UNIVERSITY OF CALGARY

On the Application of the Wilkie Model to the TSX Price Index

by

Chao Qiu

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

JULY, 2007

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled “On the Application of the Wilkie Model to the TSX Price Index” submitted by Chao Qiu in partial fulfillment of the requirements for the degree of Master of Science.

Supervisor, Dr. Rohana S. Ambagaspitiya
Department of Mathematics and Statistics

Dr. Murray D. Burke
Department of Mathematics and Statistics

Dr. Xifeng Diao
Department of Finance
Haskayne School of Business

Date

Abstract

This thesis evaluates the ability of Wilkie's stochastic investment models to predict TSX price index yield. In the Wilkie Model, correlated econometric indices were modeled through a cascade structure.

The empirical study of TSX price index yield in this thesis shows the following results with regard to the Wilkie Model:

- The advised multivariate models by the Wilkie Model do not make significantly better prediction than univariate models do.
- The dividend yield model is not suitable for prediction based on recent data analysis.
- The suggested new model structure based on vector autoregressive method in this thesis is similar to the Wilkie's structure, but not a cascade one. There are some significant feedback relationships between different components.

Therefore this thesis suggests a multidirectional model structure without dividend yield. The models are used to predict the movement of the components within the structure. The models for those variables is constructed from the linear relationship on their own lagged values and other variables in the structure.

Acknowledgements

I would like to thank in deeply and sincerely to my supervisor Dr. Rohana S. Ambagaspitiya, for his supervision, instruction, guidance for my thesis writing and relevant course studying, and thanks for his encouragement and friendship to me during my master program in the University of Calgary.

I also gratefully thanks other examination committee member Dr. Murray D. Burke, Dr. Xifeng Diao, and Dr. Anatoliy Swishchuk for their instructions and valuable suggestions to my thesis.

I would also like to thank the instruction and support received from Dr. Cheung Ka Chun, Dr. John Collins, Dr. Peter Ehlers, Dr. Xuewen Lu, Dr. David Scollnik and all other professors and faculty staff in the department of Mathematics and Statistics, University of Calgary. I gratefully thank the department giving me such an impressive and valuable experience in my master's degree education.

Table of Contents

List of Tables

Chapter 1

Introduction

For insurance companies, modeling the future values of assets and their interactions is useful for setting up asset strategies and managing reserves for liabilities. In the literature there are a number of different methods to model the return of assets. Panjer (1998) classifies those models into three categories: financial models, econometric models, and actuarial stochastic investment models. Generally we can classify them based on short term and long term application due to the argument that long term movements of asset returns could be very different from their short term movements. For example, Independent Log–normal models (ILN) can be used to represent the random walk, a type of movement for short term share price movement; ARIMA models can detect the long term trend for time series, after some transformation is made for the original series. Also, many well-know asset pricing models are based on Black–Scholes model, which is in turn based on the ILN assumption for the movements of short term asset returns. We can further classify them as univariate models and multivariate models. For example, the Wilkie Model is used to detect the movements of multiple assets and their correlation, while other models mentioned above can be used to model single variables such as TSX Total Return Index yield.

Even though there are many financial and econometric models for investment returns, U.K. actuary David Wilkie is the first person who developed stochastic investment models for the long term return of multiple assets for actuarial application. His models is known as "The Wilkie Model". Wilkie originally carried out such researches for the Maturity Guarantees Working Party of the Institute of Actuaries from 1980 to 1981. As a result of those researches, Wilkie (1986) published his cascade structured models. Those models are designed for four basic assets: inflation rate, dividend yield, dividend, and consols yield. Wilkie (1992) developed this method into modeling exchange rate. Later on Wilkie (1995) developed the models for wages and property yield, and applied those models to other countries such as Canada and the United States. Wilkie (1995) also discussed GARCH models and Vector Autoregressive models with correlated errors for inflation and wages. The application to TSX index yield based on the four basic models is the topic of this thesis. Even though Wilkie (1986) and Wilkie (1995) both discussed asset modeling, there are no significant changes for the models of those four basic assets. Therefore we use the basic assets models presented by Wilkie (1986) for the analysis.

Since the introduction of the Wilkie Model 20 years ago, there has been much research into his hierarchy structural modeling method and related structured multivariate assets pricing models. Research in this area is still ongoing. Mulvey and Thorlacius (1996) introduced a cascading set of differential equations, developed by Towers Perrin, for the future financial scenario generating based on simulating multiple assets and indices. Sharp (1992) applied the Wilkie Model into Canadian inflation and wages. Tomson (1996) applied the Wilkie Model to South Africa by developing a method to determine the structure of multiple time series models.

Beside application and development, there are also some critical voices for the Wilkie Model. Whitten and Thomas (1999) refined the Wilkie Model by introducing a non–linear Threshold Autoregressive (TAR) model to analyze investment series. They applied both threshold techniques and the ARCH modeling method in the cascade structure. Chan and Wang (1998) used the impulse function to make adjustment of outliers for the inflation rate series. They showed that this adjustment is closer to reality and decreases the variance of prediction after removing the impulses. This works better than the AR–ARCH model used by Wilkie (1995). However this methods will produce lower cumulative yield for the future. Huber (1995) criticized the Wilkie Model regarding to its data resources and manipulation, model construction and calibration, and its parameter estimation. These are valid criticism. However we still want to study the advantages and problems in the Wilkie Model in order to improve it. Some summary of the comparison of the Wilkie Model, TAR model, GARCH model and Outlier models are done by Rambaruth (2003).

With the recent development of financial markets, the series of Wilkie's models also needs to be reviewed for its validation in today's markets. That is why we investigate the validity of the Wilkie Model by testing its application to the latest TSX Total Return Index. The following is an outline of different chapters.

Chapter two is the validation of the Wilkie Model. This chapter includes two sections. The first section is a brief introduction to the cascade structure of the Wilkie Model and its multivariate models for 4 basic indices: inflation rate, dividend yield, dividend, and consols yield. The second section is an evaluation of the performance of the Wilkie Model. This evaluation is based on the prediction of TSX Total Return Index yield. In this chapter we also use other methods as a comparison. Those methods include Independent Log-normal models, univariate ARIMA – GARCH models, and Regime – Switching models.

In the third and fourth chapters we discuss some updates for the Wilkie Model. In the third chapter we investigate the validation of the component dividend yield in the Wilkie Model. After reconsidering the involvement problem for dividend yield, we try to test a similar model structure based on the Wilkie Model without dividend yield. Therefore in chapter four, the structure is constructed based on the testing of relationship between different components and their lagged values, and some rollover values. The testing includes multivariate regression models, test of the residuals of univariate time series models, and Vector Autoregressive models.

Due to the complexity of the movement for multiple indices, and the significant changes in the world markets which have strong impact on the values of financial assets, we need to keep up the calibration of the multivariate models in the future, in order to make their prediction valuable.

Chapter 2

Validation of the Wilkie Model

The key point in the Wilkie Model is to construct a multi–level hierarchy structure for multiple variables. The models are built on the linear relationships between dependent variables and predictors and their lagged values.

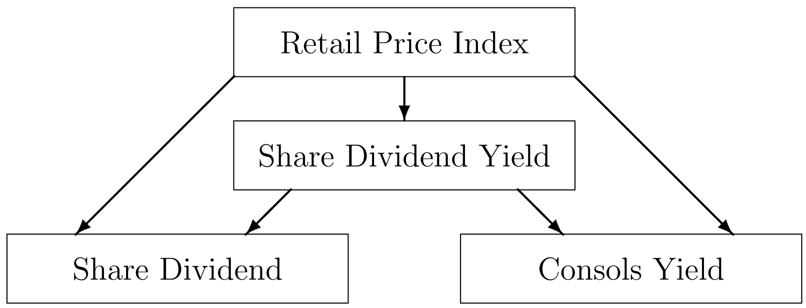
2.1 Introduction to the Wilkie Model

2.1.1 Structure of the Wilkie Model

The fundamental parts of the Wilkie Model for the UK market include four variables as follows:

- Retail price index (Q)
- Share dividends index (D)
- Dividend yield (Y) on share price index (P)
- Consols yield or long–term government interest rate (C).

The models of those four variables are built within a cascade structure such that those variables are ordered from the top level to the lower levels. The values of the lower level variables depend on the lagged value of themselves and the values of the variables in the upper levels. From the structure plot following this paragraph, we can see that inflation rate, which is calculated from the change of retail price index, depends on its own lagged values and is placed on the top layer of the structure. Its prediction is based on its own historical performance. Dividend yield is in the second layer, and its prediction is based on both the historical performance of itself and that of inflation rate. The third layer includes dividend and consols. Their prediction is based on the historical data of themselves, that of dividend yield, and that of inflation rate.



The variables of the Wilkie Model were modeled under the following procedures. Firstly the variables are modeled by the regression on the upper level variables. Secondly the models' residuals were tested and constructed through the standard Box–Jenkins univariate time series modeling method. Beside four basic variables mentioned above, Wilkie expanded his model to some other econometric factors like wages, short-term interest rate, property yield and exchange rate. Also he tested other methods such as Vector Autoregressive (VAR) modeling of two correlated variables, and GARCH modeling for the variance of residuals. There are two features of the Wilkie Model that we need to pay attention:

- Wilkie (1995) stated that there is a cointegration relationship between the logarithm of dividend $\log(D)$ and the logarithm of share price $\log(P)$. So the model uses dividend yield $Y(t)$, where $\log Y(t) = \log(D) - \log(P)$, as a stationary variable to model this two variables.

- The Wilkie Model chooses one way influence from the higher level variables to the lower level variables, even though there is some two – way interaction. Wilkie claimed that two – way interaction models in this structure would not improve the accuracy of prediction but make the models much more complicated. An advantage of the Wilkie Model is that its structure could expand easily to be a larger system with more layers and more variables in those layers. This expansion does not need to change the models for the upper level variables. However any errors from the prediction of upper level variables could also be transferred to the prediction of lower level variables; the effects of this error transferring depend on the weight given to upper level variables in modeling lower level variables. In general, the upper level variables play more important roles in the structure.

Beside the structural problems, and the problems mentioned by Huber (1995), there are some other facts to consider when applying the Wilkie Model into North American markets.

- The selection of variables is lack of direction. As we see, Wilkie did not tell us how to select the component variables into his models. So we need to check the validation of current components at first. For example dividend yield is not a good choice based on the recent stock market performance where dividend yield is not very stationary. Intuitively, dividend yield is not a listed financial product in the stock market. It is a latent variable in the Wilkie Model. Dividend and share price are two products in the financial markets. Dividend yield is a financial derivative calculated from those two products. When the stock market adjusts share price and share dividend separately, it is possible that both share price and share dividend show some stationary quality, but dividend yield does not. In another words, it is highly probable for dividend yield not be a very stationary when in the market either share dividend or share price departures from their expected movement. This fact could lead to the change of the structure of the Wilkie Model into a new one having fewer components.

- Through the comparison of the Wilkie Model with univariate models such as ILN model, ARIMA–GARCH model and Regime–Switching model, we want to test the ability of prediction based on the multi–variate Wilkie Model. The test was done by calculating the log–likelihood values of the Wilkie Model and the other models. Usually models with more parameters have high log–likelihood values. However this is not met by the Wilkie Model when compared with other univariate models.

- Compared with Wilkie's cascade structure, Vector Autoregressive models with regression part (VARX) could be more robust. The coefficients of each variables will determine whether these variables are independent, have unidirectional relationship, or have feedback relationship. Transforming VARX models into a structured VARX models [17] could end up with a structure similar to the cascade structure of the Wilkie Model. But it is possible that the components in the new structure may be latent variables created from the original indices, not the indices themselves. Similarly we could think of cointegrated VARX models for multi-variate non-stationary series. Adding some long memory components could also be a good consideration for modeling those indices with long memory quality. For simplicity, we just used rollover means and rollover variance here as variables of memory input into some models. Wilkie use the exponentially weighted average as a long memory input in his models.

2.1.2 Formulas of the Wilkie Model

In the Wilkie Model, the four basic variables are modeled as follows:

- The first model is for retail price index $Q(t)$. Its model uses a logarithmic transformation of $Q(t)$, where $\nabla \log Q(t)$ follows AR(1) process. The formula is as follows:

$$\nabla \log Q(t) = QMU + QA(\nabla \log Q(t-1) - QMU) + QSD * QZ(t)$$

$$
\begin{aligned}
where: \quad & \nabla X(t) = X(t) - X(t-1) \\
& QZ(t) \overset{i.i.d.}{\sim} N(0,1), \qquad QMU \text{ is the mean} \\
& QA \text{ is the AR coefficient}, \quad \text{with } |QA| < 1 \\
& QSD \text{ is the standard deviation of residuals}
\end{aligned}
$$

- The second model is for dividend yield $Y(t)$. This model uses logarithmic transformed dividend yield $\log Y(t)$ as the response variable. As we know, logarithmic transformation requires the non–negative value of $Y(t)$. After taking regression on $\nabla \log Q(t)$, the residuals follow MA(1) process. The formula is as follows:

$$
\begin{aligned}
\log Y(t) &= YW * \nabla \log Q(t) + YN(t) \\
YN(t) &= \log YMU + YA(YN(t-1) - \log YMU) + YE(t) \\
where: \quad & YW \text{ is the coefficient of regression of } \log Y(t) \text{on } \nabla \log Q(t) \\
& \log YMU \text{ is the mean of } YN(t) \\
& YA \text{ is the AR coefficient of } YN(t), \quad \text{with } |YA| < 1 \\
& YE(t) \text{ is the residuals}, \quad \text{with } YE(t) = YSD * YZ(t) \\
& YSD \text{ is the standard deviation of residuals} \qquad YZ(t) \overset{i.i.d.}{\sim} N(0,1)
\end{aligned}
$$

- The third model is for dividend index $D(t)$. This model uses the force of dividend $\nabla \log D(t)$ as the response variable, which has regression on the inflation rate $\nabla \log Q(t)$, and the residual of dividend yield model of the previous period $YE(t-1)$. The residuals of the regression follow MA(1) process. The formula is as follows:

$$
\begin{aligned}
\nabla \log D(t) &= DW\left(\frac{DD}{1-(1-DD)B}\right)\nabla \log Q(t) + DX * \nabla \log Q(t) \\
& \quad + DMU + DY * YE(t-1) + DE(t) + DB * DE(t-1) \\
\text{where}: \quad & \frac{DD}{1-(1-DD)B} \quad \text{is exponential weighted average}
\end{aligned}
$$

$$
\begin{aligned}
& BX(t) = X(t-1), \qquad DD \text{ is an coefficient of weight} \\
& DE(t) = DSD * DZ(t) \\
& DSD \text{ is the standard deviation of residuals} \qquad DZ(t) \overset{i.i.d.}{\sim} N(0,1) \\
& DW,\ DX,\ DY, \text{ are coefficients of regression,} \quad \text{with } DW + DX = 1 \\
& DMU \text{ is the mean factor} \\
& DB \text{ is the MA coefficient of residuals, with } |DB| < 1 \\
\text{therefore} \quad & \text{The weight for } \nabla \log Q(t-i) \text{ is } DD(1-DD)^i \qquad \text{for } i \geq 1 \\
\text{if we use} \quad & DM(t) = DD * \nabla \log Q(t) + (1-DD) * DM(t-1) \\
\text{then} \quad & \\
\nabla \log D(t) \quad = \quad & DW * DM(t) + DX * \nabla \log Q(t) \\
& + DMU + DY * YE(t-1) + DE(t) + DB * DE(t-1)
\end{aligned}
$$

- The fourth model is for the Consols yield C(t). The model is based on CN(t), which is C(t) adjusted the long memory effect of inflation rate $(\frac{CD}{1-(1-CD)B})\nabla \log Q(t)$. $\log CN(t)$ follows an AR(1) process, and has regression relationship on the current period error of dividend yield model $YE(t)$.

$$
\begin{aligned}
C(t) \quad &= \quad CW(\frac{CD}{1-(1-CD)B})\nabla \log Q(t) + CN(t), \\
\log CN(t) \quad &= \quad \log CMU + CA * (\log CN(t-1) - \log CMU) \\
& \quad\quad + CY * YE(t) + CSD * CZ(t) \\
\text{where :} \quad & \quad\quad CW,\ CY \text{ are the coefficients of regression} \\
& \quad\quad \log CMU \text{ is the mean factor,} \qquad CA \text{ is the AR coefficient of } \log CN(t) \\
& \quad\quad CSD \text{ is the standard deviation of residuals,} \qquad CZ(t) \overset{i.i.d.}{\sim} N(0,1)
\end{aligned}
$$

2.1.3 Log-likelihood of the Wilkie Model

The derivation of the log–likelihood function for total return index yield R_t from the Wilkie Model is as follows:

$$
\begin{aligned}
D_t &= \text{Dividend at time t} \\
Y_t &= \text{Dividend yield at time t} \\
T_t &= \text{Total return index at time t} \\
\text{Here, let } & f \text{ denotes density function, which is different for different variables} \\
\text{let } & \mathcal{F}_{t-1} \text{ represents the information up to time } t-1, \text{ i.e. the information} \\
& \text{we get based on the values of } D_i,\ Y_i, \text{ and } T_i. \text{ with } i \leq t-1 \\
T_t &= D_t + \frac{D_t}{Y_t} = D_t(1+\frac{1}{Y_t}) \\
f(T_t|\mathcal{F}_{t-1}) &= f((1-\frac{1}{Y_t})D_t|\mathcal{F}_{t-1}) = f(1-\frac{1}{Y_t}|\mathcal{F}_{t-1})f(D_t|\mathcal{F}_{t-1}) \\
& \text{Since } Y_t \text{ and } D_t \text{ are independent according to the Wilkie Model}
\end{aligned}
$$

$$
\begin{aligned}
f(Y_t|\mathcal{F}_{t-1}) &= \frac{1}{Y_t} f(\log(Y_t)) \\
&= \frac{1}{Y_t}\phi(\frac{\log(Y_t) - YW * \nabla \log Q(t) - \log YMU - YA * YN(t-1)}{YSD}) \\
f(D_t|\mathcal{F}_{t-1}) &= \frac{1}{D_t} f(\log(D_t)) \\
&= \frac{1}{D_t}\phi(\frac{\log(\frac{D_t}{D_{t-1}}) - DW * DM(t) - (1-DW) * \nabla \log Q(t)}{DSD} \\
&+ \frac{-DY * YE(t-1) - DMU - DB * DE(t-1)}{DSD}) \\
\text{then} & \\
f(T_t|\mathcal{F}_{t-1}) &= f(Y_t|\mathcal{F}_{t-1})f(D_t|\mathcal{F}_{t-1})|J| \\
&= \frac{1}{D_t Y_t} f(\log(Y_t)|\mathcal{F}_{t-1})f(\log(D_t)|\mathcal{F}_{t-1})|J| \\
\textit{where} \quad & |J| = |\frac{dY}{dk}|_{k=1+\frac{1}{Y}} = |\frac{1}{(k-1)^2}| = Y^2
\end{aligned}
$$

then

$$\begin{aligned} f(T_t|\mathcal{F}_{t-1}) &= \frac{1}{D_t Y_t} f(\log(Y_t)|\mathcal{F}_{t-1}) f(\log(D_t)|\mathcal{F}_{t-1}) * Y_t^2 \\ &= \frac{Y_t}{D_t} f(\log(Y_t)|\mathcal{F}_{t-1}) f(\log(D_t)|\mathcal{F}_{t-1}) \\ &= \frac{1}{P_t} f(\log(Y_t)|\mathcal{F}_{t-1}) f(\log(D_t)|\mathcal{F}_{t-1}) \end{aligned}$$

therefore

$$\sum_t \log f(T_t|\mathcal{F}_{t-1}) = \sum_t \log(Y_t) - \sum_t \log(D_t) + \sum_t \log f(\log(Y_t)|\mathcal{F}_{t-1}) + \sum_t \log f(\log(D_t)|\mathcal{F}_{t-1})$$

From the above we get the log–likelihood value of the Wilkie Model based on the calculation of the log–likelihood function of total return T_t. Since other models use index yield R_t as the variable to calculate log–likelihood values, we need further calculation.

since:

$$\begin{aligned} R_t &= \log(\frac{T_t}{P_{t-1}}) \\ f(R_t) &= P_{t-1} e^{R_t} f(T_t) = T_t f(T_t) \end{aligned}$$

therefore

$$\begin{aligned} \sum_t \log f(R_t|\mathcal{F}_{t-1}) &= \sum_t \log f(T_t|\mathcal{F}_{t-1}) + \sum_t \log(T_t) \\ &= \sum_t \log f(T_t|\mathcal{F}_{t-1}) + \sum_t \log(\frac{D_t + D_t Y_t}{Y_t}) \\ &= \sum_t \log f(T_t|\mathcal{F}_{t-1}) + \sum_t \log(D_t + D_t Y_t) - \sum_t \log(Y_t) \end{aligned}$$

therefore

$$\begin{aligned} \sum_t \log f(R_t|\mathcal{F}_{t-1}) &= \sum_t \log(D_t + D_t Y_t) - \sum_t \log(D_t) \\ &\quad + \sum_t \log f(\log(Y_t)|\mathcal{F}_{t-1}) + \sum_t \log f(\log(D_t)|\mathcal{F}_{t-1}) \end{aligned}$$

where:

$$\begin{aligned}
\sum_t \log f(\log(D_t)|\mathcal{F}_{t-1}) &= \sum_t \log f(\log(D_t) - \log(Dt-1)|\mathcal{F}_{t-1}) \\
&= \sum_t \log f(\nabla \log(D_t)|\mathcal{F}_{t-1})
\end{aligned}$$

therefore

$$\begin{aligned}
\sum_t \log f(R_t|\mathcal{F}_{t-1}) &= \sum_t \log(D_t + D_t Y_t) - \sum_t \log(D_t) \\
&\quad + \sum_t \log f(\log(Y_t)|\mathcal{F}_{t-1}) + \sum_t \log f(\nabla \log(D_t)|\mathcal{F}_{t-1})
\end{aligned}$$

Note:

- When we calculate the log–likelihood of the Wilkie Model for R_t, we use only dividend yield model and dividend model, but no inflation model. That is to say, the number of parameters in the estimation is 10, not 13.
- The difference between $\sum_t \log f(R_t|\mathcal{F}_{t-1})$ and $\sum_t \log f(\log(Y_t)|\mathcal{F}_{t-1})+$ $\sum_t \log f(\nabla \log(D_t)|\mathcal{F}_{t-1})$ is only some constant terms. Thus when we maximize the log–likelihood of dividend yield and dividend to estimate the parameters, we actually also maximize that of TSX Total Return Price yield. In another words, what we get as the maximized log–likelihood estimator (MLE) of parameters based on Wilkie's dividend yield model and dividend model are also the MLE of parameters based on the information of R_t.

Now, we get the log–likelihood function of R_t, and we can compare the maximized log–likelihood value of the Wilkie Model with those of other models, even though the estimated parameters and log-likelihood value for the Wilkie Model is not obtained from the TSX Total Return Index yield R_t directly.

2.1.4 Simulation and Prediction of the Wilkie Model

Before estimating parameters of the Wilkie Model and applying those models to simulate and to predict the Total Return Index yield R_t, we need to do some data manipulation and calculation. The data we used include both Canadian data and US data. Because the Wilkie Model is treated as a long term model, monthly data are considered to have too much noise for this long term analysis. Therefore we need to use quarterly data and yearly data for the models. Since the data set for the Canadian Market is as short as 50 years, we also use the US data set, which is over 80 years, to test the model. The data structure is as follows:

- The data used in the models for the Canadian market are as follows:
  - Inflation rate $\nabla lnQ(t) = log(\frac{CPI_t}{CPI_{t-1}})$, where CPI is the consumer price index;
  - Dividend yield and dividend. Because there are not many historical data of dividend yield available, we followed the method of Wilkie (1995) to create them. On one side, these manipulated data are not the same dividend yields as published in recent years by Standard & Poors, and there could be negative values for some countries, which is a contradiction to the reality. Fortunately no negative values were found for the dividends we get from the Canadian and US data. On the other side, those calculated dividend yields provide a meaningful representation of dividend yield for the models, and the path of the calculated yield is very close to that of real dividend yield provided by Standard & Poors, which is available for most countries for recent years. The calculation is as follows under continuous assumption of dividend yield:

$$\begin{aligned} Y'_{t-1} &= \log(\frac{T'_t}{T'_{t-1}}) - \log(\frac{P_t}{P_{t-1}}) \\ \text{where} \quad & T'_t: \text{ TSX Total Return Index}, \quad P_t: \text{TSX Price Index} \end{aligned}$$

Dividend:

$$D_t = P_{t-1} * (e^{Y'_{t-1}} - 1)$$

where Dividend is assumed to be discrete at the end of each period

Dividend Yield:

$$Y_t = \log(\frac{D_t}{P_t} + 1)$$

The calculation under discrete assumption of dividend yield is as follows:

$$\begin{aligned} \text{Dividend:} \quad D_t &= (T'_t - T'_{t-1})\frac{P_{t-1}}{T'_{t-1}} - (P_t - P_{t-1}) \\ \text{Dividend Yield:} \quad Y_t &= \frac{D_t}{P_t} = \frac{(T'_t - T'_{t-1})(P_{t-1})}{T'_{t-1}P_t} - \frac{P_t - P_{t-1}}{P_t} \end{aligned}$$

The Wilkie Model is thought to use discrete data due to its long term features, and it uses logarithm transformation for inflation rate, dividend yield and consols rate. This logarithmic transformation requires the non-negative property of the variables. For inflation, the logarithmic transformed value also assumes a force of increasing, which comes from continuously increase. Continuous assumption does not make much difference for the manipulated data when the data are small, and stationary without many extreme values. This situation is suitable for monthly yield data. But for quarterly and yearly data, we have to be careful when the yield rate is not that small. Continuous assumption for logarithmic transformation could enlarge the scale of negative values. Therefore we prefer to use discrete assumption for the calculations of yearly data, and continuous assumption for the calculation of monthly data in this thesis. For specific cases, the choice of continuous assumption or discrete assumption for data transformation will also depend on what the resource data we got.

- The data used in the models for the US market are as follows:
  - Inflation rate $\nabla \log Q(t)$
  - Dividend yield and dividend are calculated based on the monthly data for US stock market obtained from the Center for Research in Security Prices (CRSP)

* Value-Weighted Return – all distributions (rr_1)
* Value-Weighted Return – excluding dividends (rr_2)
* S&P Composite Index P_t

The calculation of dividend and dividend yield is as follows under continuous assumption of dividend yield for monthly data:

$$
\begin{aligned}
& & & \text{under constant yield assumption} \\
& Y'_{t-1} &=& e^{rr_1} - e^{rr_2} \\
& & & \text{under discrete dividend assumption} \\
\text{Dividend:} \quad & D_t &=& P_{t-1} * (e^{Y'_{t-1}} - 1) \\
& & & \text{under constant yield assumption} \\
\text{Dividend Yield:} \quad & Y_t &=& ln(\frac{D_t}{P_t} + 1)
\end{aligned}
$$

- For the quarterly US data of indices, we used the index of March, June, September and December of every year. For the yearly US data of indices, we used the index of June of every year.

- The prediction and simulation include:
  - Prediction and simulation of future inflation rate $\nabla \log Q(t)$ with AR(1) process.
  - Prediction and simulation of dividend yield $Y(t)$, which is in turn based on the prediction and simulation of $\log Y(t)$. This step also simulate and predict an intermediate variable $YN(t)$, whose residuals are used to predict dividend.
  - Prediction and simulation of dividend $D(t)$ from the prediction and simulation of $\nabla \log D(t)$.

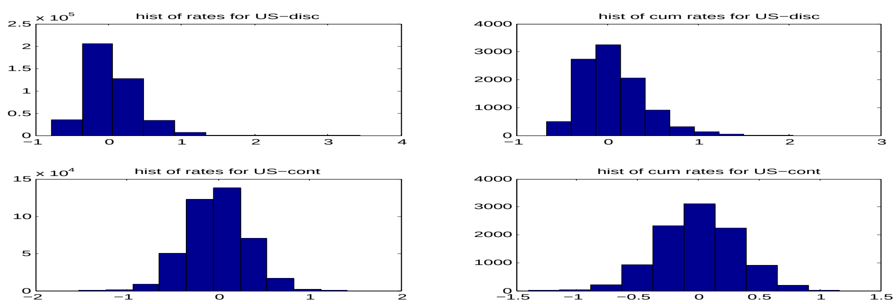
- The next step is the simulation and prediction of total return T_t, which is used to calculate the total return index yield R_t. The R_t is used to analyze the fitness of the Wilkie Model and the tails' thickness of the simulated distribution from the Wilkie Model. The calculation of T_t is as follows:

$$\begin{aligned} T_t &= D_t(1+\frac{1}{Y_t}), \quad \text{under discrete assumption for dividend yield} \\ T_t &= D_t(1+\frac{1}{e^{Y_t}-1}), \quad \text{under continuous assumption for dividend yield} \end{aligned}$$

The choice of continuous and discrete assumption for R_t is based on the approximation of the simulation to the real world values, and the errors between $\log(x)$ and x when x is small. The following is my choice of different assumptions for use:

- R_t under continuous assumption is used to compare short term models. Continuous assumption means $R_t = \log \frac{T_t}{P_{t-1}}$.
- R_t under discrete assumption is used to illustrate and compare long term results. Discrete assumption means $R_t = \frac{T_t}{P_{t-1}} - 1$
- Wilkie (1995) advised a method to create discrete data for share dividend yield. As we know, log–transformation could enlarge the scale when R_t is negative, and $|R_t|$ is large enough. So we need to be careful and to be consistent when selecting discrete or continuous assumption for R_t, and interpret them correctly. For US data we can see that, while mean rate under discrete assumption is positive, it could be negative under continuous assumption of yield rate. This is because log-transformation enlarges the effect when there are many times of sharp decrease in the index price.

Therefore when there are many extreme cases in the yield data, which makes $\log(x)$ quite different from x, we need to take the inverse of the log-transformation of x to show the result. The following are the plots for the distribution of yield based on the same price series. The top two plots are yield based on discrete assumption for yield rate, but the bottom two are growth force of index, which often represents yield, based upon continuous assumption. We can see from the plots that different assumption changed the distribution of the yield. Therefore what we discussed before for the choice of continuous assumption and discrete assumption for dividend yield also applied in calculation of R_t here. In summary, for the calculation of monthly index yield, it does not really matter which assumption you choose, because the yield rate is small; however for the calculation of yearly yield, it is better to choose discrete assumption for your calculation if you are not sure about the assumption of resource data. If you know the assumption of resource data, you can just keep the same assumption for the future calculation.



Because there are many other choices of data manipulation, due to the lack of published historical data, there would be more arguments with regard to the results. Therefore we learned that for a complicated model, the availability and reliability of data source are important for testing and implementation of the models, which was pointed out by Huber (1995).

The estimation of parameters could also be further argued for two different methods based on the dependence of multiple equations.

- In one way, parameters were estimated for each model separately. This method is used when there are no common parameters for two models, or no variables which are estimated from one model and will be transferred into another model as input, or no correlation among residuals of different models. The example model of this type is the inflation rate model. The variables in this model are published inflation rate values, which are not estimated values. Inflation rate is on the top of the structure, so no residual effect will be transferred into its model for the top-down structured multi-variate models.

- In the other way, if there are common parameters, or if there are estimated transferred variables, we better estimate parameters of the two models together, even if these models are at different levels. Common parameter is not the case in the Wilkie Model, but there exist estimated transferred variables. The examples are dividend yield model and dividend model. There is a transfer variable $YN(t)$ from dividend yield model to dividend model. Even though there is only slightly difference between these two methods for parameter estimation in dividend yield model and dividend model, the parameter estimation based on the second method is more reasonable. The difference is very slight because in this structure dividend values have only minor influence on the parameters' estimation of dividend yield model. Correlation of residuals will change the models for each variable in the Wilkie Model. Based on the assumption of the Wilkie Model, residuals are independent. Though they are not independent for many models in the Wilkie Model, which will be seen in the later sections, we did not consider it right now when estimating parameters.

From the results, we can see that our parameters' estimates are different from those of Wilkie's estimation for the Wilkie Model. This could attribute to different data sources or different time periods. Wilkie (1995) does not specify the data source for Canadian Market, and he used some combined data series from various resources for different periods. Another reason for the difference of the parameter estimates is that there are no historical data available for dividend yield. We followed the suggestion from Wilkie (1995) to manipulate the so-called dividend yield in Wilkie's models. But the calculated dividend is different from the real dividend yield published for recent years.

Comments based on our analysis are as follows:

- Based on the parameters estimated for the Canadian yearly data of 1956–2005 and for the US year data of 1926–2006, we can see, from the tables at the end of this chapter, that for the dividend yield model, the AR1 parameter of $YN(t)$ is close to 1. That is to say, the $YN(t)$ is not very stationary. It will produce many outliers of the residuals if we try to predict $YN(t)$ using stationary ARMA models. This would end up with errors in the prediction of dividend yield $Y(t)$ and other following indices.

- For the dividend, we can see that the sign of parameter DY is positive as opposed to Wilkie. Wilkie claimed that if dividend yield is high in the preceding term, it will end up with lower dividend yield in the current term, which will result in a lower dividend growth rate in the current term. However the negative sign is not necessary for DY, if we think that the dividend growth is determined by both the change of price and that of dividend yield. Since the dividend growth rate = $\frac{P_t * Y(t) - P_{t-1} * Y(t-1)}{P_{t-1} * Y(t-1)} = \frac{P_t}{P_{t-1}} * \frac{Y(t)}{Y(t-1)} - 1$, if the dividend yield goes down, the dividend growth force could also goes up if the price goes up more quickly. In another words, if previous $Y(t)$ is high, then if it leads to the decrease of current $Y(t)$, it does not necessarily result in decreasing of dividend growth force if the price rises faster. Another reason to suspect Wilkie's claim is that the dividend yield is not very stationary. This is because the price movement is not at the same speed as that of dividend. Therefore the deviation of dividend yield from the expected value, which is estimated from historical data, may not be an indicator for the deviation advised by Wilkie. In another words, when the share price grows faster, the current dividend yield will decrease, but this decrease is not because of the high level of dividend yield in previous term, and this decreasing may happen when the dividend increases faster. As a result, we see the positive relationship between dividend yield of preceding term and the dividend force of the current term.

The estimation of the MLE of parameters assumes the model residuals to be independently normally distributed. Detailed calculation is as follows based on the advice given for simulation in Wilkie (1986):

- Inflation rate model, with AR(1) process fitting inflation rate:

$$\begin{aligned}
\nabla \log Q(t) &= \frac{CPI(t)}{CPI(t-1)} \\
\log \text{L} &= \prod_{i=1}^{n} \phi(\frac{\nabla \log Q(t) - \mu_{\nabla \log Q(t)}}{\sigma_{\nabla \log Q(t)}}) \\
\mathit{where} & \\
\nabla \log Q(t) &\sim N(QMU + QA * \nabla lnQ(t-1), QSD), \quad t = 2 \ldots n \\
\nabla \log Q(t) &\sim N(QMU, \frac{QSD}{\sqrt{1-QA^2}}), \quad t = 1
\end{aligned}$$

- Dividend yield model:

$$\begin{aligned}
\log \text{L} &= \prod_{i=1}^{n} \phi(\frac{\log(Y(t)) - \mu_{\log(Y(t))}}{\sigma_{\log(Y(t))}}) \\
&= \prod_{i=1}^{n} \phi(\frac{\log(Y(t)) - YW * \nabla \log Q(t) - \log YMU - YA * YN(t-1)}{\sigma_{\log(Y(t))}}) \\
&= \prod_{i=1}^{n} \phi(\frac{YN(t) - YA * YN(t-1)}{\sigma_{\log(Y(t))}})
\end{aligned}$$

where

$$
\begin{aligned}
\nabla YN(t) &\sim N(YA * YN(t-1)), \sigma = \sigma_{\log(Y(t))} = YSD), \quad t = 2 \cdots n \\
\nabla YN(t) &\sim N(YA * YN(0), \sigma = \sigma_{\log(Y(t))} = \frac{YSD}{\sqrt{1 - YA^2}}), \quad t = 1 \\
Y(0) &: \text{ using the real values}
\end{aligned}
$$

- Dividend model:

$$
\begin{aligned}
\log \text{L} &= \prod_{i=1}^{n} \phi(\frac{\log(D(t)) - \mu_{\log(D(t))}}{\sigma_{\log(D(t))}}) \\
&= \frac{1}{D_t} \phi(\frac{\log(\frac{D_t}{D_{t-1}}) - DW * DM(t) - (1 - DW) * \nabla \log Q(t)}{\sigma_{\log(D(t))}} \\
&\quad + \frac{-DY * YE(t-1) - DMU - DB * DE(t-1)}{\sigma_{\log(D(t))}})
\end{aligned}
$$

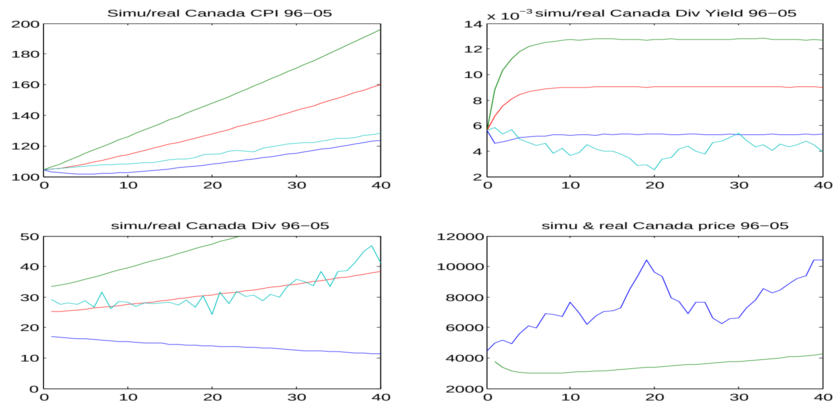
where

$$
\begin{aligned}
DM(t) &= DD * \nabla \log Q(t) + (1 - DD) * DM(t-1), \quad t = 1 \cdots n \\
DM(0) &= QMU, \quad \text{which is the neutral value of DM(t)} \\
YE(t) &= YN(t) - YA * YN(t-1) \\
&\quad \text{YE(t) is not the standardized value. YE(0) is a real value} \\
DE(t) &= \nabla \log(D(t)) - \mu_{\nabla \log(D(t))} \\
&\quad \text{DE(t) is not the standardized value. DE(0)} = 0
\end{aligned}
$$

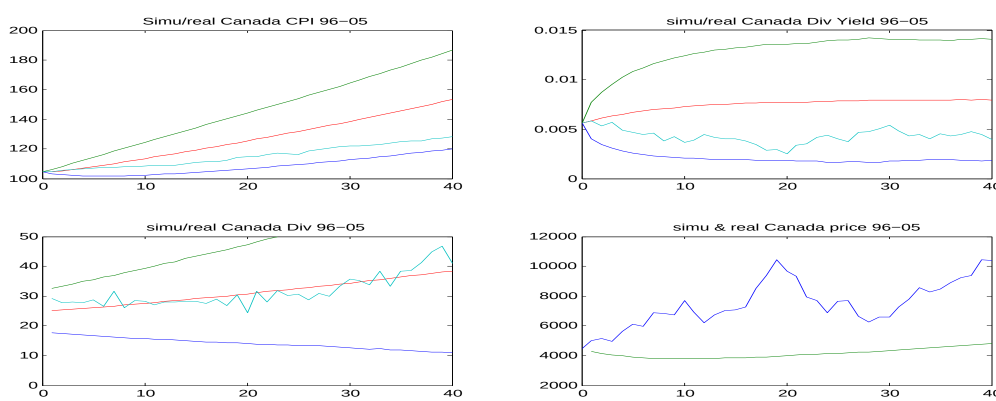
where

$$
\begin{aligned}
\sigma_{\log(D(t))} &= DSD, \quad t = 2 \cdots n \\
\sigma_{\log(D(t))} &= DSD\sqrt{(1 + DB^2)}, \quad t = 1
\end{aligned}
$$

The estimated parameters of the Wilkie Model for Canadian and US markets are listed in the tables at the end of this chapter. The following are the plots of confidence intervals of prediction and real values of quarterly Canadian inflation rate, dividend yield, dividend, and share price. The real value plots are those zigzag curves. The parameters' estimation is based on 1956–1995 quarterly data, and the prediction is for 1996–2005.



The following plots are the predictions from the updated parameters based on Canadian quarterly data during 1956–2005. The predictions are the mean values of the simulation for each year over 10 years with the initial value of October 1995. The parameters of dividend yield model and dividend model are estimated simultaneously, not one by one in order, because there is an estimated transfer variable $YN(t)$.



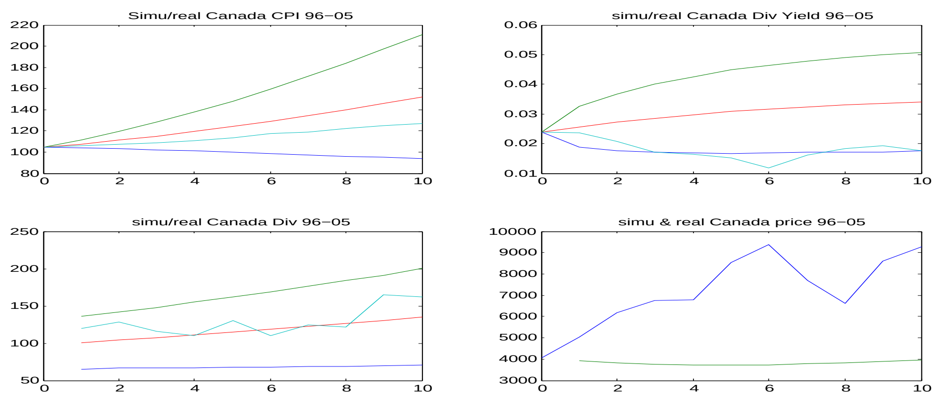
We can see some changes in the above plots as follows:

- The mean value of dividend yield $Y(t)$ decreases from 0.009 to less than 0.008. This makes the predicted share price index rise from less than 4500 up to around 5000.

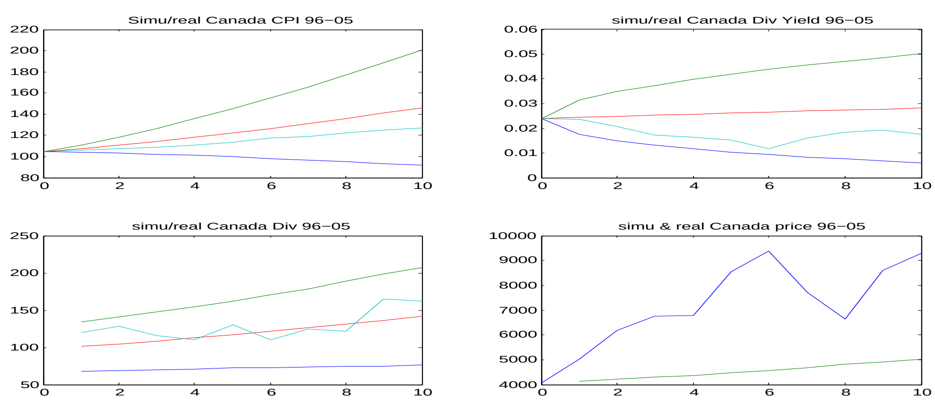
- The variance of dividend yield increase a little bit, thus it covers the values of real dividend yield.

- The predicted model for dividend is not bad, but the prediction for dividend yield and share price is really poor.

Similar results can be obtained for yearly data. The following plots are based on the yearly Canadian data. The prediction is for 10 years with the initial value of 1995.

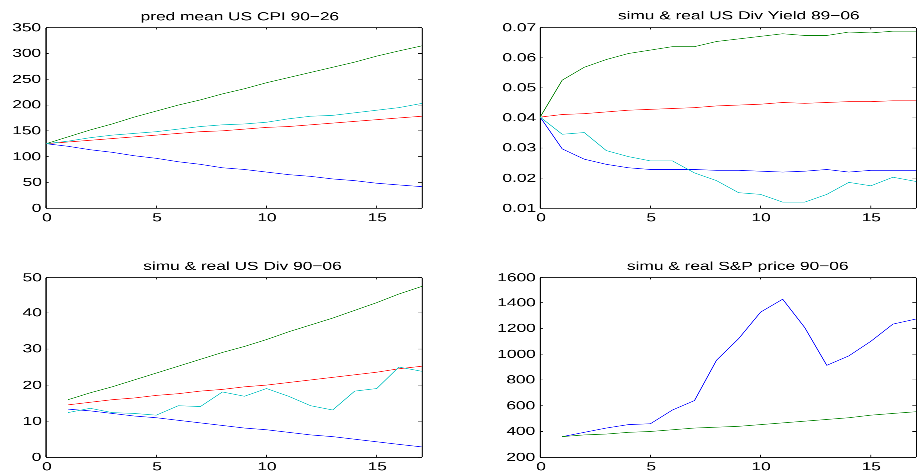


The following plots are the Wilkie Model with update parameters estimated from the yearly Canadian data during 1956–2005.

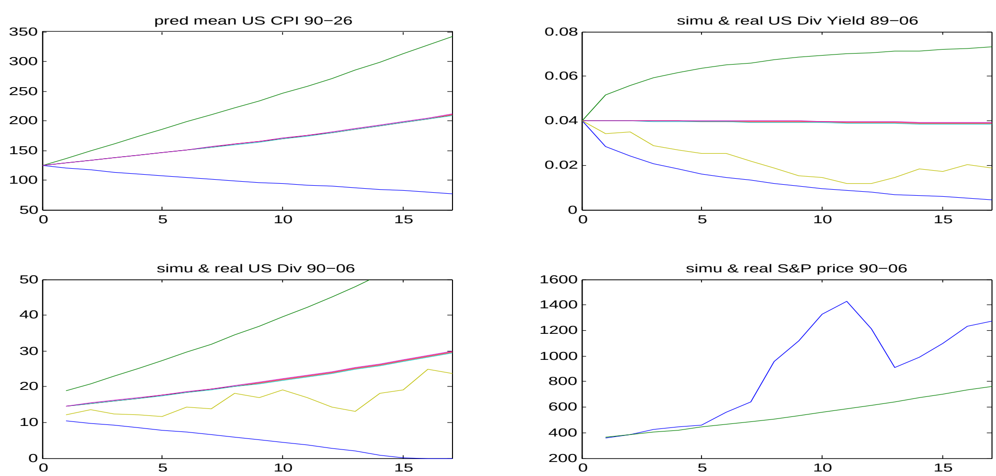


From the plots, we did not see any significant difference between using yearly data and using quarterly data for the Wilkie Model.

The following plots are the predictions and real values for 17 years in US market. The parameters' estimation is based on US data during 1926–1989, and the prediction is for the period of 1990–2006. We can see the similar results as those of Canadian data.



The following plots are based on parameters estimated from yearly data of 1926–2006. From the plots we can get the similar results as those of Canadian data.



From these plots, we can see that the share price prediction is poor as given by the Wilkie Model. The variance analysis of prediction of price is not easy to obtained due to the nonlinear relationship between price, dividend and dividend yield.

In order to predict total return index, we could use one of the following three methods:

1. Construct an univariate model for index price yield, if the yield is stationary and ergodic.

2. Use VAR method to model stationary multivariate relations, and cointegrated VAR to model non-stationary multivariate relations. In the cointegrated modeling, we construct a model for two or more closely related factors, such as share price index and total return index, and use cointegrated factors as a latent variable to transfer the prediction of one variable, e.g. share price index, into the prediction of another one, e.g. total return index. Since this requires the cointegration relationship to be very significant, which is not a good application for dividend and share price.

3. Assume the price is not stationary or not predictable, it could be determined by two or more other variables, which are stationary and predictable. This actually is not the case in the Wilkie Model , because price is even more stationary than dividend yield in the Wilkie Model.

Therefore we need to reconsider the structure of the Wilkie Model.

2.2 Other Selected Models

Regarding the predictability of the Wilkie Model for the stock market yield R_t, we conducted a comparison test with Independent Log-Normal model (ILN), fitted univariate ARIMA–GARCH model, and Regimes–Switching Log-Normal model with two regimes (RSLN2).

The data used in these tests are TSX Total Return Index T'_t. The period covered is 1956–2005. The data used are monthly data and quarterly data for different models. Quarterly data are the values of January, April, July and October of every year. The choice of quarterly data is preferred because these models are used to predict medium terms like 10-year period, not short terms. Wilkie (1995) also preferred the choice of quarterly or yearly data, because he thought monthly data brings too much noise for longer term analysis. The fitted models for TSX total return index with various methods are listed in the following subsections. Those models include ILN, ARIMA–GARCH, and RSLN2.

2.2.1 ILN Model

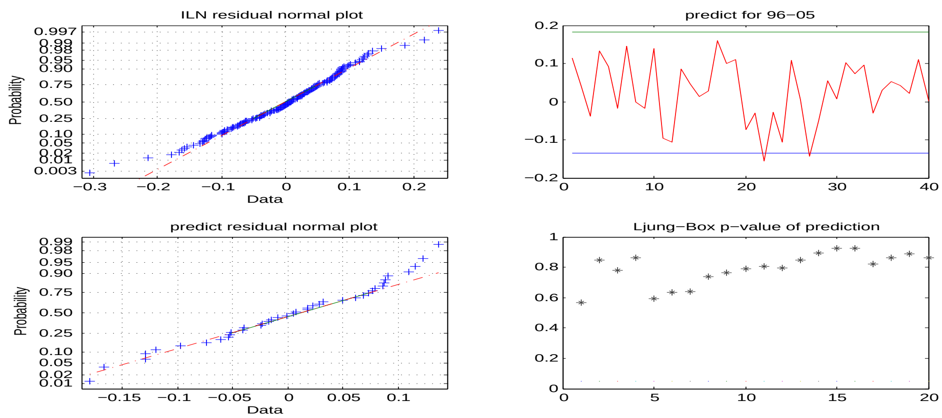
Independent log-normal model was developed for financial application when the continuous Brownian motion was used in assets pricing. The continuous Brownian motion $B(t)$ is a continuous–time stochastic process for a time–dependent variable such as TSX Total Return yield R_t. Then the TSX Total Return index is a geometric Brownian motion $T'_t = T'_{t-1} * e^{B(t)}$. Since $B(t)$ could have a drift μ and volatility σ. Then cumulatively $T'_t = T'_0 * e^{\mu t + \sigma W(t)}$, where W(t) is a standard Brownian motion, and $E(T'_t) = T'_0 * e^{\mu t + \frac{1}{2}\sigma^2 t}$.

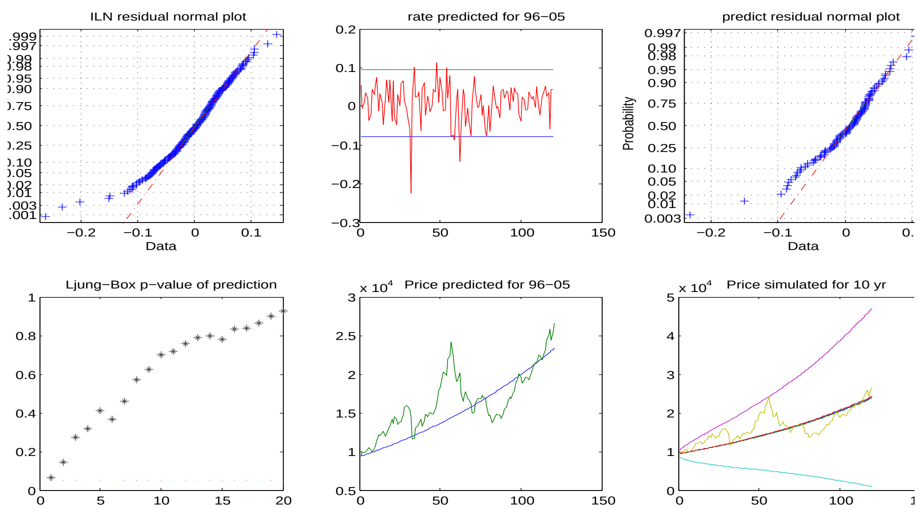
When we estimate the parameters for R_t, we use conditional density for R_t given all the information up to time $t-1$, which is denoted by $\mathcal{F}_{t-1}$. The density function for total return index yield $R_t|\mathcal{F}_{t-1}$ is as follows:

$$f(R_t) = \phi(\frac{R_t - \mu}{\sigma}), \text{ with the MLE of } \mu \text{ and } \sigma$$

The results from the fitted model by ILN are plotted as follows. The top four are plots of the fitted ILN model for quarterly TSX yield, and the bottom four are plots of the fitted ILN model for monthly TSX yield.

From the following plots, we can see that ILN method is good for modeling the trend of total return index. The prediction of index price is unbiased for real value. The real values are within the 95% confidence intervals. Ljung–Box tests for residuals are not significant. The estimated parameters are in the tables at the end of this chapter. But we can also see from the Q–Q plots that ILN does not catch the thick tail of TSX yield. However when we use ILN to model index yield, we have more concern with quarterly data. ILN model is based on the continuous assumption. Its result is the force of yield, not real yield. However in other models, it is better to use discrete assumption to model quarterly data, as in the Wilkie Model. The worst result from this comparison is that ILN enlarge the absolute values of the negative yield values. In another words, ILN has the lower percentiles than the corresponding values under the discrete assumption, and lower percentiles are just what we care about. Therefore when we compare the results of percentiles, we have to keep this error in mind. One solution for this problem in comparison is to use yield under the continuous assumption for all models, even though it does not make sense for some models. But at least the results are comparable.





2.2.2 ARIMA–GARCH Model

The ARIMA–GARCH model is used to model the dynamic structure of long term time series, which are stationary, ergodic, and dependent on lagged values. ARIMA or Autoregressive Integrated Moving Average model is used to model the conditional mean of the series. This part includes three components: difference of time series, Autoregressive models, and Moving Average models. The method of model construction is through observing the Autocorrelation Function (ACF) and the Partial Autocorrelation Function

(PACF) performance of the time series and their residuals after fitting a model. P–values of Ljung-Box Chi-Squared statistics of residuals are also used to test the dependence between residuals. GARCH or Generalized Autoregressive Conditional Heteroscedasticity model is used to model the variance of the residuals of ARMA models. The ARCH model was introduced in Engle (1982), and GARCH model was introduced in Bollerslev (1986). Ljung–Box statistics for squared residuals can be used to test the ARCH effect of residuals. Based on GARCH models, mathematicians developed many similar models. For example, Tsay (1987) suggested Conditional Heteroscedastic Autoregressive Moving Average models (CHARMA), and Nelson (1991) came up with Exponential GARCH models (EGARCH). Because our model is for the long term time series with wider time gap for two consecutive values, the GARCH effect for residual variance is not very significant. However we still test it as a modeling process. The models and the relative density function used to estimate parameters are as listed below.

- AR(1) model: $R_t = \mu + a(R_{t-1} - \mu) + \sigma\varepsilon_t$.
  Its conditional density function is $f(R_t|\mathcal{F}_{t-1}) = \phi(\frac{R_t - \mu - a(R_{t-1} - \mu)}{\sigma})$, with the MLE of μ, a and σ.

- MA(1) model: $R_t = \mu + \sigma\varepsilon_t + \theta\sigma\varepsilon_{t-1}$
  Its conditional density function is $f(R_t|\mathcal{F}_{t-1}) = \phi(\frac{R_t - \mu - \theta\sigma\varepsilon_{t-1}}{\sigma})$

- GARCH(1,1) model: $R_t = \mu + \sigma\varepsilon_t$, $\sigma_t^2 = a_0 + a_1(R_{t-1} - \mu)^2 + \beta\sigma_{t-1}^2$.
  Its conditional density function is $f(R_t|\mathcal{F}_{t-1}) = \phi(\frac{R_t - \mu}{\sqrt{a_0 + a_1(R_{t-1} - \mu)^2 + \beta\sigma_{t-1}^2}})$.

By repeated trials, we obtained different time series models fitting the TSX index yield. The final choice of the univariate ARIMA–GARCH model for TSX total return index is: ARMA(1,1) and ARMA(1,1)+GARCH(1,1) for monthly data, and ARIMA(0,0,0) x

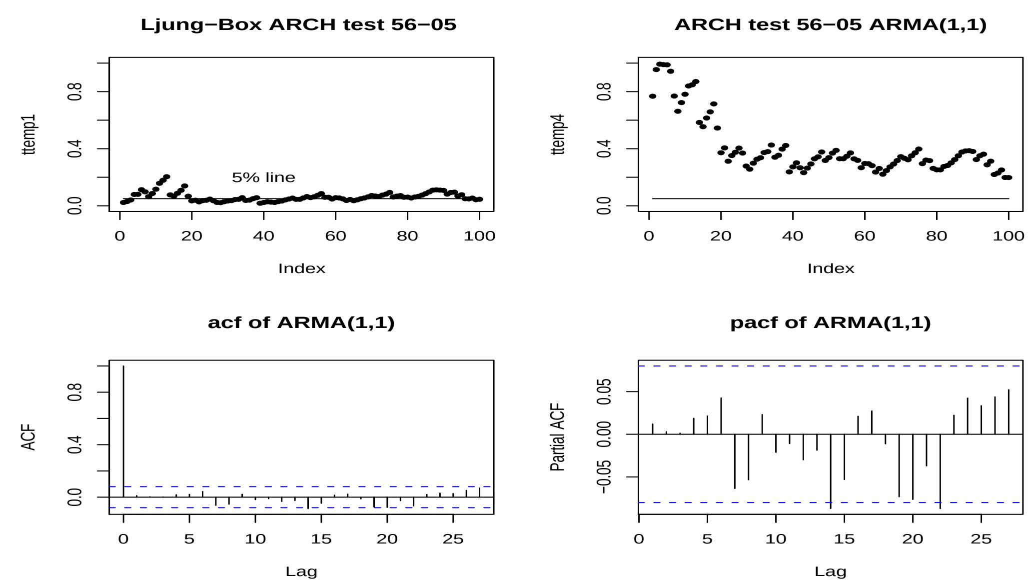
(0,0,1)[5] for quarterly data. The following is what we did for the modeling according to Tsay (2005):

- Find the stationary mean model;
- Test the ARCH effect of residuals. If there is ARCH effect, fit the ARIMA-GARCH model.

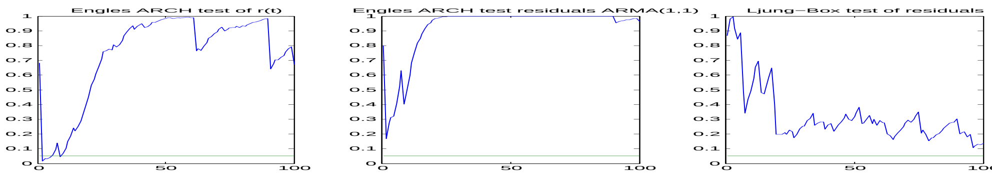
The modeling process is illustrated as follows. We can see that in the ARMA(1,1) model, parameters are significant, and the maximized log–likelihood value increases the value of 6 from the random walk model. When Box–Ljung P–value shows no further modifications are necessary to the modeling of mean part, we choose ARMA(1,1) as the fitted univariate models for the mean.

```
    # Modeling for the original TSX return rate
> Model1 = arima(rate, c(0,0,0))
    # Random Walk model for mean
    # Ljung-Box tests are significant, so there is necessary for
    # further modeling. This model's log likelihood = 1013.26
> Model4
    Call:
    arima(x = rate, order = c(1, 0, 1))
    Coefficients:
              ar1     ma1  intercept
          -0.7356  0.8309     1.0088
    s.e.   0.1874  0.1574     0.0019
    sigma^2 estimated as 0.001948:  log likelihood = 1019.12,  aic = -2030.24
    Box-Ljung test for fitted ARMA(1,1) has p-value = 0.7678
```

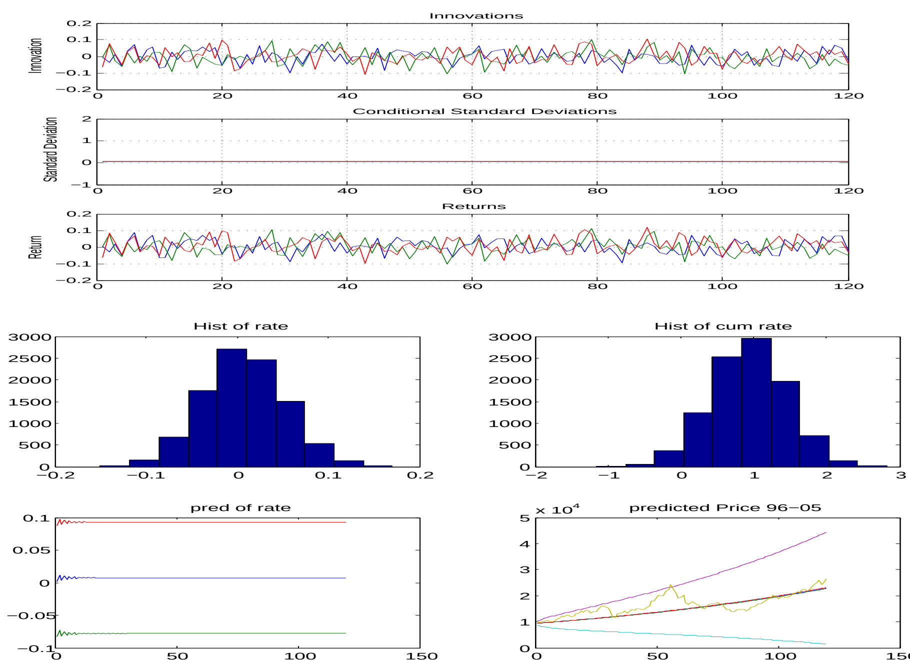
After finding the stationary mean model for the monthly data, we took the Ljung–Box test for squared residuals e_t^2 to test the ARCH effect. This method was introduced in Tsay (2005). From the following plots of test, we can see that ARCH effect is significant before modeling of mean based on the first plot. But this ARCH effect is a false one, because the volatility in variance is mainly due to the change of the mean level of the TSX yield. After the mean is stationary, we took the ARCH effect test of the residuals again based on Ljung–Box statistics for squared residual series. In this case the Ljung–Box tests for the residual series are not significant. From the third plot we can see that the Autocorrelation Function (ACF) of the residuals after fitting ARMA(1,1) model does not show any significant patterns. That means the values of residuals are independent. The Partial Autocorrelation Function (PACF) does not show any significant process patterns either. That means the variance of residuals is independent too. Therefore we use ARMA(1,1) model for the TSX yield.



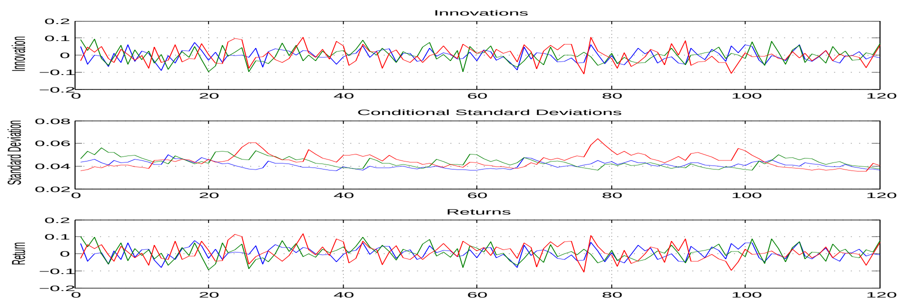
Similar results were obtained from Engle's test in Matlab, which are the plots given below. The left one is the ARCH effect test plot for the raw data; the middle on is the ARCH effect test plot after fitting ARMA(1,1) model, and the right one is the Ljung–Box test for residuals.



The following are the plots of simulation and prediction based on ARMA(1,1) model for monthly data. The simulated values were based on 10,000 times simulation. From the top three plots, we can see the standardized residuals (or innovations) and values of yield rate. From the middle four plots, we can see the histogram plots for the final value, yearly or cumulative yield, of the 10 years and the fitness of price prediction based on this model. We can see that the prediction is as good as ILN model. But we will see in the tables that the distribution of the simulated values has thicker tails for yield rate.



Even though GARCH effect is not significant under ARMA(1,1) model, the ARMA(1,1)+GARCH(1,1) model does improve some features of simulation. The following is the plot of simulation based on ARMA(1,1)+GARCH(1,1) model for monthly data. It creates thicker tail with minimum value of -0.3558, while minimum value of the simulation by ARMA(1,1) model is -0.2153. Probability of market crash is 0.0008, while that of ARMA(1,1) is 0. This part will be further explained in next section.



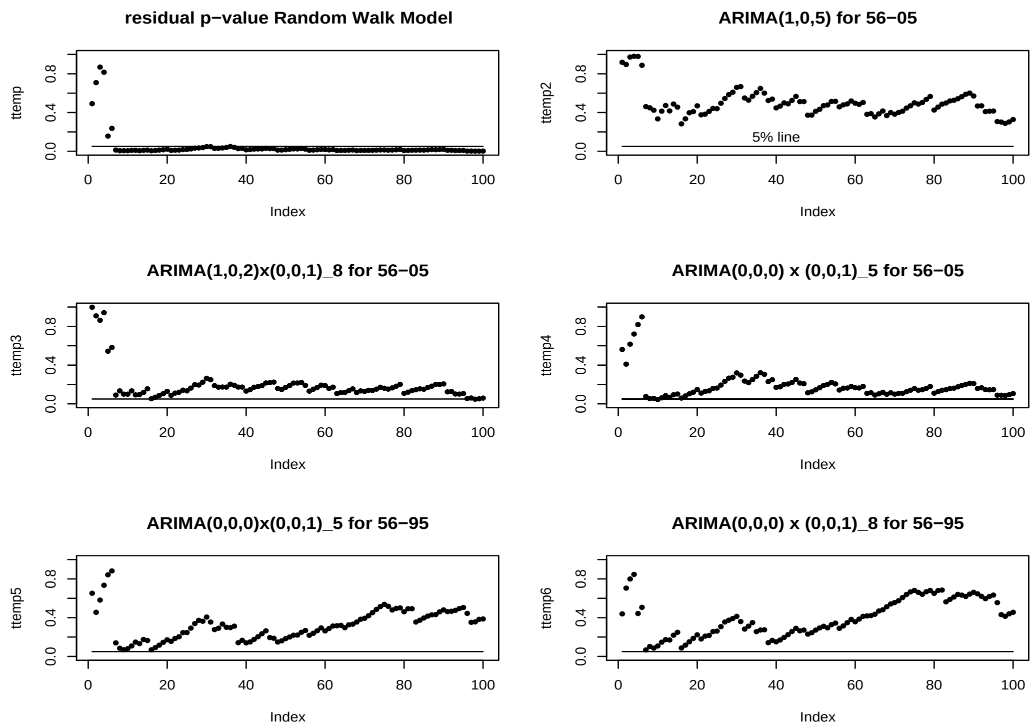
With regard to quarterly TSX total return index, the univariate ARIMA–GARCH model was obtained as follows:

- The ARIMA modeling for mean is as follows:

```
> Model0  # Random Walk Model
Coefficients:
    intercept
      1.0267
s.e.  0.0058
log likelihood = 216.23
Box-Pierce test for lag 1:  p-value = 0.4948

> Model1  # ARIMA(1,0,5)
Coefficients:
          ar1     ma1      ma2      ma3     ma4      ma5  intercept
      -0.3226  0.3846  -0.1096  -0.0713  0.0063  -0.2815     1.0265
s.e.   0.2565  0.2627   0.0863   0.1051  0.0866   0.0913     0.0039
log likelihood = 223.48

> Model2  # ARIMA(1,0,2) x (0,0,1) seasonal period is 8.
Coefficients:
         ar1      ma1      ma2    sma1  intercept
      0.8896  -0.8771  -0.1229  0.1338     1.0277
s.e.  0.0381   0.0782   0.0764  0.0625     0.0010
log likelihood = 222.31

> Model7QQ  # ARIMA(0,0,0) x (0,0,1) seasonal period is 5.
Coefficients:
          sma1  intercept
```

```
        -0.2371     1.0265
 s.e.    0.0792     0.0044
sigma^2 estimated as 0.006378: log likelihood = 220.44
```

The plots Ljung–Box tests for residuals of different fitted models are as follows:

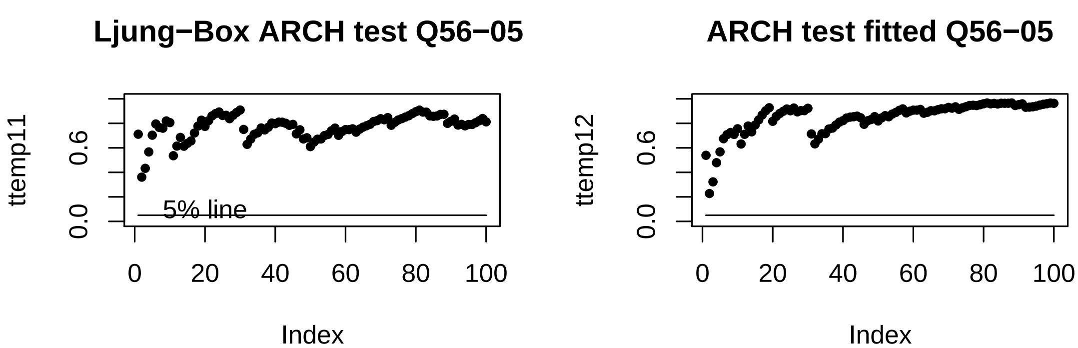


From the results we can see that if we use ARIMA model to fit the mean of quarterly data during 1956–2005, there are at least three models fit the data: ARIMA(1,0,5), ARIMA(1,0,2)x$(0,0,1)^8$, and ARIMA(0,0,0) x $(0,0,1)^5$. We can see that the best model here available is the last one ARIMA(0,0,0) x (0,0,1) with seasonal periods of 5. Although this result does not make much sense, it is statistically significant based on Ljung–Box statistic. This result may come from many reasons such as the choice of the value of certain month as quarterly data. But we can see that even if the model constructed from the data which may not be selected correctly according to financial practice, the statistically optimized model could still simulate and predict the future movement of the index pretty well.

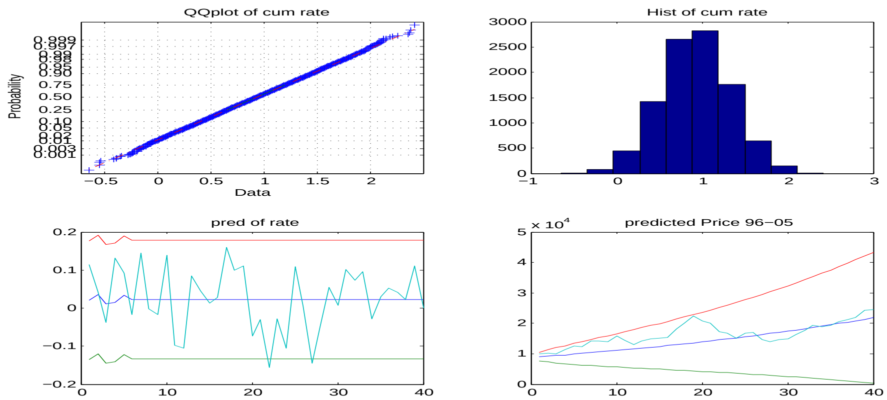
- This model was also confirmed in the prediction model based on data during 1956–1995. We can see from the following models that the model ARIMA(0,0,0) x $(0,0,1)^5$ suit the selected quarterly data best. The model data is as follows:

```
> Model0Q2 # Random Walk Model
 log likelihood = 172.84,

> Model1QQ2 # ARIMA(0,0,0) x (0,0,1) with seasonal period of 5.
Coefficients:
        sma1  intercept
      -0.1975     1.0261
s.e.   0.0870     0.0051
sigma^2 estimated as 0.006444: log likelihood = 175.34

> Model8QQ2 # ARIMA(0,0,0) x (0,0,1) with seasonal period of 8.
Coefficients:
        sma1  intercept
      0.1253     1.0258
s.e.  0.0681     0.0072
sigma^2 estimated as 0.006517: log likelihood = 174.47
```

- Test of heteroscedasticity by Ljung–Box test. The tests have high p–values for original quarterly data, and for the residuals after fitting the models ARIMA(1,0,0) x $(0,0,1)^5$, so we does not reject the constant variance assumption. If we apply GARCH(1,1) to fit the data, the parameters are not significant.



The forecasting plots for the period of 1996–2005 based on the data from 1956 to 1995 are as follows:

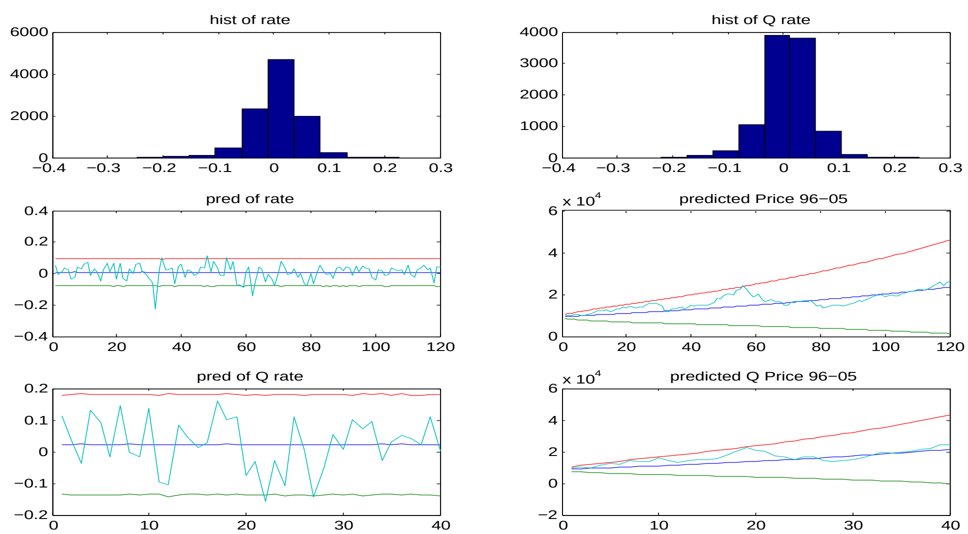


From the plots we can see that the prediction is as good as other models we have done. And the parameters are listed in the following tables at the end of this chapter. From the tables we can see that the tails of ARIMA model is a little thicker than ILN.

2.2.3 RSLN2 Model

Regime–Switching (RS) model assumes there are different regimes of distribution for the stock price yield movement. The advantage of the model is to model a regime using partial information in the historical data, and model another regime using another part of information. The switching between different regimes is assumed to be Markov. This method could create more accurate models for different parts of the historical data. As a result, this method could create thick tails in many applications. This regime-switching model was first introduced by Hamilton (1989). Log–normal Regime Switching (RSLN) model was then used by Bollen (1998). Because it can be fitted to many thick tail distribution, it is widely recommended to model equity index yield. Here we used the model RSLN2 recommended by Hardy (2003). The model of RSLN2 is as follows: $R_t|\rho_t \sim N(\mu_{\rho_t}, \sigma^2_{\rho_t})$, where ρ represents the regime. There are transition probabilities between regimes $P_{i,j} = \text{Prob}(\rho_{t+1} = j|\rho_t = i), i = 1,2 \;\; j = 1,2$.

Hardy (2003) describes a method to compute log–likelihood function for RSLN2 models. The procedure is to maximize the log–likelihood of TSX yield at time t based on the conditional distribution given information up to $t - 1$. The difficult part is that the conditional distribution involves both the distribution of regimes and the distribution of yield rate given the condition of certain regime. The estimated parameters are in the tables at the end of this chapter. The plots of prediction are as below, which are similar to what we did with the ILN and ARIMA–GARCH models. The middle two plots are the simulation of quarterly data, and the bottom two are the simulation of yearly data. The distribution of stock price yield is unbiased to the mean level and has thicker tails.



With regard to RSLN model, we still have two concerns as follows:

- The first concern is about the volume of data with respect to the total 6 parameters. In quarterly data, we have only 200 data values, which results in 30 data values for each parameter on average. That will make the model unreliable for prediction, especially in this long term situation for prediction over 10–year period.

- The second problem for the modeling is that the constructed regimes based on a Markov chain are not observable. When we modeled RSLN, the commonly used test is the LRT test, which is used to compare maximized log–likelihood values with other models such as ILN, ARIMA models. Fortunately when we compare RSLN models with other models, our purpose here is to check the estimation of variance or thickness of tails. Therefore our test is not used to find the Markov chain probability matrix, but to find the variance or tails. After taking the residual of the variance, we want to see:

- Whether the standardized residuals are white noise, which could be verified from the histogram plots of rate residuals of each regime.
- Whether the tails are significantly increased, or whether the yield distribution has lower 5% percentiles.

2.3 Model Validation

To evaluate various methods of modeling TSX index yield, we took a look at their fitness of trend, checked the behavior of residuals, and compared the measures of percentiles after simulation.

The first step of testing is to check the fitness of TSX Total Return Index yield. The tests include:

- Testing the predictability for the period of 1996–2005, total 10–year period, based on the information given by the Canadian data during 1956–1995. The plots are done in the previous sections. From the results, we can see that ILN, ARIMA and RSLN model are quite similar in predicting the moving trend of the index. Unfortunately the Wilkie Model does not provide a good prediction for the TSX Total Return Index yield based on the prediction of dividend yield and dividend.

- We also calculated and compared the values of maximized log-likelihood function, AIC, SBC, and LRT test for each model.
  - For log-likelihood function $\log L = l = \prod_{i=1}^{n} f(R_t|\mathcal{F}_{t-1})$, its maximized values of different models are computed and compared, which are used to compare the fitness of index variables.
  - Akaike Information Criterion (AIC) : $AIC = \log L - k$, where k is the number of parameters. In this criterion, the penalty of AIC value for each parameter is one.
  - Schwarz–Bayesian Information Criterion (BIC): $BIC = \log L - \frac{1}{2}k \ \log(n)$, where n is the number of observations. In this criterion, the penalty of BIC value for each parameter is $\frac{1}{2} \ \log(n)$.
  - P–value of Log–likelihood Ratio test (LRT) : $LRT = 2 * (\log L_1 - \log L_2) \sim \mathcal{X}^2_{d.f.}$, where degree of freedom (d.f.) is the difference of the number of parameters in two models. This test is used to see whether the models with more parameters are significantly better based on the p–value of $\mathcal{X}^2_{d.f.}$ distribution.

From the results shown in the tables, we can see that for monthly data, RSLN2 model is significantly better than ILN and ARIMA models. For quarterly data, RSLN also has the highest log–likelihood values. However it does not significantly perform better than ARIMA. ILN is significantly worse than RSLN and ARIMA models. This is because ILN model does not consider the dependence features of the time series. The Wilkie Model does not show any improvement compared to other models, even though it has more parameters. Actually we can see that the Wilkie Model is the worst among all these four methods in fitting the data.

The second step of testing is for the residuals after fitting the models. The calculation of residuals is listed as follows. We did not use standardized residuals, because the residuals were obtained from the same TSX index yield. Whether the residuals are standardized or not does not affect their qualitative characters such as independence and the shape of distribution. What is more, the non-standardized residuals can be used to compare the variance of each group of residuals.

- ILN model.

  residuals $= R_t - \mathrm{E}(R_t) = R_t - e^{\mu+\frac{\sigma^2}{2}}$

- ARIMA model.

  residuals $= R_t - \mathrm{E}(R_t|\mathcal{F}_{t-1})$. There is no further residual test for this method, since we know that they are white noise when we construct the models.

- RSLN model

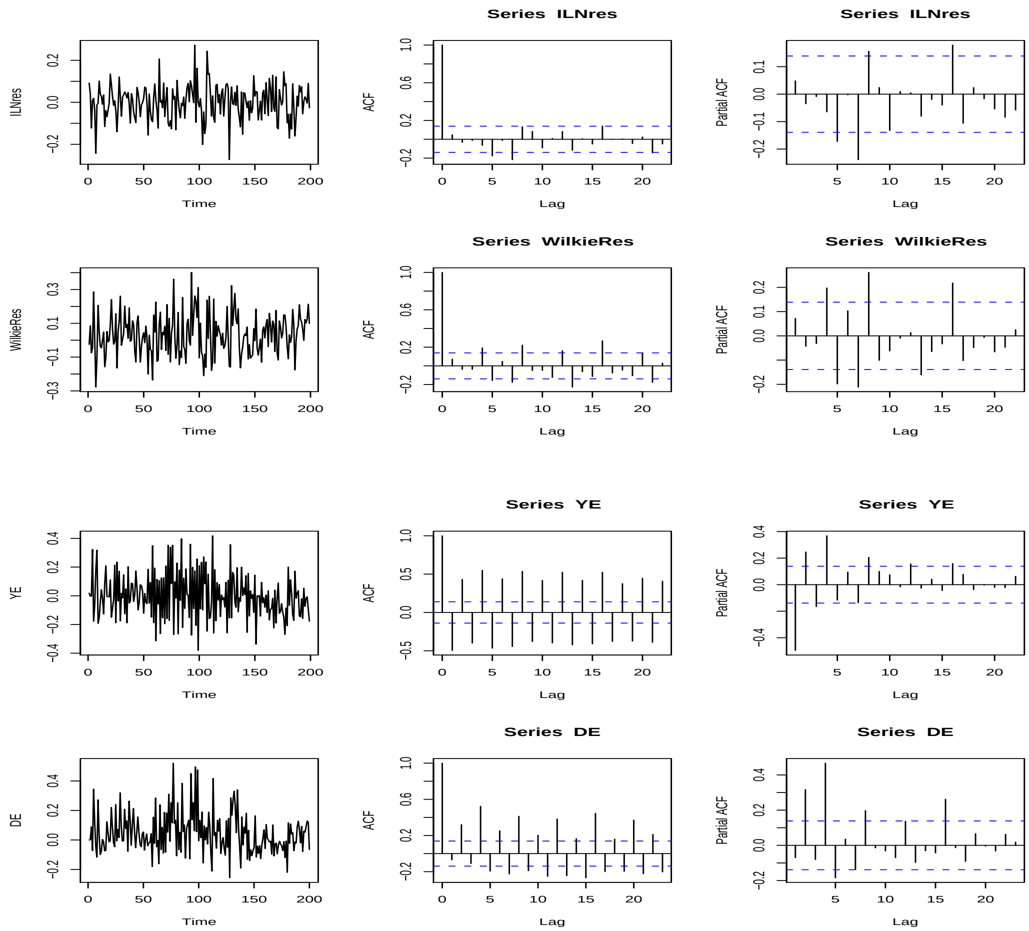
  residuals $= R_t - \mathrm{Prob}(\rho = 1) * \mu_1 - \mathrm{Prob}(\rho = 2) * \mu_2$, where ρ is the regime.

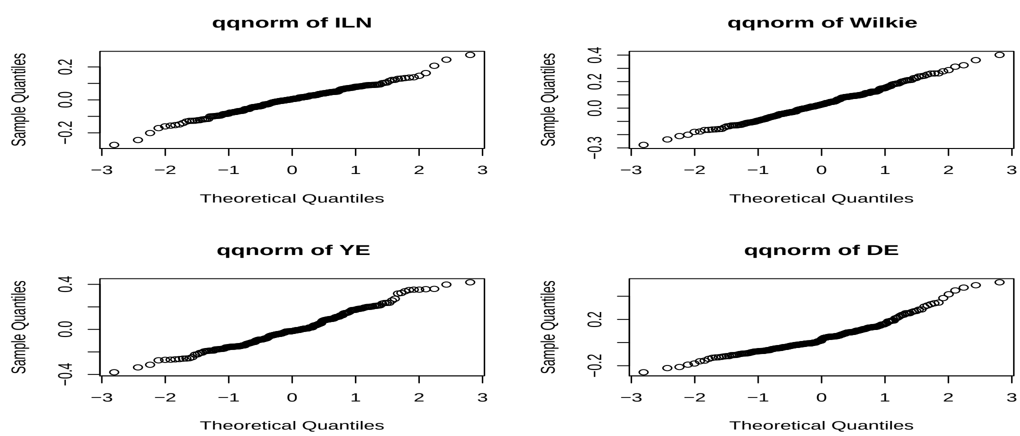
- The Wilkie Model

  residuals $= R_t - \dfrac{\text{predicted (price + dividend) at time t } |\mathcal{F}_{t-1}}{\text{real index price at time t-1}}$

- Other residuals in the Wilkie Model used to predict TSX index yields, such as $YE(t)$ from dividend yield model and $DE(t)$ from dividend model, were also obtained for the test.

The following are the plots of residuals and their ACF, PACF, and Q–Q normal plots. From the following plots, we can see that the residuals of the Wilkie Model for TSX index yield show some dependence relationship with previous residuals. $YE(t)$ and $DE(t)$ are not white noise at all. They are dependent on the previous values of their own. These results contradict the assumptions of the Wilkie Model.





The following are the correlation values for the residuals of ILN, RSLN, and the Wilkie Model, and the correlation between $YE(t)$ and $DE(t)$ in the Wilkie Model. From the correlations, we can see that:

- Residuals of RSLN and ILN model are closer to each other. That is to say, their predictions are close, while the residuals of the Wilkie Model show less correlation with them.

- The residuals $YE(t)$ and $DE(t)$ in the Wilkie Model show a strong correlation, which displays their dependence on each other, and implies some other information should be added to the models.

```
Correlations:
                    ILN       RSLN     Wilkie       YE
RSLN              0.732
                  0.000
Wilkie            0.559      0.501
                  0.000      0.000
YE               -0.067     -0.151     -0.617
                  0.349      0.033      0.000
DE                0.430      0.269      0.159      0.671
                  0.000      0.000      0.025      0.000

Cell Contents: Pearson correlation
               P-Value
```

The third testing was conducted after simulation of 10,000 times. We tested the performance of each model. The interested statistics include mean, standard deviation, skewness, kurtosis, minimum value, percentiles of 2.5%, 5%, 10% The standards of percentiles are set up by CIA SFTF report 2000 are listed as follows:

period	2.5th	5th	10th percentile
10-year	0.85	1.05	1.35

For monthly data, we also used the statistic Pr(Crash):

$\Pr(\text{Crash}) = \Pr\{\min_{1 \leq t \leq 120} Y_t \leq -0.2552\}$

The Prob(crash) is based on the event in October 1987 when the TSE index crashed with the historical low yield value of -0.2552. This test was not done for the quarterly data, because I have no crash value for quarterly yield. The calculation of skewness and kurtosis is as follows:

$$\begin{aligned}
\text{Skewness} &= \frac{\sum_{i=1}^{N}(R_i-\bar{R})^3}{(N-1)s^3} \quad \text{(0 under normal distribution)} \\
\text{Kurtosis} &= \frac{\sum_{i=1}^{N}(R_i-\bar{R})^4}{(N-1)s^4} \quad \text{(3 under standard normal distribution)}
\end{aligned}$$

From the table, we can see that the simulation of RSLN2 shows the best performance. It does not only improve the probability of crash and has lower percentile values, but it also fits the data best. Of course RSLN method has its disadvantages. It does not behave smoothly when it switches from one regime to the other one, which is a contradiction to reality.

2.4 Tables

In this section I listed all the tables we obtained for the previous models.

- Table 2.1–2.6 are the estimated parameters for each model.
- Table 2.7–2.8 are the comparisons of the values of model selection statistics for each model.
- Table 2.9–2.12 are the summary of statistics based on the simulation of 10,000 times for each model.

Table 2.1: Estimated Parameters of the Wilkie Model for Canadian Quarterly Data

Models	Excel Parameters	56–05	56–95	56–05 sim	56– sim
Inflation	QMU	0.010058	0.011161	0.010058	0.011161
	QA	0.598371	0.605274	0.598371	0.605274
	QSD	0.007225	0.007502	0.007225	0.007502
Dividend	YW	2.412561	5.022816	2.035675	4.795301
Yield	YA	0.89686	0.604646	0.906966	0.631055
	YMU	0.007261	0.008357	0.007367	0.008376
	YSD	0.159886	0.157827	0.159913	0.157874
Dividend	DW	1	1	1	1
	DD	1	1	1	1
	DMU	0.00129	0.0000705	0.00107	0.0000708
	DY	-0.15644	-0.11854	-0.16682	-0.1277
	DB	-0.66449	-0.68753	-0.66031	-0.68413
	DSD	0.1258548	0.136015	0.125749	0.135897

Table 2.2: Estimated Parameters of the Wilkie Model for Canadian Yearly Data

Models	Parameters	Wilkie's values	My values 56–95 sim	values 56–05 sim
Inflation	QMU	0.034	0.043819	0.038636
	QA	0.64	0.804559	0.821502
	QSD	0.032	0.018062	0.016788
Dividend	YW	1.17	-2.274744033	-2.616794131
Yield	YA	0.7	0.847249498	0.966969169
	YMU	0.0375	0.039639689	0.038402052
	YSD	0.19	0.129959649	0.138155021
Dividend	DW	0.19	0.860235993	0.554153424
	DD	0.26	0	0
	DMU	0.0010	0.004341127	0.007470869
	DY	-0.11	0.430619503	0.353427584
	DB	0.58	-0.969651819	-0.99
	DSD	0.07	0.127028916	0.119019924

Notes:

- The Canadian data for all variables have the same length between 1956 to 2005, or between 1956 to 1995.
- Parameters of different models are estimated one by one in order (sep), or simultaneously (sim) for dividend yield model and dividend model.

Table 2.3: Estimated Parameters of the Wilkie Model for US Quarterly Data

Models	Parameters	Data 26–89	Data 26–06	Data 26–89 sim	Data 26–06 sim
Inflation	QMU	0.007604	0.07395	0.007604	0.07395
	QA	0.625884	0.593821	0.625884	0.593823
	QSD	0.010894	0.010212	0.010894	0.010212
Dividend	YW	0.761259	1.696026	1.864281	2.519933
Yield	YA	0.694283	0.891736	0.768182	0.925786
	YMU	0.010485	0.008829	0.010471	0.009639
	YSD	0.197583	0.193611	0.198853	0.194404
Dividend	DW	1	1	1	1
	DD	1	1	0.999792	0.999902
	DMU	0.003892	0.003694	0.003282	0.001204
	DY	-0.28056	-0.29176	-0.33527	-0.32668
	DB	-0.52753	-0.52858	-0.52399	-0.5331
	DSD	0.146562	0.135818	0.144477	0.134476

Table 2.4: Estimated Parameters of the Wilkie Model for US Yearly Data

Models	Parameters	Wilkie's values 26–89	Data 26–89 sim	Data 26–06 sim
Inflation	QMU	0.030	0.030621	0.03012
	QA	0.65	0.621124	0.620844
	QSD	0.035	0.036179	0.032398
Dividend	YW	0.5	0.354067	0.389223
Yield	YA	0.7	0.820782	0.953086
	YMU	0.0430	0.043148	0.032163
	YSD	0.21	0.144368	0.145177
Dividend	DW	1	1	1
	DD	0.38	0.73576	0.75731
	DMU	0.0155	0.013279	0.014013
	DY	-0.35	0.142592	0.203263
	DB	0.50	-0.3014	-0.28749
	DSD	0.09	0.13983	0.138464

Data source:

- The US yearly data for all variables have the same length between 1926 to 2006.
- Parameters of different models are estimated one by one in order based on the one way direction cascade structure of the Wilkie Model, or simultaneously (sim) for dividend yield model and dividend model.

Table 2.5: Parameters of Fitted Models for Quarterly TSX Return of 1956-2005/1995

Models	Parameters	Data 56–05	95%CI	Data 56–95	95%CI
ILN	μ	0.0231	(0.012, 0.034)	0.0225	(0.010, 0.035)
	σ	0.0809	(0.074, 0.090)	0.0810	(0.073, 0.091)
ARIMA(0,0,0)	μ	0.0265	(0.022, 0.031)	0.0261	(0.021, 0.031)
x(0,0,1)[5]	sma1	-0.2371	(-0.392, -0.819)	-0.1975	(-0.111, -0.285)
	$\hat{\sigma^2}$	0.001948			
RSLN2	p_{12}	0.388005657		0.51686224	
	p_{21}	0.313011721		0.321146443	
	μ_1	0.05191		0.052867	
	σ_1	0.046991		0.038279	
	μ_2	-0.00011		0.003641	
	σ_2	0.093703		0.093437	

Table 2.6: Parameters of Fitted Models for Monthly TSX Return of 1956-2005

Models	Parameters	Data 56–05	95%CI	Data 56–95	95%CI
ILN	μ	0.00781	(0.004,0.011)	0.00760	(0.004, 0.012)
	σ	0.04503	(0.043, 0.048)	0.0810	(0.042, 0.047)
ARMA(1,1)	μ	0.0088	(0.007, 0.011)	0.0086	(0.005, 0.013)
	ar1	-0.7356	(-1.103, -0.368)	-0.8348	(-1.003, -0.667)
	ma1	0.8309	(0.495, 1.112)	0.9172	(0.793, 1.042)
	$\hat{\sigma^2}$	0.001948		0.001879	
ARMA(1,1)	μ	0.0148	(0.008, 0.219)	0.0137	(0.006,0.021)
+Garch(1,1)	ar1	-0.7629	(-1.042, -0.484)	-0.8426	(-1.050,-0.635)
	ma1	0.8380	(0.607,1.070)	0.9078	(0.748,1.068)
	K^*10^{-4}	2.6599	(0.160,5.160)	2.5445	(-0.544,5.633)
	Garch	0.7591	(0.589,0.930)	0.7755	(0.558,0.994)
	ARCH	0.1129	(0.044, 0.182)	0.0932	(0.014,0.172)
RSLN2	p_{12}	0.193244986		0.193901148	
	p_{21}	0.042154273		0.040173968	
	μ_1	-0.015944318		-0.012303931	
	σ_1	0.072537313		0.073023395	
	μ_2	0.013034636		0.011762245	
	σ_2	0.034086235		0.033750334	

Table 2.7: Comparison of Fitted Models for Quarterly TSX Return during 1956-2005

Statistics	ILN	ARMA(0,0)x(0,1)^5	RSLN	the Wilkie Model
# of Parameters	2	2	6	10
logL	218.43	220.44	224.28	212.58
SBC	212.17	214.17	205.48	181.24
AIC	216.43	218.44	218.28	202.58
LRT p-value	0.0197	0.104	N/A	N/A

Table 2.8: Comparison of Fitted Models for Monthly TSX Return during 1956-2005

Statistics	ILN	ARMA	ARMA–GARCH	RSLN	the Wilkie Model
# of Parameters	2	3	3	6	N/A
logL	1013.04	1019.12	1024.8	1049.22	N/A
SBC	1006.65	1009.72	1005	1034.42	N/A
AIC	1011.04	1016.12	1018.8	1043.22	N/A
LRT p-value	$< 10^{-8}$	$< 10^{-8}$	$< 10^{-8}$	N/A	N/A

Note:

- The number of parameters of the Wilkie Model for prediction of Total Return yield is 13. But when we calculated the log–likelihood of total return yield, we use only dividend yield model and dividend model, no inflation model. Therefore there are 10 parameters here for the Wilkie Model.

- The LRT test is based on the comparison with the log–likelihood value of RSLN2 model, because it is the highest among the values of all used models. The test statistics $= 2(l_{\text{RSLN}} - l_{\text{others}})$

- We can see that the Wilkie Model does not improve the maximized log–likelihood values significantly, even though it has much more parameters. The main reason is the poor fit of recent 10 years, which can be seen from the plots for the Wilkie Model.

- The maximized log–likelihood values will not be comparable when the calculation of distribution density is different, based on different methods. So in the calculation we need to keep consistent in computing the log–likelihood function based on normal density function, i.e. use the standardized normal $f(z) = \phi(z = \frac{x-\mu}{\sigma})$ in all cases, or use normal $f(x) \sim N(\mu, \sigma)$ in all cases. Do not mix these two methods.

Table 2.9: Simulation of 10–Year Monthly TSX Return under Continuous Assumption

Statistics	ILN	ARMA(1,1)	ARMA–GARCH(1,1)	RSLN
mean	0.0092	0.0074	0.0085	0.0079
std.dv	0.0456	0.0452	0.0458	0.0450
Skewness	0.1430	0.0450	0.0005	-0.6437
Kurtosis	2.9685	2.9779	3.1870	5.6144
Minimun accumulated value	-0.7563	-1.1553	-1.3169	-2.3717
2.5 percentile accumulated	-0.0401	-0.0675	-0.0108	-0.3254
5 percentile accumulated	0.1192	0.0836	0.1578	-0.0908
10 percentile accumulated	0.3446	0.2724	0.3495	0.1791
Minimun value	-0.1986	-0.2153	-0.3558	-0.3463
2.5 percentile	-0.0806	-0.0803	-0.0817	-0.0958
5 percentile	-0.0663	-0.0661	-0.0663	-0.0677
10 percentile	-0.0500	-0.0497	-0.0492	-0.0444
Pr(crash)	0	0	0.0008	0.0088

Table 2.10: Simulation of 10–Year Quarterly TSX Return under Continuous Assumption

Statistics	ILN	ARMA(0,0)x(1,1)[5]	RSLN
mean	0.0273	0.0231	0.0229
std.dv	0.0829	0.0811	0.0807
Skewness	0.2306	-0.0284	-0.4634
Kurtosis	3.1670	3.0082	3.6518
Minimun accumulated value	-0.6952	-0.6478	-1.6515
2.5 percentile accumulated	-0.0806	0.1197	-0.1366
5 percentile accumulated	0.0909	0.2608	0.0316
10 percentile accumulated	0.3053	0.4106	0.2342
Minimun value	-0.3683	-0.3700	-0.4215
2.5 percentile	-0.1357	-0.1356	-0.1586
5 percentile	-0.11010	-0.1099	-0.1258
10 percentile	-0.0808	-0.0806	-0.0864

Comments:

- From the results, we can see that ILN model has smallest tail for monthly data, and RSLN model has the thickest tail for both monthly data and quarterly data.

Table 2.11: Simulation of the Wilkie Model for Yearly TSX Return over 10 years

Statistics	Wilkie CA	CA 56-05	Wilkie US	US 26-06
mean	0.0475	0.0721	0.0747	0.0755
std.dv	0.2589	0.2127	0.3463	0.2024
Skewness	0.6996	0.6158	1.0254	0.5414
Kurtosis	3.8159	3.6865	4.7640	3.5715
Minimun accumulated value	-0.8840	-0.6556	-0.8888	-0.8093
2.5 percentile accumulated	-0.7040	-0.3241	-0.7207	-0.4976
5 percentile accumulated	-0.6454	-0.2504	-0.6616	-0.4044
10 percentile accumulated	-0.5735	-0.1623	-0.5870	-0.2942
Minimun value	-0.6433	-0.5615	-0.7419	-0.5385
2.5 percentile	-0.3701	-0.2905	-0.4512	-0.2663
5 percentile	-0.3218	-0.2441	-0.3963	-0.2233
10 percentile	-0.2614	-0.1879	-0.3242	-0.1681

Note:

- The parameters used here include Wilkie's parameters for Canadian market, parameters estimated based on Canadian data of 1956–2005, Wilkie's parameters for US market, and parameters estimated based on US data of 1926–2006.

- Updated parameters were estimated from data up to 2005 and the estimation for the parameters of dividend yield and dividend models were made simultaneously.

- The initial values are based on Canadian market 1995 data, and US market 1995 data.

- The final calculation of R_t is based on discrete assumption.

Comments:

- Updated US and Canadian models have higher TSX Total Return Index yield, which is due to the recent development of the market. Updated US modes also has smaller deviance, smaller kurtosis, smaller skewness, and thinner tails. These improvement not only comes from the more data and the development of their distribution, but also comes from the updated parameters which makes the simulation more reasonable.

- We can see that the Wilkie Model for yearly data underestimates the confidence intervals, which create too thick lower tails, even for updated parameters.

Table 2.12: Simulation of the Wilkie Model for Quarterly TSX Return over 10 years

Statistics	based on CA 56–95	CA 56–05	US 26–89	US 26–06
mean	0.0107	0.0145	0.0127	0.0139
mean2	0.0422	0.0410	0.0592	0.0517
std.dv	0.2730	0.2556	0.3498	0.3248
Skewness	0.7879	0.7146	1.0286	0.9510
Kurtosis	3.9965	3.8208	4.8726	4.7542
Minimun accumulated value	-0.7905	-0.9120.	-0.9516	-0.9848
2.5 percentile accumulated	-0.5868	-0.6500	-0.8203	-0.8718
5 percentile accumulated	-0.5250	-0.5789	-0.7699	-0.8295
10 percentile accumulated	-0.4433	-0.4739	-0.6930	-0.7598
Minimun value	-0.7171	-0.6826	-0.7964	-0.7515
2.5 percentile	-0.3948	-0.3709	-0.4650	-0.4449
5 percentile	-0.3441	-0.3220	-0.4081	-0.3888
10 percentile	-0.2785	-0.2590	-0.3357	-0.3182

Note:

- The parameters used here are parameters estimated based on Canadian data 1956–1995, on Canadian data 1956–2005, on US data 1926–1989, and on US data 1926–2006.
- Estimation for the parameters of dividend yield and dividend models were made simultaneously.
- The initial values are based on Canadian market October 1995 data, and US market October 1995 data.
- The final calculation of R_t is based on discrete assumption
- mean $= \frac{\text{average}(P_{t+1}+D_{t+1})}{\text{average}(P_t)}$
  mean2 $= \text{average}(\frac{(P_{t+1}+D_{t+1})}{P_t})$

Comments:

- We can see that parameters are different. The difference is because of different length of series and different markets. This result makes us doubt the method of linear modeling for those series.
- Compared with RSLN model, we can see that the Wilkie Model for quarterly data have bigger kurtosis, thus has wider confidence intervals, and too thick tails. Also the Wilkie Model underestimates TSX yield values and is biased.

Chapter 3

Discussion of Dividend Yield

3.1 Test of Dividend Yield

Dividend yield $= \dfrac{\text{Dividend}}{\text{Price}}$ shows the meaning of dividend yield $Y(t)$. Dividend yield could display the property of a stationary series through fitting some ARIMA models. However, it need to be more stationary than stock price such that its models has better prediction ability than stock price models, and it also needs to be an ergodic process in order to be valid for prediction.

The data is from CRSP monthly index. This database provides access to NYSE/AMEX/Nasdaq daily and monthly security prices. What I used to calculate dividend yield are two indices "Value weighted return–all distributions" rr_1, and "Value weighted return–exclude dividends" rr_2, and "S&P Composite Index" P. Since it does not matter whether we use discrete assumption and continuous assumption for monthly data, here we calculated dividend yield based on discrete assumption for dividend yield according to Wilkie (1995). The calculation is as follows:

$$\text{dividend yield for month t-1:} \quad y_{t-1} = rr_1 - rr_2 \tag{3.1}$$

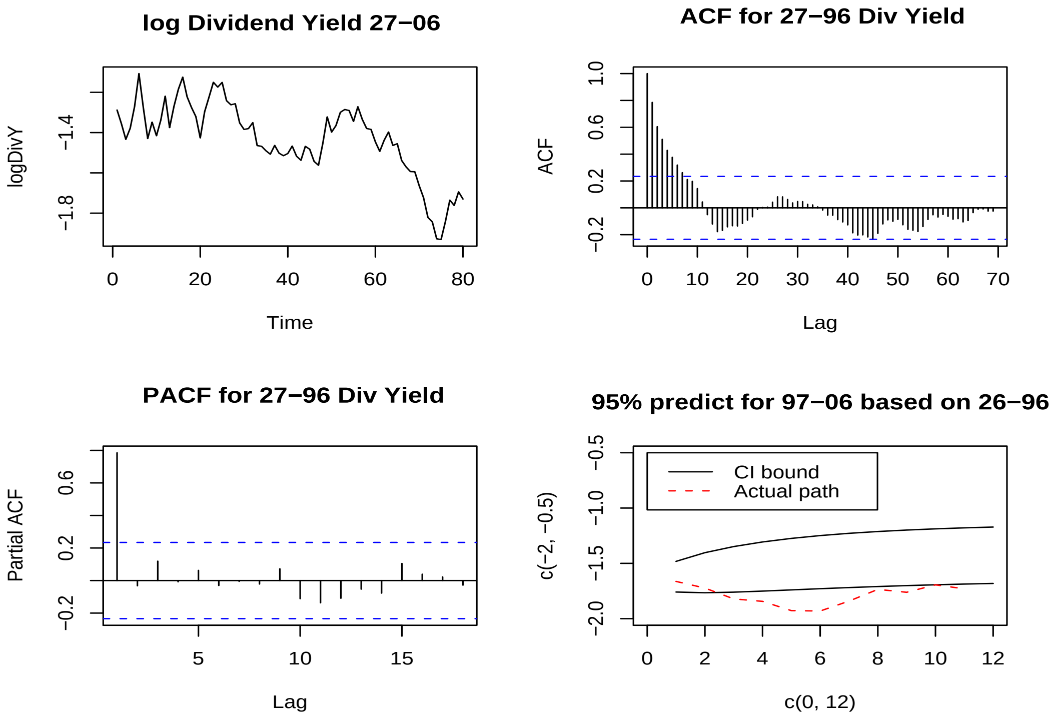
$$\text{dividend yield for month t:} \quad y_t = y_{t-1} * \frac{P_{t-1}}{P_t} \tag{3.2}$$

$$\text{Yearly dividend yield} \quad Y_t = \prod_{i=1}^{12}(1 + y_t) \tag{3.3}$$

$$\text{logarithm of dividend yield} = \log(Y_t) \tag{3.4}$$

The equation (3.2) above may be unnecessary, which actually does not make any significant difference from those data series produced by the methods without considering equation (3.2).

The following plots are based on the data series for US dividend yield $Y(t)$ of 1927 – 2006.



From the above plots we can see that:

- The time series for dividend yield is not very stationary.
- The time series from 1927 to 1996 fits very well with a AR(1) model, which is verified by the residual analysis through Ljung–Box test.
- The prediction of the period from 1997 to 2006 by the model inferred from the data 1926 to 1996 is poor based on this simple ARIMA model. The actual values are outside the 95% CI from the prediction.

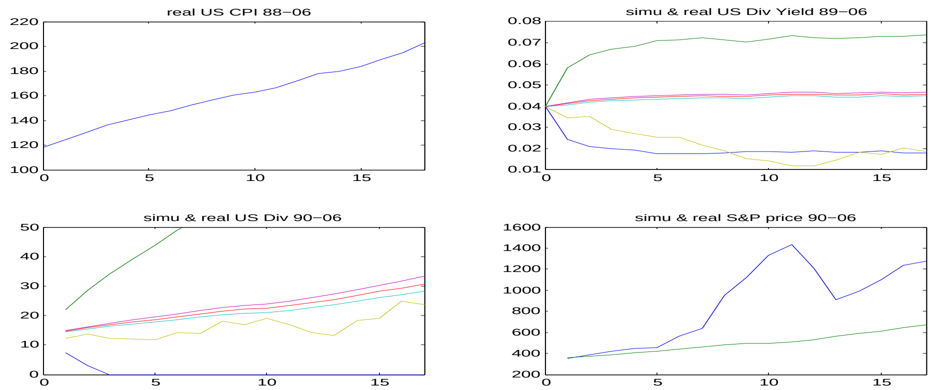
From the above information, we have to suspect the validity of using dividend yield as a stationary time series to predict price. Let us look at what is the Wilkie Model for $\log Y(t)$. It is a regression model with AR(1) residuals. The regression part is $\log(Y(t)) \sim \nabla \log Q(t)$, where $Y(t)$ is dividend yield and $\nabla \log Q(t)$ is the inflation rate. The accuracy of prediction depends on two factors. One is the prediction of inflation rate, which is independent of dividend yield in the Wilkie Model; the other is the regression model with AR(1) residuals.

We start with the regression model. That is to say, we assume the prediction of inflation rate is perfect and we use the real value of inflation as a replacement for the predicted values. The inflation index is calculated based on the Consumer Price Index from CRSP database. The price of June is chosen as the representation of annual index. The calculation of inflation rate is as follows: $\nabla lnQ(t) = log(\frac{CPI_t}{CPI_{t-1}})$.

The following are some notes for using the Wilkie Model to predict share price.

- The model parameters are estimated based on US stock market from 1927 to 1989.
- The inflation rate is from the real values of 1990 to 2006. It is an input for dividend yield and dividend model. Stock Price is defined as: $P(t) = \frac{D(t)}{Y(t)}$
- All simulated values are based on repeated 1000 simulations.

The following are the plots of the results:

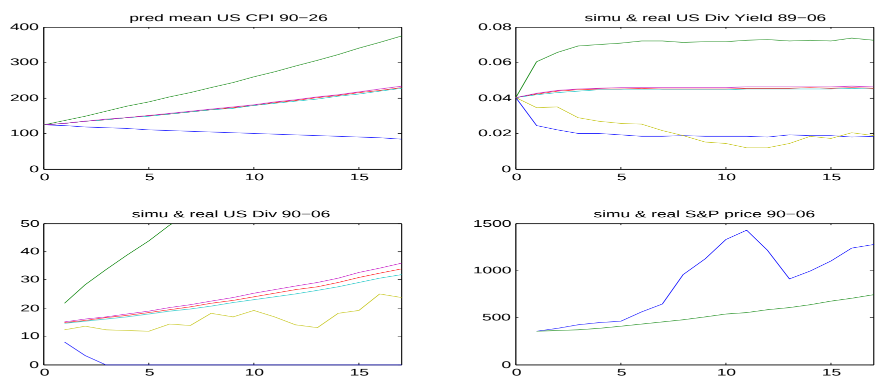


The values calculated are as follows:

```
predicted dividend level =   30.9675

real dividend = 23.76869

predicted mean Yield =    0.0459

95%CI predicted mean Yield =    0.0450    0.0467

95%CI predicted yield =    0.0185    0.0732

mean Real Yield =    0.0222

range of Real Yield =    0.0117 0.0401

real_price =  1276.7

sim_price =  672.5140
```

From the above plots and results, we can see that in the Wilkie Model, even though the prediction for dividend is good, because of the poor prediction for dividend yield $Y(t)$, the price index of S&P is poorly predicted. As a result, we doubt whether dividend yield would be a good component in the structure of the Wilkie Model.

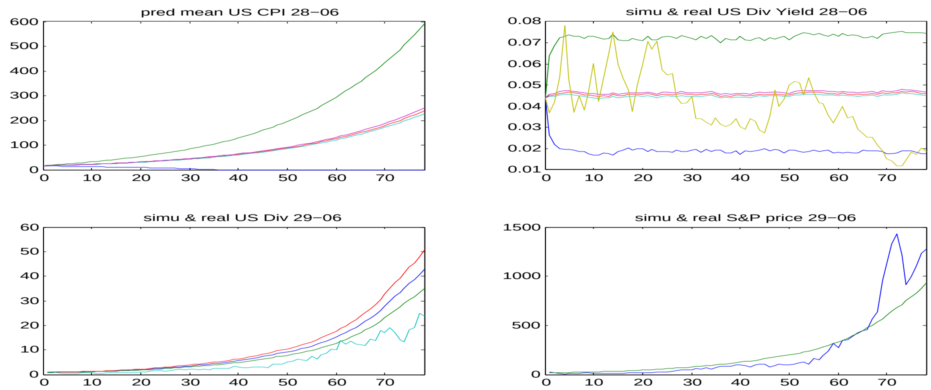
Another interesting part that we want to compare is whether the aggregated models from the Wilkie Model which includes inflation rate, dividend yield and dividend will improve the total prediction than the separate models for each single component of the structure. Unfortunately, I could not find the any improvement for the prediction based on the aggregated models of the Wilkie Model, when I use predicted inflation rate, dividend yield and dividend together to predict price. Actually his prediction for inflation rate needs an evaluation too. The following are the plots of the aggregated prediction.



The values of the 17–year predicted inflation rate from the Wilkie Model are as follows:

```
>> Thesis_US_simInf_DivY_Div_Price(1000)

    meanQ2 =  228.1643

    CImeanQ2 = ( 223.8392 , 232.4895 )

    CIQ2 = (91.3908, 364.9379 )
```

The mean of simulated results for 17 years inflation rate are close to the real values. From the plots below, we can clearly see that, for dividend yield and share price, the simulated long run values are far from the real values in recent development.



If we just treat the dividend yield as a nuisance variable, we find that the reason for the poor prediction of price is not due to the poor modeling of dividend yield. It is more likely due to the structure of the Wilkie Model. Actually the share price prediction based on the prediction of dividend and dividend yield is not a good method. Based on the parameters of the ARIMA model and of the Wilkie Model for the Canadian quarterly data during 1956–2005, we can see that the AR coefficient for stock price in the ARIMA model is -0.1975, and the AR coefficient for the dividend yield in the Wilkie Model is 0.906966. From these two numbers we can see that the stock price is more stationary than dividend yield, based on the fit of ARIMA–Regression models. In another words, using dividend yield to predict share price will not be better than predicting share price directly, with the regression on other predictors.

Why are the degrees of stationary for the price and dividend yield different? In the stock market, share price and dividend tend to be stationary, while dividend yield may have no such process. In recent years, the sharp decrease of dividend yield was due to the quick increase of stock price. In the stock market where share price determines dividend yield, it may not be a good idea to model dividend yield first, then model the price based on the output of dividend yield. However the interaction between dividend yield and stock price could be considered when modeling share price.

Another reason to discard dividend yield as a predicting factor for share price is that the dividend yield keeps going down and is at a low level already. If the percentage of dividend is at a really low level, how much weight can you give to dividend to help predict price? A reasonable one is to treat it as a trivial one. We think this is a more reasonable suggestion for the market with higher risks. That is why many financial managers try to keep the dividend payments more stationary to keep the risk feeling at a lower level. However, the stock market is a risky market, if the dividend yield reaches a low level, the power of dividend to predict the share price could decrease to a very low level. The following paired data bootstrap method was used to test the decreasing trend for the dividend yield based on the historical data, especially those in recent years.

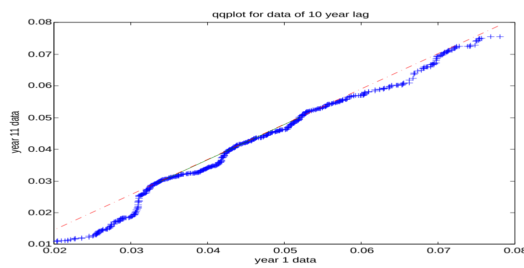
The structured paired data bootstrap method is based on monthly data for S&P dividend yield. Its process is as follows:

- Obtaining smoothed yearly dividend yield based on monthly yield rate.
- Bootstrapping the paired data of 10–year period. That is to say if data one was taken from the year 1 June index yield, then data two was taken from the year 11 June index yield. The sample size is 2000.
- Using the nonparametric test of two sample paired sign–rank test for the median values. The advantage of this test is that we do not have to know the distribution of the difference. We only need to know that they come from the same distribution. And we want to test whether the difference of the means from two index yield sets

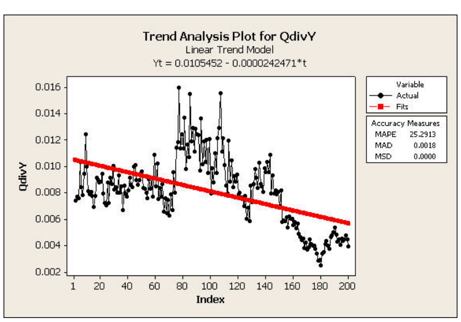
of 10–year period is zero.

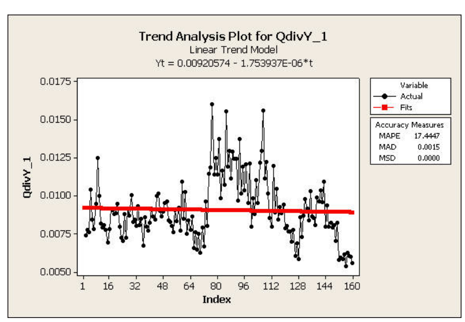
```
>> Thesis_bootstrap_DivYield_1yr()
mean_Q =     0.0399    0.0394     % 0.0399 is year 1 mean, 0.0394 is year 11 mean
median_Q =   0.0397    0.0392     % 0.0397 is year 1 med, 0.0392 is year 11 med
p =     9.7316e-007               % p-value < 0.05
h =     1                         % this means reject the H0: median of the
                                  %        difference is zero.
```

The plot below is the Q–Q plot of the paired values with dividend yield of year 11 as y coordinate, and those of year 1 dividend yield as x coordinate.



From the above paired sign–rank test results, we can see that the dividend yield shows decreasing trend after 10 years. That is to say, the weight for using dividend yield to evaluate share price is decreasing. From the above Q–Q plot, we can see that dividend yield of year 11 (y coordinate) has lower values in percentiles, compared to those of year 1 dividend yield (x coordinate). This downside trend can also be seen from the following plots. The left plot is the time series plot with trend for data from 1956 to 2005, and the right plot below is for data from 1956 to 1995.





But if we do not use dividend yield as a factor to predict share price, it will be impossible to use dividend only to predict share price. The following known as Variance Bound Test verifies this claim, even if dividend could be predicted well.

3.2 Variance Bound Test for Dividend-Price

Variance bound is introduced by Shiller, LeRoy and Porter (see LeRoy (1996)). The idea is as follows:

$$\begin{aligned} p_t &= \frac{D_t}{Y_t} = \beta E_t(d_{t+1} + p_{t+1}) \\ &= \beta E_t(d_{t+1} + \frac{D_{t+1}}{Y_{t+1}}) \\ &= E_t(\beta d_{t+1} + \beta^2 d_{t+2} + \ldots + \beta^{n+1} d_{t+n+1} + \beta^{n+1} p_{t+n+1}) \end{aligned}$$

$$\begin{aligned} where: &\quad p_t : \text{price of stock price} \quad d_t : \text{dividend} \quad \beta = max(\text{market PV rate}) \\ \text{let} &\quad p_t^* = \sum_{n=1}^{\infty} \beta^n d_{t+n} \\ \text{therefore} &\quad p_t = E_t(p_t^*) \\ \text{therefore} &\quad V(p_t) \le V(p_t^*) \end{aligned}$$

Since the calculation of P_t^*involves the evaluation of an infinite series in its final form, we use the original form $p_t = \beta E_t(d_{t+1} + p_{t+1})$. The test is based on monthly data from Jan. 1926 to Dec. 2006, S&P share price index and monthly dividends.

```
# This result shows that:

     variance of Price > variance of discounted(Dividend + Price)

> var(dividend)^0.5

[1] 0.5410521

> var(Dist_price_Dividend)  # variance of discounted(Dividend+Price)

     [1] 113076.0

> var(price) # variance of (Price)
```

```
[1] 137861.9
```

This verification will be more significant if we use higher lags of share price with the sum of related dividend in previous periods. So the price prediction must be based on another model. The new model structure will discard the component of dividend yield $Y(t)$.

Chapter 4

Discussion of New Model Structure

For the current analysis, we used Canadian data of Consumer Price Index (Q) as inflation index, 10+ years government bond yield (C) as long term interest rate, and TSX Total Return Index (T) as Price. All the items were available on various time spans, but the time period 1956.01–2005.12 is common. The data obtained is as follows:

- Consumer Price Index (Q) monthly data are from 1914.0 to 2006.11, total 1115. Its Quarterly data are from 1914.01 to 2006.10, total 372, and Yearly data are from 1914 to 2006, total 93.

- TSX Total Return Index monthly data are from 1956.01 to 2005.12, total 600, quarterly data are from 1956.01 to 2005.10, total 200, and yearly data are from 1956 to 2005, total 50.

- Consols return index monthly data are from 1950.10 to 2006.11, total 674, quarterly data are from 1950.10 to 2006.10, total 225, and yearly data are from 1951 to 2006, total 56.

- Data from the common period of 1956.01 – 2005.12 will be used for the multi–variable analysis of the three factors

- All the quarterly data are taken from the index of January, April, July and October of every year. All the yearly data use the index of June.

- Data of other periods will be used to calibrate the models for respectively component. The model construction will be based on the data from 1956.01 to 1995.12, prediction will be made for 1996.01 to 2005.12. From this idea, the model construction is still for a medium term prediction.

Since we are interested in dropping the dividend yield from the model, we need to test the relationship between each of the variables.

Before considering the new structure, we change the variable of share price by the TSX Return Index yield. The motivation for Wilkie to model the price based on the dividend model is due to the difficulty in modeling price directly and the feasibility to model dividend easily. Our suggestion is to use Total Return Index T'_t. Here in our models the dividend represents all distributions ($Dist_t$)the investors get other than capital gain from the price increasing.

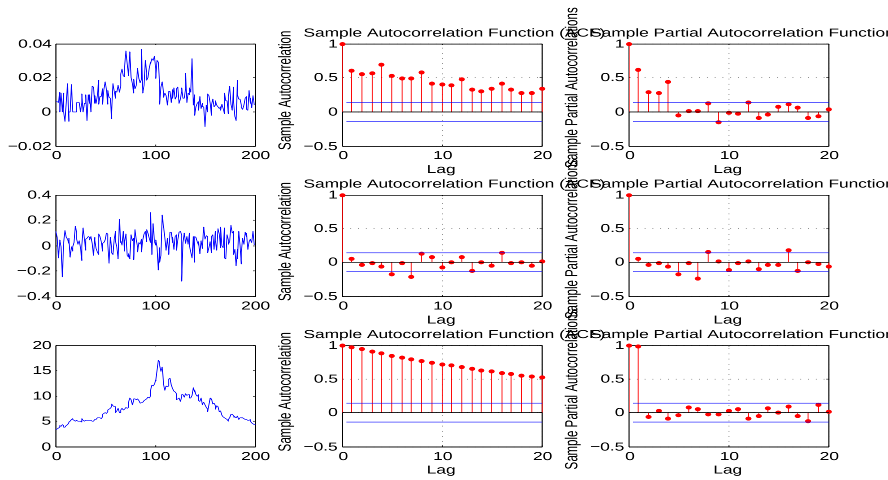
$$
\begin{aligned}
\text{since:} \qquad & \frac{T'_{t+1}}{T'_t} = \frac{P_{t+1}}{P_t} + \frac{Dist_{t+1}}{P_t} \\
\text{then:} \qquad & P_{t+1} = P_t(\frac{T'_{t+1}}{T'_t}) - Dist_{t+1} \\
& Dist_t = P_t(\frac{T'_{t+1}}{T'_t}) - P_{t+1}
\end{aligned}
$$

The reasons for us to choose T'_t instead of P_t are as follows:

- The distributions $Dist_t$ mainly consists of dividends and moves in a more stable path. Therefore I keep the idea to model dividend.

- Total return index T'_t is a martingale $T'_t = E_t(T'_{t+1}|\mathcal{F}_t)$ and can be predicted in a more predictable way.

Now we want to consider each variable before we select it into the structure. The following are the plots of inflation rate, TSX yield and Consols yield. They are ordered from the top to the bottom, and we include their ACF and PACF.



From the plots, we can see that:

- The inflation rate and the Consols yield show the long memory character with long lag effects. But the TSX yield shows a random walk property. So the models for inflation rate and Consols could include some memory factors of themselves, but the TSX yield should not.
- The Consols yield are more stable for a short period.
- Consols yield and inflation rate show some cointegration relationship, which can be seen in their covariance.

4.1 New Model Structure Based on Regressions

Here we plots eight variables:

- CPI growth rate (QCPIyield)
- TSX Return Index yield (QtseYield)

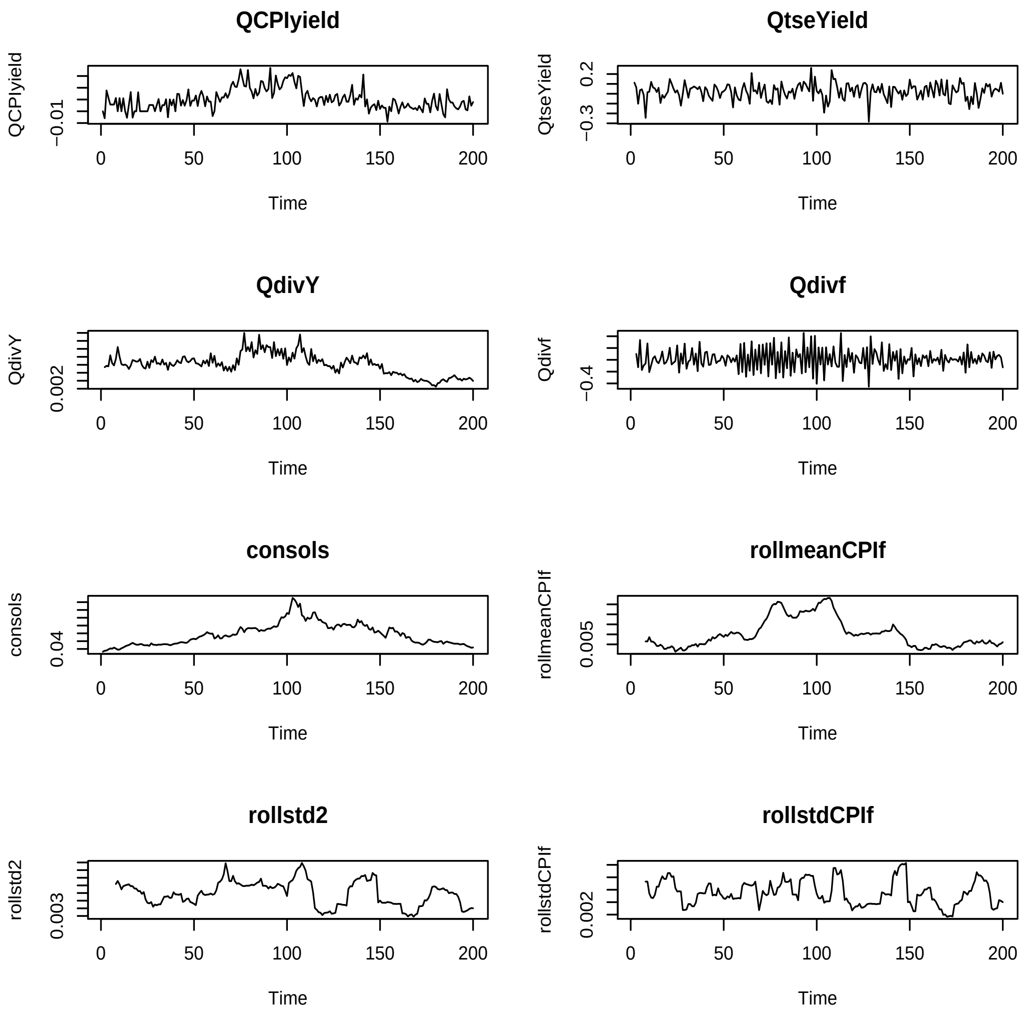
- Share dividend yield (QdivY)
- Share dividend growth rate (Qdivf)
- Consols yield (Consols)
- Rollover mean of CPI growth rate obtained from previous 2 years or 8 quarters (rollmeanCPIf). This variable works as a medium memory factor for the mean level of inflation rate.
- Rollover standard deviation of CPI growth rate from previous 2 years or 8 quarters (rollstdCPIf). This variable works as a medium memory factor of the variance of inflation rate.
- 50% rollover standard deviation of CPI growth rate for 4 years or 16 quarters (rollstd2). This variable also works as a medium memory factor of the variance of inflation rate.

Overall relationships can be built up through the correlation test among all variables. Here I use 15 variables. The data used are quarterly data. Those variables include the variables mentioned above and the following:

- Previous term TSX Return Index yield (lagTseYield)
- Previous term CPI growth rate (lagCPIyield)
- Previous term share dividend growth rate (lagQdivf)
- Previous term share dividend yield (lagQdivY)
- Previous term share Consols yield (lagConsols)
- Lagged rollmeanCPIf (lagRollmean)

- Lagged rollstdCPIf (lagRollstd)

The plots are as follows:



The correlation values are listed as follows:

```
Cell Contents: Pearson correlation
               P-Value

                 Consols     QtseYield     QCPIyield          QdivY
QtseYield         -0.046
                   0.522

QCPIyield          0.549        -0.081
                   0.000         0.258
```

QdivY	0.472	0.045	0.506	
	0.000	0.527	0.000	
Qdivf	-0.014	0.383	0.064	0.287
	0.841	0.000	0.371	0.000
lagConsols	0.980	-0.007	0.510	0.470
	0.000	0.918	0.000	0.000
lagTseYield	-0.017	0.050	0.053	-0.190
	0.813	0.484	0.461	0.007
lagCPIyield	0.571	-0.119	0.600	0.507
	0.000	0.094	0.000	0.000
lagQdivY	0.469	0.073	0.455	0.828
	0.000	0.309	0.000	0.000
lagQdivf	-0.005	-0.120	-0.007	-0.146
	0.943	0.093	0.918	0.041
rollmeanCPIf	0.749	-0.019	0.777	0.696
	0.000	0.797	0.000	0.000
rollstdCPIf	0.086	-0.001	0.114	0.310
	0.235	0.992	0.114	0.000

	Qdivf	lagConsols	lagTseYield	lagCPIyield
lagConsols	0.007			
	0.922			
lagTseYield	-0.367	-0.048		
	0.000	0.505		
lagCPIyield	-0.044	0.550	-0.081	
	0.536	0.000	0.257	
lagQdivY	-0.226	0.467	0.043	0.507
	0.001	0.000	0.546	0.000
lagQdivf	-0.742	-0.023	0.384	0.058
	0.000	0.746	0.000	0.418

rollmeanCPIf	0.019	0.732	-0.014	0.794
	0.796	0.000	0.847	0.000
rollstdCPIf	0.008	0.080	0.028	0.133
	0.910	0.269	0.701	0.066

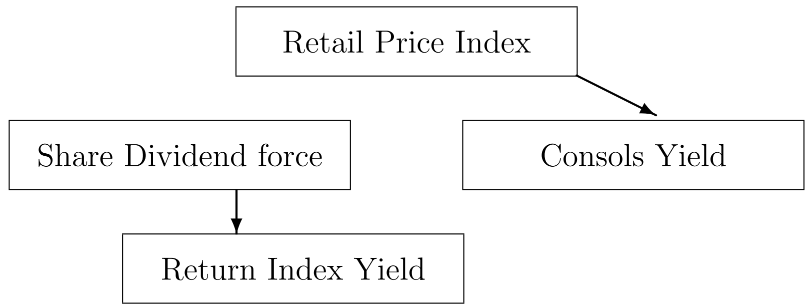
	lagQdivY	lagQdivf	rollmeanCPIf
lagQdivf	0.282		
	0.000		
rollmeanCPIf	0.683	0.029	
	0.000	0.689	
rollstdCPIf	0.318	0.007	0.258
	0.000	0.921	0.000

From the above table we can see that:

- There is close relationship among inflation rate (QCPIyield), dividend yield (QdivY), dividend force (Qdivf) and Consols.
- There is some relationship between dividend force (Qdivf) and TSX yield (QtseY).

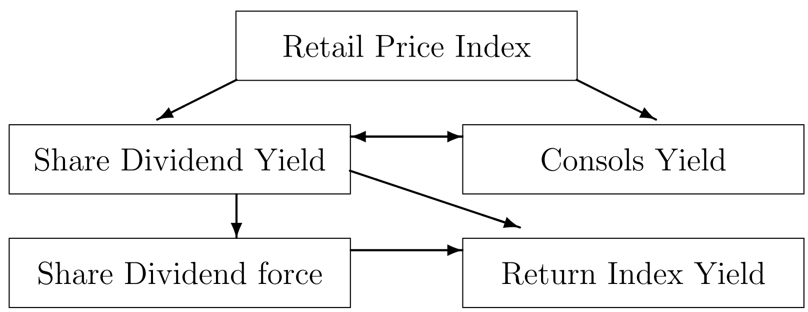
Based on the above results and the judgement to discard dividend yield, we can see the new structure for the model. This structure is further confirmed by stepwise regressions built for each variable.

- if we discard dividend yield from the structure, we get the following structure plot.



Now we can see a really simple structure for the models, but we do not connect all the variables.

- If we restore the dividend yield in the structure, we get another structure.



We can see that the dividend yield is really a complicated variable which has relationships with all other variables. Therefore, it is not easy to predict dividend yield. Also, there should not be both dividend yield and share price in the structure if we keep the dividend, because we can easily get one of them from the other.

With regard to stepwise regression, we came up with linear relationships for four dependent variables: inflation rate (QCPIyield), Consols yield, dividend growth rate (Qdivf), and TSE yield (QtseY). The predictors are all the other variables. There are some more explanation as follows:

- The predictor selection is guided by the structure.

- In modeling, except the inflation rate which is modeled based on all the other variables, we model current variables based on the lagged values of all variables and the current value of the inflation rate.

- Inflation rate is modeled based on the following predictors: Consols, QtseYield, QdivY, Qdivf, lagConsols, lagTseYield, lagCPIyield, lagQdivY, lagQdivf, lagRollmean, lagRollstd. However the final model for inflation rate is an univariate model with some medium memory factors.

- Consols yield is considered for two conditions. In one case there is no information about current TSX index yield. In this situation, the model is based on: QCPIyield , lagConsols, lagTseYield, lagCPIyield, lagQdivf, lagRollmean, lagRollstd, lagQdivY, rollmeanCPIf, rollstdCPIf . In the other situation there is information about current TSX index yield: QtseYield, Qdivf,and QdivY .

- Dividend yield is based on the information of QCPIyield, Consols, lagConsols, lagTseYield ,lagCPIyield, lagQdivf , lagRollmean, lagRollstd, lagQdivY, rollmeanCPIf, rollstdCPIf.

- Dividend growth rate is based on the same information as that of dividend yield.

- TSX Return Index yield is not necessary if we know both current dividend and current dividend yield. Therefore, we drop of the current information of dividend yield when modeling TSX Return Index yield.

The regression equations selected based on the minimum value of Mallows C–p are as follows:

Inflation rate:

$$\begin{aligned} \text{QCPIyield} &= 0.003084 + 0.8 * \text{lagRollmean} - 0.43 * \text{lagRollstd} \\ &\quad +0.127 * \ \text{lagCPIyield} + \epsilon_t \end{aligned}$$

Consols rate:

$$
\begin{aligned}
& 1. \quad \text{no information of QtseYield} \\
\text{Consols} &= 0.002133 + 0.0089 * \text{lagTseYield} + 0.945 * \text{lagConsols} \\
& \quad +0.194 * \text{QCPIyield} + \epsilon_t \\
& 2. \quad \text{given information of QtseYield} \\
\text{Consols} &= 0.002516 + 0.0096 * \text{lagTseYield} + 0.946 * \text{lagConsols} \\
& \quad -0.041 * \text{QtseYield} + 0.184 * \text{QCPIyield} + \epsilon_t
\end{aligned}
$$

Dividend yield:

$$
\begin{aligned}
\text{QdivY} &= 0.0007408 + 0.866 * \text{lagQdivY} - 0.0055 * \text{lagQdivf} \\
& \quad +0.151 * \text{rollmeanCPIf} - 0.00273 * \text{lagTseYield} - 0.108 * \text{lagRollmean} + \epsilon_t
\end{aligned}
$$

Dividend force:

$$
\text{Qdivf} = 0.02494 - 0.718 * \text{lagQdivf} - 0.19 * \text{lagTseYield} + \epsilon_t
$$

TSX yield:

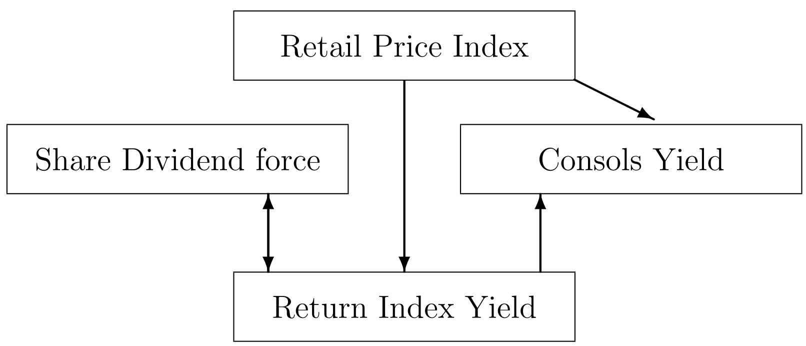
$$
\begin{aligned}
& 1. \quad \text{has information of lagQdivY} \\
\text{QtseYield} &= -0.01572 + 0.321 * \text{Qdivf} + 0.103 * \text{lagQdivf} \\
& \quad +0.197 * \text{lagTseYield} + 10.7 * \text{lagQdivY} - 1.73 * \text{QCPIyield} \\
& \quad -4.3 * \text{lagRollstd} - 1.16 * \text{lagCPIyield} + \epsilon_t \\
& 2. \quad \text{no information of lagQdivY} \\
\text{QtseYield} &= 0.02603 + 0.311 * \text{Qdivf} + 0.139 * \text{lagQdivf} \\
& \quad +0.181 * \text{lagTseYield} - 1.08 * \text{QCPIyield} + \epsilon_t
\end{aligned}
$$

Our comments of the relationship:

- The inflation rate has positive relationship with its rollover mean level and negative relationship with its rollover variance level.

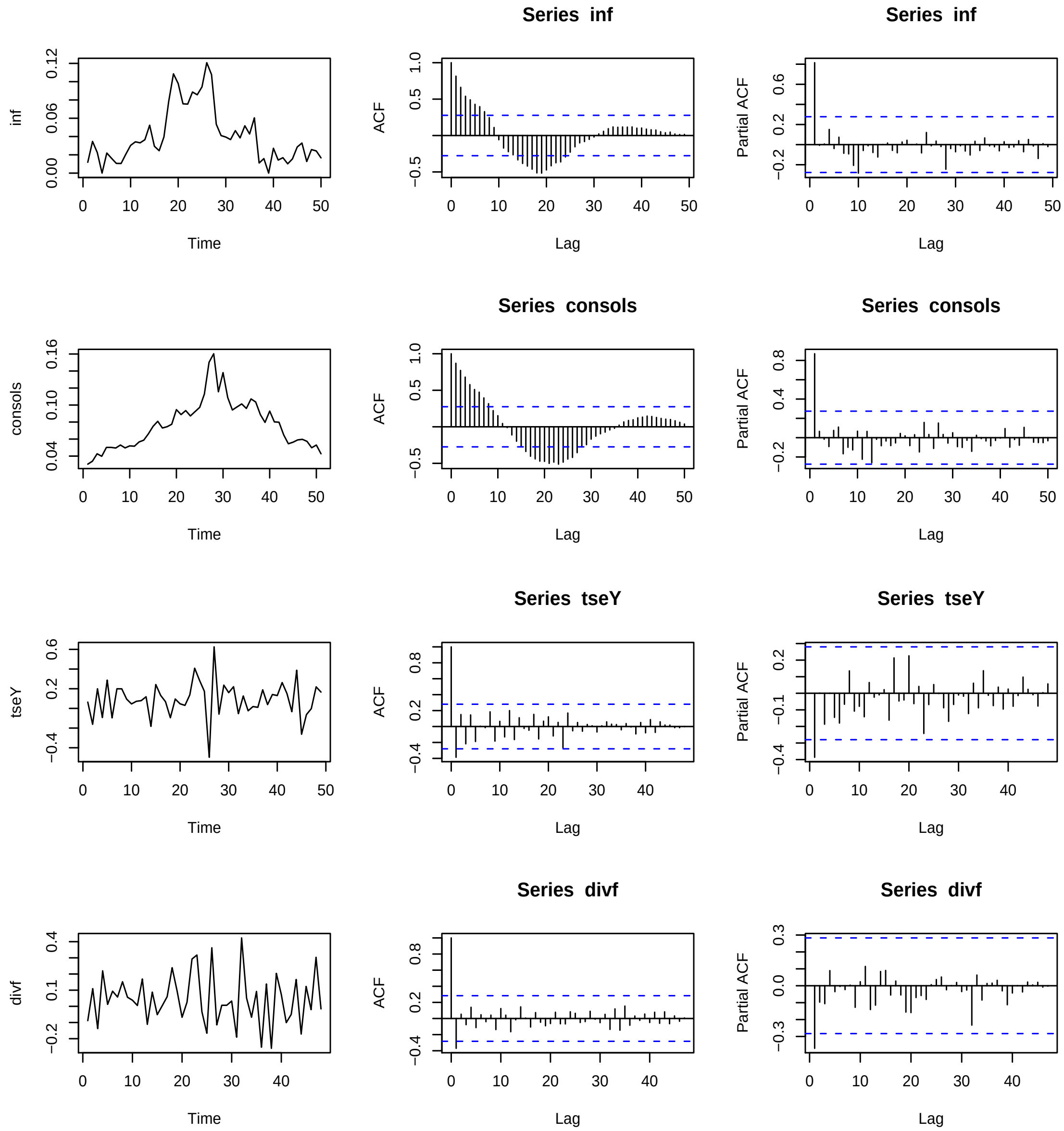
- Consols rate has a positive relationship with previous TSX yield and inflation rate, and has a negative relationship with current TSX yield.
- Dividend yield has a relationship with many other factors.
- Dividend growth force has a negative relationship with its lagged values due to its seasonal behavior. It also has a negative relationship with previous TSX yield.
- TSX Return Index yield has interaction with inflation rate and dividend force. The relationship with inflation rate is negative.

Base on the linear regression relationship, we now can simplify the structure as:



4.2 Analysis of Residuals of Univariate ARIMA Models

Because quarterly data processes need to consider seasonal effect, I use yearly data to simplify the modeling. The following are the plots of the four basic components in the structure.



From the plots we can see that the time series structure shows the pattern of AR(1) for inflation rate and Consols yield, MA(1) pattern for dividend growth rate. Those results are consistent with the Wilkie Model. Also we can see AR(1) pattern for TSX yield series. All this model are significant for each series and relevant residuals show white noise pattern. The models are as follows:

<u>Inflation rate:</u>

$$\nabla \log Q(t) = 0.0364 + 0.8242 * \nabla \log Q(t-1) + 0.017 * \epsilon_t$$

Consols:

$$C(t) = 0.0644 + 0.92 * C(t-1) + 0.012 * \epsilon_t$$

Dividend growth force:

$$\nabla \log D(t) = 0.0903 - 0.38 * \nabla \log D(t-1) + 0.165 * \epsilon_t$$

TSX yield:

$$R(t) = 0.0449 - 4.012 * \epsilon_{t-1} + 0.14 * \epsilon_t$$

we can see that:

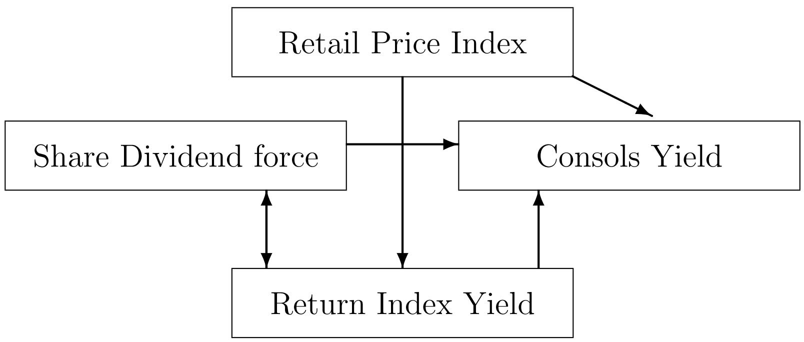
- Inflation rate and Consols yield have high value of AR coefficient, which make their behavior like random walks and create larger waves in their historical performance.
- TSX yield shows the mean reverting character with the negative AR coefficient.
- Dividend growth rate has some correlation in variance, however there is no significant ARCH effects.

After taking the univariate time series model, let us take a look at the residuals from each models. We already know that all residuals show the property of white noise. The results of their correlation are as follows:

```
 Correlations:
                  resi-Inf  resi-Consols     resi-Divf
resi-Consols         0.417
                     0.003

resi-Divf           -0.092         0.042
                     0.532         0.775

resi-TseY            0.092        -0.152         0.482
                     0.529         0.298         0.001
```

From the results we can see the correlation between inflation rate residuals and Consols rate residuals, and the correlation between TSX yield residuals and dividend force residuals. This correlation structure of residuals is exactly the same as that one based on original values.

Through the stepwise regression of residuals based on the minimum values of Mallows C–p, we get the correlation map as follows:



This structure of time series model's residuals is very close to that of original values we get based on in section 4.1. The only difference is that dividend force has influence on Consols rate. This difference does not change the flow of calculation for dividend force and the calculation for Consols rate.

4.3 VAR Analysis of The Structure

VAR analysis is an extension of AR time process into vector variables. In order to compare the results using other methods, we used yearly data to compute the relationship structure of those components. The VAR regression equations are as follows:

$$Y_t = c + \prod_1 Y_{t-1} + \cdots + \Phi X_t + G * XX_t + \epsilon_t$$

where: Y_t : represents response time series variables

X_t : represents the trend matrix, which is consist of deterministic components.

XX_t : represents the matrix of exogenous variables.

$\prod_1, \Phi, G$: represent the parameter matrices.

Even though this method is based on matrix calculation, it is fine to calculate the regression equations separately for each of the variables. The best regression for each variable is based on the minimum Mallows C–p values. After getting residuals from each regression model, we can do some correction for those models based on the correlation matrix for residuals. Thus we change the VAR model into the structured VAR model.

The variables used here are the same as what we did in section 4.1. So the response variables are TSX yield (YtseYield), dividend growth force (Ydivf), Consols rate (Consols), and inflation rate (CPIforce). The predictors are the lagged values of those variables, rollover mean, and rollover variance of inflation rate. The difference between VAR and the method in section 4.1 is that we use yearly data in this section instead of quarterly data.

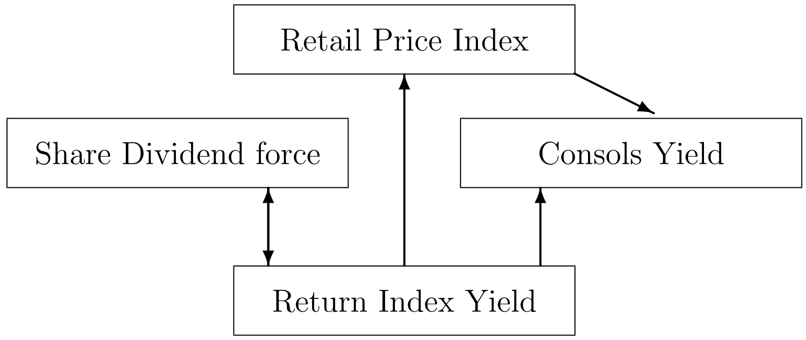
```
YtseYield = 0.126 - 0.388 lagTseY
    S = 0.170545   R-Sq = 15.0%   R-Sq(adj) = 13.1%

Ydivf = 0.0733 - 0.328 lagTseY
    S = 0.144304   R-Sq = 14.9%   R-Sq(adj) = 13.1%

Consols = 0.00737 + 0.222 lagCPIforce + 0.749 lagConsols + 0.0381*lagTseY
    S = 0.00924065   R-Sq = 89.7%   R-Sq(adj) = 89.0%

CPIforce = 0.00357 + 0.847 lagCPIforce + 0.0273 lagTseY
   S = 0.0167509   R-Sq = 71.8%   R-Sq(adj) = 70.5%

Correlations: Cell Contents: Pearson correlation
               P-Value

                 RESI1_tsx    RESI1_divf  RESI1_consol
```

```
RESI1_divf          0.493
                    0.000

RESI1_consol       -0.125        0.008
                    0.396        0.955

RESI1_CPI           0.066       -0.062        0.330
                    0.658        0.674        0.022
```

Because the correlation between dividend growth force residuals and TSX yield residuals are significant with the value of ρ being 0.493. Therefore additional term could be add up to the equation of YtseYield if we predict YtseYield based on dividend force.

$$
\begin{aligned}
S_{\text{div force}} &= 0.144304 \\
S_{\text{TSX yield}} &= 0.170545 \\
\epsilon_{TSXyield} &= 0.493 * \frac{0.170545}{0.144304} * \epsilon_{\text{dividend force}} + \epsilon_t \\
\text{where:} & \quad \epsilon_t \text{ and } \epsilon_{\text{dividend force}} \text{ are independent noise} \\
\text{Therefore:} & \\
\text{YtseYield} &= 0.126 - 0.388\text{lagTseY} + 0.493 * \frac{0.170545}{0.144304} * \epsilon_{\text{dividend force}} + \epsilon_t \\
\text{Since:} & \\
\epsilon_{\text{dividend force}} &= \text{Ydivf} - 0.0733 + 0.328\text{lagTseY} \\
\text{Therefore:} & \\
\text{YtseYield} &= 0.126 - 0.388\text{lagTseY} \\
& \quad +0.493 * \frac{0.170545}{0.144304} * (\text{Ydivf} - 0.0733 + 0.328\text{lagTseY}) + \epsilon_t \\
&= -0.083292 + 0.58265 * \text{Ydivf} - 0.19689 * \text{lagTseY} + \epsilon_t
\end{aligned}
$$

Base on the VAR models, we now can simplify the structure as:

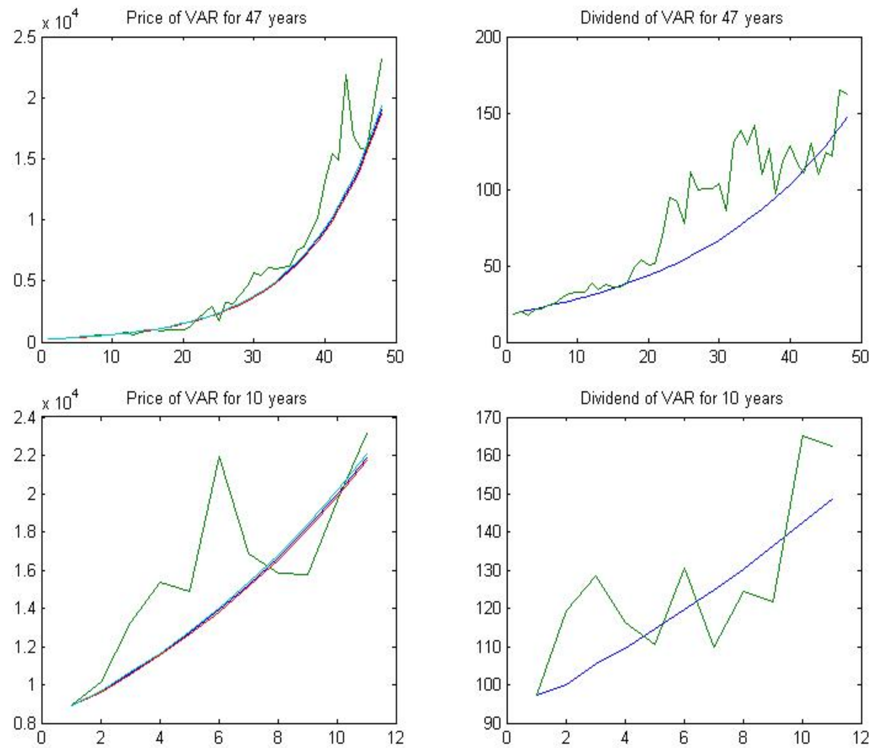


Through the Comparison of the the above plot and the structure in section 4.1, we can see that the yearly data and the quarterly data show the very similar results. This could be treated as the final structure of ours for these four variables: inflation rate, dividend, Consols rate, and TSX total return index yield.

Based on what we get from the VAR models, we did the simulation of 10 years and 47 years. The simulation procedure is as follows:

- Creating correlated random numbers for residuals. In this step, we use the Cholesky decomposition (denoted as V) of the covariance matrix $\sum$. Then for the random multi–variate normally distributed numbers (denoted as N) with identity matrix of covariance I, the new random numbers VN are multi-variate normally distributed with covariance matrix $\sum$.
- Using loops to create TSX values from VAR models
- Repeating this procedure for 10000 times

The plots of prediction are listed as below. The top two are the predictions for 47 years, 1959–2005, with the initial value of 1958, and the bottom two are the predictions for 10 years, 1996–2005, with the initial value of 1995. The zigzag curves are real values.



The relevant statistics from the simulation of year TSX yield are listed in the following table.

Statistcs	Values	Statistcs	Values
Mean	0.0898	Minimum value	-0.7535
St. dev.	0.1972	2.5 percentile	-0.2984
Skewness	-0.0252	5 percentile	-0.2351
Kurtosis	3.0202	10 percentile	-0.1622

As a comparison, we also list the results from the simulation of TSX year yield using RSLN2 model as follows:

Statistcs	Values	Statistcs	Values
Mean	0.0906	Minimum value	-1.2167
St. dev.	0.1780	2.5 percentile	-0.2660
Skewness	0.0819	5 percentile	-0.1762
Kurtosis	6.9630	10 percentile	-0.1033

From the above plots and table information, we can see that VAR models predict quite well for stock price and dividend over 10 years period and 47 years based on multiple variables. The variance of the prediction of the yearly TSX yield is greater than RLSN2 model, if we assume the independent of quarterly data. However, the its variance is less than that of the Wilkie Model. The tail of the distribution for TSX yield is not necessarily thicker than that of RSLN2 models.

Chapter 5

Conclusions

From the analysis presented, we draw the following conclusions:

- Based on the Canadian stock market, the Wilkie Model overestimates the inflation for recent years, overestimates the dividend yield for recent years, and closely predicts the dividend. As a result, the Wilkie Model underestimates the share price. Thus the Wilkie Model has too thick lower tails in its simulation.

- Univariate models could predict trend quite well, while the multivariate models could produce larger variance. Both VAR models and the Wilkie Model create larger variance for TSX yield than any of the univariate models. This is due to the aggregated variance of multiple variables.

- Neither dividend yield nor dividend is suitable for predicting TSX yield. Dividend yield is not very stationary, and is less stationary than share price in the long term, thus it is hard to predict. Dividends are much more stable than TSX yield, thus they have different behaviors. Also dividend yield has decreased to a low level, while capital gain rate increased a lot. This makes it poor for dividend yield to predict share price in a risky market.

- The monthly data and the quarterly data could have different results for modeling, even based on the same method. For using univariate ARIMA–GARCH models to model TSX yield, monthly data has ARMA(1,1) model to fit. However in the quarterly data, we can find long term pattern easier. There are a couple of different models to fit the quarterly data, and we choose ARIMA(0,0,0) x $(0,0,1)^5$. Quarterly data and yearly data have close results for prediction, except that quarterly data may have seasonal effects. In general, the continuous assumption and discrete assumption are suitable for monthly data and quarterly data respectively.

- Even a historically stationary model may not be predictable for future period, because of the non–ergodic problem. This will be seen in the future prediction of dividend yield, which is modeled using ARMA models with changing parameters.

There are still other methods for modeling TSX total return yield. For example, Gilles and LeRoys (1992) suggested predicting TSX price by adding a bubbles term onto dividend. These are still problems left for future work:

- Specifying any measurement for the degree of stationarity of time series for different periods, and testing the ergodicity of the time series.
- Improving RSLN models with smoothed regimes and for multiple variables, and providing better methods for testing RSLN effects
- Updating models by taking more consideration of realities as constraints input of the models.
- Modeling more specific information for the prediction, such as the movement patterns of those indices through different time points.

Appendix A

Program Code

1. the R code of "test of heteroscedasticity"

```
#--test of heteroscedasticity for monthly data 56-05

    data <- read.csv('cointegration2.csv',header = TRUE)

    TSEreturn <- data$TSEreturn length(TSEreturn) temp = embed(TSEreturn,2)

    rate = temp[,1]/temp[,2]  #--data of 56-05

    rate2test1 = rate^2

    Modeltest1 = arima(rate2test1, c(0,0,1))#--fit garch(0,1) for 56-05

    ttemp = c(1:100)

    for (i in 1:100){

    ttemp[i] = Box.test(rate2test1, lag=i, type = "Ljung-Box")$p.value

    }

    par(mfrow = c(2,2))

    plot(ttemp, main = "Ljung-Box ARCH test 56-05",pch=20, ylim = c(0,1))

    lines( matrix(0.05,100,1)  )

    text(40,0.15,"5% line")
```

2.. the Matlab code of simulation for TSX index yield based on Wilkie's model

```
function Thesis_simulation2_Wilkie_Y(n)

% this simulation use parameters estimated from Yearly data of Wilkie model
%-----parameters from CA 56-05 in Excel
QSD =0.016788021; QMU = 0.038636152; QA =0.821501371;

YMU = 0.038402052; YA = 0.966969169; YW =-2.616794131; YSD
=0.138155021;

DW =0.554153424 ; DD =0  ; DMU =0.007470869; DY = 0.353427584; DB =
-0.99; DSD = 0.119019924;

%-----------------------------------------
m=10; t = m; QZ = randn(t+1,n);
DlnQ = zeros(t+1,n);% delta ln(Q) values
Q = zeros(t+1,n);

%----produce delta ln(Q) values: DlnQ
temp = ones(1,n); for i = 1:t
    DlnQ(i+1,:)= QMU + QA * ( DlnQ(i,:) - QMU ) + QSD * QZ(i+1,:);
    temp = exp(DlnQ(i+1,:));
    Q(i+1,:) = Q(i,:) .* temp ;
end

%---------------------------------------model for dividend yield
DlnQ(1,:) = 0.02718614;%---initial value from Canadian 1995

%----modelling initial values for yield Y(0) and YN(0)
infForce = DlnQ; Y = ones(m+1,n); YN = ones(m+1,n);
Y(1,:) = 0.023955937;%---intial values of Y(0), using real Canadian value for year

YN(1,:) = log(Y(1,:))-YW*infForce(1,:);

%--loops for YN(t) from t=1:m
YE = YSD * randn(m,n); for t = 1:m
    YN(t+1,:)= log(YMU) + YA * (YN(t,:)- log(YMU))+ YE(t,:);
end Y(2:m+1,:)=exp(YW*infForce(2:m+1,:)+YN(2:m+1,:));

%--------------------------------------------------Dividend Model

%---set initial values
D = ones(m+1,n);
```

```
D(1,:)= 97.40454826;%---initial value from Canadian 1995
DM = ones(m+1,n);
%infForce = inflation(m,n);

DM(1,:) = infForce(1,:); YE2 = [zeros(1,n);YE]; YE2 = YE(1:m,:);

%---modelling
for t=1:m
    DM(t+1,:) = DD*infForce( t+1, :) + (1-DD) * DM(t,:);
end

infForce2 = mean(infForce')';%---for testing of mean at last

DM = DM(2:m+1,:); infForce = infForce(2:m+1,:); DE1 = DSD *
randn(m+1,n);
DE2 = DE1(1:m,:);%--DE(t-1)
DE = DE1(2:m+1,:);%--DE(t)

Dforce = ones(m,n); Dforce = DW * DM + (1-DW) * infForce + DY * YE2
+ DMU + DE + DB*DE2;

Df2 = Dforce; for i = 2: m
    Df2(i,:) = Df2(i-1,:)+Dforce(i,:);
end D2 = D(1,1)*exp(Df2);

%-----------------after get T(t) for mxn, m=10
D3 = [D(1,:);D2];
T = D3./Y; %---price index
T0 = T + D3; T1 = T0(2:m+1,:); T2 = T(1:m,:);
rate = T1./T2 -1; %---price index yield

meanRateWhole = mean(mean(rate'))
sedevRateWhole = (mean(var(rate')) + var(mean(rate')))^0.5 %--var

meanRate = mean(rate(m,:));
sedevRate = var(rate(m,:))^0.5 ;%--var of the last period only
skewRate = sum((rate(m,:) - meanRate).^3) /((10000-1)*sedevRate^3)
KurtosisValue = sum((rate(m,:) - meanRate).^4)
/((10000-1)*sedevRate^4)

temp = rate(1,:);
for i =1:10  %-----------------------------------change for yearly/quarterly
    temp = [temp,rate(i,:)];
end minRate = min(temp) percentile2_5R0 = prctile(temp,2.5)
```

```
percentile5R0 =prctile(temp,5) percentile10R0 = prctile(temp,10)

%-- values of cumulative yield
cumRate = T(m+1,:)./T(1,:)-1; meanCumRate = mean(cumRate);
sedevCumRate = var(cumRate)^0.5; minCumRate = min(cumRate)
percentile2_5C = prctile(cumRate,2.5) percentile5C
=prctile(cumRate,5) percentile10C = prctile(cumRate,10)

crash = 0; for i = 1:10000
    if min( rate(:,i)) <= -0.2552, crash = crash+1; end;
end crash probCrash = crash/10000

subplot(2,2,1), hist(temp),title('hist of rates for US-disc')
subplot(2,2,2), hist(rate(m,:)),title('hist of cum rates for
US-disc')
```

Bibliography

[1] Bollen, N.P.B. (1998). "Valuing Options in Regime Switching Models", *Journal of Derivatives*, **6**: 38–49.

[2] Bollerslev, T. (1986). "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, **31**: 307–327.

[3] Chan, W.S. and Wang, S. (1998). "The Wilkie Model For Retail Price Inflation Revisited", *British Actuarial Journal*, Vol 4, III, 637–652.

[4] Engle, R.F. (1982). "Autoregressive Conditional Heteroscedasticity with Estimates of The Variance of United Kingdom Infaltions", *Econometrica*, **50**: 987–1007.

[5] Gilles, C. and LeRoy, S. F. (1992). "Bubbles and Charges", *International Economic Review*, **33**: 323–329

[6] Hamilton, J.D. (1989). "A New Approach to The Economic Analysis of Non–Stationary Time Series", *Econometrica*, **57**: 357–384.

[7] Huber, Paul (1995). "A Review of Wilkie's Stochastic Investment Model", Presentation to the Staple Inn Actuarial Society, March 14th.

[8] Hardy, M.R. (2003). *Investment Guarantees: Modelling and Risk Management for Equity-Linked Life Insurance*, Wiley (New York).

[9] LeRoy, Stephen F. (1996). *Handbook of Statistics 14: Statistical Methods in Finance, edited by G.S¿ Maddala and C.R. Rao*, chapter 6, Stock Price Volatility.

[10] Mulvey, John M. and Thorlacius, A Eric (1996). "The Towers Perrin Global Capital Market Scenario Generation System", *CAP–Link–Proceeding AFIR, Nürnberg, Germany.*

[11] Nelson, D.B. (1991). “Conditional Heteroskedasticity in Asset Returns: A New Approach”, *Econometrica*, **59**: 347–370.

[12] Panjer, Harry H. (1998). *Financial Economics* , Oxford University Press, chapter 9, Investment Return Models.

[13] Rambaruth, Ganeshwarsing (2003). “A Comparison of Wilkie-Type Stochastics Investment Models”, Master Thesis, The City Univeristy, Faculty of Actuarial Science and Statistics.

[14] Sharp, Keith P. (1992). “Modeling Canadian Price and Wage Inflation”, University of Waterloo, Department of Statistics and Actuarial Scicence.

[15] Tomson, R. J (1996). “Stochastic Investment Modeling: the Case of South Africa”, *British Actuarial Journal*, Vol 2, 765–801.

[16] Tsay, Ruey S. (1987). “Conditional Heteroscedastic Time Series Models”, *Journal of The American Statistical Association*, **82**: 590–604.

[17] Tsay, Ruey S. (2005). *Analysis of Financial Time Series*, 2nd edition.

[18] Whitten, S.P. and Thomos, R.G. (1999). “A Non-linear Scochastic Model for Actuarial Use”, *British Actuarial Journal*, Vol 5, 919–953.

[19] Wilkie, A.D. (1986). “A Stochastic Investment Model for Actuarial Use”, *Transactions of the Faculty of Actuaries*, Vol 39, 341–403.

[20] Wilkie, A.D. (1992). “Stochastic Investment Models for XXIst Century Actuaries”, *Transactions of the 24th International Congress of Actuaries: Montreal, Canada, May 31 to June 5 1992*, Vol 5, 119–137.

[21] Wilkie, A.D. (1995). “More on a Stochastic Asset Model for Actuarial Use”, *British Actuarial Journal*, Vol 1, 777-945.

[22] Wong, Albert C.S. and Chan, Wai Sum (2005). "Mixture Gaussian Time Series Modeling of Long–Term Market Returns", *NAAJ*, October.