Actuarial Modeling with Structural Break and Intervention - A Framework for Unexpected and Quick Change in Economic Environment

Aparna Dutta *Satadru Sengupta*

Abstract: In this paper we discuss the use of intervention and structural change in scientific modeling and possible application in the context of actuarial modeling. Application of the mathematical models in business context requires an accommodation for "what-if" conditions and its flexibility of incorporating sudden unexpected changes. This is a highly desired property that makes a model realistic and usable to a real world business situation successfully. Statistical tools for predictive models (e.g. Regression or GLM) are highly available in the market place and an ordinary use of such tools without following the data and the business case within can lead to myopic decisions and severe consequences. Today's vast and highly complex financial world deviates from the underlying conditions and assumptions of most of these models; at times these deviations are quite substantial and to an extent that the usability of such models becomes truly questionable. This paper is divided in four sections: the very first section discusses the conceptual foundation of modeling a dynamic economic scenario. The following two sections introduce intervention and structural break in classical time series and Bayesian dynamic linear models. Each of them gives a short theoretical description with text book reference along with an illustrative example for each methodology. The final section shows the possible applications of intervention modeling in an actuarial context trough two examples: one involving a legislative change in reserve analysis and other on property and casualty underwriting cycle. SAS and R routines and procedures are used for programming implementation.

1. Understanding Modeling Perspective and Structure

1. I. Overview of a Modeling Process

Understanding modeling perspective is imperative for a successful modeling. Essentially there are two possible perspectives of any modeling technique viz. Learning and Prediction or Forecasting. Oxford English Dictionary provides a definition of scientific modeling: "simplified description of a system that assists calculations and predictions". This lexicographic definition matches the practical purpose of any modeling techniques. A scientific model accumulates available information and history; creates a framework for understanding the underlying system and if needed helps forecasting the outcome or state of the system in future. Without imposing a proper structure it is not possible to build and use a scientific model and in this section we will identify this structure. We need to understand that it is impossible to "predict" or "quantify" the true state of any unknown system in nature (Singpurwalla [1]) and that is not the purpose of modeling. An ideal model helps understanding a system better and helps in taking better decision in future through a systematic accumulation of history and by creating a plan of actions. In financial scenario we often encounter a dynamic system and additional properties are required in order to model such dynamic financial scenarios. Skeleton of such models should not alter frequently just to accommodate the dynamism in the system. It should have an adequate structure to accommodate continuous routine changes without hurting the basic conceptual building blocks of the model. Within this framework the model performance will be monitored and that will suggest any possible changes in the model and also a comparison among possible alternative models based on a goodness criterion. In case there is an evidence of inadequacies in the model a change in higher level will take place. One major thing that can happen to a financial scenario is subject of this paper: introducing non-routine extrinsic information into the model that will reflect a major structural change in the system under study. This is precisely where intervention or structural break comes in the picture; we will discuss the methodologies in rather detail in subsequent sections. Rest of this section is dedicated to formalizing the aforesaid modeling structure.

1. II. Model Forms: Conceptual, Qualitative and Quantitative

A sound model structure is comprised of three components: the conceptual basis (C), the qualitative form (F) and the quantitative form (Q). A detailed study is available in West and Harrison [2] and we are discussing a gist here. We can formulize the model structure using a triplet M = [C, F, Q]. This structure provides the desirable properties of a model viz. description, control and finally, robustness.

The conceptual basis (C) of the model is an abstract formulation of the model. This is where the model perspective and objectives are discussed and stated. For example, with the last 30 years of underwriting (UW) cycle (P & C) data the management at an insurance company can think of different objectives. When will be the next hard market? What drives the cycle? What should be the strategy to anticipate the UW Cycle? As we can see modeling is not necessarily predicting the future; it is learning a system to decide actions, anticipating business environment and of course prediction. And this perspective should be very clear at the conceptual level of the modeling and should be ideally unchanged for the entire process. Alternative models are based on the same conceptual level to ensure we are not comparing an apple with an orange at any stage.

The qualitative form (F) transforms the abstract form of conceptual basis to a mathematically-transferable form. At this stage of the modeling analyst will decide what would be the set variables used, how much data should be used, what are the external information available and of course exploring relations among the variables in the data. The set of variables and their relationship could be expressed in any form: in diagram, graphical, tabular, and algebraic. With the above UW cycle example: at this stage with the management objective and perspective an analyst will decide how to answer the question asked. What would be a suitable target variable? What are the explanatory variables? What are their relationships? What is the modeling technique that answers the question best? Different competing models will have the same qualitative form.

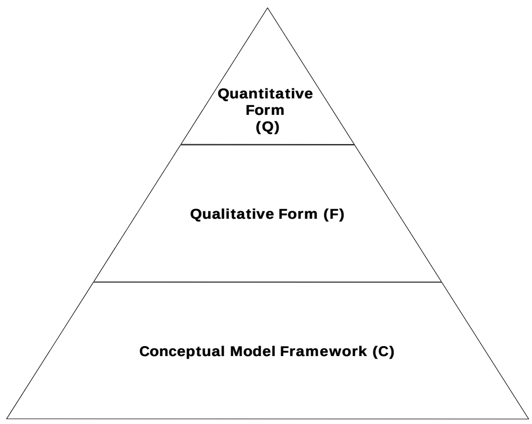
The quantitative form (Q) assumes values for parameters in the model form that's been selected in qualitative form. The values of the parameter changes with a changing set of information and along the history. With a fixed conceptual and qualitative form, these changes in parametric values are expected. Again, with our UW cycle data, if graphical presentation of the data suggests that an Regression model with Interest Rate (X1), GDP growth (X2) and CPI (X3) as explanatory variables is a suitable model (everything remain the same), then at this stage goal is to find the coefficients for the three explanatory variables. As new data comes, the coefficient changes keeping the qualitative form the same. Now the residual structure may suggest that the model is inadequate; an analyst then would go back to the qualitative level to think of a different modeling form (may be a time series model or a different set of predictors or theoretically any model under the sun). A change may come even at the conceptual level (only as an exceptional case) when management decides to have a change in objective or perspective (this could be a result of the unsupervised or supervised exploration of the data, due to the ease of answering one question over other or just because of a change in management strategy).

1. III. Model Description, Control and Robustness

Description of a model is a mathematical manifestation of the entire process; it is a complete, meaningful and readable representation of the model. There are two major required properties in the model description: parsimony and perspective. Parsimony makes a model simpler and canonical without losing any available information. Perspective talks about model's power and limitation. Description of a model enables the management with a set of action plan through a quantitative and qualitative understanding of the system.

Historically, the matrix of the realized predictive variable is called design matrix. The reason behind such a naming is it actually design the entire study based on its perspective. In the complex financial environment an insurance company gets affected by many macro level extraneous factors on which it hardly has any control. By just predicting the state of such factors the management will not get any competitive advantage. An alternative objective could be exploring the reactions and sensitivity of the organization to an adverse economic development in the financial environment where it resides. This is precisely the phase of "what-if" analysis. By controlling the design matrix in the model the management may seek to find an optimal strategy to anticipate the adverse environment. The UW cycle example is an extremely appropriate example in this context. As an insurance company it is not possible to set a hard market or soft market. So just by predicting a hard market the insurance company will not be able to improve its position in the market. With the 30 years of UW cycle data a better perspective would be finding an optimal strategy to anticipate the hard market and exploring a fairly large set of "what-if" consequences.

Main purpose of inducing robustness in modeling is to use intervention effectively and cost efficiently. This is accomplished by optimally defining the triplet M = [C, F, Q] so that at the time of exceptional changes in the system the model triplet will get minimally affected without sacrificing any information related to the changes. With robustness a drastic change in quantitative level will not make any effect at the qualitative level. A major change in qualitative level will not affect the conceptual level. Additionally, robustness is required for the components within a level. Intervention may be relevant only for one or few components; not for all the components in a level. Bringing intervention to a required component shouldn't disturb others that have been stable. In our UW Cycle example, there might be a major policy change which only affects the corresponding component without affecting the others e.g., price elasticity, catastrophic component or CPI component. Robustness assures this desired smooth transition in the model.



Building this understanding of a modeling framework was extremely important; however it has been a digression from the main topic of this paper. In the following sections we will introduce and discuss intervention and structural break in classical time series model and in Bayesian dynamic linear model. In the final section we will present two insurance scenarios where we can successfully apply intervention and structural break.

2. Structural Break and Intervention in Classical Time Series Model

2. I. Overview of Structural Break

Structural break in an economy refers to a fundamental change in the structure of the economy. This may be due to fiscal or monetary decision laid down by the government or solely an external factor. The effect of such a structural change can be seen as an "unusual" shift in the time series data. Simply stated, structural change means a situation where at least one of the underlying parameters has changed at some date called the **break date**.

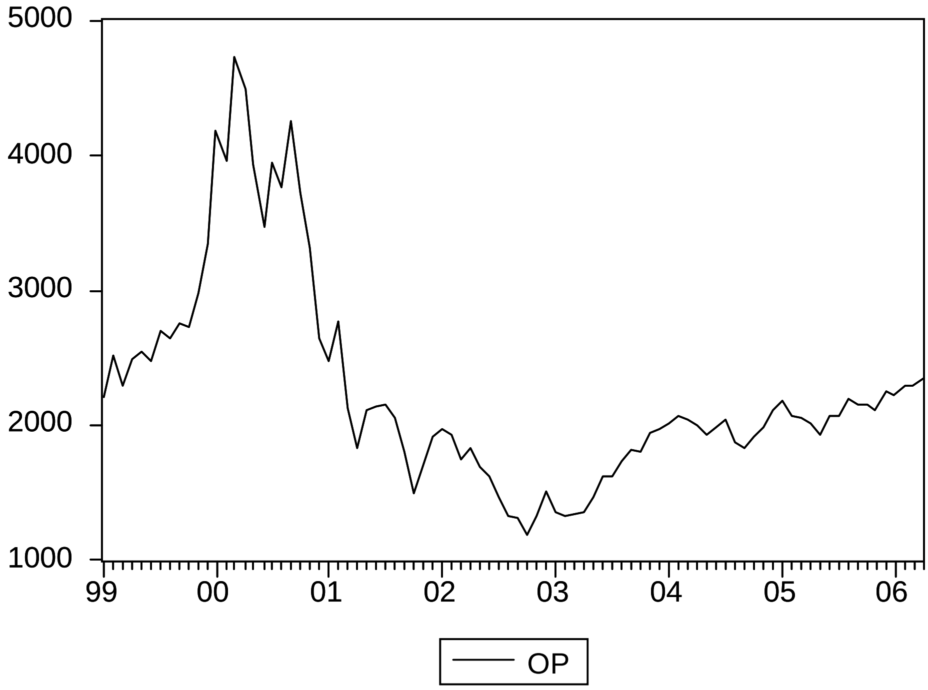


Figure: 2

Above graph plots the NASDAQ Composite Monthly Open across the years 1999 to 2006. The series has been following a fluctuating pattern. However by the end of 2002 there is a change in the pattern; graphical description of the index clearly suggests that there is a structural break by the end of 2002. The series does not have a mean reverting property so the series can be said as non-stationary. Question remains: how one can statistically test the break point.

2. II. Testing of Structural Break

Chow test: To statistically test for Structural break the most popular test used is Chow test where the data is divided into two sub periods or regimes (say (t_1,T) and (T,t_2) under the assumption that the break date is known or can be guessed and realized as T) and test for the equality of the parameters in the two regimes. Suppose that we model our data as

$$y_t = a + bx_{1t} + cx_{2t} + \varepsilon.$$

If we split our data into two regimes, then we have

Regime I (t_1,T): $y_t = a_1 + b_1x_{1t} + c_1x_{2t} + \varepsilon.$

Regime II (t_2,T): $y_t = a_2 + b_2x_{1t} + c_2x_{2t} + \varepsilon.$

The Chow test asserts that $a_1 = a_2$, $b_1 = b_2$, and $c_1 = c_2$.

Let S_C be the sum of squared residuals from the combined data, S_1 be the sum of squares from the first group, and S_2 be the sum of squares from the second group. N_1 and N_2 are the number of observations in each group and k is the total number of parameters (in this case, 3). Then the Chow test statistic is

$$\frac{(S_C - (S_1 + S_2))/(k)}{(S_1 + S_2)/(N_1 + N_2 - 2k)}.$$

The test statistic follows the F distribution with k and ($N_1 + N_2 - 2k$) degrees of freedom.

Even though Chow test is the most commonly used for testing for Structural break it suffers from 2 very important shortcomings. Firstly, the assumption that the break point is known or simply that it can be guessed and secondly that the variance of the error term in both the regimes are equal.

2. III. Overview of Intervention Analysis

Most important question at this point is the reason behind a structural break and what would its effect and consequence. Definitely a set of extraneous factors causes a structural break and due to which the primary variables and few secondary variables are affected. We need to incorporate these changes in our model following the model structure that we have discussed in Section I in detail. If the exact timing of this structural change is known then the underlying analysis is known as intervention analysis, that is, there are known changes that affect the dependent series or outliers.

We start with the following time series model: $Y_t=\Psi(B)*X_t^{(T)} + \epsilon_t$ where T is the time point at which intervention has occurred and Y_t is the observation at the t^{th} time point.

With a transfer function form, $\Psi(B)$ can be expressed as:

$\Psi(B)= \omega(B)/\delta(B) = (\omega_0 - \omega_1 B - \omega_2 B^2 - \omega_3 B^3 \ldots . - \omega_s B^s) / (1 - \delta_1 B - \delta_2 B^2 - \delta_3 B^3 - \ldots\ldots - \delta_q B^q)$

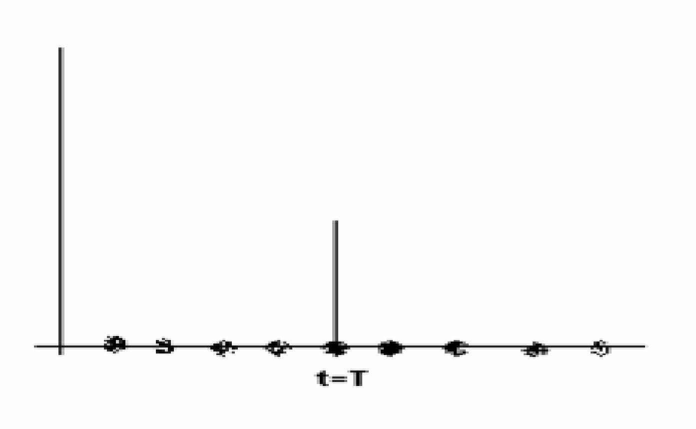
$\epsilon_t = (\theta(B)/\Phi(B))a_t, \; a_t \sim WN(0,\sigma_a^2)$ and B is the backshift operator: $By_t = y_{t-1}$

$X_t^{(T)}$ is the Intervention/dummy variable that can be of the three types:

Pulse Dummy (Point Interventions)

The point intervention is a one-time event of the form:

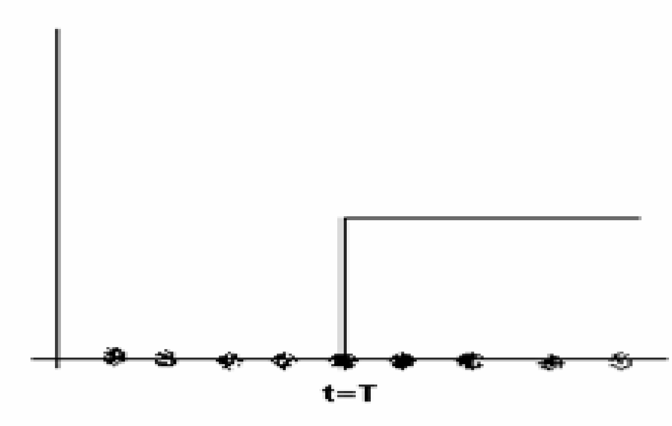
$$X_{i,t} = \begin{cases} 1, & t = t_{int} \\ 0, & otherwise \end{cases}$$



Step Dummy (Step Interventions)

Step interventions are continuing, and the input time series flags periods after the intervention and of the form:

$$X_{i,t} = \begin{cases} 1, & t \geq t_{int} \\ 0, & otherwise \end{cases}$$



Ramp Interventions

A ramp intervention is a continuing intervention that increases linearly after the intervention time, that is, proportional to time. It has the following form:

$$X_{i,t} = \begin{cases} t - t_{int}, & t \geq t_{int} \\ 0, & otherwise \end{cases}$$

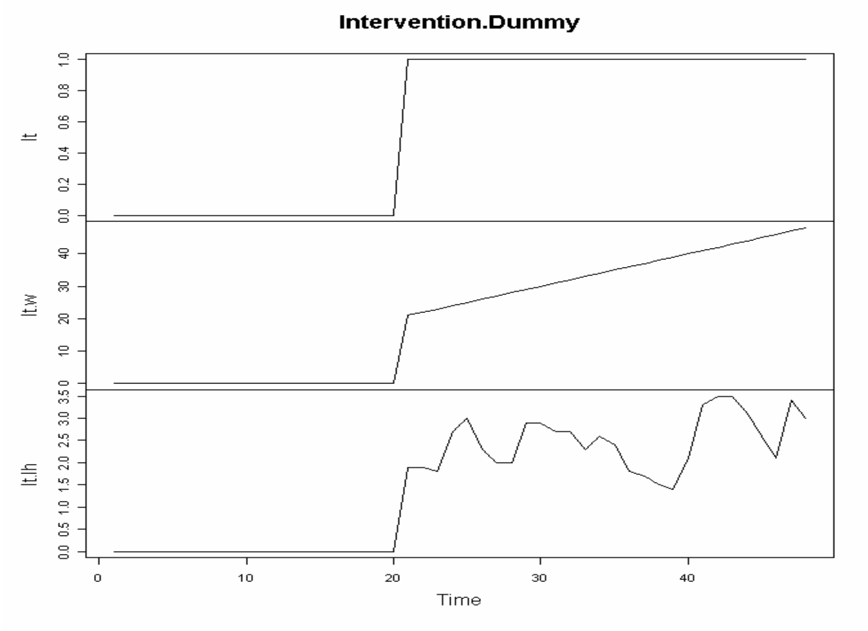
2. IV. A ready-to-do program implementation of a Time Series with Intervention

This subsection will show a routine using statistical software R that will fit a time series with intervention to a popular time series data available in R. This data is taken from P.J. Diggle (1990) Time Series: A Biostatistical Introduction. Oxford, table A.1, series 3. This is a regular time series that records the luteinizing hormone in blood samples at 10 minutes intervals from a human female, 48 samples. R base has the data namely "lh". This example will help reader to understand the different type of interventions and their application in a time series model. This program is self-contained and don't need any additional library to run in R. The time series simulated has a 48 data points and an intervention takes place at T = 20.

Code:

```
lh
plot.ts(lh)

Intervention.Dummy <- cbind(It=(1:48)>20,
                     It.w=((1:48)>20)*(1:48),
                     It.lh=((1:48)>20)*c(0, lh[-48]) )
Intervention.Dummy
plot.ts(Intervention.Dummy)

arima(lh, order = c(1,0,0), xreg=Intervention.Dummy)
```

Selected Output:




```
arima(x = lh, order = c(1, 0, 0), xreg = Intervention.Dummy)
Coefficients:
ar1           intercept            It           It.w          It.lh
0.3611  2.2527        -1.1383 0.0147  0.3589
Standard errors
0.1636  0.1429        0.6128  0.0154        0.2060
sigma^2 estimated as 0.178:  log likelihood = -26.76,  aic = 65.51
```

2. V. A Complete Financial Case Study on Time Series with Structural Break

Objective and Introduction to data: In this subsection we use the Nasdaq composite, monthly open series data. NASDAQ (originally an acronym for National Association of Securities Dealers Automated Quotations) is an American electronic stock exchange. On July 17, 1995 the NASDAQ stock index closed above the 1,000 mark for the first time. The index peaked at 5132.52 on March 10, 2000, which signaled the high point of the dot-com bubble. Within a year it had declined to less than half of its peak value. Here NASDAQ monthly opening data has been considered for the years 1999 to 2006 and has been plotted in Figure 2. The objective of the study is to analyze the data and show that the dot-com crisis has brought a structural break in the NASDAQ data.

The Dot-Com Crash (March 11th, 2000 to October 9th, 2002): The NASDAQ Composite lost 78% of its value as it fell from 5046.86 to 1114.11. The stock market downturn of 2002 (some say "stock market crash" or "the Internet bubble bursting") is the sharp drop in stock prices during 2002 in stock exchanges across the United States, Canada, Asia, and Europe. After recovering from lows reached following the September 11, 2001 attacks, indices slid steadily starting in March 2002, with dramatic declines in July and September leading to lows last reached in 1997 and 1998. The dollar declined steadily against the euro, reaching a 1-to-1 valuation not seen since the euro's introduction.

Stationarity: The first step while modeling any time series data is to check the stationarity of the series because of the following reasons-

- o The stationarity of a series can strongly influence its behavior and properties. For example, if the NASDAQ composite open data is stationary then shocks to the system will gradually die away. That is, the effect of a shock during time t will have a smaller effect in time t+1, a smaller effect still in t+2, and so on. This can be contrasted with any non-stationary series, where the persistence of the shock will be infinite, so that for a non-stationary series, the effect of time at shock t will have a similar effect in time t+1, and in time t+2, etc.
- o The use of non-stationary data can lead to spurious regression.
- o If the variables employed in the model are not stationary then the standard assumptions of the data analysis will not be valid. In other words, the usual 't-ratios' will not follow t-distribution, and the F-statistic will not follow F-stationary.

To test for existence of unit roots we consider the Dickey Fuller, Augmented Dickey Fuller test and Phillips-Perron test. The data by itself is found to be non-stationary, hence we transform the data by taking logarithm and then differencing it. Also both drift and trend for the transformed data appear to be insignificant. We confirm the stationarity of the transformed series by the following graph

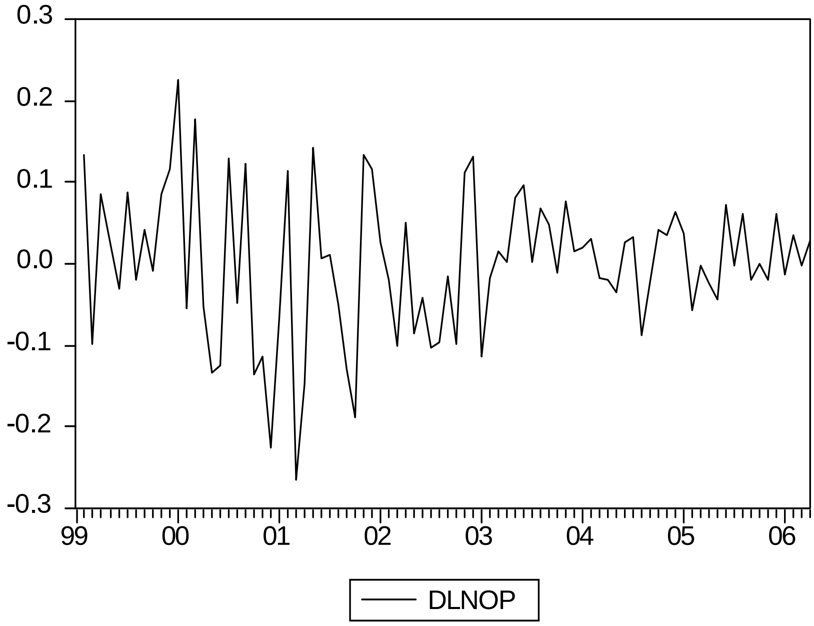


Figure: 3

The graph suggests that the series is now stationary around 0. We also check for seasonality in the series and conclude that there exists no seasonality in the data.

Estimation the model: In order to estimate the model **Box-Jenkins approach** is used and the following steps are followed.

Identification-this involves determining the order of the model required to capture the dynamic features of the data.

Estimation-This involves estimation of the parameters of the model specified in step1.

Diagnostic checking-This involves model checking-that is determining whether the model specified and estimated is appropriate.

Box and Jenkins suggests two methods: over fitting and residual diagnostics. **Over fitting** involves deliberately fitting a larger model than that required to capture the dynamics of the data as identified in stage1. **Residual diagnostics** imply checking the residuals for the evidence of linear dependence which, if present, would suggest that the model originally specified was inappropriate to capture the features of the data.

Over fitting: Here we can use AIC, BIC, AICC or Hall's Procedure to get the suitable model (if have more than one model).

Diagnostic testing with the residuals: Ljung Box Test

Q-Statistic:The Q-statistic at lag k is a test statistic for the null hypothesis that after fitting the model the error are white noise or not (white noise process has a constant mean and variance, and zero autocovariances, except at lag 0) thus Ljung-Box tests whether the noise term is uncorrelated or not and is computed as

$$Q_{LB}(k) = n(n+2)\sum r_j^2 / n - i$$

where i =1 (1) k, where r_i is the j-th autocorrelation , n is the number of observations and k is the maximum lag length. If the series is not based upon the results of ARMA estimation, then under the null hypothesis, Q is asymptotically distributed as a χ^2 with degrees of freedom equal to the number of autocorrelations. If the series represents the residuals from ARMA estimation, the appropriate degrees of freedom should be adjusted to represent the number of autocorrelations less the number of AR and MA terms previously estimated (k-p-q). There remains the practical problem of choosing the order of lag to use for the test. If you choose too small a lag, the test may not detect serial correlation at high-order lags. However, if you choose too large a lag, the test may have low power since the significant correlation at one lag may be diluted by insignificant correlations at other lags. For the given case study, i.e. the transformed data ARMA(2,2) is found to be a good fit model for the series. We also check if the residuals for the ARMA(2,2) model are white noise using Diagnostic test and confirm the selection of the model.

Testing for Structural Break:

We use a sequence of Chow tests to ensure that the limitations of Chow test is restricted and make more realistic assumptions such as the break date is unknown, there is only one break date and the variance in the two regimes are different. That means we test for a sequence of break points, in the sense of n* being all possible break dates leaving out some periods at both ends. To test the hypothesis that there is no switch the appropriate log likelihood ratio (Wald Statistic) is:

LR = -2(maximized Log-liklihood ratio under constraint – maximized Log-liklihood ratio unconstrained)*

This is the Quandt Andrews test. We plot the sequence of Wald's statistic (or LR or LM statistic) as a function of candidate break date. The candidate break dates are along x-axis and the values of the log likelihood ratio on y-axis.

After plotting these values we check whether the maximum of Wald's statistic (or LR or LM) lies above the Andrew's critical value, if it lies above the critical value we conclude that we have a break.

If break is detected for the series then we estimate where the break date has occurred. We divide the sample in two sub periods, at each possible break date and calculate the residual sum of square for the sub periods and store it. The break date estimate is the date at which the full sample sum of square error is minimized. That is calculating RSS1 and RSS2, take the break date at which RSS1+RSS2 is minimum for the entire sample, through this procedure we get the estimated break date. Here Wald statistic (W=Nx(Rss-Rss1-Rss2)/(Rss1+Rss2) where N is the total number of observations) is used and compared with Andrew's critical value. 15% of the data have been left on both sides. We find that there exists a structural break as the Wald statistic exceeds the Andrew's critical value at 1% level of significance.

We plot the value of the statistic

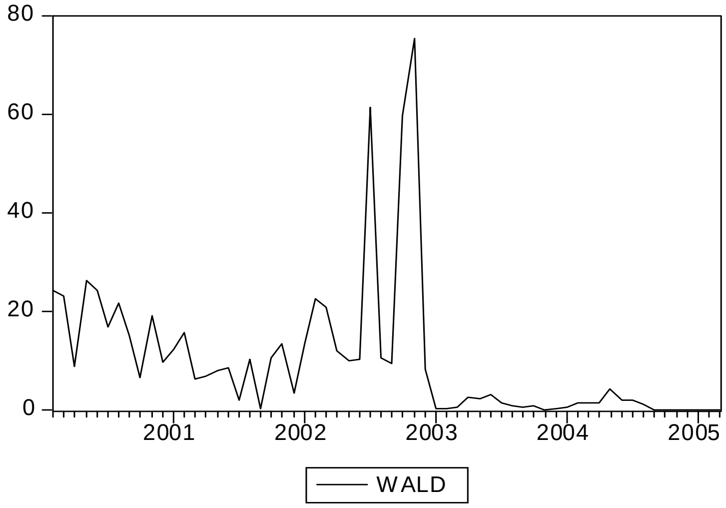


Figure: 4

To estimate the break date we subdivide the data and calculate the value of rss1 and rss2 of the sample and candidate break date would be the date where the sum is the minimum.

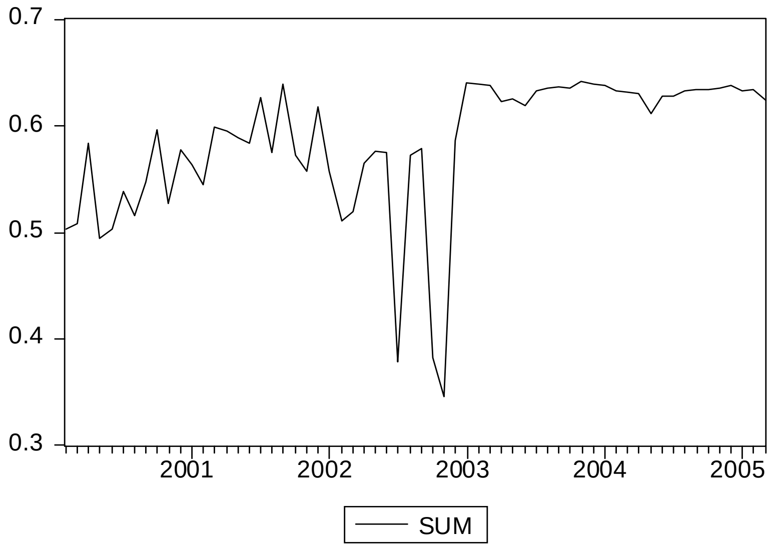


Figure: 5

This suggests that the minimum is at 2002:11.Hence there exists a structural break at 2002:11. This validates our conception that given a strong external factor the time series data does go through a structural change and it is very important to incorporate it in the model. The next step is test for unit root and to estimate the model for the two sub-periods. A complete list of final conclusions from the case study can be listed as:

1. There is non-stationarity at level values
2. There is no seasonality in the data
3. Stationarity is attained at difference of log value series
4. The best model estimated is ARMA(2,2)
5. There exists a structural break at 2002:11
6. The model changes for the second sub-series (2002:12-2006:04)

3. Concept of Intervention in Bayesian Dynamic Linear Model

3. I. An Overview of Modeling in by Levels

In the second section the idea of structural change and intervention was introduced and discussed in a classical time series set up. A hand on programming implementation was given; also a detailed case study on a financial scenario was shown. In this section we will introduce Bayesian dynamic linear model and will show how intervention along with other subjective information can be used within a mathematical model to incorporate an extraneous change.

Bayesian dynamic linear model (DLM) is a natural extension of the classical state-space model in time series. In the state-space model data generation happens through an underlying system which changes its state but remain unobserved directly. Time series data are observed as a function of the underlying state model. There are classical ways to estimate and forecast a state-space model of which Kalman filter is the most popular technique. It has a vast application in Finance. A good study of Kalman filter can be found in Shamway and Stoffer [3]. Bayesian DLM has almost the same structure as that of different classical state-space models; major difference, which is obvious, is its analysis through Bayesian inference. The most important usefulness of this kind of modeling (hierarchical, state-space or DLM) is that it helps understand the system in layers or in components. It definitely gives a better understanding of the system; moreover, in nature most of the systems are layered in nature (with some observable and some unobservable components). Bayesian aspect adds some more advantage by allowing the input of extrinsic and subjective information in the model in a formalized manner.

In the rest of the section we will introduce Bayesian DLM and will discuss the functionalities of various components of the model. This discussion is based upon the literature by West and Harrison[2] which gives an great detail of the subject.

3. II. Bayesian Dynamic Linear Model (DLM): Formulation

We will start with the mathematical form of the dynamic linear model and then will discuss the meaning and interpretations of different components in the model. A Bayesian DLM can be expressed as a quadruple $[F_t, G_t, V_t, W_t]$ with the following specifications:

Observation Equation: $Y_t = F_t \theta_t + v_t$ with $v_t \sim N[0, V_t]$
System Equation: $\theta_t = G_t\theta_{t-1} + w_t$ with $w_t \sim N[0, W_t]$
Initial Information: $(\theta_0 \mid D_0) \sim N[m_0, C_0]$

The system equation above evolves over time with its defined specification and generates the observation equation accordingly. The observation equation is expressed through a regression form with regression factor F_t, and a normal error v_t,. The system equation is also expressed in a regression form with evolution matrix G_t, and normal error w_t,. The process initiates with prior information defined in as initial information. These normal errors suggest the "linearity" in DLM while the system equation brings the "dynamic" characteristic the model.

Different type of DLM can be formed by changing F_t, and G_t. We will show few form and will keep our discussion limited to the constant F_t and G_t which will give local level polynomial model. Interested readers can look at West and Harrison [2] for a detailed formulation of different DLMs.

Model Specifications: Different Types of DLM

The local level model is specified for a unit value for both regression factor and evolution factor.

$$[F_t = 1, \ G_t = 1]$$

The corresponding DLM will be specified as:

Observation Equation: $Y_t = \theta_t + v_t$ with $v_t \sim N[0, V_t]$
System Equation: $\theta_t = \theta_{t-1} + w_t$ with $w_t \sim N[0, W_t]$
Initial Information: $(\theta_0 \mid D_0) \sim N[m_0, C_0]$

The first order polynomial model will look like:

Observation Equation: $Y_t = \theta_t + v_t$ with $v_t \sim N[0, V_t]$

System Equation: $\theta_t = \theta_{1,t-1} + \theta_{2,\ t-1} + w_{1,t}$ with $w_{i,t} \sim N[0, W_t]$

$\theta_{2,t} = \theta_{2,t-1} + w_{2,t}$

This organic extension is fascinating and can be used for more complex models. For illustration purpose we will keep ourselves confined in the constant model or local level polynomial model. Further advancement can be done to the model by introducing seasonality component through the regression factor F; we strongly encourage an interesting reader to look at West and Harrison[2] for a complete detail.

3. III. Algorithm for Estimation and Forecasting

With the above generalized description of Bayesian DLM, we will confine ourselves now only on local level polynomial model corresponding to a unit regression coefficient and same evolution factor. Before we present an algorithm for the Bayesian updating, forecasting and estimation lets recall the model again:

Observation Equation: $Y_t = \theta_t + v_t$ with $v_t \sim N[0, V_t]$

System Equation: $\theta_t = \theta_{t-1} + w_t$ with $w_t \sim N[0, W_t]$

Initial Information: $(\theta_0 \mid D_0) \sim N[m_0, C_0]$

We assume, without loss of any generality, that we have started our observation about the system at time t = 0 and at that point of time we have information D_0. As we go over the time, we gather more information and update the information set. Essentially at any time point t > 0 we can express the available information as: $D_t = [D_{t-1}, Y_t]$.

Here is a 4-step algorithm for updating and forecasting future state and observation of the system through recursive formulae:

1. Posterior for θ_{t-1}

$(\theta_{t-1} \mid D_{t-1}) \sim N[m_{t-1} , C_{t-1}]$

2. Prior for θ_t

$(\theta_t \mid D_{t-1}) \sim N[m_{t-1} , R_t]$ with $R_t = C_{t-1} + W_t$

3. 1-Step Forecast

$(Y_t \mid D_{t-1}) \sim N[f_t , Q_t]$ with $f_t = m_{t-1}$ and $Q_t = R_t + V_t$

4. Posterior for θ_t

$(\theta_t \mid D_t) \sim N[m_t , C_t]$ with,

$m_t = m_{t-1} + A_t e_t$ and $C_t = A_t V_t$

where,

Adoptive Coefficient: $A_t = R_t / Q_t$ and
1-Step Forecast Error: $e_t = Y_t - f_t$

3. IV. The Constant Model: Notion of Adoptive Coefficient and Intervention

Simply put, constant model is the model where the observation variance (variance corresponding to the observation equation) and the signal variance (corresponding to the system equation) remain constant over time. The constant $r = W / V$ measures the system variance relative to the observation variance and referred as signal-to-noise ratio. With this variance structure this models doesn't include any external information in its information set $D_t = [D_{t-1}, Y_t]$ in time t and in that sense its called "closed" model. But with Bayesian underpinning it is very easy to add subjective information with the data information. Step 4 of the Bayesian update algorithm will allow the analyst do an over-ride on the posterior variance based on external event that is significantly different from the model experience.

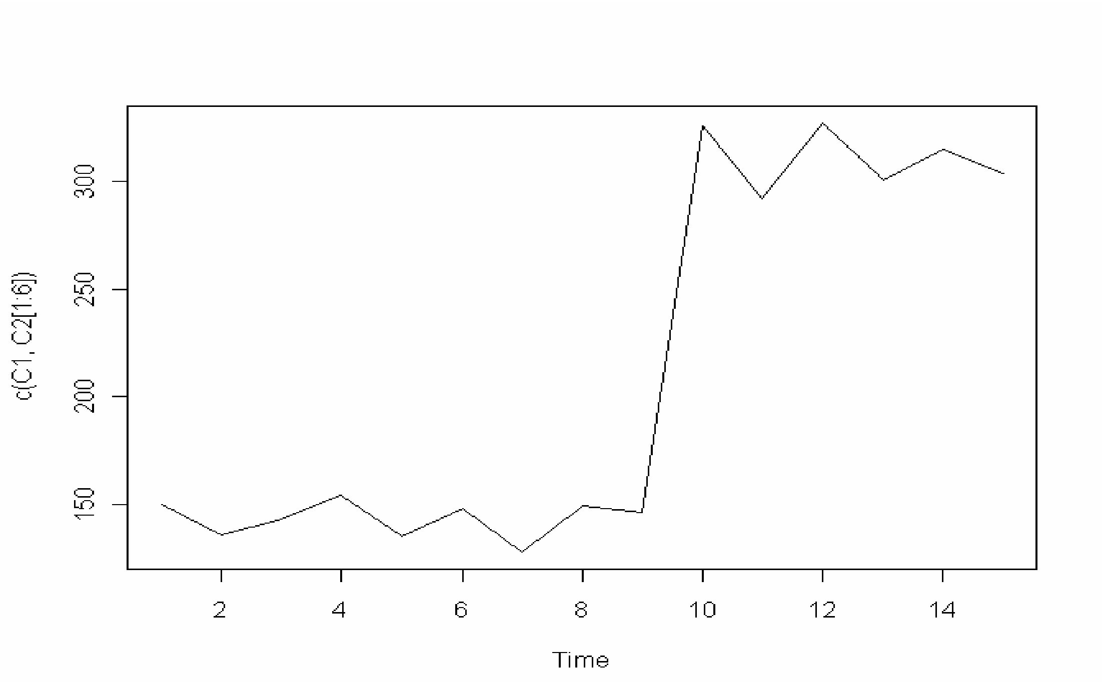
We will discuss the notion intervention and adaptive coefficient with a practical example from West and Harrison [2].

We will start with a data set Y = [Y1,Y2] where:

Y1 = 150 136 143 154 135 148 128 149 146 *< data from t=1 to 9 >*

Y2 = 326 292 327 301 315 304 *< data from t =10 to 15 >*

Time series plot for Y is:



Graph suggests that the level of the time series got changes at t = 9. Based on the background of this data we write down the constant Bayesian DLM with intervention to treat this time series data for forecasting.

At t = 0: $(\theta_0 \mid D_0) \sim N[130,400]$ < a prior information from the expert >
With this information for Y1 we have the following model:

$$Y_t = \theta_t + v_t \quad \text{with } v_t \sim N[0,100]$$
$$\theta_t = \theta_{t-1} + w_t \quad \text{with } w_t \sim N[0,5]$$
$$(\theta_0 \mid D_0) \sim N[130,400]$$

Lets recall the updating algorithm:

1. Posterior for θ_{t-1}
$(\theta_{t-1} \mid D_{t-1}) \sim N[m_{t-1}, C_{t-1}]$

2. Prior for θ_t
$(\theta_t \mid D_{t-1}) \sim N[m_{t-1}, R_t]$ with $R_t = C_{t-1} + W_t$

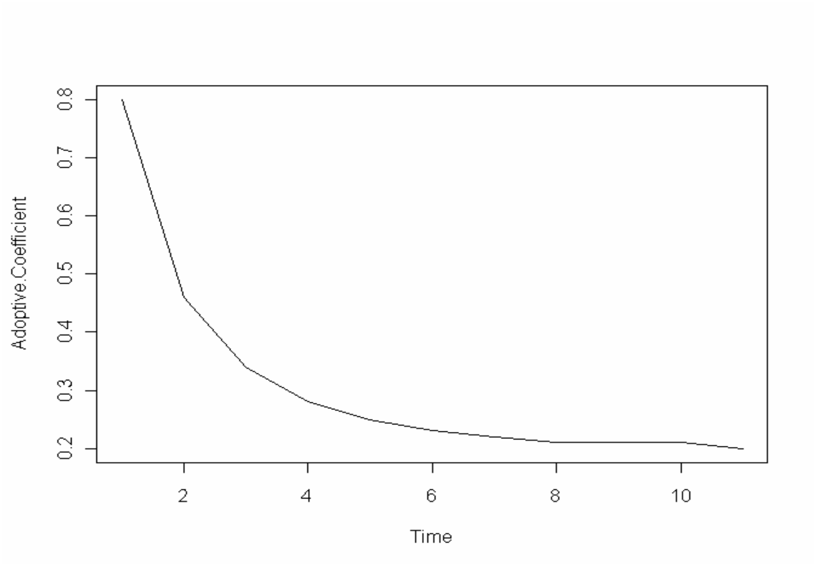
3. 1-Step Forecast
$(Y_t \mid D_{t-1}) \sim N[f_t, Q_t]$ with $f_t = m_{t-1}$ and $Q_t = R_t + V_t$

4. Posterior for θ_t
$(\theta_t \mid D_t) \sim N[m_t, C_t]$ with,
$m_t = m_{t-1} + A_t e_t$ and $C_t = A_t V_t$
where,

Adoptive Coefficient: $A_t = R_t / Q_t$ and
1-Step Forecast Error: $e_t = Y_t - f_t$

From the above recursive formula we observe that:
$m_t = A_t Y_t + (1-A_t)m_{t-1}$

Interpretation of the above equation is: the adoptive coefficient A_t is the prior regression coefficient of θ_t upon Y_t. This credibility looking formulae actually holds similar notion of that of credibility. Adoptive coefficient assumes value between 0 and 1 and it is closer to 0 when $R_t < V_t$ suggesting that the prior distribution is more concentrated than the likelihood. In the example we used for Y1, we have the following graph for the adoptive coefficient:



Adoptive Coefficient Graph

Immediate observation from this graph is: as more and more observations are received the adoptive coefficient drops from 1 towards a value closer to 0 (in deed converges to 0.20 in the example used). Its reducing value suggests more confidence to the coming observation or likelihood. With the inclusion of an intervention at point t = 9, the analyst will make correction to the system equation through subjective information or prior for θ_t and this will result in further increase of the adoptive coefficient. The equation for Y2 is:

$$Y_t = \theta_t + v_t \qquad \text{with } v_t \sim N[0,100]$$
$$\theta_t = \theta_{t-1} + w_t + \Delta\theta_t \text{ with } w_t \sim N[0,5]$$
$$(\Delta\theta_{t=10} \mid D_9 \text{ , Intervention Information}) \sim N[143,920]$$

Similar to the choice of the prior at t = 0, at this intervention stage also, the values of the normal parameters will be decided by the experts based on all information available in the environment of the system (here the environment suggests a 100% increase (=143) in the time series outcome and that added to the model at t=10).

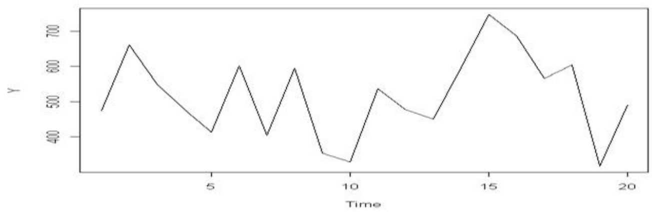
With this corrected information in prior for θ_{10} the analyst now can follow the usual updating and forecasting 4 step algorithm to get the forecast equation for $t > 10$.

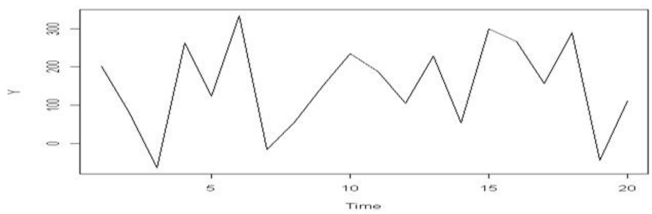
3. V. A ready-to-use simulation of Bayesian DLM

Here is a simple ready-to-use program implementation to generate a Bayesian DLM process. Reader can change the value of observation variance V and signal variance W to realize different simulations based on different signal-to-noise ratios. We have shown two simulations with a same signal-to-noise ratio which is set to be 0.05 here. This is the already discussed text book example from West and Harrison [2].

```
V <- 100
W <- 5
signal.to.noise.ratio <- W/V
mu.0 <- rnorm(1,130,400)
mu.t<-NULL
Y<-NULL
mu.t[1] <- mu.0
Y[1] <- mu.t[1] + rnorm(1,0,100)

for(i in 1:19){

      mu.t[i+1] <- mu.t[i] + rnorm(1,0,W)
      Y[i+1] <- mu.t[i+1] + rnorm(1,0,V)

  }

cat(signal.to.noise.ratio,"\n",Y)
plot.ts(Y)
```

Selected Output:



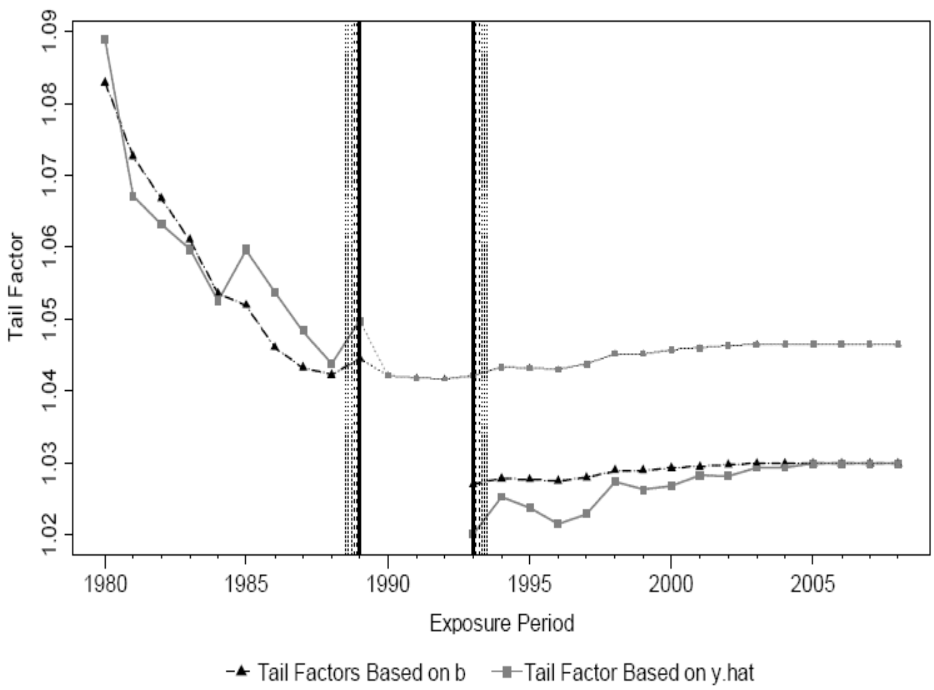


4. Suggested Applications in Actuarial Context

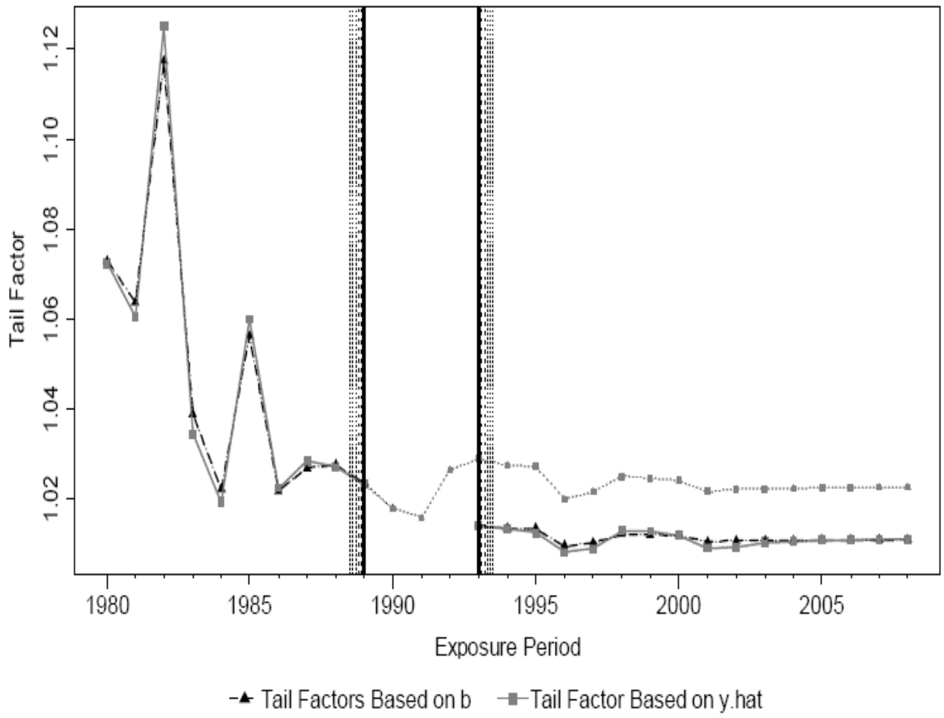
In the last three section we have discussed the required structure of scientific modeling; introduced structural break and intervention in classical time series; introduced Bayesian DLM and showed how the notion of intervention naturally embraced the Bayesian methodology under a structured modeling. The examples that we have used were for illustrative purpose and no insurance example was shown (however the case study shown is a typical financial scenario and quite similar to that of an insurance scenario). In this section we will show possible applications of intervention and structural break in two main operations of Insurance: Loss Development in the context of Reserving and Pricing of Product based on UW Cycle.

4. I. Capturing the Change in Regulations in Worker's Compensation Loss Development

This reserving example involves the application of structural break in the loss development curve. In several modern studies, loss development is considered as a time series data. This example is taken from an NCCI (National Council on Compensation Insurance, Inc.) study on Worker's Compensation Data in California. The loss development pattern in WC is highly sensitive to the regulation and legislative changes. This particular example involves a time-period when California experienced a drastic reform in WC insurance. The data expands from 1980 to 2007. Pre-reform time can be identifies as 1983 to 1989 and post-reform time as 1993 to 2004. NCCI study shows the graphs with regime identified for indemnity factors and for medical tail factors. This study by NCCI is a novel application of usage of intervention in actuarial modeling. A detail of this study (along with the program implementation) can be found in the CLRS presentation by Frank Schmid [4].



Indemnity Tail Factors

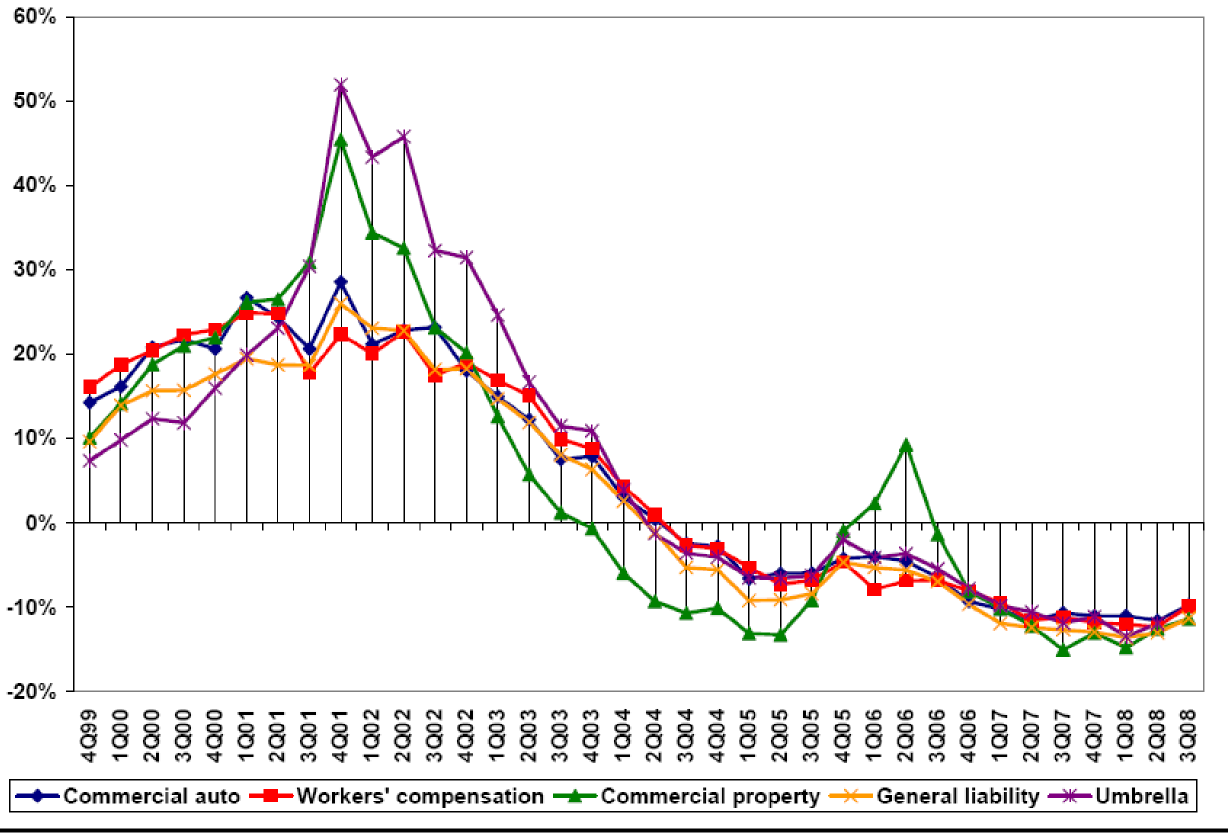


Medical Tail Factors

4. II. Switching between Hard-market and Soft-market: Insurance Pricing through UW Cycle

Underwriting cycle is a phenomenon experienced by the property and casualty insurance business. During the soft-market insurance companies tend to price the product low and during hard market the price jumps up. UW Cycle is defined to be these up-and-downs of the product pricing in P & C Insurance Industry. What causes UW Cycle? That is a long-unanswered question in the industry. To an insurance company it is, as we have discussed earlier, an uncontrollable system. It is an environment factor and an insurance company will have to price its product based on the state of environment. In that sense it is extremely important to understand the cycle and analyze as many "what-if" situations as possible.

In this context we are going to use quarterly report [5] from the Council of Insurance Agents and Brokers on Commercial P & C premium change in US from 4th quarter 1999 to 3rd quarter 2008. This entire time period definitely experienced one hard market and one soft market. Till the date of the publication of this report on 22nd October, 2008, the soft market was in flow with declining premium. Many things happen in the second half of 2008 and now recession being officially announced, insurance companies are speculating a hard market for coming quarters. A purely statistical analysis of this graph will not be able to forecast a soft market or hard market for coming quarters as there is no systematic cyclic nature in the data that can be captured by traditional time series or regression model. It is an extremely ideal scenario where Bayesian statistics can come into the picture and expert opinion and extrinsic information can be used to switch between the regime of hard market and soft market.



Source: The Council of Insurance Agents & Brokers. Chart prepared by Barclays Capital Equity Research.

As discussed in the section 3: a Bayesian DLM model can be fitted on the above data with intervention being applied at the time of transitions from hard market to soft market and vice versa. Decisions upon priors including volatility, time of intervention and usage of external information are upon the management of the insurance company. A toy code to model the above phenomenon is available with the authors and available upon request.

4. III. Conclusions

Many industries have embraced the science of Bayesian Modeling and input of subjective information in the modeling in recent past. The modeling technique that we discussed in this paper gives tremendous flexibility to incorporate subjective information and actuarial judgment in the mathematical modeling. Insurance industry in the whole world (and recently in India also) has started using Predictive Modeling techniques to find answers to different insurance problems. But it is not wise to use predictive modeling without putting an importance to business expertise. Modeling discussed in this paper gives a tool kit to combine both systems: data driven predictive models and experience based Actuarial judgments. Additionally in Indian market it makes much more sense because of two main reasons:

- Lack of insurance data: Success of traditional predictive modeling is highly dependent on data. Insufficient information always brings a possibility of over-fitting the data. Understanding the system and building a model using expert opinion (oppose to completely ask data to fit a model) can give a solution to a market with insufficient data.
- Rapid Changing Industry: Insurance industry in India is changing in a phenomenal rate. Understanding the system in this changing environment is extremely important. Information on newer regulations, reforms, competitions in the environment besides company's own experience are needed to be modeled. Intervention and subjective information should be induced using expert opinion whenever needed. In this rapid changing market Bayesian models and intervention-based classical models are natural solution.

Primary intention of this paper is to bring the discussed modeling techniques, specifically the Bayesian technique, to a greater audience in Indian Insurance Industry. The Authors wish to continue this study in a sub-sequent paper on an end-to-end application of Bayesian modeling in an insurance related modeling in Indian market.

Disclaimer:

The views expressed in this paper are of the authors and not of their employer.

Reference:

[1] ND Singpurwalla, Reliability and Risk: A Bayesian Perspective, Wiley Series in Probability and Statistics, 2006

[2] M. West and P.J. Harrison, Bayesian Forecasting and Dynamic Models, Springer-Verlag, New York, 1997. (2nd Edn. - 1989 1st Edn.)

[3] Shumway, Robert H., Stoffer, David S., Time Series Analysis and Its Applications
With R Examples, Springer Texts in Statistics, 2nd ed., 2006

[4] Frank Schmid, Loss Development in Workers Compensation in the Presence of Legislative Reform, NCCI, CLRS, 2008

[5] the Council of Insurance Agents and Brokers , Commercial P/C Premiums Continue to Drop in Third Quarter-2008, 22nd October, 2008

About the Authors:

Aparna Dutta: Aparna received Master's in Quantitative Economics from the Indian Statistical Institute in Kolkata in 2007. She is student member of the Institute of Actuaries in UK and has passed 4 examinations till date. She is working with the Predictive Modeling Team of a leading US based consulting firm in its Hyderabad location. Her work experiences include handling and econometric analysis of large dimensional insurance dataset for various US based clients.

Satadru Sengupta: Satadru received Masters in Statistics from George Washington University in Washington, DC in 2007 and Masters in Actuarial Science from Illinois State University in 2005. He works with a leading property and casualty insurance company in US in its Product Management Research Team in Boston. He is working towards the Fellowship of Casualty Actuarial Society (CAS) in US. Satadru has presented his work in the CAS Annual Meeting in San Francisco in 2006 and has received the first ever student award for CAS-COTOR challenge.