Turkish Journal of Forecasting vol. 01 no. 2 (2017) pp. 66-71

Content list available at JournalPark

Turkish Journal of Forecasting

Journal Homepage: tjforecast.xyz

BIST 100 Index Estimation Using Bayesian Regression Modelling

C. Ozturk[1,*], D. Efendioglu[1], N. Gulec[1]

[1]*Ankara Yıldırım Beyazıt University, Engineering and Natural Sciences Faculty, Department of Industrial Engineering, Etlik Campus, Ankara, Turkey*

A R T I C L E I N F O

Article history:
Received 04 October 2017
Accepted 25 December 2017
Available online 29 December 2017

Keywords:
BIST 100 index
Bayesian regression method
Forecasting

A B S T R A C T

Identification of factors, determining the fluctuations of stock indices in the market, possesses great importance for the capital market actors. Not only specifying the factors and market but also explaining the relationship between them correctly, will reduce the amount of exposure financial risk and bring this topic into the limelight of market actors. The most important indicator of good progress in the economic cycle can be observed as the stock market index which is also used as macroeconomic indicators for developed economies. It is generally observed that all the stock prices rise or fall in the same period and this question gives us the impression that there are some factors that have an influence in this period. Having the properties of dynamic, complex and non-linear structure makes this analysis tough and solving the problem requires very complex and difficult processes. Due to high uncertainty and volatility, while estimating the stock price behaviour, stock investments carry more risk than any other investment. Moreover, BIST index possesses very high chaotic structure, it is not possible to conclude the long-term predictability. Our study will address the portion of the macro factors affecting the stock index. Interest rates, exchange rates, money supply, inflation, gold, oil prices are investigated among the macro factors. For the study data was obtained using CBT from the Electronic Data Dissemination System monthly frequencies. The factors influencing the stock index was determined by the method of regression relationship between them, but the Bayesian method is used with regression for the estimation.

1. Introduction

Identification of factors that determine the fluctuations of share certification in the market possesses great importance for the capital market actors. Specifying the factors and market with also the relationship between them correctly will reduce the amount of exposure financial risk, brings into this topic the limelight of market actors. The most important indicator of good progress in the economic cycle for the economy can be observed as the stock market index which is also used as macroeconomic indicators for developed economies. It is generally observed that all the stock price rises or falls in the same period and this case gives the impression that there are some factors that have an influence in this period.

Our study will address the portion of the macro factors affecting the stock. Finding a meaningful and systematic data related to micro factors can take too much time. Interest rates, exchange rates, money supply, inflation, gold, oil prices are investigated among the macro factors to observe a relationship between stock market.

Macroeconomic factors may be in relation to the different degree and direction of stock prices. Although the causes of changes in economic factors can be very dissimilar, changes occurring in the macroeconomic and

* Corresponding author.
E-mail addresses: cozturk@ybu.edu.tr (Cihat Öztürk), taha.deniz.fb@gmail.com (Deniz Efendioğlu), gulecnurullah@gmail.com (Nurullah Güleç),

development affects all businesses in an economy due to these factors. Attempts to characterize stock return predictability generally resulted in little consensus on the important conditioning variables, giving rise to model uncertainty. Stock Indices may be influenced by many factors such as political, economic, environmental and commercial factors.

Being dynamic, complex and non-linear structure makes this analysis tough and requires very complex and difficult process. Due to high uncertainty and volatility, making it difficult to estimate the stock price behaviour, stock investments carry more risk than any other investment. Because of possessing BIST index very high chaotic structure, it is not possible to conclude that the long-term predictability.

2. Background & Literature Review

Bańbura, Giannone and Reichlin (2010) shows that vector autoregression (VAR) with Bayesian shrinkage is an appropriate tool for large dynamic models. When the degree of shrinkage is set in relation to the cross-sectional dimension, the forecasting performance of small monetary VARs can be improved by adding additional macroeconomic variables and sectoral information [1].

Chen, Leung and Daouk, H. (2003) investigate that the probabilistic neural network (PNN) is used to forecast the direction of index return after it is trained by historical data. Statistical performance of the PNN forecasts is measured and compared with that of the generalized methods of moments (GMM). Moreover, the forecasts are applied to various index trading strategies, of which the performances are compared with those generated by the buy-and-hold strategy as well as the investment strategies guided by forecasts estimated by the random walk model and the parametric GMM models [2].

Cremers (2002) introduce a new methodology that explicitly incorporates model uncertainty by comparing all possible models simultaneously and in which the priors are calibrated to reflect economically meaningful information. Our approach minimizes data snooping given the information set and the priors. We compare the prior views of a sceptic and a confident investor [3].

Maitre, Bührer, Thomson and Stanski (1991) work on NONMEM does not provide the analyst with individual subject parameter estimates. As a result, the relationship between pharmacokinetic parameters and demographic factors such as age, gender, and body weight cannot be sought by plotting demographic factors vs. kinetic parameters. To overcome this problem, they devised a three-step approach. Specifically, for step 2 consists of individual Bayesian regressions using the measured drug concentrations for each subject and the population pharmacokinetic parameters obtained in step 1. The Bayesian parameter estimates of the individual subject can be plotted against the demographic factors of interest. From the scatter plots, it can be seen which are the demographic factors that appear to affect the pharmacokinetic parameters [4].

3. Methodology & Work Study

Predictive equations are very important tools for the pavement management systems. However, databases to support the developments and updating of these models are lacking. These databases are often inadequate in sample size, noisy, or incomplete. Conventional statistical modelling tools, such as classical regression analysis, may have limited success in these applications [5]. A promising solution lies in the use of Bayesian regression, which explicitly allows experts to be used to supplement poor quality data [6]. Bayesian regression methodology was adopted by the Canadian Strategic Highway Research Program (C-SHRP) for the Canadian Long-Term Pavement Performance (C-LTPP) monitoring program.

The interpretation and conclusion drawn from the experimental data can be quite different depending on what other evidence exists on the subject at hand. However, this difference in interpretation does not simply mean biasing a result. Interpretation of results using Bayes' Theorem is a mathematically consistent way to interpret evidence/ information [6].

In its simplest sense, Bayesian regression is a specialized adaption of the Bayes' theorem involving the development of multivariate regression models which explicitly consider two disparate sources of information:

1. A prior information, i.e. information that is known prior to an experiment
2. Experimental data, i.e. information that is derived from an experiment.

Table 1. Summary statistics (Monthly from 2013 to 2014)

Indices	CPI	UFA	Oil Prices	Exchange Rate	Interest Rates
85721,13	8,62	5,33	97,49	17713,00	8,23
78783,47	8,33	4,72	92,05	17778,00	7,97
79333,67	8,08	4,23	97,23	18111,00	7,76
85898,99	7,66	3,74	93,46	17991,00	7,67
86046,04	7,51	3,27	91,97	18301,00	7,56
85990,01	7,47	3,18	96,56	18999,00	7,53
76294,51	7,47	3,23	105,03	19342,00	7,63
73337,45	7,42	3,39	107,65	19635,00	7,86
66394,41	7,32	3,58	102,33	20210,00	8,17
74486,56	7,32	3,93	96,38	19913,00	8,45
77620,37	7,39	4,10	92,72	20267,00	8,83
75748,27	7,49	4,48	98,42	20667,00	9,21
67801,73	7,53	5,22	97,49	22272,00	9,35
61858,21	7,60	6,11	102,59	22139,00	9,63
62553,32	7,70	6,95	101,58	22191,00	9,91
69736,34	7,97	7,89	99,74	21294,00	10,20
73871,54	8,23	8,66	102,71	20934,00	10,48
79289,80	8,31	9,03	105,37	21210,00	10,38
78489,01	8,35	9,26	98,17	21230,00	10,27
82156,87	8,46	9,55	95,96	21634,00	10,12
80312,94	8,54	9,84	91,96	22130,00	9,80
74937,81	8,65	10,11	80,54	22587,00	9,51
80579,66	8,80	10,32	66,15	22380,00	9,36
86168,66	8,85	10,25	53,73	22969,00	9,34

The Bayes approach calculates a meaningful and credible answer without relying solely on a small database. In doing so, the Bayes technique allows decisions to be made in the short term while improvements to the data, judgement and the model continue to be made [6].

In assembling information for Bayesian regression, data collected in the traditional manner is supplemented with prior knowledge. The “prior” may be drawn from expert judgement, "old" data sets, or knowledge that is generally accepted in the field. Expert judgement can also be encoded by polling experts and asking them to estimate the value of the dependent variable for a combination of contributory variables. Once collected, the experts' observations are interpreted similarly to the traditional data.

4. Computational Study & Analysis

The mean and standard deviation under “node statistics” are the empirical posterior mean and standard deviation from the MCMC simulated posterior distribution for each parameter. These values are used as the estimate and standard error (respectively) of the associated parameters. Moreover, the val2.5pc and val97.5pc correspond to the 2.5 and 97.5 percentiles which can be used as the 95% credible interval. The “posterior density” plots are histograms of the same 16,001 draws.

Posterior summaries and densities, after running the MCMC algorithm for 14000 iterations, are provided in Table 2 and Figure 1 respectively.

Table 2. Node statistics

	Mean	SD	MC_Error	val2.5pc	Median	Val97.5pc	Start	Sample
Beta0	7,75	99,09	0,7506	-186,8	7,405	202,1	4000	16001
Beta1	68,68	99,83	0,7932	-126,2	67,900	263,6	4000	16001
Beta2	59,69	100,00	0,8233	-138,3	59,750	254,4	4000	16001
Beta3	655,60	54,02	0,7825	534,2	660,600	744,3	4000	16001
Beta4	167,20	99,87	0,9597	-28,9	167,500	362,1	4000	16001
Beta5	70,61	99,40	0,8506	-125,9	70,560	265,6	4000	16001
Mu	40950	3488,00	48,8800	32970,0	41340,00	46560,0	4000	16001
Tau	3,97E-06	1,58E-06	2,20E-06	1,35E-06	3,81E-06	7,52E-09	4000	16001

Descriptive analysis of the posterior distribution of R2B indicates a considerable improvement of the precision (posterior mean equal to 0.95) in the prediction of delivery times when including in the model covariates cases and distance.

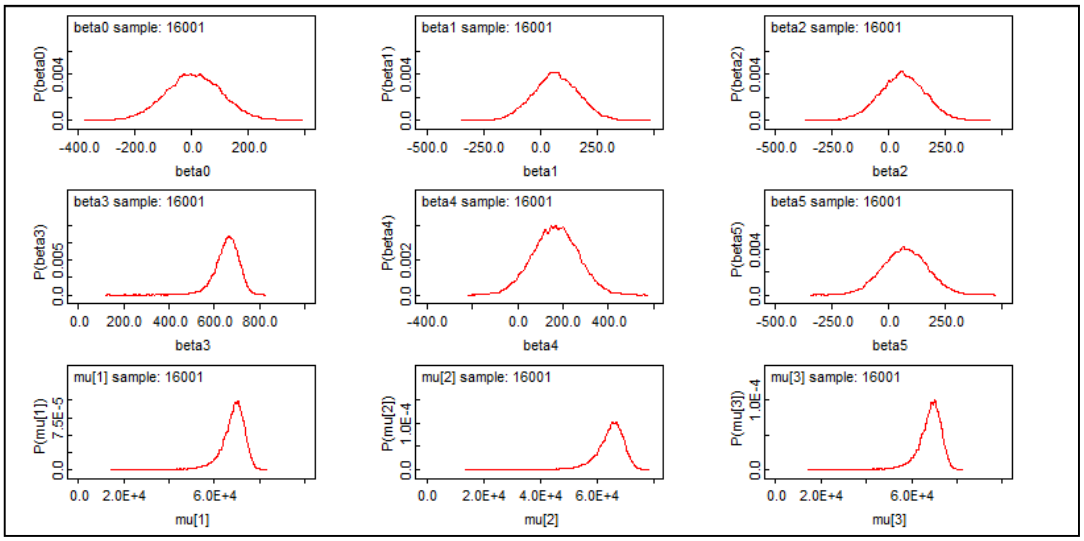

Figure 1. Posterior density

Considering as point estimates the posterior means, we end up with model;

$$Stock\ Share\ Exc. = \beta_0 + \beta_1 * CPI + \beta_2 * PPI + \beta_3 * OP + \beta_4 * EXC + \beta_5 * INTEREST \quad (1)$$

$$Stock\ Share\ Exc. = 7.8 + 68.8 * CPI + 59.7 * PPI + 655.6 * OP + 167.2 * EXC + 70.6 * INTEREST \quad (2)$$

Diagnostics for assessing convergence of the MCMC method can be done by examining the "history" plot of the samples at each iteration and looking for random scatter. We have found a consistent result from regression model by looking at auto-correlations graphs. Observing all parameters, we can infer that the effect of all explanatory variables (CPI, PPI, OP, EXC) have an important contribution to the prediction of BIST 100 index. However, the interest rate did not seem to have a significant effect on the BIST 100 index. The two-year evaluation made here can be expanded and examined in a wider range of different studies.

One of the ways that you can assess convergence is with visual examination of the trace plot, which is a plot of the sampled values of a parameter versus the sample number. Figure 2 displays trace plots to parameters. Beta0, Beta1, Beta2, Beta4 and Beta5 plots show good mixing. The samples stay close to the high-density region of the target distribution; they move to the tail areas but quickly return to the high-density region. Also, Figure 3 shows auto-correlations of Beta values.

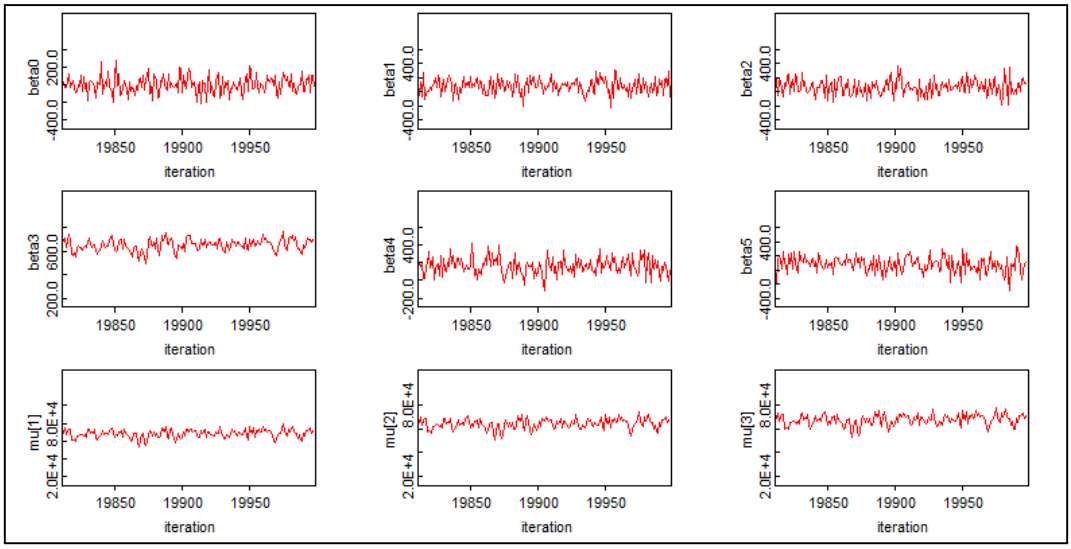

Fig 2. Dynamic trace of the data set

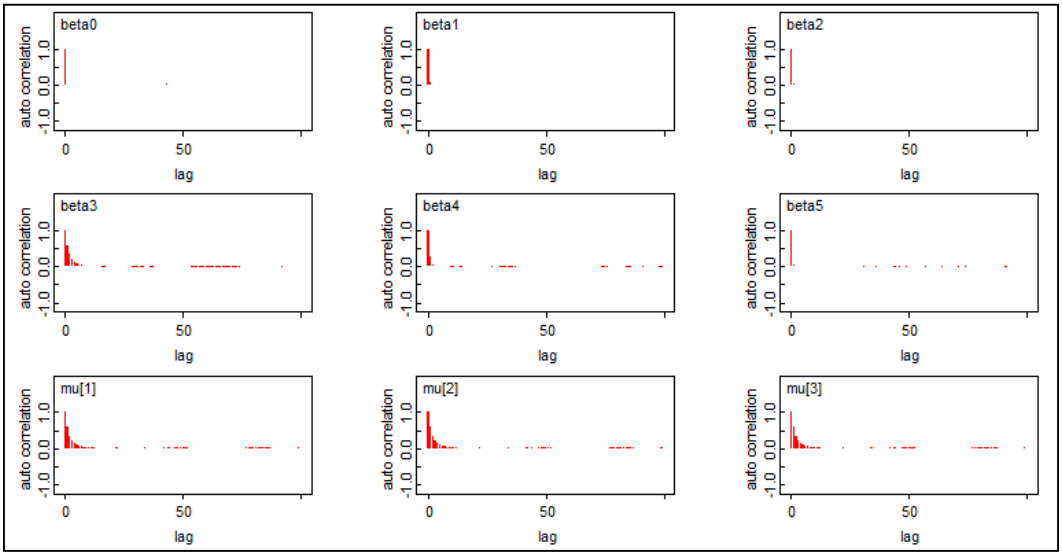

Fig 3. Auto-correlations

5. Conclusion

Minor changes are observed in the regression equation if posterior medians used as point estimates instead. Observing all parameters, we can infer that the effect of both explanatory variables (cases and distance) have an important contribution to the prediction of delivery time. All summary statistics and posterior densities indicate that zero is far away from posterior distribution with posterior probability of having positive association between each X_j and Y equal to one.

Furthermore, for each additional case determined by the result of BIST Index, the index is a posterior expected to increase by 655.6 points. The increase in expected index value for each additional case lies between 520 and 750 points with probability 95%. For every increase of consumer price rate by one point, the index expected to increase by 68.68. The increase in expected index value for each additional case lies between -200 and 270 points with probability 95%. For every increase of producer price index by one point, the index expected to increase by 59.69.

The increase in expected index value for each additional case lies between -150 and 260 points with probability 95%. For every increase of exchange rate by one point, the index expected to increase by 167.2. The increase in expected index value for each additional case lies between -120 and 260 points with probability 95%. For every increase of interest rate by one point, the index expected to increase by 70.61. The increase in expected index value for each additional case lies between -30 and 400 points with probability 95%.

Parameter β_0 has no sensible interpretation in this study since the zero value is non-sense for both explanatory variables. For this reason, no interpretation of this parameter attempted. We only observe that zero value lies at the left tail of the posterior distribution within the range of the 95% posterior interval.

Since the interpretation for β_0 is meaningless, we can focus on the predicted value for a typical or representative delivery route. According to the posterior summaries of node typical. *y*, a typical index point will take 66840 on average and will range from 54620 to 74380 with probability 95%.

References

[1] Bańbura, Marta, Domenico Giannone, and Lucrezia Reichlin. "Large Bayesian vector auto regressions." Journal of Applied Econometrics 25.1 (2010): 71-92.

[2] Chen, An-Sing, Mark T. Leung, and Hazem Daouk. "Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index." Computers & Operations Research 30.6 (2003): 901-923.

[3] Cremers, KJ Martijn. "Stock return predictability: A Bayesian model selection perspective." The Review of Financial Studies 15.4 (2002): 1223-1249.

[4] Maitre, P. O., et al. "A three-step approach combining Bayesian regression and NONMEM population analysis: application to midazolam." Journal of pharmacokinetics and biopharmaceutics 19.4 (1991): 377-384.

[5] Kajner, Lyle, Marian Kurlanda, and Gordon Sparks. "Development of Bayesian regression model to predict hot-mix asphalt concrete overlay roughness." Transportation Research Record: Journal of the Transportation Research Board1539 (1996): 125-131.

[6] Hossain, Mustaque, and Tanveer Chowdhury. Use of falling weight deflectometer data for assessing pavement structural evaluation values. No. K-TRAN: KSU-96-4,. NTIS, 1999.