uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Alireza Gazerani

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (System Science)

GRADE / DEGREE

School of Information Technology and Engineering

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Empirical Estimation of Affine Gaussian Term Structure Model

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. Chen Guo

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. Anna Dodonova

Dr. Yuri Khoroshilov

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

EMPIRICAL ESTIMATION OF AFFINE GAUSSIAN TERM STRUCTURE MODEL

Alireza Gazerani

Thesis Submitted to
The Faculty of Graduate and Post Doctoral Studies
In Partial Fulfillment of the Requirements
For M.Sc. Degree in System Science

Telfer School of Management
University of Ottawa

Supervised by Prof. Dr. Chen Guo

Your file Votre référence
ISBN: 978-0-494-50881-7
Our file Notre référence
ISBN: 978-0-494-50881-7

Table of Contents

List of Figures

Chapter 1 Introduction

"A man, a dream, the courage to go for it, and the will to always try one more time; this is all it takes to succeed." Andres Lord --- Argentina

The above quota may describe this thesis form a general point of view, and for learning more about the whole work, its importance, achievements as well as details; this explanatory chapter has provided a short glance.

Initially, the motifs which brought the idea of this thesis to existence are mentioned and then the objectives of the project are explained; then the structure of the thesis and the order of the chapters with a brief explanation of their contents and finally a list of achievements are offered. As H. Jackson Brown says "In the confrontation between the stream and the rock, the stream always wins, not through strength but by perseverance"; so here are the details.

1.1 Motivation

As it is clearly mentioned in the vision and mission statement of the system science program, it is aimed to prepare experts interested in the analysis and modeling (mathematical and computer) of natural and man-made systems, while providing the skills and knowledge required to understand, control, predict and optimize behavior in a variety of fields from engineering and computer science to management and applied economics. The financial engineering is one of the fields with many occasions for the system science applications. Besides this was a great opportunity to widen the author's view and strengthen his skills in a rather less discovered field in his engineering background. Thus, the Multifactor models which have been developed for term structure of interest rates since almost three decades ago were brought to the attention, focusing on the lack of simplicity and convenient calculation methods for upgrading them to higher dimensions; were enough encouraging to search for a practical and user-friendly multifactor model which is not only able to explain the detailed features of the market model but also designed and structured to be estimated without sophisticated and complex methods and hardware.

1.2 Thesis Objective

The thesis is basically a successful try to expand the existing essentially affine Gaussian term structure model from three dimensions to higher dimensions, by overcoming the computational obstacle of dealing with large scale panel data trough implementing a separated method of handling the panel date (primarily as cross section and then as times series) and effective combination of nonlinear optimization algorithm (cyclic coordinate along with EMM – SNP method to estimate the model parameters that reflect complicated specification of market prices of risk. The fact that the computational barriers have made complexity and limitations to be inseparable factors for the models in the field, made it quite desirable to develop a model with calculations not beyond the traditional level, and privileged higher precision and more accuracy by utilizing numerous parameters of a higher dimension. This would be great help to provide a tool appropriate for market practitioners with strong theoretical fundamentals to satisfy academicians.

1.3 Thesis Organization

The rest of the chapters are arranged as follows:

- Background and literature review; a quick look over the related work in the field emphasizing the existing affine Gaussian models and their pros and cons
- The proposed model, its theoretical basis and key formulations along with necessary assumptions and limitations; the influential factors, parameters and variables.
- Simulation and modeling explanations; utilized computational methods and procedures, the interlinking and reciprocal data transfer between them and cyclic phenomena of the optimization procedure as well as sample results
- Discussions and explanations, interpretation of achieved results and their financial descriptive details
- Conclusions, comments and suggestions for future works based on this model.

1.4 List of contributions

Based on the objectives and the vision of the thesis the following contributions could be considered:

- Implementing the idea of partitioning the panel data for separation between the cross section and the time series in calculations, to reduce the dimension of matrices enough to make the task feasible for traditional computers.
- Achieving precise estimates for parameters of a higher dimension model which is more accurate and closer to the reality of market procedures.
- Proper combination of existing computational methods to make a robust and resilient optimization routine while remaining enough user-friendly.

Chapter 2 Literature Review

2.1 Introduction

After taking a short glance at the history of the term structures model developments and famous works in the field the focus will be on Gaussian models, especially the literature related directly to the development of the model established in this work.

2.2 Background

Besides Durand (1942) that provided yield curve data for corporate and government bonds respectively, McCulloch (1971, 1975b) utilized a spline interpolation method to obtain an after tax discount function. Further researches brought up other functional forms of the term structure by Schaefer (1981), Chambers Carlton and Waldman (1984), Jordan (1984), Nelson and Siegel (1985), Shea (1984, 85) and Vasicek and Fong (1982).

The oldest hypothesis about the term structure is the relation between slope of the term structure and the expectation about future interest rates that is discussed academically by Fisher (1896). More discussion on expectations hypothesis are done later by Fisher (1930), Williams (1938), Lutz (1940), and Hicks (1946). As a common sense people will buy long term bonds when interest rates are decreasing. Therefore downward slope of term structure shows the expectation of a decreasing interest rate and vice versa. Different researchers found fault with the expectation hypothesis like Macaulay (1938), Hickman (1942) Culbertson (1957) while others judged it by considering if the forward rates are in accordance with the expectations like Meiselman (1962) Kessel (1965). In fact the early researchers did not considered the forward rates other than actual future spot rates shifted by a constant factor. A very interesting assumption in expectation hypothesis was risk neutrality to help justification of the hypothesis by Meiselman (1962), Malkiel (1966), Bierwag and Grove (1967) and Richard (1978).

Meiselman (1962) brought up the idea of error learning hypothesis which implies that after each period of time the agents revise their expectations. Diller (1969) and Nelson (1970a) noticed that error learning principle belongs to optimal linear forecasts and could be a restriction. Instead Friedman (1979) used survey expectation data 1969-1978 for market expectations. Kane and Malkiel (1967) run their survey on banks and life insurance companies as well as non- financial corporations.

Vasicek (1977), Brennan and Schwartz (1979), Langetieg (1980), Cox, Ingersoll, and Ross (CIR) (1985), defined the term structure as discount function of some state factors and formulated the no-arbitrage condition as a second-order partial differential equation (PDE). As Gou (2003) has mentioned, the existing term structure models may be roughly divided into two classes. One is the affine class of Duffie and Kan (1996), which summarizes most of the equilibrium models, such as Vasicek (1977) and Cox, Ingersoll, and Ross (CIR) (1985). The other is the arbitrage-free class suggested by the Heath, Jarrow, and Morton (HJM) (1992) methodology, which summarizes most of the arbitrage-free models that take the current term structure as inputs rather than outputs. It is known that some of the single-factor models, such as Vasicek (1977) and CIR (1985), belong to both classes. Langetieg (1980) extended the Vasicek (1977) model to multivariate by decomposing the spot rate into a weighted average of a number of unobservable state factors that follow a joint elastic random walk. However, Langetieg (1980) model is, in fact, a collection of Gaussian term structure models. Heath, Jarrow, and Morton (HJM) (1992) criticized the traditional methodology, with particular reference to the Langetieg (1980) model, as generally inconsistent with no-arbitrage. They pointed out that the spot rate process and the market prices of risk cannot be both exogenously and separately specified. However, HJM (1992) did not state if there could be any multivariate Gaussian model that could be consistent with the HJM methodology.

2.3 Gaussian Term Structure Models

This was Vasicek (1977) which should be credited as the first milestone for the Gaussian models. In his paper few important assumptions are made; spot rate follows a diffusion process i.e. Markovian and continuous, the discount bond price merely depends on the spot rate over its term and market is efficient i.e. no transaction fee, information available to all at the same time and agents are rational. Vasicek uses similar argument as Black Scholes (1973) for pricing to develop his model in continuous time.

Vasicek model is considered single factor in sense that the spot rate is the only state variable for the whole term structure. This led to prefect correlation of instantaneous returns of bonds with different maturities; so the short bond and just one other bond span

the whole of the term structure, meanwhile the spot rate is supposed to stochastic while Markovian so its future developments is just based on its present value and time continuous without any instantaneous jump.

Vaiscek expresses his model in closed form by stochastic partial differential equation which results in a stochastic representation of the bond price and in general he take the market price of risk as constant and assuming the spot rate follows Ornstein-Uhlenbeck process (elastic random walk or Wiener process) as proposed by Merton (1971)

The next milestone could be considered the famous paper of Duffie and Kan (1996) by introducing the a multifactor model with Markov diffusion process and stochastic volatility and yields of zero coupon bond yields are taken to be maturity dependent affine combination of a selected basis set of yields.

Hence the factors of their model are the observable yields of zero coupon bonds of various fixed maturities. They have considered their model similar to multifactor version of CIR (1985a) as well, while its special case is a Gaussian-Markov with constant volatility like Vasicek (1977) and comparable with El Karoui – Lacoste (1992) within HJM setting. Besides, Jegadeesh – Pennacchi (1996) also have made a successful attempt in empirical implementation of a two factor Gaussian model for Eurodollars utilizing Kalman filter.

Duffie – Kan (1996) suggested yield factor model is offered both as single non-stochastic volatility factor and two factors (short and long rate) stochastic volatility model. They described disadvantage of their model the inconsistency of its given parameterization with every initial yield curve, suggesting calibration as remedy.

Dai – Singleton (2000) has made comprehensive explanations on affine term structure models which are initially set forth by Vasicek (1977) and CIR (1985), provided strong empirical evidences for multifactor model's efficiency and coverage.

Conceptually they describe the trade off between volatilities of risk factors and flexibility of in modeling the conditional correlations or in other words rich econometric representation of state variables and burden of pricing and estimation an essential motif for researchers to focus on affine models. Dai – Singleton (2000) compares evaluates the Gaussian models and consider their main shortcoming to be the strong assumption of constant conditional variances while their complete flexibility regarding signs and

magnitudes of conditional and unconditional correlations among the state variables as great advantage.

In their attempt to describe the comparative properties of three factor affine models, the Gaussian model is described by the condition that none of the state variables affect the volatility of state vector which follows three dimension Gaussian diffusion and the state variables are homoskedastic (the errors have same distribution). Moreover, in the Gaussian model the state variables are free to be negative or positive (an important desirable feature) and the conditional variances of zero coupon bond yields are free to be determined by common factors. Dai – Singleton (2000) estimate parameters for three factor models using EMM-SNP programs, initially suggested by Gallant and Tauchen (1989) and developed afterwards (This package is used partially in estimating mean reversion parameters of our model and will be explained more in detail in the next chapter).

In their next paper, Dai – Singleton (2002); they prove the ability of richer dynamic term structure models (higher number of factors) to explain stylized facts like linear projection of returns on the slope of yield curves or variation in expected returns across the time and maturities. The brilliant point in the results is multifactor Gaussian models are matching very well to the linear projection yield and this is due to their constant market price of risk. In their attempts to resolve the mis-pricing the at the very short end of maturities (e.g. one month) by adding the fourth factor to the Gaussian model, the successful resolving results shows clearly that there are still room for more factors in Gaussian model.

Duffee (2002) discusses the essentially affine models show merit in fitness for either the first or second moment of yields and not both since the flexibility of the models in correlation structure pf the state variables leads to prevention of conditional variation in yields volatilities. This gives the lead to the Gaussian models in satisfying full set of moment condition. Therefore, this may conclude building a model to fit the data sufficiently, which is going to be covered by error minimizing approach, and sufficiently smooth functioning spot rate curve over the time series by optimization procedure has critical importance. Most implemented affine term structure models suffer from significant complexity of computation while large number of parameters is considered.

Higher flexibility of Gaussian model to restrictions allows me to modifying the canonical form to a diagonal form, and reduce the model parameters enough for making increase in the number of state variables which optimally to less measurement errors . Because the sufficient number of state variables almost eliminates the measurement errors, parameter estimation will be simpler and more accurate.

2.4 Estimation Method

Guo (2001) take $m > n$ be the number of observed zero coupon bond yields and K be the sample size and multidimensional search algorithm to find the optimal set of parameters is developed. He suggests that for each candidate set of parameters in the parameter space Θ, the state vector can be estimated from the yield equation by cross-sectional ordinary least square (OLS):

$$R(t,\tau_j) = a(\tau_j) + \sum_{i=1}^{n} b_i(\tau_j) Y_i(t)] + u_{jt},\ j = 1,\cdots,m,$$

where the measurement errors u_{jt} are assumed to be zero mean and independent and identically distributed (i.i.d.) in cross section, but they plays virtually no role in parameter estimation because the purpose of the cross-section fitting is simply to convert the observable yields into the non-observable (latent) state variables based on a given set of parameters. When the parameters are optimal, the least square of the measurement errors of the state variables calculation will be minimized.

Guo (2001) mentions, theoretically, the measurement errors can be eliminated if the Gaussian model is specified with exactly m factors; however, because the reported yields are based on the average bid-ask quotes collected from a number of bond dealers, data errors naturally exist, and human errors in data recording and handling also occur occasionally. Since the nature of the data errors has nothing to do with the dynamics of the term structure, perfect-fit would carry the irrelevant errors into the dynamics of the state variables forcefully, which would not be plausible. Hence, the number of factors should be slight less than the number of observed yields in order to allow some measurement errors.

Optimizing state variables is accomplished by searching for the optimum parameters one at a time utilizing golden section linear search inside a procedure

designed based on the Cyclic Coordinate optimization algorithm. At the beginning of estimation, the instantaneous covariance matrix Σ is initially set to zero. Based on the estimation of the state variables $Y(t)$s the EMM – SNP program package optimizes the set of mean reversion parameters, thus:

$$\varepsilon_{it} = Y_i(t+1) - e^{-\alpha_i \delta t} Y_i(t) - \gamma_i,$$

Where

$$\gamma_i = (1 - e^{-\alpha_i \delta t})\theta_i$$

are constants in time series.

2.5 Golden section search method

The golden section search method (an open source mathematical concept) is basically developed to find the minimum of a uni-modal continuous function over an interval without using derivatives. This is a great advantage of this search method especially when taking derivative is too complicated or not feasible. This method applies as well for finding the maximum.

Taking the function f over the interval [a; b]. It is assumed that f(x) is continuous over [a; b]; and that f(x) is uni-modal over [a; b], i.e.: f(x) has only one minimum in [a; b]; this condition is simply satisfied by making the search interval narrow enough. The other approach in this category is the bisection method but Golden Section is faster and more recommended since an optimal reduction factor for the search interval and minimum number of function calls is desirable.

The bisection method starts with computing the midpoint $m = (a + b)/2$ and to evaluate at x_1 and x_2, defined by $x_1 = m - \delta/2$ and $x_2 = m + \delta/2$, for some small value δ for which $f(x_1) \neq f(x_2)$. If $f(x_1) < f(x_2)$, then the search interval is [a; x_1], otherwise it will be[x_2; b]. While this halves the search interval in each step, two new function evaluating must be taken in each step which is not optimal. To reduce this to only one new function evaluation in each step and having a constant reduction factor, say c, for the size of the interval.

For x_1 and x_2 somewhere in [a; b], there are two cases:

1. If $f(x_1) < f(x_2)$, then [a; b] := [a; x_2], with interval size reduction

$x_2 - a = c(b - a) \Rightarrow x_2 = a + cb - ca \Rightarrow x_2 = (1 - c)a + cb$

2. If $f(x_1) > f(x+2)$, then [a; b] := [x_1; b], with interval size reduction

$b - x_1 = c(b-a) \Rightarrow -x_1 = cb - ca - b \Rightarrow x_1 = ca + (1-c)b$

Therefore, knowing c, we know the location of x_1 and x_2.to find c, consider the case $f(x_1) < f(x_2)$. Without loosing generality and seeking ease of calculation, let [a; b] = [0; 1].

If $f(x_1) < f(x_2)$, then we take $x_1 = 1 - c$ and have to determine where to evaluate next, either at the left, or at the right of 1 - c.

1. With a function evaluation at the left of $x_1 = 1 - c$, then x_1 is the right point of the interval [0; c], and x1 could be written in two ways (using the formula for x_2 derived above with a = 0; b = c):

$1 - c = (1-c)0 + cc \Rightarrow c^2 + c - 1 = 0$

The positive root leads to $c = (-1 + \sqrt{5})/2$, which equals approximately 0.6180.

2. With a function evaluation at the right of $x_1 = 1 - c$, then x_1 is the left point of the interval [0; c], then x1 could be written in two ways (using the formula for x_1 derived above with a = 0; b = c):

$1 - c = (c)0 + (1-c)c \Rightarrow (1-c)^2 = 0$

Hence the root of this equation is 1, leads to no reduction, so it is excluded.

Thus two rules are in effect. If $f(x_1) < f(x_2)$, we keep x_1, which becomes x_2, and compute a new x_1 using $x_1 = ca + (1-c)b$.

If $f(x_1) > f(x_2)$, x_2 is kept, which becomes x1, and a new x2 will be calculated by $x_2 = (1-c)a + cb$.

2.6 Cyclic Coordinate Method

Among several methods of optimization developed for non-differentiable objective functions the coordinate descent method is well known as an effective and popular approach. This method suggests the objective function to be minimized along every coordinate direction, for each of iterations, using a search method (e.g. golden section) and the cycle has to be repeated until finding the final optimum point. Hence, if the order of optimization is cyclic the procedure is called cyclic coordinate method. The general

formula of the procedure for moving from i^{th} coordinate of the vector x^{k+1} could be written as $x_i^{k+1} = \arg \underset{\xi \in \Re}{Min} f(x_1^{k+1}, x_2^{k+1}, ..., x_{i-1}^{k+1}, \xi, x_{i+1}^{k}, ..., x_n^{k})$.

Using this method, during iteration, all the variables are kept constant except for one and then a new value for this variable that reduces (or optimizes) the objective function is found. The next iteration will do the same with the next variable and so far so forth, till all the variables are tried. Then the cycle repeats beginning with the first variable again. The convergence speed is faster in the earlier cycles and slowed down when the optimum point get closer.

Its overall speed is slower than steepest descent method but the advantage of being free of calculations for gradient of objective function, compensate it in complicated functional forms especially if the variables are loosely coupled.

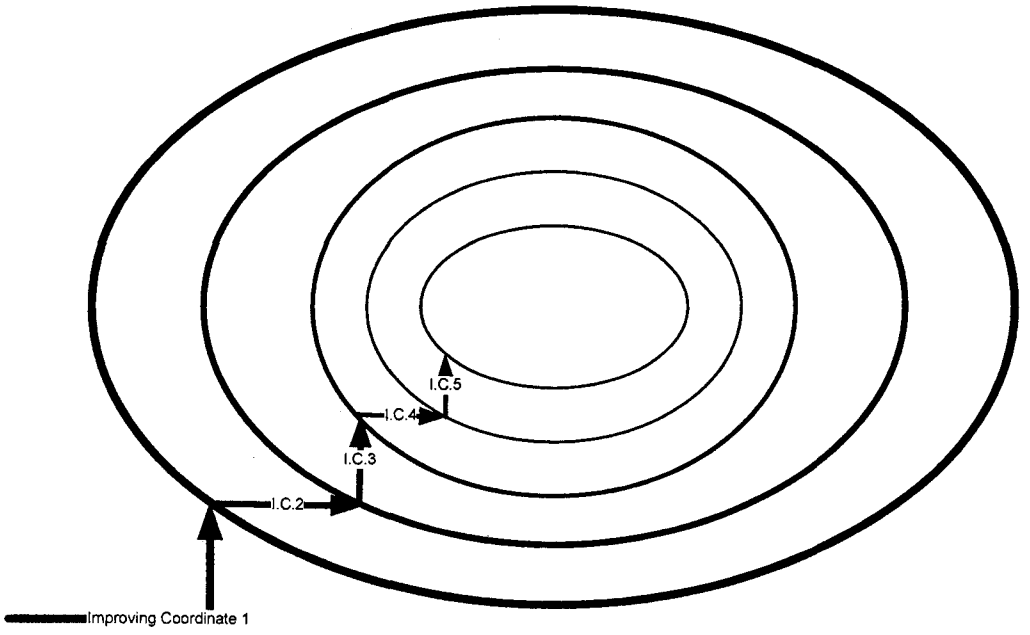

Figure 1 - Cyclic Coordinate Method

The advantages of this method are duality (it can be used for maximization as well), restricted minimization (x^i could be bounded so $\zeta \in \Re$ will be changed to the restriction statements) and parallel computation as well. This great advantage is very helpful when different sets of coordinates are not coupled together.

As a convergence condition a minimum should be attained along each coordinate. This method could be slower than derivative approaches but nevertheless it is quite effective.

Chapter 3 Proposed Model

3.1 Introduction

Guo (2001) explains that existing affine term structure models are formulated in a canonical form, in which the drift of the state vector process is an upper triangle matrix. This form is convenient for nesting both the Gaussian and the square-root processes, but the large number of parameters causes significant computational complexity, thus severely restricts the dimension of the state space.

On the other hand, Guo (2001) suggests the small number of state variables is insufficient to fit the observed bond yields, and accuracy of the estimation could be questionable. Since the square-root processes have not been found satisfactory in empirical studies, the focus is on the Gaussian processes. The non-restrictiveness of the Gaussian processes allows modification of the canonical form to a diagonal form. This reduces the model parameters substantially, thus makes room for substantial increase in the number of state variables. Because the sufficient number of state variables almost eliminates the measurement errors, parameter estimation will be simpler and more accurate.

3.2 Data Preparation

By $P(X(t),t,T)$, Guo (2001) expresses the price of a discount bond as a linear combination of $n+1$ "hypothetical" present values of the cash flow of the discount bond, discounted by various "discount rates", (i.e., the exponential parameters). Coupon bond is a linear combination of zero coupon bonds, and keeps the same linearity.

Guo (2001) considered $P_k(T_k - t, c_k)$ to denote the observed price of the k^{th} bond at time t for maturity at T_k, which has \$100 face value, and an annual coupon c_k with semiannual interest payments. Taking AI_k its accrued interest, the observed bond price can be expressed as

$$P_k = \left(1 - \sum_{i=1}^{n} x_i\right) P_k^0 + \sum_{i=1}^{n} x_i P_k^i + \varepsilon_k,$$

Thus P_k^0 is the present value of the cash flows of the bond discounted by the rate λ_0, and P_k^i is the present value of the bond discounted by the rate λ_i.The U.S. Federal

Reserve website, as one of the most reputable sources, provides 8 key rates with fixed maturities of 3M, 6M, 1Y, 2Y, 3Y, 5Y, 7Y, 10Y. The first 3 rates are yields for Treasury bills of 3 months, 6 months and 1 year. Hence, they can be converted directly into prices. The rest are the coupon rates of par-valued (arbitrarily set at $100) Treasury notes (2Y to 10Y) and Treasury bonds (20Y and 30Y). Since they have been adjusted to fixed maturities, the accrued interests are all zero. Supposing these key interest rates are all correct, the corresponding bond prices would be error free. Calculation of par coupons from bond price state variables for weighted difference calculations. It takes $Y_1, Y_2, \ldots, Y_6, \alpha_1, \alpha_2, \ldots, \alpha_6$ and maturity and using the following formulas return the par coupon values while τ is the maturity on semiannual basis:

$$ParCoupon = 200\frac{1-\sum_j\left(1-\sum_k Y_k\right)e^{-0.5\mathring{a}\tau}+Y_j e^{-0.5(\alpha_j+\mathring{a})\tau}}{\sum_i\sum_j\left(1-\sum_k Y_k\right)e^{-0.5\mathring{a}i}+Y_j e^{-0.5(\alpha_j+\mathring{a})i}}, \tau \succ 2 \text{ and}$$

$$ParCoupon = 200\left(e^{\left.-Log\left|1-\sum_j\left(1-\sum_k Y_k\right)e^{-0.5\mathring{a}\tau}+Y_j e^{-0.5(\alpha_j+\mathring{a})\tau}\right|\right/\tau}-1\right), \tau \le 2$$

3.3 Brief Theoretical Description of the model

Guo (2001) considered the term structure of interest rates as a pricing function, $P(Y(t),\Theta,\tau)$, that represents discount bond price at time t with arbitrary maturity $\tau \geq 0$, in which $Y(t)$ is a vector of n unobservable (latent) state variables, and Θ is a large set of unknown parameters. Guo (2001) emphasizes; most of the existing dynamic term structure models can be classified into the affine term structure models. Duffie and Kan (1996) as well as Dai and Singleton (2000), have nested and generalized the path-breaking single-factor models of Vasicek (1977) and Cox, Ingersoll, and Ross (CIR 1985). Then maximum likelihood estimation has been applied to a three-factor family of affine models□symbolized by the canonical form of Dai and Singleton (2000) as $A_m(3)$ with m CIR square-root processes and $3-m$ Gaussian processes, with impressive successes as in Dai and Singleton (2000 and 2002), Duffee (2002), Duarte (2003).

Guo (2001) indicates, the major problem with the exiting empirical studies is the low dimensionality, or too few state variables. Typically, the term structure is represented by eight treasury securities with maturities of 3 months, 6 months, 1 year, 2 years, 3 years, 5 years, 7 years and 10 years, which have sufficient liquidity ("on-the-run"). Because three factors are insufficient to fit eight yields, the above studies have selected three yields to be fitted without error, and let other yields to have measurement errors. Consequently, the objective function is modified with an additional likelihood function of the measurement errors, resulting in a Quasi Maximum Likelihood (QML).

Some authors, for example, Brandt and He (2002), have criticized this methodology as it implies that different choice of the three yields would imply different term structure dynamics. These authors suggest that all the yields should be fitted with error, and the impact of the measurement errors on the estimation of the state variables can be corrected by either a Kalman filter like Babbs and Nowman (1999), Duan and Simonato (1999), or by simulated likelihood estimation as Brandt and He (2002). However, the measurement error problem is clearly a missing variable problem, or model misspecification, because it is caused by insufficient number of state variables. It should be corrected by increasing the number of state variables, but the computational

complexity becomes an obstacle. For example, Duffee (2002, pp 418) admitted that "adding another factor would make this investigation impractical".

So, the large number of parameters in the affine models is partially attributable to the canonical form, which is designed to accommodate the CIR square-root process. The CIR square-root process is theoretically more appealing than the Gaussian process because it allows stochastic volatility, but numerous empirical studies have found that the pure Gaussian model ($m = 0$) performs much better in interest rate forecasting (Duffee 2002) and matching the observation (Dai and Singleton 2002). Because the CIR process requires strictly positive state variable and does not allow it to be correlated with other processes, the parameters of the affine model have to be restricted in order to accept one or more CIR components. The canonical form is designed to accommodate this "admissibility" issue.

Therefore Guo (2001) concludes, the simplicity, non-restrictiveness, and superb empirical performance make the pure Gaussian affine model the first choice for high-dimensional expansion. Without the interference of the CIR processes, the canonical form of the affine model can be simplified to a diagonal form, in which the number of parameters can be substantially reduced.

3.4 Model Formulation

Guo (2001) assumes that the spot rate is a linear combination of state factors, where the coefficients of the factors have been normalized to unity, as in Dai and Singleton (2002, pp 424):

$$r(t) = a_0 + \sum_{i=1}^{n} Y_i(t).$$

He explains that state vector process under the physical probability measure is mean-reverting with constant volatility:

$$dY(t) = \alpha(\theta - Y(t))dt + \sigma\, dW(t),$$

where a α is $n \times n$ diagonal, θ is $n \times 1$; standard deviation σ is $n \times n$, and $\sigma\sigma' = \Sigma$ is the instantaneous variance-covariance matrix which is not diagonal because the elements of the Brownian motion vector W are correlated. The existing models set α

as an upper triangle matrix, then impose it to be diagonalizable. Guo (2001) defines the risk-neutral probability measure as

$$d\tilde{W}(t) = -\lambda dt + dW,$$

where the market prices of risk in the essentially affine Gaussian model (Duffee 2002 and Dai-Singleton 2002) are specified as

$$\lambda = \sigma^{-1}(\lambda^{0} + \lambda^{1} Y(t)),$$

where λ^0 is $n \times 1$ and Guo (2001) specified λ^1 as $n \times n$ diagonal. The state vector under the risk-neutral measure becomes

$$dY(t) = \tilde{\alpha}(\tilde{\theta} - Y(t))dt + \sigma\, d\tilde{W},$$

where $\tilde{\alpha} = \alpha + \lambda^1$ which is also diagonal, and $\tilde{\theta} = \tilde{\alpha}^{-1}(\alpha\theta - \lambda^0)$. The bond pricing function can be solved as

$$P(t,\tau) = \exp[-A(\tau) - B(\tau)'Y(t)],$$

and the yield function can be expressed by definition as

$$R(t,\tau) = a(\tau) + b'(\tau)Y(t),$$

where $b_i(\tau) = \dfrac{B_i(\tau)}{\tau} = \dfrac{1 - e^{-\tilde{\alpha}_i \tau}}{\tilde{\alpha}_i \tau}$, and

$$a(\tau) = \frac{A(\tau)}{\tau} = a_0 + \sum_{i=1}^{n} [1 - b_i(\tau)]\tilde{\theta}_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\rho_{ij}\sigma_i\sigma_i}{\tilde{\alpha}_i\tilde{\alpha}_j}\left[1 - \frac{(1 - e^{-\tilde{\alpha}_i\tau})}{\tilde{\alpha}_i\tau} - \frac{(1 - e^{-\tilde{\alpha}_j\tau})}{\tilde{\alpha}_j\tau} + \frac{(1 - e^{-(\tilde{\alpha}_i + \tilde{\alpha}_j)\tau})}{(\tilde{\alpha}_i + \tilde{\alpha}_j)\tau}\right]$$

Guo (2001) considers this model to be substantially simpler than the existing canonical form because function $b_i(\tau)$ involves only one exponential term with parameter $\tilde{\alpha}_i$. In contrast, because the matrices α and $\tilde{\alpha}$ in the canonical form are not diagonal, the bond pricing function is much more complicated. Since both matrices α and $\tilde{\alpha}$ are required to be diagonalizable, they can be set diagonal at the beginning, i.e., at the stochastic differential equation level, so the bond pricing can be reduced to diagonal form.

Chapter 4 Simulation and Modeling

4.1 Computational System

The major difficulties that previous works were facing with the panel data has been the dimension of the covariance matrix of state variables which makes the calculations non feasible. The initiative in this new work is dividing the calculation and optimization in two separate stages; initially the cross section part provides the optimized model parameters and subsequently the time series optimizes the mean reversion parameters. In a reciprocal cycle the two parts converge to the final values in addition to practical observation of convergence of elements of the covariance matrix of the state variables in a couple of iterations. Therefore project is structured as a two part system. The first part is the cross section optimization utilizing VB program and the second part is time series optimization by Fortran codes. The VB program has a central module and a few distinct modules which are called by the main program.

Besides the EMM–SNP program is in Fortran codes developed by Ronald Gallant and George Tauchen and primarily described in the paper titled "Which Moments to Match" (Gallant and Tauchen, 1996a) and utilized to optimize the mean reversion parameters as well as covariance matrix of the state variables which will be described more in detail later.

4.2 Modular Programming

The VB codes as well as Fortran are made of a combination of different modules each responsible for specific tasks while interacting with each others results. All functions used in the structure of models are tested and verified separately in excel in form of macros that make every step of process observable and convenient for review.

The programs provided for accomplishment of the tasks have modular structure. In both VB and Fortran codes each module has its own input and output which provides enough recorded results to have every detail necessary for analysis and investigation.

4.2.1 Cross Section Estimation Module

The cross section estimation module is designed to perform calculations for the objective function and return it to the optimization module for optimization purpose. Every time it is called by the optimization module it takes all the necessary data, parameters and

variables in from of function arguments as well as files and produces the value of the objective function for the fed amounts in addition to the new series of the state variables for this specific set. Here is the block diagram of estimation module:

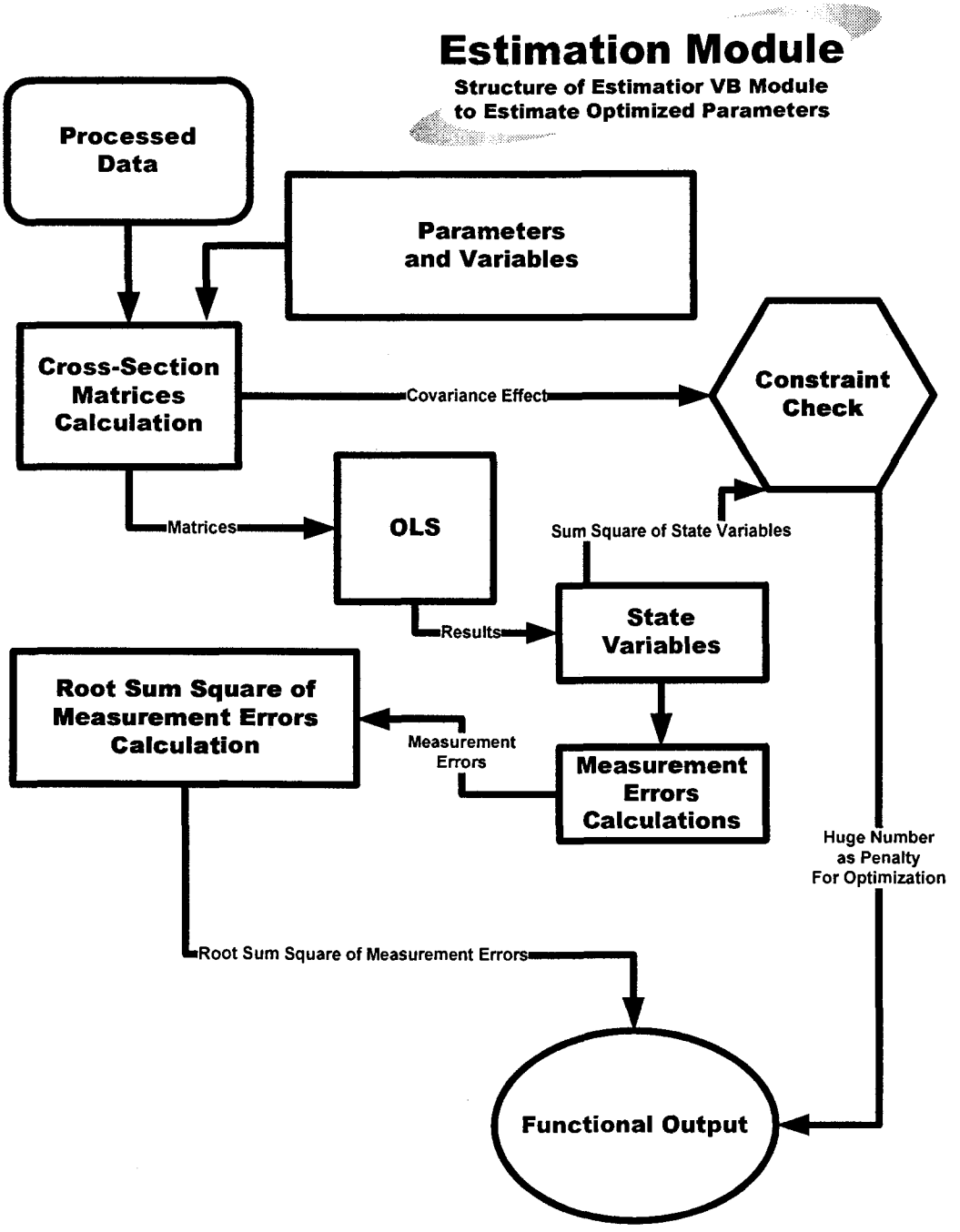

Figure 2 - Estimation Module

The module primarily takes in the processed data (8 weekly zero coupon yields time series as a text file), in addition to all other necessary parameters and variables obtained by the model; the optimum mean reversion parameters $\lambda_1, \lambda_2, \ldots, \lambda_5$ produced by EMM–SNP modules for the Gaussian state transition process (as function arguments) as well as the optimum mean of the state transition process $\varphi_1, \varphi_2, \ldots, \varphi_5$ and covariance matrix calculated by their standard deviation and correlation matrix of optimum state variables (both as text files). The function input arguments are: $\lambda_1, \lambda_2, \ldots, \lambda_5, \alpha_1, \alpha_2, \ldots, \alpha_5, Rightboundary, \overset{\circ}{a}, samplesize, frequency, Y\lim in$ which $\lambda_1, \lambda_2, \ldots, \lambda_5$ are mean reversion parameters, $\alpha_1, \alpha_2, \ldots, \alpha_5$ are model parameters (Y estimation), *Rightboundary* limit for parameters, $\overset{\circ}{a}$ asymptotic forward rate, *samplesize* the number of observation in the sample, *frequency* the time interval (week) and *Ylmit* the boundary for sum of Y^2s'; optimum mean of measurement errors $\mu_1, \mu_2, \ldots, \mu_8$ (8 scalars), is also taken into the module as text file.

After taking inputs and setting up the values an ordinary regression is run over the processed data, with the minimum least square as criteria in consideration for output to optimize measurement errors subsequently in the optimization module.

Basically, the general regression has the general format of $y - x\beta = \varepsilon$ so the regression calculation is conducted as $[Y] = [B^T B]^{-1} \cdot [B^T] \cdot [C]$

while $B_{ij} = 1 - \dfrac{e^{-\alpha_j T_i}}{\alpha_j T_i}$,

$C_{ij} = A_{ij} - \overset{\circ}{a} + Co\,\text{var}\,ianceTimelyEffect$

and $A_{ij} = 2\log(1 + ZeroCouponYield_{ij}/200)$

in which:

$$\text{Covariance Timely Effect} \sum_{k}^{9}\sum_{j}^{5}\sum_{i}^{5} \frac{\sigma_{ij}}{2\sigma_i\sigma_j T_k} * \left[T_k + \frac{1 - e^{-(\alpha_i+\alpha_j)T_k}}{\alpha_i + \alpha_j} - \frac{1 - e^{-\alpha_i T_k}}{\alpha_i} - \frac{1 - e^{-\alpha_j T_k}}{\alpha_j} \right]$$

and $Co\,\text{var}\,ianceFactor = \sum_{j}^{5}\sum_{i}^{5} \dfrac{\sigma_{ij}}{2\alpha_i\alpha_j}$ as well as $LongrateR = \overset{\circ}{a} - Co\,\text{var}\,ianceFactor$

The regression results in $[Y]$ as well as measurement errors η by $[U]=[C]-[B]\cdot[Y]$ which are the eight time series for calculation of objective function and its value as the output:

$$ObjctiveFunction=\sum_{j}^{8}\sum_{i}^{N}\frac{\eta_{ij}^{2}}{8}$$

to be minimized during the cycles in optimization module. Two indications are designed to flag out negative amounts of Covariance Factor as well as exceeding the sum square of state variables $\sum_{i} y^2$ over preset value of *Y limit*; in either case a big number is put to the function output value to represent a penalty in optimization process and neglect the results obtained under such circumstances.

In addition to the function arguments, the cross section estimation module output files including state variables $Y_1, Y_2, \ldots, Y_5$ time series, measurement errors $\eta_1, \eta_2, \ldots, \eta_5$ time series, summary of estimation module calculations containing annualized covariance matrix of measurement errors and their correlation coefficients, model parameters $\alpha_1, \alpha_2, \ldots, \alpha_5$, mean reversion parameters $\mu_1, \mu_2, \ldots, \mu_8$, Long rate R, asymptotic forward rate $\overset{\circ}{a}$ and covariance factor values are stored after every run of estimation module. Nonetheless, the single scalar output of function; root mean square of measurement errors is the most important value which is used in the optimization module as the main criteria for optimization.

4.2.2 Optimization Module

The optimization module is calling the estimation module as its function in a master-slave format. The exchange of the values and parameters are trough the function arguments as well as the data files already mentioned. The main duty of the optimization module is performing cyclic coordinate procedure. For every coordinate (parameter $\alpha_1, \alpha_2, \ldots, \alpha_5$ and $\overset{\circ}{a}$), the line search is performed using the golden section method by calling the estimation module. The other important task, is recording the results in the correct moment. At the end of line search for each parameter the result is saved in appropriate files if any improvement is observed in the value of the objective function of the least square and then the search in the direction of the next coordinate continues.

The following block diagram makes a visual image of structure of the model:

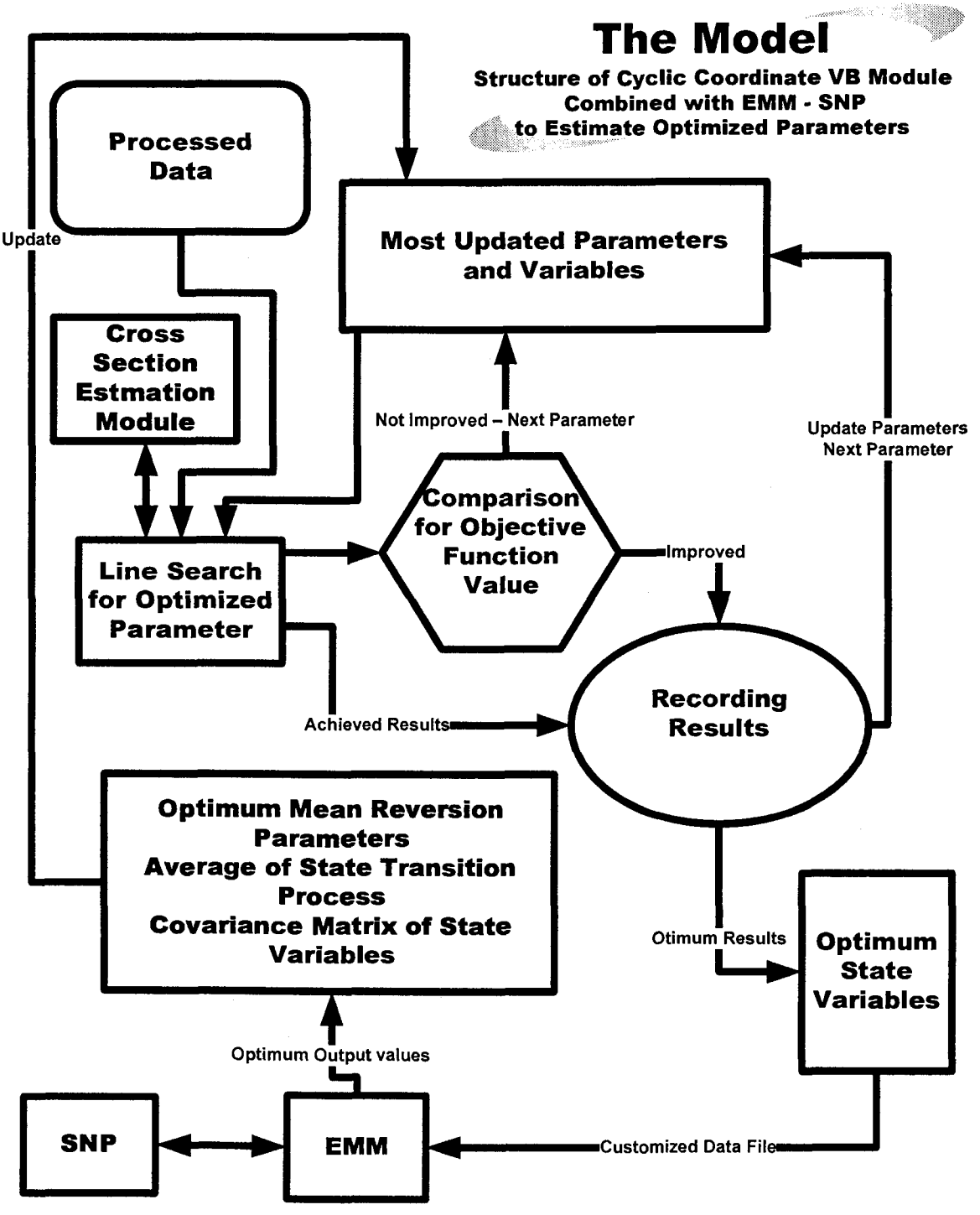

Figure 3 - The Model

The main structure of the model is built as cycling optimizer that utilizes specific technique (Golden section line search) to provide the best results after running enough throughout the data. Firstly all the variables used in the program are introduced by their attributes as double precision or integer. All the inputs taken by optimization module are trough data files, although some of the variables are introduced already, here are optimization inputs in detail: Optimum model parameters $\alpha_1, \alpha_2, \ldots, \alpha_5$ are bond pricing parameters (Y estimation), *Rightboundary* limit for parameters, $\overset{\circ}{a}$ asymptotic forward rate, *samplesize* the number of observation in the sample, *frequency* the time interval (week) and *Ylmit* the boundary for sum of Y^2 s

Optimum mean reversion parameters (lambdas) $\lambda_1, \lambda_2, \ldots, \lambda_5$

Optimum covariance matrix of state variables

Optimum means of state transition process or phi array $\varphi_1, \varphi_2, \ldots, \varphi_5$

Optimum mean of measurement errors $\mu_1, \mu_2, \ldots, \mu_8$

Optimum long rate term

State variables of regression procedure

Temporary long rate

And these three files are taking in form EMM–SNP results:

Optimum mean reversion parameters $\lambda_1, \lambda_2, \ldots, \lambda_5$,

Optimum state variables covariance matrix

Optimum means of state transition process or phi array $\varphi_1, \varphi_2, \ldots, \varphi_5$

The output files are as follows:

Optimum model parameters $\alpha_1, \alpha_2, \ldots, \alpha_5$, *Rightboundary*, $\overset{\circ}{a}$, *samplesize*, *frequency*, *measurement errors minimum least square*

Optimum and temporary means of measurement errors $\mu_1, \mu_2, \ldots, \mu_8$

Optimum and temporary long rate

Temporary mean of measurement errors $\mu_1, \mu_2, \ldots, \mu_8$

The procedure of the optimization module starts after all the parameters and variables are initialized by their optimum amount. The number of the cycles could be set arbitrary and

is set by an internal variable (kj) and each cycle has six stages (each for one of the parameters). At the end of each stage in case of any observable improvement in the value of the objective function, an update of all optimum values are recorded for the next cycles as well as history of the process. After the entire number of cycles are accomplished similarly to the end of each stage the recording and storage task will be completed. The following table summarizes the variables/ parameters mentioned before:

	VARIABLE/ PARAMETER	DEFINITION	CALCULATED/OPTIMIZED BY
1	$\alpha_1, \alpha_2, \ldots, \alpha_5$	Model Parameters	VB Program
2	$\lambda_1, \lambda_2, \ldots, \lambda_5$	Mean Reversion Parameters	EMM – SNP
3	$\varphi_1, \varphi_2, \ldots, \varphi_5$	Average of State Transition Process	EMM – SNP
4	$\mu_1, \mu_2, \ldots, \mu_8$	Means of Measurement errors	VB Program
5	$\overset{\circ}{a}$	Asymptotic Forward Rate	VB Program
6	*Rightboundary*	Upper limit for parameters value	Set Manually
7	*samplesize*	Total number of Observation	Equal to 1000
8	Y	State Variables	VB Program
9	$Y \lim it$	Constraint on Y^2	Set manually
10	$[\sigma]_{5\times5}$	Covariance Matrix	EMM – SNP
11	η	Measurement Error	VB Program
12	R	Long Rate	VB Program

13	T	Maturity time	Predetermined by data
14	τ	Time to maturity	Data Preparation
15	P	Bond Price	Data Preparation

After cross sectional optimization of the model parameters and state variables the state variables time series are transferred to the EMM – SNP programs for optimization of mean reversion parameters and calculation of the elements of the covariance matrix. These programs' explanations are followed.

4.2.3 EMM Introduction (Efficient Method of Moments)

EMM Package is introduced in the preface of the EMM 1.7 user's guide (October 11, 2004) as a program implementing the simulation estimator described in Gallant and Tauchen (1996), using the algorithm by Chernozhukov, Victor, and Han Hong (2003), "An MCMC Approach to Classical Estimation," 115, 293--346. Supposedly it could be used for maximum likelihood estimation, Bayesian estimation and simulation estimation. The program is in FORTRAN and has to be edited by the user to suits to perform the intended tasks; besides this program is free software; and could be redistributed and/or modified it under the terms of the GNU General Public License as published by the Free Software Foundation. The efficient method of moments approach systematically utilized by Gallant and Tauchen (1996a, 2002a) to generating moment conditions for the generalized method of moments (GMM) estimator of the parameters of a structural model, as an alternative to the common practice of selecting a few low order moments on an ad hoc basis and then proceeding with GMM.

They explain basic idea is using the expectation under the structural model of the score (derivative of the log density of the auxiliary model with respect to the parameters of the auxiliary model or score generator) from an auxiliary model as the vector of moment conditions. Therefore, the moment conditions depend upon both set of the parameters of the auxiliary and structural model. The efficiency of the method is discussed more in detail in Gallant and Tauchen, 1996a, 2001, 2002a, Tauchen 1997, Gallant and Long, 1997.

The recent versions of SNP permit a GARCH specification for the conditional variance of the leading term of score generator; hence the results obtained with EMM are featuring GARCH as well.

4.2.4 SNP Program Introduction (Semi Non-Paramétrique)

A. Ronald Gallant and George Tauchen in their SNP user's guide (December 1990) describe SNP as a method of nonparametric time series analysis, to approximate the conditional density of a multivariate process, employing an expansion in Hermite functions which is a nonlinear nonparametric model that directly nests the Gaussian VAR model, the semi-parametric VAR model, the Gaussian ARCH model, the semi-parametric ARCH model, the Gaussian GARCH model, and the semi-parametric GARCH. It uses conventional maximum likelihood along with a model selection strategy which determines the appropriate order of expansion.

It initially proposed by Gallant and Tauchen (1989) for an asset pricing application. Estimation of SNP models entails using a standard maximum likelihood procedure together with a model selection strategy that determines the appropriate degree of the expansion. It is claimed that under reasonable regularity conditions, the estimator is consistent.

The name of the method is coming from the term Semi Non Parametric, due to its position in the categories of the procedures. As it was mentioned before the Hermite series expansion is the foundation of SNP method. Both the Gaussian component of the Hermite expansion (for modeling VAR, ARCH and GARCH) and computational ease of evaluation and differentiation are considerable and Hermite density practically facilitates simulation.

4.2.5 SNP Program

As it is mentioned already the details about SNP or semi-nonparametric program could be found to some extent in the user's guide provided by its authors Gallant and Tauchen (1990 – last version 2007); however explanations about the necessary changes for tailoring the program to fit the specific purpose of join procedure with VB program would be reviewed here. SNP authors have mentioned that their program has fundamentally based on Hermite function series expansion for estimation of a conditional

density. The non parametric feature makes the program enable to have higher order terms to capture the differences from an original suggested parametric model and overcome the shortcomings of kernel and spline approaches. The set of parameters in various arrangements provide coverage for many leading time series models like VAR, ARCH and GARCH.

In the available package of SNP program two main folders are SNPRUN, NPSOL and CODE. NPSOL is the folder containing 11 program files for the non parametric calculations including the functions, subroutines and error handlers. The CODES folder keeps the main modules of the program for the purpose of simulation and also the data handling procedure. NPSOL is not supposed to have any change while the only file needed to be touched in CODE folder is SNPDATA.F for tailoring it suitable for the input data; since SNPDATA.F is the file in charge for taking the data in so this has to be set according to the structure of the data which is going to be fed to the program. It is very important to assure the data file structure is exactly the same as it is declared in the SNPDATA.F file; any discrepancy will result in error and will stop the program immediately. The data has to be set in columns and every row has to begin with date in format of yy-mm-dd and then the data values will follow. Usually three characters would be enough for the first three columns indicating the date but for the data column the number of digits and the decimals has to be set according to the data coming from the VB program.

There are other files in the CODE folder like SNPOPEN.F for opening and closing the files, SNPOPT8.F for evaluation of the objective function and constraints, SNPSUB8.F keeping set of subroutines for SNP program, SNPVER8.F for latest changes in version 8 and UTILITY.F containing all kind of useful matrix operations subroutines.

Unless drastic changes get necessary for the structure of the program further changes to the files in the two mentioned folders won't be needed. The SNPRUN is the folder that keeps the main module SNP.EXE as well as the data file, control file and parameter files, input and output.

Each line of the CONTROL.DAT file sequentially gives the input file, the output file, values for the perturbation of the parameters, number of iteration, and initial seed (for replicating data just in case). Although this is a very important file for SNP but working

under EMM optimizer will leave it idle and give the control to EMM with same format for the content. The parameters files are input and output distinguished by their extensions as .in0 and .fit while their structures are identical to be replaceable for cyclic iterations. The only item which needed to be changed to enable it works with VB is the number of observations (1000) in the Data Description part.
The fact that SNP has been used under EMM control leaves the necessary changes at this level and the rest of alterations have to be considered for EMM program.

4.2.6 EMM Program

It is already mentioned that very detail of the EMM program is provided by its creators Gallant and Tauchen in their User's guide for version 1.7 therefore it would be convincible here to focus merely on the parts that have been touched for VB program involvement. EMM (efficient method of moments) is an improved form of original approach generalized method of moments by utilizing the minimum Chi-Square estimator. The improvement is basically made by defining an auxiliary model which produces a score with respect to every set of parameters; the auxiliary model is the SNP that generates the scores for EMM to optimize the parameters. The optimization is taking the approach of maximizing the pseudo-maximum-likelihood of auxiliary model resulting quasi-maximum-likelihood estimates and these estimates are replacing the parameters of the auxiliary model every round of iterations, while the vector structuralized parameters (named RHO in the program codes and their length is LRHO) is estimated with the help of endogenous variables and their lags by the score generator each iteration. EMM shares the same set of folders with SNP plus one more folder called EMMRUN. Very much like SNP program, EMM uses the folder for similar type pf files; NPSOL for the non-parametric solution modules, CODES contains the modules including the one which has the user supplied routines EMMUSR.F as well as the common blocks EMMUNIC0.F and EMMUNIC1.F while EMMRUN holding the input and output files, parameters, controls, and the main module of the program EMM.EXE inside. In CODE folder EMMINCL0.F shows the variables and parameters available to all EMM routines while EMMINCL1.F does the same for variables and parameters for file operations, interface to SNP, related to EMM objective value and parameter vector and NPSOL. EMMINCL2.F declares the

variables and parameters related to EMM standard error computation. EMMOPT.F has the subroutines which take care of the optimization with non parametric solution. The CONTROL.DAT with following contents:

```
xxxxxxxx.xxx yyyyyyyy.yyy  0.10E+00 -0.10E+00        20                454589
```

which is explained previously for SNP program appears in EMMRUN folder, however it has no importance but should exist due to programming structural specifications; instead EMMCNTL.DAT is in charge of managing the parameters' files (named like PARM1.IN or PARM2.IN etc);

```
parm1.in      parmhold.000 parm1.000          0.e00 654321 0000
parm1.in      parmhold.001 parm1.001          8.e-1 654321 0010
```

in each line, the name of the input, hold and output parameters file is followed by the perturbation factor (named TWEAK in program codes), seed and number of trials for the randomly perturbed start values.

The details of the optimization procedure could be reviewed in DETAIL.DAT and the input data SAMPLE.DAT should be stored exactly the same fashion which has suited for SNP program already. Although the input file could be renamed, but such change should be reflected accordingly in SNPOPEN.F and SNPDATA.F as well. The content of input data is $(Y_1, Y_2, \ldots, Y_5)$ the state variables obtained from optimization by VB program plus three additional columns for the date in form of (yy mm dd).

There are three user supplied subroutines that should be changed according to the needs and specification of the desired model; setsim, gensim and simdat all located in the file EMMUSR.F; the subroutine setsim is taking the initial values of the parameters based on the files determined by EMMCNTL.DAT for simulation, gensim produces the simulated data based o the parameters and the selected model and simdat feeds SNP with the simulated data for score generation.

Two very important subroutines are located in the file EMMUOTHR.F which has the main structural role in defining the model; drift and diffuse. It is very important to set up these two subroutines according to the model's specifications carefully $dX_t = (\theta + (-\lambda)X)dt + \sigma dw$:

The input parameters files located in EMMRUN folder are distinguishable with .FIT extension and their contents are used by SNP for simulation and generating score.

The file SUMMARY.DAT is keeping a summary record of optimization process. At the end of a round of iterations based on the information collected in the file SUMMARY.DAT including the objective function values (initial and final) the best candidate of the parameter vector should be selected by operator and the contents of the chosen parameters file have to replace the contents of the PARM#.IN input parameters file for the next round of iteration. The chosen input parameter file for the Gaussian model has non-linear non-parametric abilities which make it the right candidate for the job.

After making all the mentioned alterations and changes the state space variables are the input data for EMM-SNP program. VB writes them in simple text file in five columns however the date will be added as the three first columns and special care needed to make sure all the digits are in the right place. Then the program will run just by calling EMM in command line and SNP will be called by EMM itself. When the program stops, initially the EMMSUMRY.DAT should be checked. This file shows the change in the value of the objective function during the process of optimization as well as indication of using non-linear solution. At this stage the operator has an important role in picking the right choice for the next round of the iterations. Although while the convergence is occurring and the change in the value of the objective function is small this human factor vanishes. After picking the best set of the parameters among the parameter files (named PARM#.000 till PARM#.009) and transfer its content to PARM#.IN and again call the EMM to begin a new round of optimization. Needless to say that the data input set will remain untouched during the whole cycle of EMM-SNP. The file SUMMARY.DAT will keep the record of all these rounds of iterations for any further review. The end of the EMM-SNP cycle will be reached as soon as the no change in the value of the objective function observed in the EMMSUMRY.DAT and the final values of the cycle will be ready to transfer to the VB program as the mean reversion parameters $\lambda_1, \lambda_2, \ldots, \lambda_5$ and $\sigma_1, \sigma_2, \ldots, \sigma_5$ as well as means of the transition process $\varphi_1, \varphi_2, \ldots, \varphi_5$ while σ_i is the standard deviation that makes the covariance matrix for the cross-sectional procedure accomplishes by QUQ module in the main VB program. To obtain the optimized time series of state variables as well as main parameters $\alpha_1, \alpha_2, \ldots, \alpha_5$ which are optimized by the main program while repeatedly using its value as objective function.

4.2.7 EMM-SNP and VB Cyclic Procedure

The EMM-SNP model is basically intended to be the optimizer for state transition process $\varepsilon_{it} = Y_i(t+1) - e^{-\alpha_i \delta t} Y_i(t) - \gamma_i$ and estimation of mean reversion parameters $\lambda_1, \lambda_2, \ldots, \lambda_5$ which results in obtaining $\phi_1, \phi_2, \ldots, \phi_5$ and standard deviations. So, as it is partly described before, in VB program the inputs are taken in and all initial amounts for variables and parameters are set. A counter for the loop measures the number of cycles while a an incremental instrument variable prevent the program to trap in a saddle point

Then the parameter optimization begins with $\overset{\circ}{a}$ after setting an initial amount for the objective function of QUQ module and also determining a domain small enough to allow run searching for an extremum point in it. During each round of the optimization for the parameter the rest will remain untouched. At the end of the each round all the results and values stored for record as well using for optimizing the next variable.

This will result optimizing the main parameters of the model $\alpha_1, \alpha_2, \ldots, \alpha_5$; then the mean reversion parameters $\lambda_1, \lambda_2, \ldots, \lambda_5$ are optimized by transferring the optimum state variables as an input file appropriate for EMM–SNP modules. Then the state transition process $ds_{it} = (a_{is} + a_{iss} s_{it})dt + (b_{i2s})dw$ will be optimized and $a_{is}, a_{iss}, b_{i2s}, i = 1,2,\ldots,5$ are estimated accordingly as $\varphi_i, \lambda_i, \sigma_i$. Therefore φ_i, λ_i are considered the optimum values for the next iteration of VB program and σ_i produces the optimum covariance matrix using the correlation matrix of state variables which are remained unchanged during the process in EMM–SNP:

$$OptimumCo\,\mathrm{var}\,ianceMartix = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{15}\sigma_1\sigma_5 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{25}\sigma_2\sigma_5 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{15}\sigma_1\sigma_5 & \rho_{25}\sigma_2\sigma_5 & \cdots & \sigma_5^5 \end{bmatrix}$$

The VB program has the ability of running several iterations with automatic perturbation of parameters for achieving better results faster, however the process of choosing the optimum parameters based on the scores generated by SNP needs operator.

Therefore the reciprocal exchange of the results between VB and EMM – SNP is part of the cyclic coordinate method implemented.

4.3 Sample Results

The following are the sample results of several rounds of the iterations to converging to the final results. The initial values for the first run was taken directly from the EMM – SNP program package suggested for weekly (Fridays) 3-Month Treasury Bill Rate as raw data fro a two factor with one observable factor and suits quite well for initial values of the data taken from VB program. The parameter file used for the purpose is PARM#.IN; the best candidates are the set of parameters which are neither too far from the optimum amount in their starting point nor above it.

This should be kept in mind that when the contents of the above file is going to be transferred to PARM1.IN file none of the last four tables should be included to keep the format correct. The more detailed results of the optimization are in appendix 1, and here the summarized table for the ease of review, is in appendix 1, so here is the results of the last trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.02923102	0.02401625	0.02903608	0.02229597	0.02192011
WaldStd.Err	0.00820557	0.00917974	0.00676444	0.00316966	0.00793021
T Ratio	3.5623383	2.61622363	4.29246112	7.03418249	2.76412834
λ	-0.0067062	-0.00564351	-0.00562409	-0.00412537	-0.00503671
WaldStd.Err	0.00205311	0.0026373	0.00110426	0.00039045	0.00187485
T Ratio	-3.26635338	-2.13987887	-5.09308337	-10.5656368	-2.68645369
σ	0.0785083	0.06709397	0.09145392	0.09620178	0.07169768
WaldStd.Err	0.00956685	0.00331985	0.01868939	0.01045432	0.0087991
T Ratio	8.20628714	20.20994082	4.89335967	9.20211089	8.14829405

The comparison between the last two trials proves occurrence of the convergence clearly.

4.3.1 Data Handling and Transfer between Programs

The time series file produced by VB program is the optimum state variables (OptStateVariables.txt) which have the following format (the table contents are the first and last five lines of 1000 lines of state variables time series):

Y1	Y2	Y3	Y4	Y5
0.229394	-0.14784	-0.12929	0.357497	-0.33591
0.256895	-0.1829	0.0974	-0.12922	-0.06291

0.253454	-0.18133	0.242946	-0.54334	0.22065
0.233696	-0.15863	0.094795	-0.22054	0.042348
0.257213	-0.18728	0.162373	-0.3722	0.148426
0.134449	-0.1509	-9.92E-02	0.098477	-7.51E-02
0.128986	-0.14359	-0.11961	0.152531	-0.11094
0.163032	-0.18941	7.98E-02	-0.33443	0.190159
0.157829	-0.17785	1.63E-02	-0.19579	0.105619
0.146835	-0.16497	8.79E-03	-0.19246	0.10919

This form of data is not suitable for the EMM – SNP program and should be converted to the following format in (SAMPLPLE.DAT):

82	1	6	0.229394	-0.14784	-0.12929	0.357497	-0.33591
82	1	13	0.256895	-0.1829	0.0974	-0.12922	-0.06291
82	1	20	0.253454	-0.18133	0.242946	-0.54334	0.22065
82	1	27	0.233696	-0.15863	0.094795	-0.22054	0.042348
82	2	3	0.257213	-0.18728	0.162373	-0.3722	0.148426
1	4	25	0.134449	-0.1509	-9.92E-02	0.098477	-7.51E-02
1	5	2	0.128986	-0.14359	-0.11961	0.152531	-0.11094
1	5	9	0.163032	-0.18941	7.98E-02	-0.33443	0.190159
1	5	16	0.157829	-0.17785	1.63E-02	-0.19579	0.105619
1	5	23	0.146835	-0.16497	8.79E-03	-0.19246	0.10919

The accepted format for the above mentioned file is three fields of integers (three digits each) for date and then five fields with floating point (eleven digits including three digits for decimals) for the data. And this should be kept and maintained using excel. Fortran is quite sensitive in format of read and write commands and any mismatch may cause an error.

However, the covariance matrix returning to the VB program which is calculated based on the results of EMM – SNP is more flexible in format and has to be just simply a comma delimited five by five matrix like following table:

0.006164	-0.00412	-0.00085	0.000525	-0.00037
-0.00412	0.004502	0.000872	-0.00056	0.000258
-0.00085	0.000872	0.008364	-0.00865	0.006253
0.000525	-0.00056	-0.00865	0.009255	-0.00684
-0.00037	0.000258	0.006253	-0.00684	0.005141

As it was explained before the calculations of the covariance matrix elements are done using the following formula

$$OptimumCo\,\mathrm{var}\,ianceMartix = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{15}\sigma_1\sigma_5 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \cdots & \rho_{25}\sigma_2\sigma_5 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{15}\sigma_1\sigma_5 & \rho_{25}\sigma_2\sigma_5 & \cdots & \sigma_5^5 \end{bmatrix}$$ taking the $\rho_{i,j}$ s

from the correlation matrix of the optimum state variables:

1	-0.78286	-0.11797	0.06955	-0.06587
-0.78286	1	0.142135	-0.0867	0.053599
-0.11797	0.142135	1	-0.98315	0.953709
0.06955	-0.0867	-0.98315	1	-0.99134
-0.06587	0.053599	0.953709	-0.99134	1

4.3.2 Complete Results

Complete results could be found in appendix.

Chapter 5 Analysis and Briefed Results

While it is already described and explained in detail, two major programs are involved in this project; hence, they are working one after the other and the result of the latter is based on the performance of the former. The followings are results for the data and main models. The results are covering the whole sample from 1882 to 2002 and obtained by running model over the prepared data. Since the best results have been achieved with five factor model it has been the final choice for the number of the factors in model selection. Minimum least square has guaranteed the minimum measurement errors and the fittest parameters simultaneously. The following tables and charts are showing the obtained results and some comments are added where necessary.

5.1 Optimum Objective function Value

The optimum amount reached in cross section part by VB program for the minimum least square of measurement errors is 0.00028890586685987 or less than 0.3 basis point.

5.2 Optimum Asymptotic Forward Rate ($\overset{\circ}{a}$)

The value of the optimum $\overset{\circ}{a}$ asymptotic forward rate is equal to 1 basis point.

5.3 Optimum Model Parameters (α)

Optimum parameters achieved by VB program

α_1	α_2	α_3	α_4	α_5
0.022316	0.068533	1.523435	2.153859	2.641968

5.4 Optimum Mean Reversion Parameters (λ)

Optimum Lambda Parameters obtained by EMM–SNP

λ_1	λ_2	λ_3	λ_4	λ_5
-0.00671	-0.00564	-0.00562	-0.00413	-0.00504

5.5 Optimum State Variables (time series)

The optimum state variables (the ten first and last elements of each time series)

Y1	Y2	Y3	Y4	Y5
0.229394	-0.14784	-0.12929	0.357497	-0.33591
0.256895	-0.1829	0.0974	-0.12922	-0.06291
0.253454	-0.18133	0.242946	-0.54334	0.22065
0.233696	-0.15863	0.094795	-0.22054	0.042348
0.257213	-0.18728	0.162373	-0.3722	0.148426
0.2488	-0.17517	0.120404	-0.2722	0.088626
0.232543	-0.16133	0.140455	-0.28038	0.08318
0.240017	-0.17764	0.261049	-0.56251	0.22549
0.241834	-0.1826	0.310856	-0.70716	0.326256
0.253261	-0.19534	0.296856	-0.70541	0.342943
0.131613	-0.15734	5.23E-02	-0.21395	0.102037
0.133854	-0.15692	5.30E-02	-0.24612	0.131202
0.136422	-0.15981	-7.16E-03	-9.15E-02	3.39E-02
0.141682	-0.16412	9.63E-03	-0.11852	4.18E-02
0.142464	-0.1647	2.68E-02	-0.17882	8.34E-02
0.134449	-0.1509	-9.92E-02	0.098477	-7.51E-02
0.128986	-0.14359	-0.11961	0.152531	-0.11094
0.163032	-0.18941	7.98E-02	-0.33443	0.190159
0.157829	-0.17785	1.63E-02	-0.19579	0.105619
0.146835	-0.16497	8.79E-03	-0.19246	0.10919

The optimum state variables descriptive statistics

STATE VARIABLES	Y1	Y2	Y3	Y4	Y5

Mean	0.176041	-0.16829	0.074829	-0.29367	0.141422
Median	0.169841	-0.16466	0.105069	-0.34382	0.162905
Maximum	0.260178	-0.08288	0.392359	0.469452	0.687253
Minimum	0.081522	-0.24285	-0.2907	-1.14094	-0.33591
Std. Dev.	0.038811	0.029301	0.129779	0.283574	0.164089
Skewness	0.025427	-0.01816	-0.80509	0.548421	-0.358
Kurtosis	2.073636	2.852254	3.284018	3.215599	3.215052
Jarque-Bera	35.864	0.964488	111.3887	52.06432	23.28706
Probability	1.63E-08	0.617396	6.49E-25	4.95E-12	8.78E-06
Observations	1000	1000	1000	1000	1000

5.6 Correlation Matrix

Optimum State Variables correlation matrix:

	Y1	Y2	Y3	Y4	Y5
Y1	1	-0.78286	-0.11797	0.06955	-0.06587
Y2	-0.78286	1	0.142135	-0.0867	0.053599
Y3	-0.11797	0.142135	1	-0.98315	0.953709
Y4	0.06955	-0.0867	-0.98315	1	-0.99134
Y5	-0.06587	0.053599	0.953709	-0.99134	1

5.7 Covariance Matrix

Optimum Covariance Matrix:

	Y1	Y2	Y3	Y4	Y5
Y1	0.006164	-0.00412	-0.00085	0.000525	-0.00037
Y2	-0.00412	0.004502	0.000872	-0.00056	0.000258
Y3	-0.00085	0.000872	0.008364	-0.00865	0.006253
Y4	0.000525	-0.00056	-0.00865	0.009255	-0.00684
Y5	-0.00037	0.000258	0.006253	-0.00684	0.005141

5.8 Optimum Means of State Transition Process

Optimum means of state transition process array by EMM–SNP

φ_1	φ_2	φ_3	φ_4	φ_5
0.029231	0.024016	0.029036	0.022296	0.02192

5.9 Comparison with peer group models

As a contribution to the New Palgrave Dictionary of Economics, (2nd ed. 2005) Brandt and Chapman described that usually the short rate is the single-factor in the affine model of bond prices by specifying a continuous-time process for the short rate and one of the suggested forms of the risk premium function. Besides the idea of bringing up the multifactor models idea, in case of insufficiency of single-factor models to explain the observed term structure, the number of factors needed for the model dynamics is raised. They describe multifactor affine term structure models by following three major components:

A linear relation between the short rate and the factors; the factor dynamics conform to an affine diffusion and the specification of the market prices of risk.

If no elements of Y affect the conditional volatility, then it is a multifactor generalization of the Vasicek model. If m < N factors affect the conditional volatility, then the multifactor affine model is a mixture of the CIR and Vasicek forms. Dai and Singleton (2000) define different classes of affine models by the number of factors that affect the

conditional factor volatilities, with A_m (N) being the general notation for an N-factor model with m-factors driving conditional volatilities.

Brandt and Chapman (2005) explain how Duffee (2002), Dai and Singleton (2002) and themselves have tried multifactor affine term structure models to U.S. bond data with A_m (3) models (for m = 0, 1, 2, 3) and all found that a Gaussian version (an A_0 (3) model) can rationalize the risk premiums revealed by yield change regressions. They also have mentioned to Duffee (2002) that shows an A_0 (3) model with the expanded risk premium specification can produce meaningful multistep forecasts of Treasury yields at different maturities.
Brandt and Chapman (2005) are not clear whether or not their latent factors can be connected in any meaningful way to structural (macroeconomic) explanations of term structure dynamics.

Ang and Piazzesi (2003) and Ang, Dong, and Piazzesi (2005) have recently tried to connect no-arbitrage term structure models with elements of the macroeconomics.
They have provided partitioned structure by merging macroeconomic and latent factors in an affine Gaussian model in the first step (2003) and then improved it by letting the macroeconomic factors to send a feed back to the latent factors.

The fact that the developed model here has removed calculation obstacle of simultaneous work with time series and panel data, has given it a unique status quite different with other previously accomplished experiments, on the other hand, the utilized techniques will show their efficiency advantage in higher number of factors and not necessarily at the same level of traditional approaches for three factors.

Chapter 6 Conclusions

6.1 Discussion and Explanations

There are points and informative hints about either of the two parts of the model which is going top be explained and discussed further more.

6.1.1 Comments on Cross Section

In the cross section part, the VB programs show considerable strength in helping the operator overcoming any unexpected technical obstacle. Ease of working in visual studio and flexibility of exchange data and variable files with other applications makes it an ideal choice for being substrata of the model. This has brought more convenience for editing or making any necessary improvement for operator. The automatic perturbation of parameters prevents any possible jump cut while satisfactorily helps th3 conduct of optimization all the ay long. The time consumption is comparatively reasonable and keeps its pace in consecutive iterations.

6.1.2 Comments on the Time Series

The acclaimed program developed by Gallant and Tauchen have an old structure apparently due to their development under UNIX system. They are not fit to commonly available environments and have to run in DOS mode. The EMM works as the optimizer while SNP is the score generator in a master-slave relationship. Although SNP could be used independently as an distinct score generator for any other fitting application. EMM takes the initial parameters in specific files and produces a set of nine suggested parameter sets that each can bring the objective function value to its local maximum in that iteration. However it remains the operator's duty to pick the most suited set and transfer its contents to the parameter intake file for the next iteration. The optimization will be achieved when the in the summary file the objective function value in the end shows equal amount with its starting value. It took between two to four iterations for optimum state variables to reach to the optimum estimated parameters. The final results are found among the above mentioned parameters files and should be picked carefully and recorded for any possible future iterations (i.e. in case of new state variables series)

The value of the objective function does not imply any meaningful statement and merely prepares a comparison basis for making the right choice among candidates.

6.2 Decision Making & Recommended Solutions

Upon the results obtained and analysis performed in the previous parts, for different cases the optimum approach would be a multifactor model that not only enabled to communicate with macroeconomic models but also provide better results with the real world data. Besides the available mathematical applications and computer facilities it is more convenient to make fine tuning and make modifications in special cases that makes it more flexible to special needs and applicable to diversified financial claims.

As a main general group of multifactor models that allow for flexible functional forms while retaining tractability, are generalized as the affine multifactor models in which the instantaneous interest rate is modeled as an affine function of the states and quadratic functions which allows for more flexible description of risk premia; and the quadratic term structure model, in which the instantaneous rate is modeled as a quadratic function of the state variables. Either of them has allowed flexibility in the correlation structure of the state variables and the sources of conditional volatility of yields.

Some of the models do not necessarily have closed-form likelihood functions and instead they are estimated by various forms of moments or approximate likelihood estimators. Therefore the models that try to approximate the unobservable dynamics are bound to have specification errors and subject to compare along the consistency of the implied structural parameters with the data. In the correctly specified models, the parameters implied by bond yields should be consistent with those implied by the time-series and cross-sectional data.

In addition to that, matching with the stylized facts of the market (e.g. US treasury bonds) has a crucial importance. Surprisingly, all the multifactor essentially affine models fail to simultaneously match the conditional mean and the conditional volatilities of holding period returns. Duffee (2002) argues that the essentially affine models emphasize flexibility in the correlation structure of the state variables by shutting down the conditional variation in yield volatilities. So Dai and Singleton (2000), and Duffee (2002) show that the essentially affine models can fit either the first moment or the

second moment of yields, but not both. Therefore, the Gaussian quadratic model, seem to be the natural winner by consistently satisfying the full set of moment conditions.

6.3 Critiques on Methodologies & Assumptions

As it has been already stated, each approach has strength and weakness that should be carefully considered in interpretation and explanation of the results obtained form it. Neglecting this important point could lead to nonrealistic conclusions and therefore inefficient strategy and policy making. Since it was more efficient to have critical points explained along with each model, this has been already accomplished in fourth chapter.

However, in general there are few points that may improve the performance of the models as well as keeping them in fashion for coming years. One of them could be expansion towards a web-base application; although it does not need to be done in near future due to slower rate of data production limited availability of the data, but it may change drastically with the next generation of Internet II using protocols version 6.

The other improvement would be object oriented language transformation. Although it is not so crucial at present, the trend of technology shows tendency towards this category of languages and VB seems stand besides the field. Also surfacing faster processors and larger storage devices are encouraging for development of higher precision and larger areas of optimization, which leads to more confidence and accuracy.

6.4 Suggestions for Future Works

Blooming computational facilities in terms of software and hardware has developed a suitable environment to improve current approaches as well as development of new methods to cover the short comes and weak points. Proper suggestions and innovative ideas would be offered in this regard specifically about the number of factors in the multivariate model. As it is already explained, the models with small number of factors have not been compared with higher numbers just to keep up with Occam razor's philosophy. Although it may sound reasonable but it has never been rejected that to what extent it could be true and if it is necessary to utilize more detailed data and necessary procedures to suppress the noise and remove trend's effects. These besides the

improvements that mentioned in the last chapter will make vision substrata, however for future work these could be done perhaps as a Ph.D. thesis.

References

Alphabetically sorted references:

1. Andrew Ang, Monika Piazzesi, (2003) "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables", *Journal of Monetary Economics* #50, pp. 745–787
2. Ang, A., S. Dong, and M. Piazzesi. (2005) "No-Arbitrage Taylor Rules" Manuscript.Columbia University and University of Chicago.
3. Babbs, Simon H., and K. Ben Nowman, (1999), "Kalman filtering of generalized Vasicek term structure models", *Journal of Financial and Quantitative Analysis* 34, No. 1, 115-130.
4. Balduzzi, Pierluigi, Das, Sanjiv Ranjan, Foresi, Silverio, Sundaram, Rangarajan (1996) "A simple approach to three-factor affine term structure models", *The Journal of Fixed Income*. New York: Dec 1996. Vol. 6, Issue 3; pg. 43, 11 pgs
5. Bierwag, G.O and M.A. Grove (1967) "A model of term structure of interest rates", *Review of Economics and statistics*, 49; 50-62
6. Brandt, Michael and Ping He (2002) "Simulated likelihood estimation of affine term structure models from panel data", Working paper, Wharton School.
7. Brennan, M J and E.S Schwartz (1979) "Conditional Productions of bond prices and returns", *Journal of Finance*, 35; 405-417
8. Chambers, D R, W.T Carlton and D.W Waldman (1984) "A new approach to the estimation of the term structure of interest rates", *Journal of Financial and Quantitative analysis,* 19; 233-252
9. Chernozhukov V. and Han Hong (2003) "An MCMC approach to classical estimation", *Journal of Econometrics* 115 293-346.
10. Cox J C, J E. Ingersoll Jr. and S.A. Ross (CIR) (1985) "An Intertemporal general equilibrium model of asset prices", *Econometrica*, 53; 363-384
11. Cox, J.C., J.E. Ingesoll, and S.A. Ross. (1985) "A theory of term structure of interest rates", *Econometrica* 53, pp. 384-407
12. Culbertson, J.M (1957) "The term structure of interest rates", *Quarterly Journal of Economics,* 71; 485-517

13. Dai Qiang and K. J. Singleton, (2000) "Specification Analysis of Affine Term Structure Models", *Journal of Finance*, 55, pp. 1943-1978.
14. Dai, Qiang and Kenneth J. Singleton, (2002), "Expectation puzzles, time varying risk premia, and affine models of the term structure", *Journal of Financial Economics* 63, 415-441.
15. Diller, S (1969) "Expectations and the term structure of interest rates", in J. Mincer Ed., *Economic forecasts and expectations,* New York, National Bureau of Economic Research.
16. Duan, J.C. and J.G. Simonato, (1999), "Estimating and testing exponential affine term structure models by Kalman filter", *Journal of Business and Economic Statistics* 18, 300-314.
17. Duarte, Jefferson (2003) "The relevance of parameterization of market price of risk in affine term structure models", *Review of Financial Studies* 17, No. 2, 379-404.
18. Duffee, G. R. (2002) "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, No 57 – 1, pp. 405-433.
19. Duffee, G. R., and R. H. Stanton (2001) "Estimation of Dynamic Term Structure Models", work paper, 2001, Haas School of Business, University of California at Berkeley.
20. Duffie, D. and R. Kan (1996) "A Yield-Factor Model of Interest Rates", *Mathematical Finance*, No 6, 1996, pp. 370-406.
21. Durand, D (1942) "Basic yields of corporate bonds", 1900-1942, Technical paper No 3, National Bureau of Economic Research.
22. Fisher, I (1896) "Appreciation and interest", Publications of American Economic association, pp 23-29and 88-92
23. Fisher, I (1930) *Theory of Interest*, New York; McMillan
24. Fong, H.G. and Oldrich A. Vasicek (1992) "Omission impossible", *Risk*, Vol.5, No. 2, 1992, pp. 62-65
25. Friedman B M and F H Hahn, (1990) *Handbook of Monetary Economics*, Vol. I, Elsevier Science Publisher
26. Friedman, B M. (1979) "Interest rate expectations versus forward rates: Evidence from an expectations survey", *Journal of Finance,* 34; 965-973

27. Gallant, A Ronald & Tauchen, George (1996) "Which Moments to Match?" Econometric Theory, Cambridge University Press, vol. 12(4), pages 657-81
28. Gallant, A. Ronald & Tauchen, George (2002) "Simulated Score Methods and Indirect Inference for Continuous-time Models", Working Papers 02-09, Duke University, Department of Economics
29. Gallant, A. Ronald, and Jonathan R. Long (1997) "Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square", *Biometrika*, 84, 125-141
30. Gallant, Ronald & Tauchen, George (1989) "Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications", *Econometrica*, Econometric Society, vol. 57(5), pages 1091-1120
31. George Tauchen, 1997 (printed in 1998) "The Objective Function Of Simulation Estimators Near The Boundary Of The Unstable Region Of The Parameter Space", *The Review of Economics and Statistics*, MIT Press, vol. 80(3), pages 389-398
32. Guo C (2001) "A path independent solution to the term structure of interest rates", working paper, University of Ottawa, August 2001
33. Guo C (2003) "A unique multivariate Gaussian term structure model from Heath-Jarrow-Morton methodology", working paper, University of Ottawa, April 2003
34. Heath David, Robert Jarrow and Andrew Morton (1992) "Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation", *Econometrica*, Vol.60, No. 1, 1992, pp.77-105
35. Hickman, W B (1942) "The term structure of interest rates; an exploratory analysis", National Bureau of Economic Research, Results in Kessel (1965, Appendix A. pp103-105)
36. Hicks, J R. (1946) *Value and Capital*, 2nd ed. Oxford, Oxford University Press
37. Jordan, J V (1984) "Tax effects I term structure estimation", *Journal of Finance*, 39; 393-406
38. Kane, E.J and B G Malkiel (1967) "The term structure of interest rates: an analysis of survey of interest rate expectations", *Review of Economics and Statistics*, 49; 343-355

39. Kessel, R A (1965) *The cyclical behavior of the term structure of interest rates*, New York, National Bureau of Economic Research.
40. Lutz, F A (1940) "The structure of interest rates", *Quarterly Journal of Economics*, 55; 36-63
41. Macaulay, F.R (1938) *Some theoretical problems suggested by the movements of interest rates, bond Yields and stock prices in the United States since 1856,* New York, National Bureau of Economic Research.
42. Malkiel, Burton G. (1966) *The term structure of interest rates: Expectations and behavior patterns*, Princeton; Princeton University Press.
43. Malkiel, Burton G. and Edward J. Kane (1969) "Expectations and Interest rates: A cross sectional test of the error-learning hypothesis ", *Journal of Political Economy*, Jul/Aug 1969, pp. 453-470
44. McCulloch, J. H (1971) "Measuring the term structure of interest rates", *Journal of Business*, 44; 19-31
45. McCulloch, J. H (1975b) "The tax adjusted yield curve", *Journal of Finance* , 30; 811-830
46. Meiselman, D (1962) *The term structure of interest rates* Englewood Cliffs; Prentice Hall
47. MichaelW. Brandt; David A. Chapman (2005) "Affine Term Structure Models', *New Palgrave Dictionary of Economics*, 2nd edition, Palgrave Macmillan, 2005
48. Nelson, C R (1970a) "A critique of some recent empirical research in the explanation of the term structure of interest rates", *Journal of political economy*, 78; 764-767
49. Nelson, C R and A F Siegel (1985) "Parsimonious modeling of yield curves for US treasury bills", National Bureau of Economic Research Working Paper 1594
50. Richard, Scott F. (1978) "An arbitrage model of the term structure of interest rates, *Journal of Financial Economics*, No. 6, 1978, pp. 33-57
51. Schaefer, S. M (1981) "Measuring a tax-specific term structure of interest rates in the market for British Government securities", *Economic Journal*, 91; 415-438

52. Shea, G. S (1984) "Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximations", *Journal of Financial and Quantitative Analysis*, 19; 253-269
53. Shea, G. S (1985) "Interest rate term structure equations with exponential splines: A note", *Journal of Finance*, 40; 319-325
54. Vasicek Oldrich A. and H. Gifford Fong (1982) "Term structure modeling using exponential splines", The Journal of Finance, Vol. 37, No. 2, May 1982, pp. 339-348
55. Vasicek, Oldrich A, (1977) "An Equilibrium Characterization of the term structure", *Journal of financial economics*, No 5 (1977), pp.177-188.
56. Williams, J B (1938) *The theory of investment value*, Cambridge, Mass Harvard University Press.

Appendices

Appendix 1: Detailed Results

The detailed results of the optimization in time series part by EMM – SNP:

First Trial:

	Y1	Y2	Y3	Y4	Y5
φ	1.14914202	0.31031954	0.11179444	1.50604953	1.49145866
WaldStd.Err	2.36708645	0.00381321	0.00065880	0.01056184	0.01629031
T Ratio	0.48546686	81.38009784	169.69445023	142.593422	91.5549866
α	0.11085308	0.04200621	0.02052247	3.69333711	0.27037033
WaldStdErr.	0.23898768	0.00015066	0.00007665	0.00478538	0.00239300
T Ratio	0.46384432	278.80617693	267.75504124	771.795485	112.983732
σ	0.37105847	2.56833253	1.11305408	6.91782288	2.03647134
WaldStdErr.	0.02121878	0.04253743	0.00331420	0.01757099	0.01752982
T Ratio	17.48726988	60.37817577	335.84399874	393.706989	116.171849

Second trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.01301301	0.00143600	0.01301300	0.00377785	0.02390020
WaldStd.Err	0.00088828	0.00015136	0.00088831	0.00034317	0.00571092
T Ratio	14.64967200	9.48740209	14.64912021	11.00872309	4.18499901
λ	0.00340253	0.00021120	0.00340253	0.00084527	0.00444475
WaldStd.Err	0.00015987	0.00001514	0.00015987	0.00004721	0.00121278
T Ratio	21.28317597	13.94966635	21.28327413	17.90345380	3.66492048
σ	0.06195090	0.04612204	0.06195089	0.04988340	0.08819326
WaldStd.Err	0.00339349	0.00308629	0.00339366	0.00310684	0.00949734
T Ratio	18.25582091	14.94415039	18.25491203	16.05598194	9.28610032

Third trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.01301301	0.001436	0.0214989	0.00374345	0.0239002
WaldStd.Err	0.00088828	0.00015132	0.00048117	0.00009867	0.00571077
T Ratio	14.64961170	9.48982351	44.68002128	37.94031941	4.18510641
λ	0.00340253	0.0002112	0.00408768	0.00086792	0.00444474
WaldStd.Err	0.00015987	0.00001514	0.00011764	0.00003117	0.00121275
T Ratio	21.28318276	13.95309929	34.74602680	27.84526187	3.66499871
σ	0.0619509	0.04612201	0.11951793	0.05041306	0.08819341
WaldStd.Err	0.00339350	0.00308564	0.00180719	0.00012112	0.00949706

T Ratio	18.25574618	14.94732138	66.13460424	416.23725386	9.28638651

Fourth trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.01316044	0.00140750	0.02173450	0.00380800	0.02849781
WaldStd.Err	0.00096934	0.00002226	0.00597816	0.00012662	0.00949774
T Ratio	13.57672056	63.23864082	3.63565002	30.07460580	3.00048298
λ	0.00343791	0.00020536	0.00344755	0.00087592	0.00632117
WaldStd.Err	0.00025027	0.00000262	0.00053197	0.00001893	0.00200274
T Ratio	13.73660035	78.25801949	6.48073256	46.27798995	3.15626790
σ	0.06219157	0.04599619	0.11193346	0.05068351	0.08066631
WaldStd.Err	0.00407038	0.00082414	0.02405581	0.00087004	0.01318713
T Ratio	15.27907554	55.81119540	4.65307434	58.25436333	6.11704680

Fifth trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.01310319	0.00142044	0.02021021	0.00379424	0.02787403
WaldStd.Err	0.00117672	0.00006816	0.00444786	0.00007498	0.00856102
T Ratio	11.13531004	20.83974382	4.54380556	50.60371598	3.25592366
λ	0.00342253	0.00024968	0.00307462	0.00087464	0.00616437
WaldStd.Err	0.00024847	0.00000926	0.00029209	0.00002070	0.00191131
T Ratio	13.77425969	26.97734247	10.52641700	42.25193218	3.22521308
σ	0.06214815	0.04663705	0.11305154	0.05058729	0.08027531
WaldStd.Err	0.00455274	0.00107168	0.02246296	0.00066304	0.01041471
T Ratio	13.65072740	43.51772212	5.03279843	76.29646082	7.70787685

Sixth trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.01313726	0.02381157	0.02353076	0.02785881	0.02283760
WaldStd.Err	0.00153228	0.01513753	0.00147483	0.00527682	0.00806639
T Ratio	8.57369271	1.57301525	15.95494952	5.27947071	2.83120469
λ	0.00343497	0.00565115	0.00385617	0.00602525	0.00518951
WaldStd.Err	0.00035710	0.00375386	0.00018047	0.00156879	0.00176273
T Ratio	9.61902870	1.50542258	21.36690312	3.84071112	2.94402345
σ	0.06230691	0.06608245	0.12262466	0.08225284	0.07324330
WaldStd.Err	0.00545082	0.00817674	0.00740378	0.00626169	0.01046647
T Ratio	11.43074734	8.08176022	16.56244182	13.13588953	6.99790071

Seventh trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.02425503	0.02336682	0.02879577	0.02231367	0.02188960
WaldStd.Err	0.00949900	0.01529734	0.00321431	0.00376610	0.00725456
T Ratio	2.55343045	1.52750892	8.95861484	5.92486697	3.01735586
λ	0.00618489	0.00555811	0.00487942	0.00411792	0.00502632
WaldStd.Err	0.00268791	0.00369564	0.00081009	0.00062358	0.00171409
T Ratio	2.30100265	1.50396516	6.02329931	6.60372591	2.93235965
σ	0.06816505	0.06540147	0.11242224	0.09642440	0.07173577
WaldStd.Err	0.00342666	0.00849684	0.00729744	0.00765319	0.00741301
T Ratio	19.89258228	7.69715633	15.40571502	12.59924576	9.67700369

Eighth trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.02423263	0.02347715	0.0299549	0.02243029	0.02106726
WaldStd.Err	0.00784738	0.0164543	0.01050049	0.00358364	0.00662892
T Ratio	3.0879892	1.42680955	2.8527156	6.25907973	3.17808484
λ	0.00613891	0.00571486	0.00569753	0.00414853	0.00489166
WaldStd.Err	0.00244765	0.00416324	0.00167759	0.00045826	0.00166495
T Ratio	2.50807865	1.37269648	3.39626091	9.05283293	2.93801623
σ	0.06849305	0.06388055	0.09421372	0.09641698	0.07041878
WaldStd.Err	0.00107034	0.00820408	0.01956519	0.00974856	0.00551337
T Ratio	63.99165952	7.78643836	4.81537558	9.89037908	12.77236495

Last trial:

	Y1	Y2	Y3	Y4	Y5
φ	0.02923102	0.02401625	0.02903608	0.02229597	0.02192011
WaldStd.Err	0.00820557	0.00917974	0.00676444	0.00316966	0.00793021
T Ratio	3.5623383	2.61622363	4.29246112	7.03418249	2.76412834
λ	-0.0067062	-0.00564351	-0.00562409	-0.00412537	-0.00503671
WaldStd.Err	0.00205311	0.0026373	0.00110426	0.00039045	0.00187485
T Ratio	-3.26635338	-2.13987887	-5.09308337	-10.5656368	-2.68645369
σ	0.0785083	0.06709397	0.09145392	0.09620178	0.07169768
WaldStd.Err	0.00956685	0.00331985	0.01868939	0.01045432	0.0087991
T Ratio	8.20628714	20.20994082	4.89335967	9.20211089	8.14829405

The comparison between the last two tables proves occurrence of the convergence clearly.