ON COMPARING VARIOUS MODELLING SCHEMES: THE CASE OF THE PRAGUE STOCK EXCHANGE INDEX

Radek Hendrych

Abstract

This year the PX index, the key price index of the Prague Stock Exchange, celebrated twenty years of its existence. Therefore, one could be indeed interested in analysing historical daily closing quotes of this stock index from an econometric point of view. In detail, the aim of this contribution is to introduce a particular class of discrete-time state space models and demonstrate that this class is appropriate for such a univariate financial time series. Particularly, it involves regularly applied econometric modelling instruments; it combines a local level model and a linear ARMA process together with conditionally heteroscedastic innovations. Moreover, the suggested modelling framework is examined in different settings of parameters. The final model is selected with respect to standard information and prediction criteria; it is further investigated and statistically verified by inspecting prediction residuals. Its empirical performance is compared with other commonly applied methods, i.e. with linear ARMA or benchmark GARCH models.

Key words: ARMA, GARCH, local level model, PX index, state space methods.

JEL Code: C01, C51, C58.

Introduction

This year the PX index, the key stock index of the Prague Stock Exchange, celebrated twenty years of its existence. Thus, one could concentrate on exploring its historical evolution from an econometric perspective. In detail, this contribution introduces a particular class of discrete-time state space models that are able to describe historical daily closing quotes of the PX index. Such an analysis can find a broad spectrum of applications in portfolio and risk management or in technical analysis of various financial instruments.

An extensive body of academically and practically oriented literature exists in this field of research. For instance, Al-Loughani & Chappell (2001) or Koutmos, Negakis, & Theodossiou (1993) elaborated on modelling stochastic behaviour of particular stock indices.

Vega & Smolarski (2012) constructed forecasts of the FTSE index reflecting movements of global stock markets. Last but not least, Sutcliffe (2006) comprehensively explained general properties of stock indices and delivered the list of related studies.

This paper is organized as follows. Section 1 introduces the stock index PX in detail. Section 2 concerns methodological issues. In particular, it presents a discrete-time state space modelling class suitable for the financial time series of the PX index historical closing quotes. Moreover, it explains procedures of its calibration and statistical verification. Section 3 reviews empirical results and analysed them. Section 4 compares the finally selected model with several simple schemes such as linear ARMA or GARCH models.

1 PX index

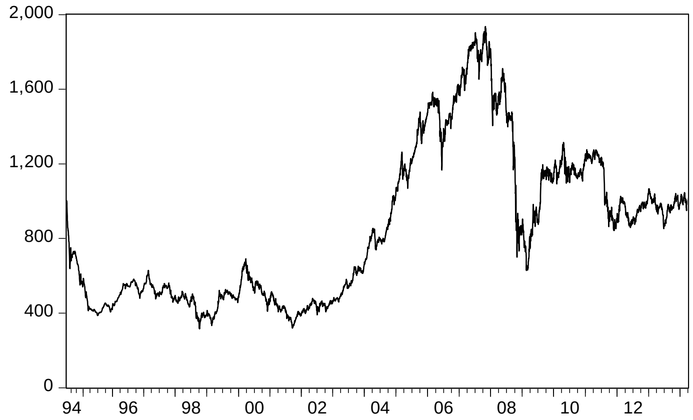
The PX index (ISIN XC0009698371) is an official market-cap weighted stock index composed of the most liquid shares traded on the Prague Stock Exchange. In particular, it is a price index of blue chips issues weighted by market capitalization calculated in real-time. Dividends are not considered. A new value of the PX index is delivered by a particular formula reflecting each single price change of index constituents. The maximum weight for a share issue is 20% on a decisive day. A portfolio of basic issues is variable and it can be restructured quarterly (Wiener Borse, 2014).

The PX index was launched on 5th April 1994 (originally known as PX-50). Its base was composed of the fifty most important share issues operating on the Prague Stock Exchange. The opening base value was fixed on 1000. The number of basic issues has been variable since December 2001. In March 2006, the PX index was officially introduced. It took over the whole history of the replaced index PX-50 continuing in its development. In March 2014, the PX base contained fourteen issues. The top five stocks had approximately 85% share of market capitalization in the portfolio. The majority of capitalization was allocated in banking, energy, and insurance sectors. Further details (including historical data) can be found on the official web pages of the Prague Stock Exchange (Prague Stock Exchange, 2014).

Figure 1 presents all historical daily closing quotes of the PX index until 4th April 2014 (i.e. 4952 observations). The minimal value 316 occurred on 8th October 1998 after the Russian financial crisis. The maximal observation 1936 was achieved 29th October 2007. It is visible that the crisis year 2008 was truly exceptional in terms of highly volatile prices. The augmented Dickey-Fuller test finding unit roots delivers statistics -1.134 with the declared p-value 0.705, i.e. the presence of the unit roots cannot be rejected at 5% level (Tsay, 2010).

This feature concerning the considered time series is likely expectable. It should be taken into account e.g. by introducing the logarithmic returns of the PX index prices, which should lead to stationarity (the augmented Dickey-Fuller test statistics is -63.212 with the p-value 0.0001).

Fig. 1: Historical closing quotes of the PX index.



Source: Author.

2 State space model of the PX index

The PX index historical quotes can be analysed by means of a particular state space modelling class that involves various regularly applied econometric modelling instruments altogether. Concretely, a local level model, a linear ARMA process, and conditionally heteroscedastic innovations are jointly considered in investigating the log returns of the PX index prices.

In detail, the following form of the discrete-time state space model is proposed:

$$\begin{aligned} r_t &= \mu_t + \theta_{r-1} x_{t-r+1} + \ldots + \theta_0 x_t + \sqrt{g_t}\,\varepsilon_t, \\ \mu_{t+1} &= \mu_t + \sqrt{h_t^{\mu}}\,\eta_t^{\mu}, \\ x_{t+1} &= \varphi_r x_{t-r+1} + \ldots + \varphi_1 x_t + \sqrt{h_{t+1}^{x}}\,\eta_{t+1}^{x}, \; t = 1, 2, \ldots \end{aligned} \tag{1}$$

In the first (signal) equation of the system (1), r_t denotes the log returns of the observed PX prices, μ_t stands for the local level, x_t is the process fully described by the third equation in (1), θ_0, ..., θ_{r-1} are the real parameters, finally the innovation term is assumed to be a GARCH (P_g, Q_g) type process. The second (state) equation in (1) represents the local level with the non-negative finite variances changing in time; the innovation term is again driven by a GARCH (P_μ, Q_μ) type process. The third (state) equation in (1) specifies evolution of the process x_t. It contains the real parameters φ_1, ..., φ_r and GARCH (P_x, Q_x) type errors. More precisely, the formulated model should be accompanied by the following set of artificial (state) equations: $x_{t-r+i} = x_{t-r+i}$, $i = 1, ..., r$-1. In summary, the whole considered state space modelling system (1) has r+1 state equations. Further, denote $r = max\,(p,q+1)$, p, q are nonnegative integers, and put $\varphi_j = 0$ for $j > p$, $\theta_j = 0$ for $j > q$, and $\theta_0 = 1$, respectively. All model disturbances, i.e. ε_t, η^{μ}_t, and η^{x}_{t+1}, are mutually and serially independent i. i. d. Gaussian random variables with zero mean and unit variance. They are also uncorrelated with the initial state vector $(\mu_1, x_{1-r+1}, ..., x_1)^{\mathrm{T}}$ which has an expected value $\mathbf{a}_1$ and a covariance matrix $\mathbf{P}_1$. Frequently, one sets $\mathbf{a}_1 = \mathbf{0}$ and $\mathbf{P}_1 = \kappa\mathbf{I}$, where κ is a large positive number.

The model (1) can be rewritten more concisely as

$$\begin{aligned}[r_t - \mu_t] &= \sum_{i=1}^{p} \varphi_i [r_{t-i} - \mu_{t-i}] + \sqrt{h_t^x}\,\eta_t^x + \sum_{j=1}^{q} \theta_j \sqrt{h_{t-j}^x}\,\eta_{t-j}^x + \sqrt{g_t}\,\varepsilon_t - \sum_{k=1}^{p} \varphi_k \sqrt{g_{t-k}}\,\varepsilon_{t-k}, \\ \mu_{t+1} &= \mu_t + \sqrt{h_t^{\mu}}\,\eta_t^{\mu}, t = 1, 2, \dots\end{aligned} \tag{2}$$

According to (2), one can concludes that the proposed model (1) of the log returns r_t is composed of the local level, the linear ARMA process, and the conditionally heteroscedastic disturbances with the additive error terms. It contains $p+q+P_g+P_{\mu}+P_x+Q_g+Q_{\mu}+Q_x+3$ unknown parameters; however, this number can be reduced considering innovations with instant variances. Under some modelling circumstances, it can be comfortably calibrated by the numerically effective Kalman recursive formulas associated with the linear Gaussian state space models. The unknown parameters are then estimated by the corresponding maximum likelihood procedure (Commandeur & Koopman, 2007; Durbin & Koopman, 2001).

Moreover, the calibrated model should be statistically verified in a proper way. The prediction residuals have the key role in this context. Particularly, the prediction residuals of the introduced model are given as

$$v_t = r_t - \hat{\mu}_t + \tilde{\theta}_{r-1}\hat{x}_{t-r+1} + \dots + \tilde{\theta}_0 \hat{x}_t, \tag{3}$$

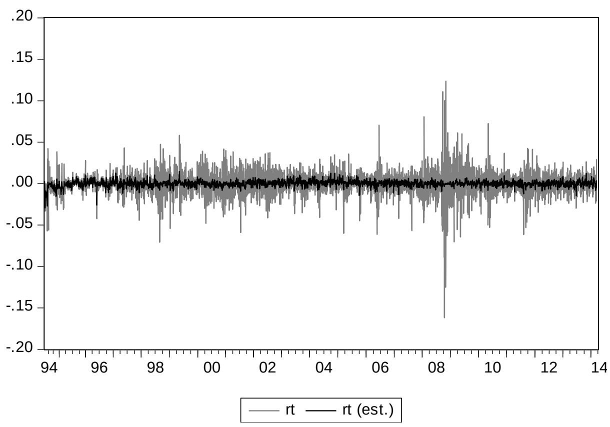
where the hats correspond to the best linear predictions of the associated unobservable states based on information accumulated by the observed time series until and including time t-1; the tildes denote the estimated counterparts of the unknown parameters. In respect to the theoretical background of state space methods (Durbin & Koopman, 2001), the prediction residuals (after their standardization) follow the Gaussian white noise with zero mean and unit variance. Therefore, one can test the adequacy of a fitted model by means of common statistical techniques (Commandeur & Koopman, 2007).

Furthermore, auxiliary residuals for the signal and states can be investigated for outliers and structural breaks. In particular, the auxiliary residuals are simply the smoothed standardized disturbances of the assumed state space model (1). It can be shown that they are autocorrelated (Durbin & Koopman, 2001). A relatively large smoothed observation (signal) error indicates the presence of an outlier while a relatively large smoothed state error indicates a structural break. In practice, the diagnostic auxiliary residual checking procedures are carried out using a conservative significance level since one is interested only in serious outliers and structural breaks. This is comparable with the issue of the studentized residuals in linear regression models (Durbin & Koopman, 2001).

3 Empirical results

The suggested model of the PX index expressed by the equations in (1) was examined in the framework of various settings of parameters p, q, and P_g, P_μ, P_x, Q_g, Q_μ, Q_x. All computations have been performed in EViews and R (Tusell, 2011; Van der Bossche, 2011). The calibrated models have been compared using: (i) the Akaike information criterion (AIC) and (ii) the root mean squared error of one-step-ahead predictions (RMSE).

Fig. 2: Log returns of the PX index prices with the smoothed estimates delivered by (1).



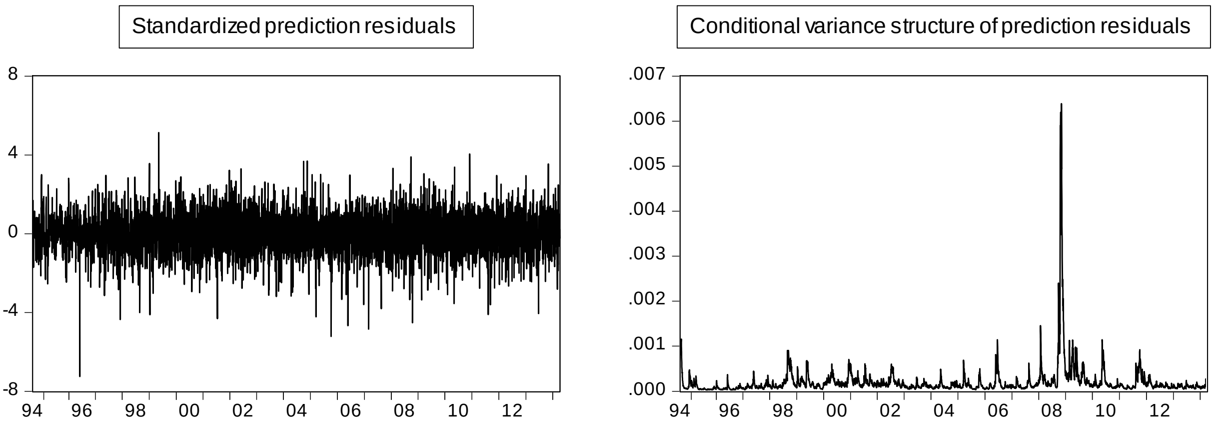
Source: Author.

According to these criteria, the model corresponding to the particular choice of parameters $p=q=P_g=P_\mu=P_x=Q_g=Q_\mu=Q_x=1$ was selected with AIC=-6.0776 and RMSE=0.01396. Figure 2 shows the log returns r_t together with their smoothed estimates. The graph including the original series of the PX index prices with their smoothed estimates is not inserted since the original dataset and the delivered estimates are not distinguishable.

Figure 3 shows the standardized prediction residuals and the variance structure of the prediction residuals (compare with the GARCH variances in Figure 5).

Furthermore, the fitted model has been verified thoroughly using frequent statistical methods. As was mentioned above, the standardized prediction residuals should follow the Gaussian white noise with zero mean and unit variance. Several econometric tests have been applied to examine these assumptions.

Fig. 3: Standardized prediction residuals and variance structure of prediction residuals.



Source: Author.

Firstly, the zero mean has been tested by the common *t*-test (resulting in the test statistics 0.0237 and the corresponding p-value 0.9811) and the Wilcoxon signed rank test (resulting in the test statistics 0.8266 and the corresponding p-value 0.4084). Thus, the null hypothesis of the zero mean is not rejected at the 5% level.

Secondly, the assumption of the uncorrelated standardized prediction residuals has been investigated by the following statistical procedures (Tsay, 2010): (i) the Durbin-Watson test finding first order autocorrelations, (ii) the commonly used Ljung-Box test, (iii) the robustified portmanteau test with automatic lag selection (Escanciano & Lobato, 2009), and finally (iv) the BDS independence test, a portmanteau test examining the null against a variety of possible deviations from independence including linear dependence, non-linear dependence, or chaos. The resulting values of the test statistics are respectively: (i) 1.9555 (indeed near 2), (ii) 5.8720 with the p-value 0.4377 (8 lags) and 19.3367 with the p-value 0.1525 (16 lags), (iii) 2.4260 with the p-value 0.1193, and (iv) 0.0013 (the achieved p-value 0.2169 for ε=1.5s, m=2, where s is the sample standard deviation of the standardized prediction residuals), 0.0015 (the achieved p-value 0.3690 for ε=1.5s, m=3) or 0.0023 (the achieved p-value 0.3126 for ε=2s, m=20).

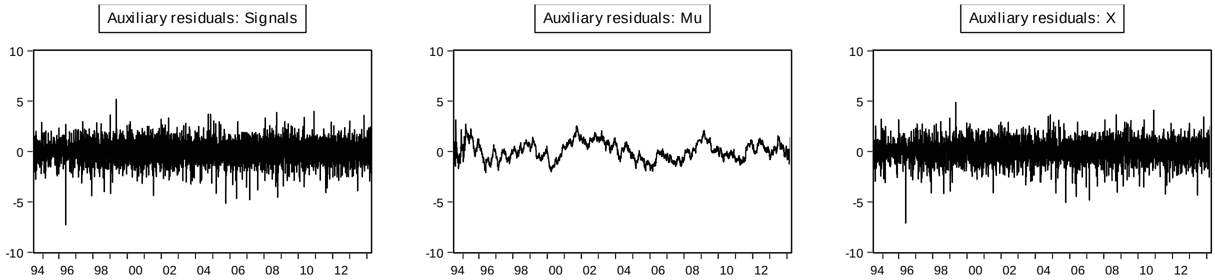
It should be noted that the declared p-values corresponding to the Ljung-Box test were calculated with the more conservative choice of degrees of freedom (Van der Bossche, 2011). In particular, they were computed as k-m+1, where k is the number of lags in the Ljung-Box statistic and m=r+1 is the number of diffuse priors, see above. In general, this correction leads to more frequent rejecting of the null hypothesis. Moreover, the BDS independence test was applied under various settings. However, these resulted in analogous conclusions (see above). Concretely, different configurations of the embedding dimensions m and the proximity parameters ε were compared (Kanzler, 1999). In summary, there is no principal evidence to reject the null hypothesis of uncorrelated standardized prediction residuals at the 5% level.

Thirdly, the homoscedasticity of the standardized prediction residuals has been verified. The ARCH LM test (Tsay, 2010) with eight lags delivered the statistics 4.1149 together with the corresponding p-value 0.8466, i.e. the null hypothesis cannot be rejected at the 5% level.

Lastly, the Jarque-Bera test (Tsay, 2010) provided the statistics 829.6438 with the p-value 0.0000, i.e. the Gaussian distribution of the standardized prediction residuals is rejected at the 5% level. Thus, one must rely on asymptotic features.

Figure 4 investigates auxiliary residuals of the signal and state equations. These are relatively strongly autocorrelated; see above. The Ljung-Box statistics with eight lags are respectively: 41.5950 (for the signal series), 35419.6773 (for the state μ_t), and 590.4719 (for the state x_t). Comparing Figure 3 and 4, one summarizes that there is an evidence of several questionable observations. For instance, it is possible to identify at least two important outliers in the left (and right) graph of Figure 4 (i.e. the first one in 1Q 1996 and the second one in 2Q 1999), or the problematic segment in the beginning of the PX index as it is visible in the centre graph of Figure 4.

Fig. 4: Auxiliary residuals for the signal and the states μ_t and x_t.



Source: Author.

4 Model comparison

The previously examined model can be compared with commonly applied modelling schemes. Particularly, one obviously considers a linear ARMA or benchmark GARCH model investigating logarithmic returns (Tsay, 2010). Firstly, assume the GARCH (1,1) model with normally distributed innovations, i.e.

$$r_t = c + \sqrt{\sigma_t}\varepsilon_t, \sigma_t = \omega + \alpha(r_{t-1} - c)^2 + \beta\sigma_{t-1}, \varepsilon_t \text{ are i. i. d. } N(0,1). \tag{4}$$

Secondly, a standard linear ARMA (p,q) model with a nonzero mean is supposed (Tsay, 2010); the appropriate order is identified using information criteria.

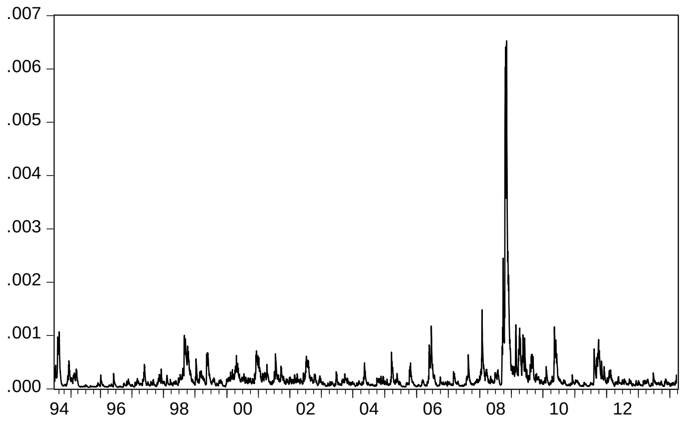
Table 1 summarizes the results introducing several characteristics which verify the fitted models, i.e. the Akaike information criterion (AIC), the root mean squared error (RMSE), the Ljung-Box statistics with eight and sixteen lags (Q (8) and Q (16)), the ARCH LM test with eight lags (ARCH (8)), and finally the Jarque-Bera test statistics. Figure 5 introduces the GARCH (1,1) variances; compare with Figure 3.

Tab. 1: Results of the GARCH (1,1) model and the ARMA (1,2) model.

Model	***AIC***	***RMSE***	***Q (8) (p-value)***	***Q (16) (p-value)***	***ARCH (8) (p-value)***	***Jarque-Bera (p-value)***
GARCH (1,1)	-6.0735	0.0141	92.7204 (0.0000)	120.2721 (0.0000)	3.6693 (0.8857)	709.7203 (0.0000)
ARMA (1,2)	-5.7063	0.0139	14.6138 (0.0121)	30.6686 (0.0038)	1155.7688 (0.0000)	28987.2700 (0.0000)

Source: Author.

Fig. 5: Estimated GARCH (1,1) variances.



Source: Author.

It is visible that the model tested in Section 3 delivers results that are more suitable. On the one hand, the GARCH model ignores the conditional mean structure of the PX log returns. On the other hand, the linear ARMA model does not respect the conditionally heteroscedastic variances of the PX log returns. Consequently, the model (1) combining these two perspectives can be more appropriate in this modelling context.

Conclusions

The key price index of the Prague Stock Exchange was investigated by means of state space methods. In particular, a class of discrete-time state space models comprising the local level model, linear ARMA terms, and conditionally heteroscedastic innovations has been introduced. The suggested approach based on combining the mentioned econometric concepts altogether demonstrated its empirical performance. More precisely, it outperformed common modelling schemes, i.e. the benchmark GARCH (1,1) model and the linear ARMA model, since it has been able to jointly model the conditional mean and variance structure of the PX index log returns. Further research will be focused on possible extensions of the considered modelling class and on examining various analogical datasets applying this scheme.

Acknowledgment

This research was supported by the grant GA P402/12/G097.

References

Al-Loughani, N., & Chappell, D. (2001). Modelling the day-of-the-week effect in the Kuwait Stock Exchange: a nonlinear GARCH representation. *Applied Financial Economics*, 11(4), pp. 353-359.

Commandeur, J., & Koopman, S. (2007). *An introduction to state space time series analysis*. New York: Oxford University Press.

Durbin, J., & Koopman, S. (2001). *Time series analysis by state space methods*. New York: Oxford University Press.

Escanciano, J., & Lobato, I. (2009). An automatic Portmanteau test for serial correlation. *Journal of Econometrics*, 151(2), pp. 140-149.

Kanzler, L. (1999). *Very fast and correctly sized estimation of the BDS statistic*. Retrieved from: http://ssrn.com/abstract=151669

Koutmos, G., Negakis, C., & Theodossiou, P. (1993). Stochastic behaviour of the Athens stock exchange. *Applied Financial Economics*, 3(2), pp. 119-126.

Prague Stock Exchange. (2014). Retrieved from: http://www.pse.cz

Sutcliffe, C. (2006). *Stock index futures*. (3rd ed.) Burlington, VT: Ashgate.

Tsay, R. (2010). *Analysis of financial time series*. (3rd ed.) Hoboken: Wiley.

Tusell, F. (2011). Kalman filtering in R. *Journal of Statistical Software*, 39(2), pp. 1-27.

Van der Bossche, F. (2011). Fitting state space models with EViews. *Journal of Statistical Software*, 41(8), pp. 1-16.

Vega, J., & Smolarski, J. (2012). Forecasting FTSE Index Using Global Stock Markets. *International Journal of Economics and Finance*, 4(4), pp. 3-12.

Wiener Borse. (2014). Retrieved from: http://www.wienerborse.at/

Contact

Radek Hendrych
Charles University in Prague
Faculty of Mathematics and Physics
Dept. of Probability and Mathematical Statistics
Sokolovská 83, 186 75 Prague 8, Czech Republic
Radek.Hendrych@mff.cuni.cz