has shown that when the step *(dt)* tends to 0, the equation of h_t converges to a particular form of stochastic volatility:

$$dh = [\omega - \varphi h]\, dt + \psi h dz \tag{26}$$

Therefore, the GARCH(1,1) model corresponds to a mean-reverting process. More precisely, equation (26) can be rewritten as follows

$$dh = \varphi \left[\frac{\omega}{\varphi} - h\right] dt + \psi h dz \tag{27}$$

According to this last equation, the conditional variance reverts to its long-term level $\frac{\omega}{\varphi}$ at speed φ.

To establish the equivalence between the parameters of equations (25) and (27), we can rewrite equation (25) as follows

$$h_{t+1} - h_t = \beta_0 + [1 - \beta_2 E(\varepsilon^2) - \beta_1]\, h_t + \beta_2 h_t [\varepsilon_t^2 - E(\varepsilon^2)] \tag{28}$$

We can thus state the following equivalence between the coefficients of a GARCH(1,1) process (the equation of h_t) and those of an equivalent diffusion process (the equation of dh):

$$\lim_{dt \to 0} (dt)^{-1} \beta_0 = \omega \tag{29}$$

$$\lim_{dt \to 0} (dt)^{-1} \beta_1 = \varphi \tag{30}$$

$$\lim_{dt \to 0} h^{-\frac{1}{2}} \sqrt{2} \beta_2 = \psi \tag{31}$$

After having estimated the equation of h_t, we can compute the parameters of the equation of d_h:

$$\hat{\varphi} = \frac{1 - \hat{\beta}_1 - \hat{\beta}_2}{dt} \tag{32}$$

$$\frac{\hat{\omega}}{\hat{\varphi}} = \frac{\left[\dfrac{\hat{\beta}_2}{1 - \hat{\beta}_1 - \hat{\beta}_2}\right]}{dt} \tag{33}$$

$$\hat{\psi} = \frac{\sqrt{2}\hat{\beta}_2}{\sqrt{dt}} \tag{34}$$

Note that we assume that the coefficient of kurtosis is equal to 3 when computing $\hat{\psi}$, which means that we assume that the distribution of the innovation is normal. Otherwise, $\hat{\psi}$ can be rewritten as:

$$\hat{\psi} = \frac{\sqrt{\iota\text{-}1}\ \hat{\beta}_2}{\sqrt{dt}} \tag{35}$$

where ι is the coefficient of kurtosis[8].

According to Fornari and Mele (2006), the sequence $(\varsigma)_{n=1}^{\infty} \equiv \left[\varepsilon_n^2 - E(\varepsilon^2)\right]_{n=1}^{\infty}$ which appears in equation (28), is an iid sequence of centered chi-square variables of one degree of freedom and represents the discretization of the Brownian increments *dW*. Furthermore, the $\sqrt{2}$ which appears in the $\hat{\psi}$ equation can be explained by the fact that $\varsigma = \varepsilon^2 - E(\varepsilon^2) = \varepsilon^2 - 1$ is a chi-square variable with one degree of freedom and with variance equal to 2. Furthermore, the normality hypothesis is not required to obtain convergence.

Using another approach, Nelson and Foster (1994) show that the ARCH models converge to a continuous diffusion processes. To show this, we use again the previous GARCH(1,1) model

$$y_t = c + \xi_t \tag{36}$$

$$h_t = \beta_0 + \beta_1 h_{t-1} + \beta_2 \xi_{t-1}^2 \tag{37}$$

The following recursive equation is a generator for h_t, which holds for the whole set of ARCH models

$$\hat{z}_{t+dt} = \hat{z}_t + \left[dt \times \hat{\kappa}\left(y_t, \hat{z}, t, dt\right)\right] + \sqrt{dt}\left[g\left(\xi_{y,t+dt}, y_t, \hat{z}, t, dt\right)\right] \tag{38}$$

$\hat{\xi}$ is a residual that can be obtained using $\hat{c}$ as an estimate of c. We know that the GARCH(1,1) process converges to the continuous one

$$dh = [\omega - \varphi h]\, dt + \psi h dz \tag{39}$$

If we compare the equations (38) and (39), we realize that the GARCH(1,1) process makes a mapping between *z* and *h*, the conditional variance. Furthermore, it also establishes the following mapping between the discrete and continuous processes of *z* and *dh*.

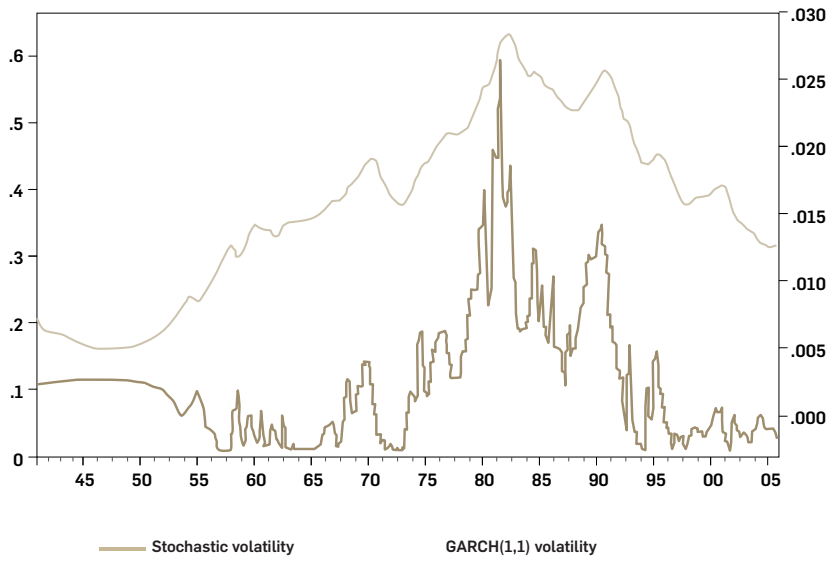
$$k = \omega - \varphi h \quad (40) \qquad\qquad g = \left(\hat{\xi}^2 - \hat{h}^2\right) \quad (41)$$

The recursion given by the *z* equation is general enough to encompass more general ARCH models as, for instance, the Nelson's EGARCH (1990).

[8] *On this matter, see Engle and Lee, in: Rossi (1996), chap. 11.*

In Figure 5, we compare the conditional volatility associated with the GARCH(1,1) model to the stochastic volatility model computed previously. We note that the dynamics of the two categories of volatilities are related, although the stochastic volatility is generally higher than the conditional volatility associated to the GARCH(1,1) model. We also notice that the stochastic volatility is less erratic than the GARCH(1,1) model. We can note further that the conditional volatility related to the GARCH(1,1) model has jumped more in the inflationary surge at the end of the 1970s and at the beginnings of the 1980s. Note that the GARCH-procedure seems more sensitive to noise or unexpected shocks compared to the stochastic volatility model which uses the Kalman filter as an estimation procedure. By analogy with the Hodrick-Prescott filter, the Kalman filter which smooths recent past is more able to capture business cycles than a GARCH(1,1) which is conditioned by recent unexpected information.

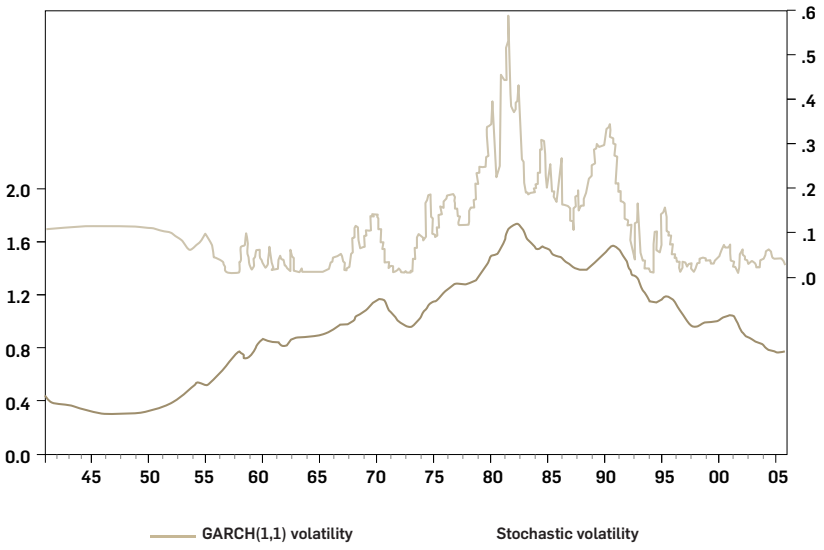
Annualized, stochastic volatility and GARCH(1,1) volatility, Canadian T-Bill yield, 1941-2005.



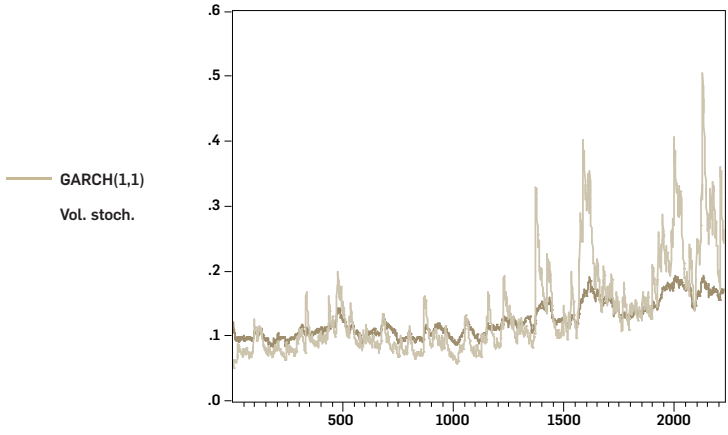
As noticed earlier, the omission of the constant η_0 in the equation of $log(x_t^2)$ might have some perverse effects. In Figure 6, we have modified Figure 5 by omitting the constant. There is an obvious scaling problem. The stochastic volatility is much too high compared to the GARCH(1,1) volatility. The estimated values related to the stochastic volatility are no longer relevant. They are overestimated compared to the annualized historical volatility of the T-bill yield which is about 0.14. Adding up a constant to the equation of $log(x_t^2)$ thus deals with this problem.

We have repeated the same exercise for the S&P/TSX return for a period ranging from 1992 to 2000. The results can be found in Figure 7. Once again, the dynamics of the two volatilities which are annualized are similar and show that the volatility of the TSX has a tendency to increase for this period but in this case, the GARCH(1,1) has a much more erratic behaviour compared to stochastic volatility.

Annualized stochastic volatility and GARCH(1,1) volatility, Canadian T-Bill yield, 1941-2005 (without a constant in the $log(x_t^2)$ equation).



Annualized stochastic volatility and GARCH(1,1) volatility, S&P/TSX return, 1992-2000.

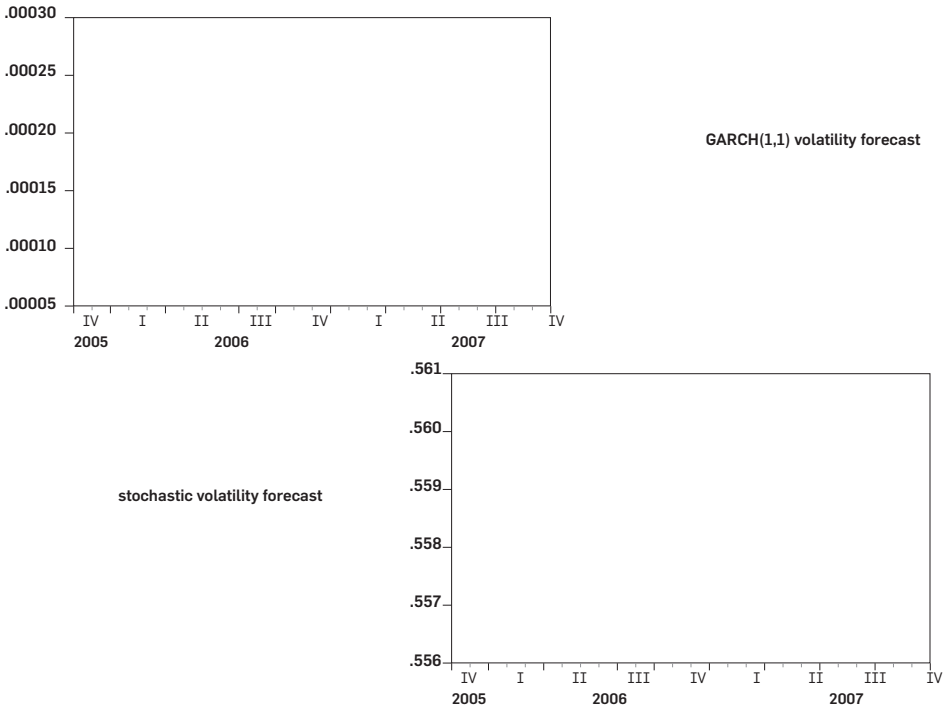


SOURCE: RACICOT AND THÉORET (2006)

3.2. Forecasting stochastic volatility

Based on our previous developments, we can now make a stochastic volatility forecast based on our stochastic volatility model, because it is strictly recursive. We use again our example of the Canadian Treasury bill yield. We start the forecasting exercise at period 834, the last observation being 833, and then we make the forecast up to period 850. The results of the forecast are shown in Figure 8. We note that the stochastic volatility model forecasts a decrease in volatility compared to the GARCH(1,1) which forecasts an increase in volatility. But we must bear in mind that the stochastic volatility was at the beginning much higher than the one resulting from the GARCH(1,1) model. Consequently, the two volatilities have a tendency to converge to each other.

Volatility forecast, stochastic model and the GARCH(1,1) model, Canadian Treasury bill yield.



These results show that when we forecast volatility, it would be incorrect to give a point forecast. As shown by our example, the volatility forecast can vary a lot from one model to another. Thus it is relevant to define a confidence interval of the forecast to give a better idea of the risk to the users of this method. But, according to our experiments, it seems more appropriate to give more weight to stochastic volatility in our forecast since it is based on a sophisticated smoothing algorithm that by nature provide optimal projections based on new information arrival but put less emphasis on incoming information than the GARCH(1,1) procedure which relies on a maximum-likelihood based technique.

Assume that the *(P/E)* ratio is modeled by a rational expectations model:

$$\left(\frac{P}{E}\right)_{t+1} - \left(\frac{P}{E}\right)_t = \beta\left[\left(\frac{P}{E}\right)_t - \left(\frac{P}{E}\right)_t^*\right] \tag{42}$$

The forecast error, which is the spread between the observed and forecasted (*P/E*) values at time *t* entails an update in the expectations. In order to filter this ratio, we rewrite this equation as follows[9]:

$$\left(\frac{P}{E}\right)_t = c(1)\left(\frac{P}{E}\right)_{t-1} + (1\text{-}c(1))\ SV1 \tag{43}$$

where *SV1* stands for the*(P/E)* long-term forecast. If $t = t\text{-}1$, we have:

$$\frac{\overline{P}}{E} = SV1 \tag{44}$$

The last estimation of *SV1* is thus particularly important since it represents the value to which the (*P/E*) converges. However, *SV1*, being a forecasted value, is unknown. We assume that this state variable follows a random walk, as usually assumed in this kind of setting:

$$SV1 = SV1(-1) + \varepsilon \tag{45}$$

Table 3 provides the EViews code used to filter the (*P/E*) ratio. We assume that the two equations, which are respectively the observation equation and the state equation, embed an innovation term and a variance, and that there exists a covariance between the two equations errors terms.

EViews Kalman filtering of the (P/E) ratio

```
pe=c(1)*pe(-1)+(1-c(1))*sv1+[ename=e1,var=exp(c(2))]
@state sv1=sv1(-1)+[ename=e2,var=exp(c(6))]
@evar cov(e1,e2)=c(4)
@param c(1) 0.90 c(2) 0.2
```

Using equations shown at Table 3, we filter the monthly S&P500 (*P/E*) ratio for the period running from January 1881 to May 2005[10]. Table 4 provides the estimation result.

[9] *We overlook the stationarity of (P/E) here.*

[10] *This series is built using : Robert J. Shiller : http://www.econ.yale.edu/~shiller/.*

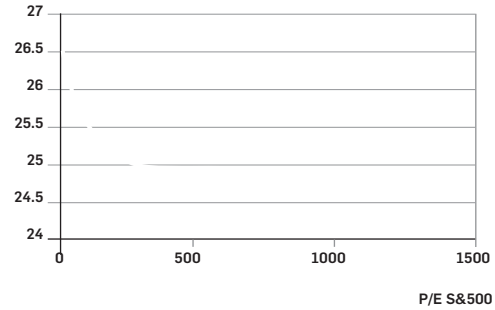
(P/E) ratio Kalman filtering, January 1881- May 2005

SSspace: SS04
Method: Maximum likelihood (BHHH)
Sample: 1 1500
Included observations: 1500
Valid observations: 1490
Failure to improve Likelihood after 239 iterations

	Coefficient	Std. Error	z-Statistic	Prob.
C(1)	0.971980	0.008887	109.3744	0.0000
C(2)	-0.130224	1878.492	-6.93E-05	0.9999
C(4)	-3.052600	58852.88	-5.19E-05	1.0000
C(6)	-2.683399	0.910691	-2.946551	0.0032
	Final State	Root MSE	z-Statistic	Prob.
SV1	24.94194	10.89715	2.288851	0.0221
Log likelihood	-1724.737	Akaike info criterion		2.320452
Parameters	4	Schwarz criterion		2.334698
Diffuse priors	1	Hannan-Quinn criterion		2.325761

As shown in Table 4, the autoregressive coefficient ($C(1)$), at 0.97, is close to 1, as expected, since the denominator of the (P/E) ratio is a moving average computed on earnings per share. According to the $SV1$ estimation, the (P/E) ratio long-term value is equal to 24.94. Moreover, in May 2005, which is the series last observation, this ratio quoted 26.48. A downturn of the US stock market was thus expected. Figure 9 provides a plot of the convergence of the (P/E) towards its long-term value starting in May 2005.

(P/E) simulation from its value in May 2005 to its steady-state value



SOURCE: RACICOT AND THÉORET (2006)

Whatever the model used, forecasting return volatility is a complicated process. Since we have to rely on stochastic models, model risk is important. We must bear in mind this issue when performing a volatility forecast. Analysts must also rely on a priori information and educated guesses to formulate their forecasts.

However, we show in this article that Kalman filter is a valuable tool to forecast stochastic volatility. It must therefore be added to the forecaster toolbox whose work is rendered increasingly difficult by the exploding number of exotic contingent claims. Statistical distributions used to forecast stochastic volatility are very close to the payoffs of short positions on plain vanilla put options. In that respect, these payoffs are similar to the ones of many financial instruments, and particularly the shares issued by Hedge funds. Integrating options theories and stochastic volatility modelling in a forecasting framework is a challenge for future financial research.

In other respects, we found that a misspecification of a stochastic volatility model could generate a biased volatility forecast. Indeed, the specification of a model in the Kalman filter setting is very sensitive to the initial assumptions. Model building requires specific assumptions but these assumptions must respect the parsimony principle.

Finally, we have presented a simple forecasting application of the Kalman Filter to the S&P 500 *P/E*. To do so, we used a basic rational expectations model. Obviously, we could elaborate this model which might prove useful to forecast stock market data.

Andersen, T. G. and Benzoni, L. (2010), Stochastic volatility, WP[2010-10], CREATES.

Andersen, T. G., Bollerslev, T., Christoffersen, P.F. and Diebold, F. (2005). Volatility Forecasting, WP [05-011], PIER.

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

Fornari, F. and Mele, A. (2006). Approximating volatility diffusions with CEV-ARCH models, *Journal of Economic Dynamics & Control,* **30**, pp. 931-966.

Gregoriou, G. N. (2009) Book Review of: The Econometric Analysis of Hedge Fund Returns: An Errors-in-Variables Perspective, *Journal of Wealth Management,* **12**, pp. 138-139.

Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge (UK).

James, J. and Webber N. (2000). *Interest Rate Modelling*, Wiley, New York.

Mills, T. C. (1999). *The Econometric Modelling of Financial Time Series*, 2nd edition, Cambridge University Press, Cambridge (UK).

Nelson, D. (1990). ARCH models as diffusion approximations, *Journal of Econometrics,* **45**, pp. 7-38.

Nelson, D. B. and Foster, D. P. (1994). Asymptotic Filtering Theory for Univariate ARCH models, *Econometrica,* **62**, pp. 1-41. Reprinted in Rossi (1996), chap. 8.

Racicot, F. É. and Théoret, R. (2001). Traité d'économétrie financière, Presses de l'Université du Québec (PUQ), Québec.

Racicot, F.-É. and Théoret, R. (2006). *Finance computationnelle et gestion des risques*, 2ième édition, Presses de l'Université du Québec (PUQ), Québec.

Racicot, F. É. and Théoret, R. (2007a). A study of dynamic market strategies of hedge funds using the Kalman filter, *Journal of Wealth Management,* **10**, pp. 94-106.

Racicot, F. É. and Théoret, R. (2007b). Book Review of: Option Pricing Models & Volatility Using Excel-VBA (by F. D. Rouah and G. Vainberg), *Journal of Derivatives* & *Hedge Funds*, **13**, pp. 181-183.

Racicot, F. É. and Théoret, R. (2008). *The Econometric Analysis of Hedge Fund Returns: An Errors-in-Variables Perspective*, Netbiblo, A Coruña.

Racicot, F.É. and Théoret, R. (2009). Modeling hedge fund retunrs using the Kalman filter: An errors-in-variables perspective, WP[2009-06], Chaire d'information financière et organisationnelle, ESG-UQAM.

Racicot, F. É. and Théoret, R. (2010). Hedge fund returns, Kalman filter, and errors-in-variables, *Atlantic Economic Journal,* **38**, pp. 377-378.

Rouah, F. D. and Vainberg, G. (2007). *Option Pricing Models* & *Volatility Using Excel-VBA*, Wiley, New York.

Rossi, P. E. (1996). *Modelling Stock Market Volatility*, Academic Press, San Diego, California.

Taylor, S. J. (1986). *Modelling Financial Time Series*, Wiley, New York.

Taylor, S. J. (1994). Modelling stochastic volatility: A review and comparative studies, *Mathematical Finance,* **4**, pp. 183-204.

Théoret, R., Rostan, P. and Zabré, L. (2004). Difficultés de calculer les cotes des swaps de volatilité, *Insurance and Risk Management,* **72**, pp. 301-320.

Wang, P. (2003). *Financial Econometrics: Methods and Models,* Routledge, London.